



UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA
DEPARTAMENTO DE ELECTRÓNICA

Rango de matrices de covarianzas no asintóticas y la matriz de Hankel, para estimar el orden de un sistema

Tesis de grado presentada por
Cristian Fernando Guajardo Yévenes

como requisito parcial para optar al título de
Ingeniero Civil Electrónico
mención Control Automático e Instrumentación
y al grado de
Magíster en Ingeniería Electrónica

Profesor guía
Dr. Ricardo Rojas R.

Diciembre 2008

TÍTULO DE LA TESIS:

Rango de matrices de covarianzas no asintóticas y la matriz de Hankel, para estimar el orden de un sistema

AUTOR:

Cristian Fernando Guajardo Yévenes

TRABAJO DE TESIS, presentado en cumplimiento parcial de los requisitos para el título de Ingeniero Civil Electrónico y el grado de Magister en Ingeniería Electrónica de la Universidad Técnica Federico Santa María.

Dr. Ricardo Rojas R.

Dr. Juan Hernández S.

Dra. Ingeborg Mahla A.

Valparaíso, Chile. Diciembre 2008.

Resumen

EL modelado de fenómenos físicos es un problema recurrente en el campo de la ingeniería, el que se aborda principalmente mediante el ajuste de un conjunto de mediciones a una estructura de modelos seleccionada con anterioridad, estimándose de esta manera sus coeficientes o parámetros. Ejemplos de estas estrategias de modelado y ajuste son los métodos clásicos de *cuadrados mínimos* y *variables instrumentales*, los que precisan del conocimiento previo de una estructura adecuada, para obtener modelos consistentes.

Recientemente, se han desarrollado diferentes enfoques que abordan el tema de identificación estructural, en su mayoría basados en matrices de covarianzas, que buscan *exhaustivamente* la estructura correcta del sistema entre un conjunto de modelos candidatos. Si se considera un entorno ruidoso se aplican de antemano algunas técnicas de reducción de ruido: variables instrumentales, reescalamiento de la entrada y la salida según la potencia del ruido, y técnicas de ajustes de datos.

Con el objeto de optimizar los métodos anteriores, en esta tesis se mejora el marco de trabajo para identificación estructural basada en matrices de covarianza, especificándose claramente condiciones suficientes para la señal de entrada, de manera que el rango de las diferentes matrices de covarianza sea conocido en forma exacta, permitiendo así desarrollar métodos *no exhaustivos* de identificación estructural. Además, en el caso de existir ruido de medición, se desarrolla un criterio que permite estimar la estructura del sistema, sin necesidad de conocer la distribución probabilística del ruido, mediante el uso de variables instrumentales. Finalmente se establecen relaciones entre estas matrices de covarianzas y las matrices de *Hankel del sistema* y de *Yule-Walker*, donde estas últimas heredan las propiedades de rango de las matrices de covarianza.

Palabras clave: identificación de sistemas, identificación estructural, matrices de correlación, matrices de covarianza, persistencia de excitación, suficiencia de información, matriz de Yule-Walker, matriz de Hankel, rango efectivo, valores singulares, valores singulares de Hankel.

Abstract

MODELING physical phenomena is a common problem in engineering, which is normally accomplished by fitting measured data to a previously selected model structure, through parameter estimation. Examples of these modeling and fitting techniques are the well-known methods of *least squares* and *instrumental variables*, which depend on the proper choice of the model structure, to obtain consistent models.

Recently, several approaches for structure estimation have been developed, mostly based on covariance matrices, which search exhaustively for the system structure among the possible models. When considering a noisy environment, some noise reduction techniques are applied in advance, for example: Instrumental variables, re-scaling the input by an input-output power ratio and data fitting techniques.

In order to optimize the previously mentioned methods, this thesis improves the framework for structural identification, based on covariance matrices, by establishing sufficient conditions for the input signal, so that the rank of covariance matrices can be exactly known, allowing the development of *non-exhaustive* methods for structural identification. Besides, a criterion for estimating the system structure is developed, when considering a noisy environment, even if the probability distribution of the output noise is unknown. Finally, relations between covariance matrices, the *system Hankel matrix* and the *Yule-Walker matrix* are obtained, where the last two inherit the rank properties of the covariance matrices.

Keywords: system identification, model structure identification, correlation matrices, covariance matrices, persistency of excitation, informative data, Yule-Walker matrix, Hankel matrix, effective rank, singular values, Hankel singular values.

*Al la memoro de miaj avoj
Fernando kaj Pedro*

Agradecimientos

ESCRIBIR estas líneas significa, la verdad, mucho para mí. Es, en primer lugar, el final de un paciente andar, por momentos amargo, que ha culminado con dulces frutos. Es también, el paso inicial hacia un futuro, en el que tengo puesta toda mi fe y espíritu, y que realmente anhelaba dar.

Mucha gente ha esperado junto conmigo este momento, entregándome su cariño, su fuerza, brindándome su apoyo y consejos. Cada una de estas personas, ha contribuído a formar la personalidad, carácter y temple que definen mi esencia. Es a ellos a quienes quiero agradecer, de todo corazón, en estas sencillas líneas.

En primer lugar, a mi familia, en especial a mis padres Fernando y Elizabeth, y mis hermanas Natalia y Lorena, quienes me han guiado y acompañado toda la vida, y a quienes quiero agradecer especialmente por su apoyo y comprensión incondicional.

A mi amigo de toda la vida Felipe C. y a mis amigos y compañeros del magíster Juan Carlos y Elena, cuyo apoyo ha sido fundamental durante estos dos últimos años. Quiero agradecer especialmente, a mis amigos Jaime, Juan, Luis, Felipe S., Felipe P., Tamara, Fernanda, Josey, Pablo, Francisco, Cristian, Daniel y Víctor, quienes realmente han sido mi familia durante estos años de universidad.

De igual forma agradezco a tres personas que han pasado a formar parte importante de mi vida, a Gerald, Rita y Parin, a quienes también considero parte de mi familia.

Finalmente, a los profesores del departamento de electrónica, y en especial a Manuel Olivares, Jaime Glaría, Ricardo Rojas, Mario Salgado y Juan Hernández, por su cercanía y entrega no sólo en lo referente a lo académico.

Al vi, korajn dankojn

Valparaíso, Chile
Diciembre de 2008

Índice general

Resumen	I
Abstract	III
Agradecimientos	VII
Introducción	XIX
1. Preliminares	1
1.1. Probabilidades y procesos estocásticos	1
1.1.1. Definiciones y conceptos	1
1.1.2. Procesos estocásticos	2
1.1.3. Propiedades de correlación	3
1.2. Predictores	3
1.2.1. Caso sin ruido	3
1.2.2. Caso con ruido	4
1.3. Señales eficaces para identificación	5
1.3.1. Excitación local y excitación persistente	5
1.3.2. Suficiencia de información	8
1.4. Comentarios del capítulo	14
2. Matrices de correlaciones para identificación estructural	15
2.1. Matriz de producto	16
2.1.1. Construcción	16

2.1.2. Búsqueda secuencial	17
2.2. Matriz de producto instrumental	18
2.2.1. Construcción	18
2.2.2. Indicaciones para el diseño del filtro instrumental	21
2.2.3. Búsqueda secuencial	23
2.3. Matriz generalizada de producto instrumental	25
2.3.1. Construcción	25
2.3.2. Indicaciones para el diseño del filtro instrumental	28
2.3.3. Búsqueda secuencial	29
2.4. Rango genérico para la matriz GIPM	34
2.5. Comentarios del capítulo	38
3. Perfiles de valores singulares para matrices de correlación	39
3.1. Caso sin ruido	39
3.1.1. Construcción del método	39
3.1.2. Determinación del rango	41
3.2. Caso con ruido	44
3.2.1. Construcción del método	44
3.2.2. Determinación del rango	44
3.3. Comentarios del capítulo	52
4. Relación con la matriz de Yule-Walker y la matriz de Hankel	53
4.1. Propiedades de valores singulares y valores propios	53
4.1.1. Caso de la matriz GPM	53
4.1.2. Caso de la matriz GIPM	53
4.2. Método de perfiles para la submatriz GIPM	55
4.3. Relación con la matriz de Yule-Walker	59
4.4. Relación con la matriz de Hankel del sistema	65
4.5. Comentarios del capítulo	67

<i>ÍNDICE GENERAL</i>	XI
5. Conclusiones generales	69
A. Notación	71
B. Propiedades de convolución y correlación	73
B.1. Convolución	73
B.2. Correlación	73
B.3. Correlación en sistemas	74
B.3.1. Caso estacionario	74
B.3.2. Caso no estacionario	75
B.3.3. Caso de ruido blanco	75
B.4. Algunas correlaciones de señales conocidas	75
C. Ergodicidad	77
C.1. Media ergódica	77
C.2. Correlación ergódica	79
C.3. Caso de ruido blanco	81
D. Matrices	83
D.1. Descomposiciones	83
D.2. Valores propios y valores singulares	84
D.3. Teoría de perturbaciones	85
D.4. Propiedades de rango	87
D.5. Matrices especiales	88
E. Demostraciones complementarias	89
E.1. Capítulo 1	89
E.2. Capítulo 2	91
F. Programas	93
F.1. Funciones comunes	93
F.1.1. <code>funciones.sci</code>	93

F.1.2. <code>dsp.sci</code>	94
F.2. Ejemplos capítulo 1	95
F.2.1. Ejemplo 1.2: <code>persis.sce</code>	95
F.2.2. Ejemplo 1.4: <code>sufinf-sr.sce</code>	96
F.2.3. Ejemplo 1.5: <code>sufinf-cr.sce</code>	97
F.3. Ejemplos capítulo 2	98
F.3.1. Ejemplo 2.1: <code>Lxux.sce</code>	98
F.3.2. Ejemplo 2.2: <code>Lyuy.sce</code>	99
F.3.3. Ejemplo 2.3: <code>Lzuy.sce</code>	100
F.3.4. Ejemplo 2.4: <code>Lxux2.sce</code>	102
F.3.5. Ejemplos 2.5 y 2.6: <code>Lzuy2.sce</code>	103
F.4. Ejemplos capítulo 3	105
F.4.1. Ejemplos 3.1 y 3.2: <code>Lxux.sce</code>	105
F.4.2. Ejemplos 3.3, 3.4 y 3.5: <code>Lzuy.sce</code>	106
F.5. Ejemplos capítulo 4	108
F.5.1. Ejemplos 4.1, 4.2, 4.3 y 4.4: <code>Gzx.sce</code>	108
F.5.2. Ejemplos 4.5, 4.6 y 4.7: <code>hsv.sce</code>	109
G. Publicaciones basadas en esta tesis	111

Índice de figuras

1.1.	Relaciones entre los diferentes tipos de convergencia.	1
1.2.	Diagrama total del sistema, considerando ruido coloreado $v(k)$. La actuación es $u(k)$ y la salida medible es $y(k)$	4
1.3.	Señal cuadrada $s(k)$ de periodo 10 y su espectro $ S(e^{j\theta}) $	6
1.4.	Componentes espectrales de la señal $u(k)$, con $p = 9$	11
1.5.	Componentes espectrales de la señal $u(k)$, con $p = 7$	13
2.1.	Componentes espectrales de la señal $u(k)$, con $p = 14$	18
2.2.	Rango de las matrices $\Omega_{uz}^M(\hat{n})$ and $\Omega_{ux}^M(\hat{n})$	21
2.3.	Cota para la correlación muestral $\varphi_{zx}^\infty(\tau)$, donde la exponencial izquierda está asociada a $G(q)$, y la derecha, asociada a $F(q)$	22
2.4.	En gris: región donde $\Lambda_{xux}^M(\hat{n}_b, \hat{n}_a)$ tiene rango completo.	26
2.5.	En gris: región donde $\Omega_{uz}^M(\hat{n}_b, \hat{n}_a)$ tiene rango completo. Achurada: región donde $\Omega_{ux}^M(\hat{n}_b, \hat{n}_a)$ tiene rango completo.	27
2.6.	Cota para la correlación muestral $\varphi_{zx}^\infty(\tau)$, donde la exponencial izquierda está asociada a $G(q)$, y la derecha, asociada a $F(q)$	29
3.1.	Búsqueda de (n_b, n_a) en el caso $n_b < n_a$. En gris: región donde Λ_{xux}^M es de rango completo. En blanco: región donde Λ_{xux}^M es deficiente de rango.	39
3.2.	Búsqueda de (n_b, n_a) en el caso $n_b > n_a$. En gris: región donde Λ_{xux}^M es de rango completo. En blanco: región donde Λ_{xux}^M es deficiente de rango.	40
3.3.	Con “o” los valores singulares de $\Lambda_{xux}^M(\hat{n}_b, \hat{n}_a)$ y en rojo su ϵ_σ . Con “x” los valores singulares de $\tilde{\Omega}_{ux}^M(\hat{n}_b, \hat{n}_a)$ y en azul su ϵ_σ	42
3.4.	Con “o” los valores singulares de $\Lambda_{xux}^M(\hat{n}_b, \hat{n}_a)$ y en rojo su ϵ_σ . Con “x” los valores singulares de $\tilde{\Omega}_{ux}^M(\hat{n}_b, \hat{n}_a)$ y en azul su ϵ_σ	43

- 3.5. Con “o” los valores singulares de $\Lambda_{zuy}^M(\hat{n}_b, \hat{n}_a)$, con “x” los valores singulares de $\Lambda_{zux}^M(\hat{n}_b, \hat{n}_a)$ y en rojo ε_σ . Cada gráfico muestra 50 realizaciones con $M = 3 \cdot 10^3$ y SNR de salida $\approx 20[dB]$. 49
- 3.6. Con “o” los valores singulares de $\Lambda_{zuy}^M(\hat{n}_b, \hat{n}_a)$, con “x” los valores singulares de $\Lambda_{zux}^M(\hat{n}_b, \hat{n}_a)$ y en rojo ε_σ . Cada gráfico muestra 50 realizaciones con $M = 3 \cdot 10^3$ y SNR de salida $\approx 10[dB]$. 50
- 3.7. Con “o” los valores singulares de $\Lambda_{zuy}^M(\hat{n}_b, \hat{n}_a)$, con “x” los valores singulares de $\Lambda_{zux}^M(\hat{n}_b, \hat{n}_a)$ y en rojo ε_σ . Cada gráfico muestra 50 realizaciones con $M = 3 \cdot 10^3$ y SNR de salida $\approx 9[dB]$. 51
- 4.1. Búsqueda de $(n_b, n_a) = (5, 3)$. Cada gráfico muestra 50 realizaciones de los valores singulares de $\Gamma_{zx}^M(\hat{n}_a)$ con $M = 10^3$ 57
- 4.2. Búsqueda de $(n_b, n_a) = (5, 3)$. Cada gráfico muestra 50 realizaciones de los valores singulares de $\Gamma_{zx}^M(\hat{n}_a)$ con $M = 10^5$ 57
- 4.3. Búsqueda de $(n_b, n_a) = (4, 6)$. Cada gráfico muestra 50 realizaciones de los valores singulares de $\Gamma_{zx}^M(\hat{n}_a)$ con $M = 10^3$ 58
- 4.4. Búsqueda de $(n_b, n_a) = (4, 6)$. Cada gráfico muestra 50 realizaciones de los valores singulares de $\Gamma_{zx}^M(\hat{n}_a)$ con $M = 10^5$ 58
- 4.5. Búsqueda de $(n_b - d_g, n_a)$ en el caso $n_b - d_g < n_a$. En gris: región donde Γ_{zx}^∞ es de rango completo. En blanco: región donde Γ_{zx}^∞ es deficiente de rango. 60
- 4.6. Búsqueda de $(n_b - d_g, n_a)$ en el caso $n_b - d_g > n_a$. En gris: región donde Γ_{zx}^∞ es de rango completo. En blanco: región donde Γ_{zx}^∞ es deficiente de rango. 60
- 4.7. Búsqueda de $(n_b - d_g, n_a) = (0, 3)$ caso ARMA. Cada gráfico muestra 50 realizaciones de los valores singulares de $\Gamma_{zx}^M(\hat{n}_a)$ con $M = 10^3$ 62
- 4.8. Búsqueda de $(n_b - d_g, n_a) = (0, 3)$ caso ARMA. Cada gráfico muestra 50 realizaciones de los valores singulares de $\Gamma_{zx}^M(\hat{n}_a)$ con $M = 10^5$ 62
- 4.9. Búsqueda de $(n_b - d_g, n_a) = (4, 6)$ caso ARMA. Cada gráfico muestra 50 realizaciones de los valores singulares de $\Gamma_{zx}^M(\hat{n}_a)$ con $M = 10^3$ 64
- 4.10. Búsqueda de $(n_b - d_g, n_a) = (4, 6)$ caso ARMA. Cada gráfico muestra 50 realizaciones de los valores singulares de $\Gamma_{zx}^M(\hat{n}_a)$ con $M = 10^5$ 64

Índice de tablas

1.1. Singularidad de la matriz $\Pi_u^M(p_0)$, para $u(k)$ con $p = 1, \dots, 10$ componentes espectrales. . . .	7
1.2. A la izquierda: En negro las combinaciones en que $\Lambda_{xux}^M(\hat{n}_b, \hat{n}_a)$ es no singular, en azul y rojo las combinaciones en que $L_{ux}^M(\hat{n}_b, \hat{n}_a, i)$ es no singular y singular respectivamente. A la derecha: El menor valor propio de $\Lambda_{xux}^M(\hat{n}_b, \hat{n}_a)$ para $\min(\hat{n}_b - n_b, \hat{n}_a - n_a) \leq 0$ y el menor valor propio de $L_{ux}^M(\hat{n}_b, \hat{n}_a, i)$ para $\min(\hat{n}_b - n_b, \hat{n}_a - n_a) \geq 0$	12
1.3. A la izquierda: Combinaciones en que $\Lambda_{yuy}^\infty(\hat{n}_b, \hat{n}_a)$ es no singular. A la derecha: El menor valor propio de $\Lambda_{yuy}^M(\hat{n}_b, \hat{n}_a)$ con $M = 500$	14
2.1. Singularidad de $\Lambda_{xux}^M(\hat{n})$, para $M = 7$ y $L = 7$	18
2.2. Singularidad de $\Lambda_{yuy}^M(\hat{n})$, para $M = 10^3$ y $\text{SNR} = 5, 7[dB]$	19
2.3. Singularidad de $\Lambda_{yuy}^M(\hat{n})$, para $M = 5 \cdot 10^3$ y $\text{SNR} = 5, 92[dB]$	19
2.4. Singularidad de $\Lambda_{yuy}^M(\hat{n})$, para $M = 10^4$ y $\text{SNR} = 5, 84[dB]$	19
2.5. Singularidad de $\Lambda_{yuy}^M(\hat{n})$, para $M = 5 \cdot 10^4$ y $\text{SNR} = 5, 78[dB]$	19
2.6. Singularidad de $\Lambda_{zuy}^M(\hat{n})$, para $M = 10^3$, $\text{SNR} = 5, 70[dB]$	24
2.7. Singularidad de $\Lambda_{zuy}^M(\hat{n})$, para $M = 5 \cdot 10^3$, $\text{SNR} = 5, 92[dB]$	24
2.8. Singularidad de $\Lambda_{zuy}^M(\hat{n})$, para $M = 10^4$, $\text{SNR} = 5, 84[dB]$	24
2.9. Singularidad de $\Lambda_{zuy}^M(\hat{n})$, para $M = 5 \cdot 10^4$, $\text{SNR} = 5, 78[dB]$	24
2.10. Singularidad de $\Lambda_{zuy}^M(\hat{n})$, con $M = 10^4$, $\text{SNR} = 5, 84[dB]$	25
2.11. Tabla de singularidad con el indicador $\text{IS}(\hat{n}_b, \hat{n}_a) = \lambda_{\min}(\Lambda_{xux}^M(\hat{n}_b, \hat{n}_a))$	30
2.12. Tabla de singularidad con el indicador $ \lambda _{\min}(\Lambda_{zuy}^M(\hat{n}_b, \hat{n}_a))$, considerando $\text{SNR} \approx 20[dB]$. . .	32
2.13. Tabla de singularidad con el indicador $ \lambda _{\min}(\Lambda_{zuy}^M(\hat{n}_b, \hat{n}_a))$, considerando $\text{SNR} \approx 10[dB]$. . .	33
4.1. Valores singulares de $\Gamma_{zx}^M(\hat{n})$ con $\hat{n} = 10$, para planta de orden 3.	66

4.2. Valores singulares de $\Gamma_{zx}^M(\hat{n})$ con $\hat{n} = 10$, para planta de orden 6.	66
4.3. Valores singulares de $\Gamma_{zx}^M(\hat{n})$ con $\hat{n} = 10$, para planta de orden 7.	67

Índice de algoritmos

2.1. Búsqueda secuencial de n a través de la matriz $\Lambda_{xux}^M(\hat{n})$	18
2.2. Búsqueda secuencial de n a través de la matriz $\Lambda_{zuy}^M(\hat{n})$	23
2.3. Búsqueda secuencial de (n_b, n_a) a través de la matriz $\Lambda_{xux}^M(\hat{n}_b, \hat{n}_a)$	30
2.4. Búsqueda secuencial de (n_b, n_a) a través de la matriz $\Lambda_{zuy}^M(\hat{n}_b, \hat{n}_a)$	31
3.1. Búsqueda del par (n_b, n_a) con el método de perfiles de valores singulares. Caso sin ruido. . . .	41
3.2. Búsqueda del par (n_b, n_a) con el método de perfiles de valores singulares. Caso con ruido. . . .	45
4.1. Búsqueda del par (n_b, n_a) con el método de perfiles para $\Gamma_{zx}^M(\hat{n}_a)$	56
4.2. Búsqueda del par $(n_b - d_g, n_a)$ con el método de perfiles para $\Gamma_{zx}^M(\hat{n}_a)$ en el caso ARMA. . . .	61

Introducción

Por qué identificar la estructura de un sistema

La obtención de un modelo, a partir de un conjunto de datos medidos, es un problema común en ingeniería, y se lleva a cabo mediante la estimación de parámetros (o constantes) de alguna estructura de modelos, seleccionada con anterioridad. La selección de esta estructura (número de polos y ceros en el caso de un sistema lineal e invariante) es un paso importante previo a la estimación de parámetros ya que, por ejemplo, métodos como cuadrados mínimos (LS), variables instrumentales (IV), u otros más sofisticados como [22], requieren al menos el conocimiento del orden del sistema. Además, la consistencia de los ya conocidos métodos LS e IV depende fuertemente de la elección correcta del orden del modelo, [19], [20] y [14].

La elección de una estructura de modelos apropiada puede abordarse principalmente desde dos ángulos diferentes:

1. Basada en las leyes que rigen al sistema
2. Basada en datos experimentales (viendo al sistema como caja negra)

Si se tiene información de las leyes que rigen al sistema, es posible obtener, de manera directa, una estructura de modelos. Por ejemplo, considérese la posición $x(t)$ de una masa, en un sistema masa-resorte-amortiguador, la cual es excitada por una fuerza externa $u(t)$. Este sistema de tiempo continuo, puede modelarse utilizando una estructura lineal e invariante (\hat{n}_b, \hat{n}_a) que hace referencia al número de ceros y polos del modelo

$$\frac{d^{\hat{n}_a} x(t)}{dt^{\hat{n}_a}} + a_{\hat{n}_a-1} \frac{d^{(\hat{n}_a-1)} x(t)}{dt^{(\hat{n}_a-1)}} + \dots + a_0 x(t) = b_{\hat{n}_b} \frac{d^{\hat{n}_b} u(t)}{dt^{\hat{n}_b}} + b_{\hat{n}_b-1} \frac{d^{(\hat{n}_b-1)} u(t)}{dt^{(\hat{n}_b-1)}} + \dots + b_0 u(t)$$

Dado que el sistema masa-resorte-amortiguador está gobernado por leyes físicas conocidas, entonces se puede obtener directamente

$$m \frac{d^2 x(t)}{dt^2} + b \frac{dx(t)}{dt} + \kappa x(t) = u(t)$$

donde m es la masa, b es el coeficiente de roce viscoso y κ es la constante elástica del resorte. De esta manera, una estructura apropiada para el modelo corresponde a $(\hat{n}_b, \hat{n}_a) = (0, 2)$, es decir una estructura lineal de solamente dos polos

$$x(t) = G_m(\rho)u(t), \quad G_m(\rho) = \frac{1}{m\rho^2 + b\rho + \kappa}$$

donde ρ es el operador derivada temporal. Este tipo de razonamiento puede hacerse, de manera análoga, para sistemas de tiempo discreto, donde el modelo $G_m(q)$ dependería del operador de adelanto q (o equivalentemente del operador retardo q^{-1}), en lugar del operador derivada temporal ρ .

Supóngase ahora que no se tiene ninguna información sobre las leyes que rigen al sistema, o que ellas son difíciles de dilucidar, como por ejemplo en un circuito electrónico RLC encapsulado en una caja que no se puede abrir, el cual es excitado con un voltaje de entrada $u(k)$ (generado con un computador) y que entrega un voltaje de salida muestreado $x(k)$. Dado que este sistema es lineal e invariante, y que el número de condensadores, resistencias e inductancias no se conoce, se puede decir que la estructura de modelos (\hat{n}_b, \hat{n}_a) para este sistema debe ser de la forma

$$x(k) = G_m(q)u(k), \quad G_m(q) = \frac{b_0 + b_1q^{-1} + \dots + b_{\hat{n}_b}q^{-\hat{n}_b}}{1 + a_1q^{-1} + \dots + a_{\hat{n}_a}q^{-\hat{n}_a}}$$

donde la estructura correcta de este sistema, dada por (n_b, n_a) , depende en general de la cantidad de condensadores e inductancias

$$x(k) = G(q)u(k), \quad G(q) = \frac{b_0 + b_1q^{-1} + \dots + b_{n_b}q^{-n_b}}{1 + a_1q^{-1} + \dots + a_{n_a}q^{-n_a}} \quad (1)$$

Dado que sólo se tiene acceso a la entrada $u(k)$ y a la salida $x(k)$ de este circuito, y que no se puede ver su interior, entonces sólo queda una opción para poder determinar su estructura y es a través del análisis de datos experimentales de entrada y salida.

Esta última situación es tratada en esta tesis, donde se mejoran algunos resultados existentes, los que permiten determinar una estructura apropiada de modelos, para sistemas lineales e invariantes de tiempo discreto, con la premisa de que las leyes que rigen al sistema en cuestión son desconocidas o difíciles de determinar.

Algo de historia

En la literatura existen principalmente dos tipos de problemas de identificación estructural, basados en análisis de datos experimentales

1. Sólo la salida se encuentra disponible para medición (caso de modelos AR, MA y ARMA)
2. La entrada y la salida del sistema están disponibles para medición

Para cada uno de estos dos casos, el problema de identificación estructural puede ser abordado en forma *paramétrica* o *no paramétrica*. La forma *paramétrica* necesita calcular los parámetros o coeficientes del modelo, para luego validar si el orden del modelo es correcto. La forma *no paramétrica* no precisa del cálculo de los coeficientes, para obtener una estructura de modelos apropiada.

Dentro de los primeros intentos de identificación estructural se encuentra el *error final de predicción* (FPE) [17] desarrollado por Akaike (1969) únicamente para sistemas AR. Este método es del tipo *paramétrico*, ya que necesita estimar previamente la varianza de los residuos $\hat{\sigma}_{\hat{n}}^2$ para poder validar si el orden \hat{n} del modelo es correcto, considerando N mediciones de la salida $x(k)$

$$\text{FPE}(\hat{n}) = \frac{N + \hat{n}}{N - \hat{n}} \hat{\sigma}_{\hat{n}}^2$$

Así el orden más adecuado consiste en encontrar \hat{n} que minimiza $\text{FPE}(\hat{n})$. Posteriormente, Akaike desarrolló otro criterio más general para modelos ARMA, llamado *criterio de información de Akaike* (AIC), el cual también es un método paramétrico, ya que necesita estimar la varianza de los residuos.

En años posteriores se han desarrollado métodos *no paramétricos* para identificación de sistemas ARMA, los que están basados en pruebas de singularidad sobre matrices de covarianzas. Entre estos métodos se encuentran los citados a continuación

- [4, section VII] Cadzow (1982), en que se propone que el orden AR de un proceso ARMA puede ser estimado a partir del rango de una matriz de covarianzas, el que se calcula utilizando descomposición en valores singulares.
- [6] Chan y Wood (1984), en que se estima los órdenes AR y MA de un sistema ARMA, a través de un método de dependencia lineal de las columnas de una matriz de covarianzas. La dependencia lineal entre las columnas es verificada mediante el proceso de ortonormalización de Gram-Schmidt.
- [21] Stoica (1981) y [9] Fuchs (1987) establecen que la matriz de covarianzas es singular si los órdenes AR y MA de un modelo ARMA sobrepasan a los órdenes verdaderos; en caso contrario la matriz de covarianzas es no singular.
- [25] Zhang y Zhang (1993), en que se utiliza el método de Cadzow para determinar el orden AR de un sistema ARMA, y además se propone un método para determinar el orden MA, el que establece que el rango de una matriz de covarianzas disminuye cuando el orden MA del modelo sobrepasa al verdadero.

Paralelamente se desarrollaron también métodos *no paramétricos*, basados en matrices de covarianzas, para sistemas donde la entrada y la salida están disponibles para medición, como en las siguientes referencias

- [23] Wellstead (1978), y luego [24] Wellstead y Rojas (1982), en que se propone un método basado en matrices de covarianzas instrumentales, el que consiste en aumentar secuencialmente el orden de los modelos y verificar si las matrices asociadas a ellos son singulares.
- [19, ejemplo 4.1 y teorema 4.1] Söderström y Stoica (1983), [14, sección 8.6] Ljung (1999), en que se establece que la matriz de covarianzas es singular si la cantidad de ceros y polos sobrepasan el número verdadero de ceros y polos del sistema. Este mismo resultado se extiende para el caso de la matriz de covarianzas instrumentales.
- [5] Cadzow y Solomon (1986), [13] Liang et alii (1993), [7] Davila y Chiang (1994), proponen que si el número de ceros y polos del modelo sobrepasa al número de ceros y polos del sistema verdadero, entonces la matriz de covarianzas tiene una cantidad definida de valores propios nulos.
- [1] Al-Smadi y Wilkes (1996), [2] (2002), en que se propone el mismo resultado que [13], pero aplicado a matrices de covarianzas de 3.^{er} orden. Este método tiene la ventaja de poder atenuar ruido gaussiano de salida.

Los resultados más importantes aquí mencionados se encuentran detallados a continuación.

Algunos resultados importantes

[21] P. Stoica (1981); [4] J. Cadzow (1982); [9] J. Fuchs (1987); [25] X. Zhang y Y. Zhang (1993). En estos artículos se considera que el sistema, cuya estructura se desea identificar, es un proceso de tipo ARMA, regido por la ecuación (1) y cuya entrada $u(k)$ es ruido blanco. El procedimiento para determinar la estructura consiste básicamente en hacer pruebas de singularidad sobre la matriz

$$R_{xx}(\hat{n}_b, \hat{n}_a) = E \begin{bmatrix} x(k - \hat{n}_b - 1) \\ \vdots \\ x(k - \hat{n}_b - \hat{n}_a) \end{bmatrix} \begin{bmatrix} x(k - 1) \\ \vdots \\ x(k - \hat{n}_a) \end{bmatrix}^T = \begin{bmatrix} r_{xx}(\hat{n}_b) & \cdots & r_{xx}(\hat{n}_b - \hat{n}_a + 1) \\ \vdots & \ddots & \vdots \\ r_{xx}(\hat{n}_b + \hat{n}_a - 1) & \cdots & r_{xx}(\hat{n}_b) \end{bmatrix}$$

donde $r_{xx}(\tau) = Ex(k)x(k - \tau)$, la cual tiene las siguientes propiedades [21], [9]

$$\det R_{xx}(\hat{n}_b, \hat{n}_a) \begin{cases} \text{genéricamente } \neq 0, & \text{si } \min(\hat{n}_b - n_b, \hat{n}_a - n_a) < 0 \\ \neq 0, & \text{si } \min(\hat{n}_b - n_b, \hat{n}_a - n_a) = 0 \\ = 0, & \text{si } \min(\hat{n}_b - n_b, \hat{n}_a - n_a) > 0 \end{cases}$$

donde (n_b, n_a) es la estructura correcta del sistema. Es más, en [4] se propone un valor para el rango de $R_{xx}(\hat{n}_b, \hat{n}_a)$ bajo ciertas condiciones para \hat{n}_b y \hat{n}_a

$$\text{rango } R_{xx}(\hat{n}_b, \hat{n}_a) = n_a, \quad \text{si } \hat{n}_b - n_b \geq \hat{n}_a - n_a \geq 0$$

además, según [25] se tiene también

$$\text{rango } R_{xx}(\hat{n}_b, n_a + 1) = \begin{cases} n_a + 1, & \text{si } \hat{n}_b = n_b \\ n_a, & \text{si } \hat{n}_b > n_b \end{cases}$$

Dado que $r_{xx}(\tau)$ no se puede obtener experimentalmente, entonces debe ser aproximada por promedios temporales, por lo que la matriz $R_{xx}(\hat{n}_b, \hat{n}_a)$ debe ser estimada mediante

$$\hat{R}_{xx}(\hat{n}_b, \hat{n}_a) = \frac{1}{N} \sum_{k=1}^N \begin{bmatrix} x(k - \hat{n}_b - 1) \\ \vdots \\ x(k - \hat{n}_b - \hat{n}_a) \end{bmatrix} \begin{bmatrix} x(k - 1) \\ \vdots \\ x(k - \hat{n}_a) \end{bmatrix}^T$$

[19, example 4.1 y theorem 4.1] T. Söderström y P. Stoica (1983); [14, sección 8.6] L. Ljung (1999). En ambos libros se obtiene que la siguiente matriz de covarianzas

$$R_{xux}(\hat{n}_b, \hat{n}_a) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N \omega_{ux}^{\hat{n}_b, \hat{n}_a}(k) \omega_{ux}^{\hat{n}_b, \hat{n}_a}(k)^T$$

$$\omega_{ux}^{\hat{n}_b, \hat{n}_a}(k)^T = [u(k) \quad \cdots \quad u(k - \hat{n}_b) \quad x(k - 1) \quad \cdots \quad x(k - \hat{n}_a)]$$

construida con mediciones obtenidas de la entrada y salida del sistema (1), satisface las siguientes condiciones de singularidad

$$R_{xux}(\hat{n}_b, \hat{n}_a) \begin{cases} \text{no singular,} & \text{si } \min(\hat{n}_b - n_b, \hat{n}_a - n_a) \leq 0 \\ \text{singular,} & \text{si } \min(\hat{n}_b - n_b, \hat{n}_a - n_a) > 0 \end{cases}$$

cuando la señal de entrada $u(k)$ es persistentemente excitante¹ de orden $\max(\hat{n}_b + n_a + 1, n_b + \hat{n}_a + 1)$, y considerando que la estructura correcta corresponde a (n_b, n_a) . Además se obtienen resultados similares para la matriz instrumental

$$R_{zux}(\hat{n}_b, \hat{n}_a) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N \omega_{uz}^{\hat{n}_b, \hat{n}_a}(k) \omega_{ux}^{\hat{n}_b, \hat{n}_a}(k)^T$$

donde la secuencia instrumental $z(k) = F(q)u(k)$ es generada a partir del siguiente filtro

$$F(q) = \frac{l_0 + l_1 q^{-1} + \cdots + l_{m_l} q^{-m_l}}{1 + k_1 q^{-1} + \cdots + k_{m_k} q^{-m_k}}$$

obteniéndose que

$$R_{zux}(\hat{n}_b, \hat{n}_a) \begin{cases} \text{singular,} & \text{si } \min(\hat{n}_b - n_b, \hat{n}_a - n_a) > 0 \text{ ó } \min(\hat{n}_b - m_l, \hat{n}_a - m_k) > 0 \\ \text{no singular,} & \text{en cualquier otro caso} \end{cases}$$

para lo cual es necesario (pero no suficiente) que la señal de entrada $u(k)$ sea persistentemente excitante de orden $\max(\hat{n}_b + n_a + 1, n_b + \hat{n}_a + 1)$.

Este resultado instrumental se puede extender también al caso en que la salida del sistema $x(k)$ se encuentra contaminada con ruido coloreado $v(k)$, es decir $y(k) = x(k) + v(k)$, ya que si $u(k)$ y $v(k)$ no están correlacionados, entonces

$$R_{zuy}(\hat{n}_b, \hat{n}_a) = R_{zux}(\hat{n}_b, \hat{n}_a)$$

¹Ver definición 1.11 en el capítulo 1

Sin duda, los resultados obtenidos para las matrices $R_{xux}(\hat{n}_b, \hat{n}_a)$ y $R_{zux}(\hat{n}_b, \hat{n}_a)$ permiten la existencia de criterios de identificación estructural, pero lamentablemente poseen la desventaja de considerar condiciones asintóticas en la cantidad de mediciones.

[5] J. Cadzow y O. Solomon (1986); [13] G. Liang et alii (1993); [7] C. Davila y H. Chiang (1994). En estos tres artículos se muestra aproximadamente el mismo resultado, el que señala que la matriz de covarianzas

$$\bar{R}_{xux}(\hat{n}_b, \hat{n}_a) = \frac{1}{N} \sum_{k=1}^N \bar{\omega}_{ux}^{\hat{n}_b, \hat{n}_a}(k) \bar{\omega}_{ux}^{\hat{n}_b, \hat{n}_a}(k)^T$$

$$\bar{\omega}_{ux}^{\hat{n}_b, \hat{n}_a}(k)^T = [u(k) \quad \cdots \quad u(k - \hat{n}_b) \quad x(k) \quad \cdots \quad x(k - \hat{n}_a)]$$

construida con mediciones de entrada y salida tomadas desde (1), tiene s valores propios iguales a cero

$$s = 1 + \min(\hat{n}_b - n_b, \hat{n}_a - n_a), \quad \text{si } \min(\hat{n}_b - n_b, \hat{n}_a - n_a) \geq 0 \quad (2)$$

donde (n_b, n_a) corresponde a la estructura correcta del sistema.

Con este resultado se puede obtener una estructura apropiada, a través de pruebas de singularidad realizadas sobre la matriz $\bar{R}_{xux}(\hat{n}_b, \hat{n}_a)$. Por ejemplo en [13] se utiliza el índice de singularidad

$$J(\hat{n}_b, \hat{n}_a) = \lambda_{\min}(N^{1/N})^{\hat{n}_b + \hat{n}_a}$$

que es proporcional al menor valor propio de $\bar{R}_{xux}(\hat{n}_b, \hat{n}_a)$, es decir λ_{\min} , y penaliza además según la cantidad de mediciones utilizadas N y el tamaño del modelo (\hat{n}_b, \hat{n}_a) . Con este índice $J(\hat{n}_b, \hat{n}_a)$ es posible construir una tabla con \hat{n}_b variando en las filas y \hat{n}_a variando en las columnas, así la estructura más apropiada se obtiene para la menor esquina (\hat{n}_b, \hat{n}_a) donde $J(\hat{n}_b, \hat{n}_a)$ es cero, es decir, donde $\bar{R}_{xux}(\hat{n}_b, \hat{n}_a)$ es singular.

Cuando las señales se encuentran contaminadas con ruido, se aplican técnicas para poder reducirlo, por ejemplo en [5] y [7] se considera que el ruido es gaussiano, y se disminuye su efecto utilizando un reescalamiento de las señales según la potencia del ruido, la cual debe ser conocida.

El resultado obtenido en (2), para la matriz $\bar{R}_{xux}(\hat{n}_b, \hat{n}_a)$, tiene la ventaja de que considera una cantidad finita de mediciones; sin embargo, el tipo de señales de entrada $u(k)$, que produce este resultado, no se encuentra bien determinado. Por ejemplo, en [5] se menciona que $u(k)$ debe ser tal que las siguientes matrices sean de rango completo

$$U(\hat{n}_b) = \begin{bmatrix} u(1) & \cdots & u(1 - \hat{n}_b) \\ \vdots & & \vdots \\ u(N) & \cdots & u(N - \hat{n}_b) \end{bmatrix}, \quad X(\hat{n}_a) = \begin{bmatrix} x(1) & \cdots & x(1 - \hat{n}_a) \\ \vdots & & \vdots \\ x(N) & \cdots & x(N - \hat{n}_a) \end{bmatrix}$$

Sin embargo no se especifica qué tipo de señales cumple esta condición. En [7] sólo se especifica que $u(k)$ debe ser persistentemente excitante de orden $n_b + n_a + 1$ para el caso de la matriz $\bar{R}_{xux}(n_b, n_a)$. Sin embargo esto no es del todo cierto, ya que el concepto de *persistencia de excitación* está definido para una cantidad asintótica de mediciones, lo que no se aplicaría en este caso ya que N es finito.

[1] A. Al-Smadi y D. Wilkes (1996); [2] A. Al-Smadi y D. Wilkes (2002). En estos artículos se desarrolla un método basado en covarianzas de 3.º orden, el cual considera que la salida $x(k)$ de (1) está contaminada con ruido gaussiano $v(k)$, resultando así que la salida medible es $y(k) = x(k) + v(k)$.

La idea básica de estos artículos radica en que las covarianzas de 3.º orden son capaces de eliminar el ruido gaussiano, ya que, por ejemplo

$$\begin{aligned} c_{yyy}(n, m) &:= E \{y(k)y(k+n)y(k+m)\} = E \{x(k)x(k+n)x(k+m)\} \\ c_{uyy}(n, m) &:= E \{u(k)y(k+n)y(k+m)\} = E \{u(k)x(k+n)x(k+m)\} \end{aligned}$$

Ahora, con estas covarianzas se construye la siguiente matriz

$$C(\hat{n}_b, \hat{n}_a) = \begin{bmatrix} C_{yyy} & C_{uyy} \end{bmatrix}$$

donde $C(\hat{n}_b, \hat{n}_a)$ depende del tamaño del modelo (\hat{n}_b, \hat{n}_a) , y las matrices C_{yyy} y C_{uyy} están formadas por las covarianzas $c_{yyy}(n, m)$ y $c_{uyy}(n, m)$. Así, se demuestra en ambos artículos que la matriz

$$R(\hat{n}_b, \hat{n}_a) = C(\hat{n}_b, \hat{n}_a)^T C(\hat{n}_b, \hat{n}_a)$$

posee s valores propios nulos, donde s está dado en (2), por lo que puede usarse el mismo método que en [13] para determinar la estructura más apropiada.

El método desarrollado en [1] y [2] tiene la ventaja de que puede eliminar bastante bien el ruido gaussiano, sin embargo tiene la desventaja de que las covarianzas $c_{yyy}(n, m)$ y $c_{uyy}(n, m)$ no pueden ser obtenidas en la práctica y deben ser estimadas con promedios temporales

$$\hat{c}_{yyy}(n, m) := \frac{1}{N} \sum_{k=1}^N y(k)y(k+n)y(k+m), \quad \hat{c}_{uyy}(n, m) := \frac{1}{N} \sum_{k=1}^N u(k)y(k+n)y(k+m)$$

Así los estimadores $\hat{c}_{yyy}(n, m)$ y $\hat{c}_{uyy}(n, m)$ y las covarianzas $c_{yyy}(n, m)$ y $c_{uyy}(n, m)$ coinciden sólo en el caso ergódico y cuando se toman infinitas mediciones.

Breve descripción de la tesis

En esta tesis se resolvieron los principales problemas de contrucción presentados anteriormente

- En algunos métodos basados en matrices de covarianzas estadísticas no se pueden aplicar directamente sus resultados, ya que requieren de estimadores muestrales (basados en promedios) para implementarlos [21], [4], [9], [25], [1] y [2].
- Algunos métodos basados en covarianzas muestrales necesitan un número infinito de mediciones para aplicar sus resultados, aún cuando se pueden medir datos de entrada y salida [19], [14].
- Para los métodos que basan sus resultados en valores propios/singulares de matrices de covarianzas [4], [13], [1], [2], no existe un criterio que permita determinar si un valor propio/singular es “nulo”, cuando la matriz de covarianzas se ve perturbada con ruido.
- En [5] se establecen ciertas condiciones sobre $u(k)$, de manera que los resultados obtenidos para las matrices de covarianzas sean ciertos. Sin embargo, no queda claro qué tipo de señales satisfacen aquellas condiciones. Por otra parte, en [7] se especifica que la señal de entrada debe ser persistentemente excitante de cierto orden sólo para el caso en que la estructura del modelo coincide con el sistema verdadero, además la definición de persistencia de excitación no es válida cuando se utiliza un número finito de mediciones.
- Algunos métodos como [1],[2] basan sus resultados en fuertes suposiciones sobre la distribución probabilística del ruido de medición.

También se pretende mejorar algunos resultados ya conocidos

- Se conoce en qué casos de la región $(\hat{n}_b, \hat{n}_a) \in \mathbb{N}^2$, la *matriz de Yule-Walker* $R_{xx}(\hat{n}_b, \hat{n}_a)$ es singular [21], [9], e incluso se conoce su rango para algunos de estos casos [4], [25]; sin embargo, puede ser de utilidad conocer el rango de esta matriz para toda la región $(\hat{n}_b, \hat{n}_a) \in \mathbb{N}^2$.

- De la misma manera, se conoce la región $(\hat{n}_b, \hat{n}_a) \in \mathbb{N}^2$ en que la *matriz de producto* $R_{xux}(\hat{n}_b, \hat{n}_a)$ y la *matriz de producto instrumental* $R_{zux}(\hat{n}_b, \hat{n}_a)$ son singulares [19], [14], más aún, se conoce el rango de la *matriz de producto* $R_{xux}(\hat{n}_b, \hat{n}_a)$ cuando se utiliza un número finito de mediciones [5], [13], [7]. Así, podría ser de utilidad conocer el rango de la *matriz de producto instrumental* $R_{zux}(\hat{n}_b, \hat{n}_a)$ para toda la región $(\hat{n}_b, \hat{n}_a) \in \mathbb{N}^2$.
- Dado que la matriz de Hankel del sistema [12] es bastante utilizada en el campo de reducción de modelos, entonces es posible que pueda relacionarse con estos métodos de estimación de estructura.

Finalmente esta tesis aborda todos estos temas en el siguiente orden:

En el capítulo 1 se introduce brevemente algo de notación y conceptos de probabilidades y procesos estocásticos a utilizar durante todo el documento. Posteriormente se estudia el contenido informacional de señales, mediante los conceptos de *persistencia de excitación local* (considera un número finito de mediciones), *persistencia de excitación* (definición tradicional, que considera infinitas mediciones) y *suficiencia de información*. Estas últimas definiciones serán de mucha utilidad para saber qué tipo de señales utilizar al momento de construir las matrices de covarianzas.

En el capítulo 2 se hace una revisión de las propiedades de rango de la *matriz de producto*, considerando un número finito e infinito de mediciones y estableciendo condiciones suficientes que debe cumplir la señal de entrada. De manera similar, se extienden estas propiedades de rango a la *matriz de producto instrumental*, la cual tiene la ventaja de atenuar el ruido de medición, aun si no se conoce su distribución probabilística. Finalmente se ejemplifican algunos métodos exhaustivos de identificación estructural, basados en pruebas de singularidad sobre la *matriz de producto* y la *matriz de producto instrumental*, similares a los utilizados en [13], [1] y [2].

Con el fin de optimizar los resultados del capítulo anterior, el capítulo 3 plantea un método no exhaustivo de identificación estructural, basado en el conocimiento de los rangos de la *matriz de producto* y de la *matriz de producto instrumental*. Para el cálculo experimental de los rangos se utilizan los valores singulares de estas matrices, desarrollándose para ello criterios que determinan si un valor singular es cero, cuando estas matrices se ven perturbadas con errores numéricos o con ruido de medición.

Finalmente, en el capítulo 4 se plantea la existencia de una relación entre la *matriz de producto instrumental* y la *matriz de Yule-Walker*. Esta relación permite conocer completamente el rango de esta última matriz, generalizando así los resultados en [21] y [9]. Además se obtiene una relación entre la *matriz de producto instrumental* y la matriz de Hankel del sistema, lo que permite estimar los valores singulares de Hankel antes de identificar el sistema en cuestión.

Capítulo 1

Preliminares

EN este capítulo se tratan brevemente algunas definiciones y convenciones de probabilidades y procesos estocásticos, además de definir algunas operaciones entre señales. Finalmente se pretende estudiar el contenido informacional de ciertas señales, a través de los conceptos de *persistencia de excitación* y *suficiencia de información*.

1.1. Probabilidades y procesos estocásticos

1.1.1. Definiciones y conceptos

Teorema 1.1 Desigualdad de Chebyshev. Sea X una variable aleatoria de media μ y varianza σ^2 , entonces se tiene que

$$\forall \varepsilon > 0, \quad P(|X - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}$$

Ver demostración complementaria E.1.

Definición 1.1 Convergencia. En probabilidades existen diferentes clases de convergencia, según la métrica que se desee utilizar, así una sucesión de variables aleatorias X_n converge:

1. siempre (definición usual de convergencia)

$$X_n \rightarrow X \Leftrightarrow (\forall \varepsilon > 0)(\exists N)(\forall n > N) |x_n - x| < \varepsilon$$

2. con probabilidad 1

$$X_n \xrightarrow{p1} X \Leftrightarrow P(X_n \rightarrow X) = 1$$

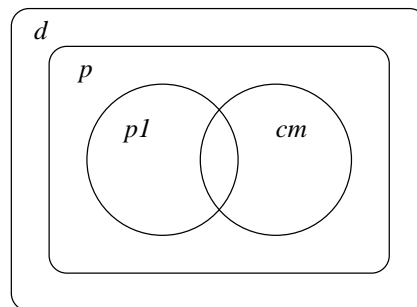


Figura 1.1: Relaciones entre los diferentes tipos de convergencia.

3. en sentido cuadrático medio

$$X_n \xrightarrow{cm} X \Leftrightarrow E \left\{ |X_n - X|^2 \right\} \rightarrow 0$$

4. en probabilidad

$$X_n \xrightarrow{p} X \Leftrightarrow P(|X_n - X| \geq \varepsilon) \rightarrow 0$$

5. en distribución

$$X_n \xrightarrow{d} X \Leftrightarrow F_n(x) \rightarrow F(x)$$

donde $F_n(x)$ y $F(x)$ son funciones de distribución acumuladas.

Las relaciones entre estos tipos de convergencia son ilustrados en la figura 1.1.

Definición 1.2 Independencia. Las variables aleatorias X_1, \dots, X_n son (mutuamente) independientes si los siguientes eventos $\{X_1 \leq x_1\}, \dots, \{X_n \leq x_n\}$ son independientes. Así se tiene que

$$f(x_1, \dots, x_n) = f_{X_1}(x_1) \cdots f_{X_n}(x_n)$$

donde $f_{X_1}(x_1), \dots, f_{X_n}(x_n)$ son funciones de densidad de probabilidad.

1.1.2. Procesos estocásticos

Definición 1.3 *Proceso estocástico.* Un proceso estocástico de tiempo discreto $s(k)$ es una sucesión, de tal forma que $\forall k \in \mathbb{Z}$ $s(k)$ es una variable aleatoria.

Para los procesos estocásticos $x(k)$ e $y(k)$ se pueden definir los siguientes momentos

1. Media

$$m_x(k) := E \{x(k)\}$$

2. Correlación

$$r_{xy}(k_1, k_2) := E \{x(k_1)y(k_2)^*\}$$

3. Covarianza

$$c_{xy}(k_1, k_2) := E \{[x(k_1) - m_x(k_1)][y(k_2) - m_y(k_2)]^*\}$$

4. Coeficiente de correlación (de Pearson)

$$\rho_{xy}(k_1, k_2) := \frac{c_{xy}(k_1, k_2)}{\sqrt{c_{xx}(k_1, k_1)c_{yy}(k_2, k_2)}}$$

Si los segundos momentos involucran sólo un proceso estocástico, $r_{xx}(k_1, k_2)$ y $c_{xx}(k_1, k_2)$ reciben el nombre de autocorrelación y autocovarianza respectivamente.

Véase [16, ecuaciones 10-7, 10-22, 10-25 y 10-26].

Nota 1.1 En muchos casos, en el campo de identificación de sistemas, se utiliza el término de “covarianza” para referirse indistintamente a $r_{xy}(k_1, k_2)$ y $c_{xy}(k_1, k_2)$. Sin embargo, en este texto se utilizará la nomenclatura de [16] para enfatizar la diferencia.

Definición 1.4 *Procesos ortogonales y no correlacionados.* Los procesos estocásticos $x(k)$ e $y(k)$ son ortogonales si el producto interno entre estos dos elementos es cero, es decir

$$\forall k_1, k_2 \in \mathbb{Z}, \quad r_{xy}(k_1, k_2) = 0$$

Los procesos estocásticos $x(k)$ e $y(k)$ no están correlacionados si su coeficiente de correlación es cero, es decir

$$\forall k_1, k_2 \in \mathbb{Z}, \quad \rho_{xy}(k_1, k_2) = 0$$

Véase [16, ecuaciones 10-31 y 10-32].

Definición 1.5 *Ruido blanco.* El proceso estocástico $e(k)$ es un proceso de ruido blanco si para todo k_1, k_2 se tiene que su autocovarianza cumple

$$c_{ee}(k_1, k_2) = \alpha(k_1)\delta_K(k_1 - k_2), \quad \alpha(k_1) \geq 0$$

Si además $e(k_i)$ y $e(k_j)$ son independientes $\forall k_i \neq k_j$, entonces este proceso se llama ruido blanco estricto.

Véase [16, ecuación 10-34].

Definición 1.6 *Proceso estacionario.* El proceso estocástico $x(k)$ es “estacionario en sentido amplio” (e.s.a.) si

1. $E \{x(k)\} = m_x$ es independiente de k
2. $E \{x(k_1)x(k_2)^*\} = r_{xx}(k_1 - k_2)$ es acotado y sólo depende de la diferencia $k_1 - k_2$

Definición 1.7 *Procesos conjuntamente estacionarios.* Se dice que $x(k)$ e $y(k)$ son “conjuntamente estacionarios en sentido amplio” (c.e.s.a.) si

1. $x(k)$ e $y(k)$ son estacionarios en sentido amplio por separado
2. $E \{x(k_1)y(k_2)^*\} = r_{xy}(k_1 - k_2)$ es acotado y sólo depende de la diferencia $k_1 - k_2$

Considérese el siguiente operador promediador

$$\overline{\langle x(k) \rangle}_{k=a}^b := \frac{1}{b-a+1} \sum_{k=a}^b x(k) \quad (1.1)$$

Una estimación de $2M+1$ muestras para la correlación se puede obtener a partir de la siguiente correlación muestral

$$\varphi_{xy}^M(-\tau_1, \tau_2) := \overline{\langle x(k-\tau_1)y(k-\tau_2)^* \rangle}_{k=-M}^M \quad (1.2)$$

ya que

$$E\varphi_{xy}^M(-\tau_1, \tau_2) = r_{xy}(\tau_2 - \tau_1)$$

Nótese también que si $M \rightarrow \infty$, la correlación muestral depende sólo de $\tau_2 - \tau_1$

$$\varphi_{xy}^\infty(-\tau_1, \tau_2) = \varphi_{xy}^\infty(\tau_2 - \tau_1)$$

Ahora, con ciertas condiciones en los procesos e.s.a., el valor de la estimación muestral $\varphi_{xy}^\infty(\tau)$ coincide con el valor estadístico de $r_{xy}(\tau)$. A esta propiedad se le conoce con el nombre de ergodicidad.

Definición 1.8 Media ergódica. Se dice que el proceso estocástico $x(k)$ tiene media ergódica si es e.s.a y además

$$E \{x(k)\} = \overline{\langle x(k) \rangle}_{k \in \mathbb{Z}}$$

Definición 1.9 Correlación ergódica. Se dice que los procesos estocásticos $x(k)$ e $y(k)$ tienen correlación ergódica si son c.e.s.a. y además

$$r_{xy}(\tau) = \varphi_{xy}^\infty(\tau)$$

Para más detalles sobre ergodicidad véase el apéndice C.

1.1.3. Propiedades de correlación

Sean las señales $z = f * u$ y $v = h * e$, donde

$$(f * u)(k) := \sum_{l \in \mathbb{Z}} f(l)u(k-l) = \sum_{l \in \mathbb{Z}} f(k-l)u(l) \quad (1.3)$$

es la operación de convolución discreta. Entonces se tiene que la correlación muestral φ_{zv}^∞ satisface

$$\varphi_{zv}^\infty = f * \varphi_{ue}^\infty * \mathcal{J}h^* \quad (1.4)$$

o en su equivalente frecuencial

$$\Phi_{zv}^\infty(e^{j\theta}) = F(e^{j\theta})\Phi_{ue}^\infty(e^{j\theta})H(e^{j\theta})^* \quad (1.5)$$

donde \mathcal{J} es el operador de inversión temporal

$$(\mathcal{J}x)(k) := x(-k) \quad (1.6)$$

y $\Phi_{zv}^\infty(e^{j\theta})$ es la transformada de Fourier de tiempo discreto de $\varphi_{zv}^\infty(\tau)$, definida como

$$X(e^{j\theta}) = \mathcal{F}_d \{x\}(\theta) := \sum_{k \in \mathbb{Z}} x(k)e^{-j\theta k} \quad (1.7)$$

La demostración de la ecuación (1.4) es directa, si se consideran los resultados (B.6) y (B.7) del apéndice B, ya que

$$\varphi_{zv}^\infty = \varphi_{(f*u)(h*e)}^\infty = f * \varphi_{u(h*e)}^\infty = f * \varphi_{ue}^\infty * \mathcal{J}h^*$$

Estas propiedades son igualmente aplicables a la covarianza y a la correlación, considerando que $u(k)$ y $e(k)$ son procesos c.e.s.a., y $f(k)$ y $h(k)$ son señales determinísticas,

$$\begin{aligned} c_{zv} &= f * c_{ue} * \mathcal{J}h^* \\ C_{zv}(e^{j\theta}) &= F(e^{j\theta})C_{ue}(e^{j\theta})H(e^{j\theta})^* \end{aligned} \quad (1.8)$$

$$\begin{aligned} r_{zv} &= f * r_{ue} * \mathcal{J}h^* \\ R_{zv}(e^{j\theta}) &= F(e^{j\theta})R_{ue}(e^{j\theta})H(e^{j\theta})^* \end{aligned} \quad (1.9)$$

para más detalles véase el apéndice B.

Nótese además que mediante estas propiedades de correlación es posible obtener la potencia de la señal $z = f * u$, en términos del espectro de $u(k)$

$$\begin{aligned} \overline{\langle z(k)^2 \rangle}_{k \in \mathbb{Z}} &= \varphi_{zz}^\infty(0) = \frac{1}{2\pi} \int_{-\pi}^{\pi} |F(e^{j\theta})|^2 \Phi_{uu}^\infty(e^{j\theta}) d\theta \\ E \{z(k)^2\} &= r_{zz}(0) = \frac{1}{2\pi} \int_{-\pi}^{\pi} |F(e^{j\theta})|^2 R_{uu}(e^{j\theta}) d\theta \end{aligned}$$

1.2. Predictores

El uso de predictores es bastante frecuente en el área de identificación, ya que se encuentra relacionado con bastantes métodos, por ejemplo: el método de cuadrados mínimos (LS) y el método del error de predicción (PEM), entre otros. Al igual que con los métodos anteriores, también tiene relación con las condiciones de singularidad de las matrices de momento de producto (PM), tratadas en esta tesis, así como con las propiedades de riqueza espectral asociadas a ellas.

Por este motivo, en las próximas secciones, se introduce algo de notación, relacionada con sistemas y predictores.

1.2.1. Caso sin ruido

Considérese el siguiente sistema generador de datos o *planta*

$$x(k) = G(q)u(k), \quad G(q) = \frac{B(q)}{A(q)} \quad (1.10)$$

$$\begin{aligned} B(q) &= b_0 + b_1q^{-1} + \dots + b_{n_b}q^{-n_b} \\ A(q) &= 1 + a_1q^{-1} + \dots + a_{n_a}q^{-n_a} \end{aligned}$$

donde

$$\begin{aligned} b_{n_b} \text{ y } a_{n_a} &\text{ son diferentes de cero} \\ B(q) \text{ y } A(q) &\text{ son coprimos} \\ n := \max(n_b, n_a) &\text{ es el orden del sistema} \end{aligned}$$

El sistema puede ser escrito en su forma de predictor,

despejando la salida actual en función de las demás variables

$$\begin{aligned} x(k) &= S(q)\zeta_{ux}(k) \quad (1.11) \\ S(q) &= \begin{bmatrix} b_0 + b_1q^{-1} + \dots + b_{n_b}q^{-n_b} \\ -a_1q^{-1} - \dots - a_{n_a}q^{-n_a} \end{bmatrix}^T \\ \zeta_{ux}(k) &= \begin{bmatrix} u(k) \\ x(k) \end{bmatrix} \end{aligned}$$

Además, este predictor cuenta con la particularidad de que puede ser escrito como regresión lineal según

$$\begin{aligned} x(k) &= \omega_{ux}^{n_b, n_a}(k)^T \beta \quad (1.12) \\ \omega_{ux}^{n_b, n_a}(k) &= \begin{bmatrix} u(k) \\ \vdots \\ u(k - n_b) \\ x(k - 1) \\ \vdots \\ x(k - n_a) \end{bmatrix}, \quad \beta = \begin{bmatrix} b_0 \\ \vdots \\ b_{n_b} \\ -a_1 \\ \vdots \\ -a_{n_a} \end{bmatrix} \end{aligned}$$

Ahora considérese un *modelo* arbitrario para la *planta* de la ecuación (1.10)

$$\begin{aligned} x_i(k) &= G_i(q)u(k), \quad G_i(q) = \frac{B_i(q)}{A_i(q)} \quad (1.13) \\ B_i(q) &= b_0^i + b_1^i q^{-1} + \dots + b_{\hat{n}_b}^i q^{-\hat{n}_b} \\ A_i(q) &= 1 + a_1^i q^{-1} + \dots + a_{\hat{n}_a}^i q^{-\hat{n}_a} \end{aligned}$$

donde i no denota una potencia, sino que simplemente un superíndice. Para fines de identificación, es usual anotar este modelo como predictor, en función de los datos obtenidos de la planta, resultando así

$$\begin{aligned} \hat{x}_i(k) &= W_i(q)\zeta_{ux}(k) \quad (1.14) \\ W_i(q) &= \begin{bmatrix} b_0^i + b_1^i q^{-1} + \dots + b_{\hat{n}_b}^i q^{-\hat{n}_b} \\ -a_1^i q^{-1} - \dots - a_{\hat{n}_a}^i q^{-\hat{n}_a} \end{bmatrix}^T \\ \zeta_{ux}(k) &= \begin{bmatrix} u(k) \\ x(k) \end{bmatrix} \end{aligned}$$

Este modelo cuenta además con la particularidad de poder escribirse como regresión lineal según

$$\begin{aligned} \hat{x}_i(k) &= \omega_{ux}^{\hat{n}_b, \hat{n}_a}(k)^T \gamma_i \quad (1.15) \\ \omega_{ux}^{\hat{n}_b, \hat{n}_a}(k) &= \begin{bmatrix} u(k) \\ \vdots \\ u(k - \hat{n}_b) \\ x(k - 1) \\ \vdots \\ x(k - \hat{n}_a) \end{bmatrix}, \quad \gamma_i = \begin{bmatrix} b_0^i \\ \vdots \\ b_{\hat{n}_b}^i \\ -a_1^i \\ \vdots \\ -a_{\hat{n}_a}^i \end{bmatrix} \end{aligned}$$

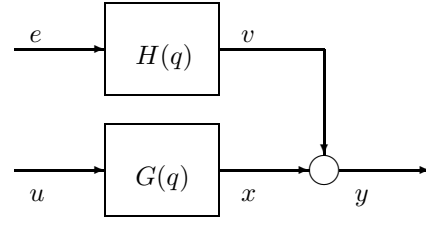


Figura 1.2: Diagrama total del sistema, considerando ruido coloreado $v(k)$. La actuación es $u(k)$ y la salida medible es $y(k)$.

1.2.2. Caso con ruido

Considérese el siguiente sistema generador de datos o *planta*, en que la salida medible $y(k)$ está afectada por ruido coloreado

$$y(k) = x(k) + v(k) \quad (1.16)$$

$$x(k) = G(q)u(k), \quad G(q) = \sum_{i=0}^{\infty} g(i)q^{-i}$$

$$v(k) = H(q)e(k), \quad H(q) = \sum_{j=0}^{\infty} h(j)q^{-j}$$

donde $e(k)$ es ruido blanco de media cero y varianza $a_e^2 > 0$, y $g(k)$ y $h(k)$ son la respuesta a impulso de los filtros $G(q)$ y $H(q)$ respectivamente. Esta situación puede esquematizarse en la figura 1.2.

Dado que la transformada de Fourier de la autocovarianza de $v(k)$ está dada por la expresión

$$C_{vv}(e^{j\theta}) = |H(e^{j\theta})|^2 C_{ee}(e^{j\theta}) = |H(e^{j\theta})|^2 a_e^2$$

debido a la ecuación (1.8), entonces el término $a_e > 0$ puede incluirse como parte del filtro generador de ruido $H(q)$, así, sin pérdida de generalidad, puede considerarse que $a_e^2 = 1$.

Para este sistema se propone el siguiente modelo arbitrario

$$y_i(k) = x_i(k) + v_i(k) \quad (1.17)$$

$$x_i(k) = G_i(q)u(k), \quad v_i(k) = H_i(q)e(k)$$

donde $g_i(k)$ y $h_i(k)$ son la respuesta a impulso de los modelos $G_i(q)$ y $H_i(q)$ respectivamente.

Suponiendo que $u(k)$ es conocida de antemano, ya que en muchos casos es posible diseñarla, entonces la obtención de un predictor asociado al modelo

está relacionada directamente con la predicción del ruido

$$\hat{y}_i(k) = G_i(q)u(k) + \hat{v}_i(k) \quad (1.18)$$

Para obtener una predicción de $v(k)$ considérese primero su modelo

$$v_i(k) = H_i(q)e(k) = [H_i(q) - h_i(0)]e(k) + h_i(0)e(k)$$

Suponiendo que se pueden medir los valores pasados de $e(k)$, para así estimar el valor actual del ruido $v(k)$, el predictor resulta

$$\hat{v}_i(k) = [H_i(q) - h_i(0)]e(k)$$

ya que depende sólo de valores pasados de $e(k)$. Supóngase ahora que no se puede medir $e(k)$, pero en cambio se tiene acceso a $v(k)$. Entonces, el predictor del ruido sería

$$\hat{v}_i(k) = [H_i(q) - h_i(0)]H(q)^{-1}v(k)$$

Sin embargo, $H(q)$ no se conoce, por lo que la mejor suposición que se puede hacer es $H_i = H$, resultando

$$\hat{v}_i(k) = [1 - h_i(0)H_i(q)^{-1}]v(k)$$

Supóngase ahora que $v(k)$ no se puede medir directamente, y que en cambio se tiene acceso sólo a $u(k)$ e $y(k)$. Entonces el predictor de ruido resultaría

$$\hat{v}_i(k) = [1 - h_i(0)H_i(q)^{-1}][y(k) - G(q)u(k)]$$

Dado que se desconoce $G(q)$, la mejor suposición que se puede hacer es $G_i = G$, por lo que el predictor de ruido resulta

$$\hat{v}_i(k) = [1 - h_i(0)H_i(q)^{-1}][y(k) - G_i(q)u(k)] \quad (1.19)$$

Finalmente, el predictor asociado a (1.17) se obtiene al reemplazar (1.19) en (1.18), resultando

$$\begin{aligned} \hat{y}_i(k) &= \bar{H}_i(q)^{-1}G_i(q)u(k) + [1 - \bar{H}_i(q)^{-1}]y(k) \\ \bar{H}_i(q) &= \frac{H_i(q)}{h_i(0)} \end{aligned}$$

o en forma más compacta

$$\begin{aligned} \hat{y}_i(k) &= W_i(q)\zeta_{uy}(k) \\ W_i(q) &= \begin{bmatrix} \bar{H}_i(q)^{-1}G_i(q) & 1 - \bar{H}_i(q)^{-1} \end{bmatrix} \\ \zeta_{uy}(k) &= \begin{bmatrix} u(k) \\ y(k) \end{bmatrix} \end{aligned} \quad (1.20)$$

Nota 1.2 Obsérvese que el predictor (1.20) requiere que el filtro $H_i(q)$ sea invertible. Es por esto que se debe tener $h_i(0) \neq 0$, es decir, $H_i(q)$ debe ser un filtro bipropio, independientemente de si $H(q)$ lo es o no.

Nota 1.3 Obsérvese también que si

$$\bar{H}_i(q)^{-1} = P(q)A_i(q), \quad G_i(q) = \frac{B_i(q)}{A_i(q)}$$

donde $P(q)$, $A_i(q)$ y $B_i(q)$ son polinomios, entonces el predictor (1.20) define una regresión lineal. Este es el caso de la estructura ARX, donde

$$\hat{y}_i(k) = \omega_{uy}^{\hat{n}_b, \hat{n}_a}(k)^T \gamma_i \quad (1.21)$$

$$\omega_{uy}^{\hat{n}_b, \hat{n}_a}(k) = \begin{bmatrix} u(k) \\ \vdots \\ u(k - \hat{n}_b) \\ y(k - 1) \\ \vdots \\ y(k - \hat{n}_a) \end{bmatrix}, \quad \gamma_i = \begin{bmatrix} b_0^i \\ \vdots \\ b_{\hat{n}_b}^i \\ -a_1^i \\ \vdots \\ -a_{\hat{n}_a}^i \end{bmatrix}$$

$$B_i(q) = b_0^i + b_1^i q^{-1} + \dots + b_{\hat{n}_b}^i q^{-\hat{n}_b}$$

$$A_i(q) = 1 + a_1^i q^{-1} + \dots + a_{\hat{n}_a}^i q^{-\hat{n}_a}$$

$$\bar{H}_i(q) = \frac{1}{A_i(q)}$$

1.3. Señales eficaces para identificación

A veces es necesario someter un sistema a experimentación, a fin de modelar su comportamiento. Para lograr lo anterior, se debe diseñar una actuación con la suficiente riqueza, para que así el sistema entregue una respuesta con información importante acerca de él. Las siguientes secciones tratan estas propiedades de señales, para el diseño de experimentos.

1.3.1. Excitación local y excitación persistente

Definición 1.10 Sea la señal $s(k)$, definida al menos $\forall k \in [-M - p + 1, M]$. Entonces $s(k)$ es "localmente excitante" (l.e.) de orden p en el intervalo $\forall k \in [-M, M]$, si la matriz

$$\Pi_s^M(p) := \frac{1}{2M + 1} \sum_{k=-M}^M v_s^p(k) v_s^p(k)^T$$

es definida positiva, donde $2M + 1 \geq p$ y

$$v_s^p(k)^T = [s(k) \quad \cdots \quad s(k-p+1)]$$

Lema 1.1 Sea la señal $s(k)$, definida al menos $\forall k \in [-M-p+1, M]$. Entonces $s(k)$ es l.e. de orden p $\forall k \in [-M, M]$ si y sólo si para cualquier filtro

$$M_p(q) = m_0 + m_1 + \cdots + m_{p-1}q^{-p+1}$$

se cumple la siguiente implicancia

$$\left(\forall k \in [-M, M], M_p(q)s(k) = 0 \right) \Rightarrow M_p(q) = 0$$

Demostración. Es claro que $m^T \Pi_s^M(p)m \geq 0$ para todo $m \in \mathbb{R}^{p-1}$, donde $m = [m_0 \quad \cdots \quad m_{p-1}]$. Además se cumplen las siguientes equivalencias

$$\begin{aligned} m^T \Pi_s^M(p)m &= 0 \\ \Leftrightarrow \left(\forall k \in [-M, M], m^T v_s^p(k) = 0 \right) \\ \Leftrightarrow \left(\forall k \in [-M, M], M_p(q)s(k) = 0 \right) \end{aligned}$$

Finalmente $m = 0 \Leftrightarrow M_p(q) = 0$, lo que completa la demostración. **QED.**

Corolario 1.1 Si la señal $s(k)$ es l.e. de orden p en el intervalo $[-M, M]$, entonces $s(k)$ es l.e. de orden $n < p$ en el intervalo $[-M, M]$.

El lema 1.1 señala que $s(k)$ es l.e. de orden p en el intervalo $[-M, M]$, si aún después de filtrarla con cualquier filtro FIR de $p-1$ ceros (o menos), se puede obtener una respuesta diferente de cero dentro del intervalo finito $[-M, M]$, es decir, si la señal $s(k)$ resiste localmente (en el intervalo $-M \leq k \leq M$) la acción de un filtro FIR de $p-1$ ceros. Esto da algunas indicaciones para saber qué señales pueden ser l.e. de orden p en el intervalo $[-M, M]$. Por ejemplo, se puede deducir que es necesario que $s(k)$ tenga al menos p componentes espectrales; sin embargo, esta condición no es suficiente.

Ejemplo 1.1 Supóngase que se desea encontrar una señal $s(k)$ que sea l.e. de orden 4 en el intervalo $[-2, 2]$, y supóngase que se propone la siguiente señal

$$s(k) = 1 + \cos(\theta_0 k)$$

definida $\forall k \in \mathbb{Z}$, que tiene 3 componentes espectrales. Claramente, esta señal $s(k)$ no es l.e. de orden 4 en

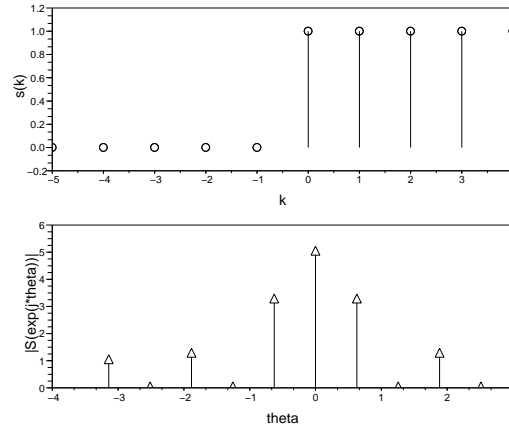


Figura 1.3: Señal cuadrada $s(k)$ de periodo 10 y su espectro $|S(e^{j\theta})|$.

el intervalo $[-2, 2]$, ya que existe un filtro FIR (no nulo) de 4-1 ceros

$$M_4(q) = \frac{(q-1)(q-e^{j\theta_0})(q-e^{-j\theta_0})}{q^3}$$

que es capaz de filtrar completamente la señal $s(k)$ en el intervalo $[-2, 2]$, es decir

$$\forall k \in [-2, 2], \quad M_4(q)s(k) = 0$$

Supóngase que se propone la señal $s(k)$ de la figura 1.3, la cual es una señal cuadrada, de periodo 10, definida $\forall k \in \mathbb{Z}$. Claramente esta señal tiene más de 4 componentes espectrales. Sin embargo, tampoco es l.e. de orden 4 en el intervalo $[-2, 2]$, ya que el filtro $M_4(q) = q^{-3}$ es capaz de eliminarla completamente en dicho intervalo.

Ahora, sólo basta encontrar algunas señales que sean l.e. de orden p en el intervalo $[-M, M]$. Así, el siguiente teorema indica qué condiciones son suficientes para un tipo específico de señales, de manera que sean l.e. de orden p sobre un intervalo de tiempo finito.

Teorema 1.2 Considérese la señal $u(k)$ definida para todo $k \in \mathbb{Z}$

$$u(k) = \sum_{i=1}^p U_i e^{j\theta_i k}, \quad \theta_\ell - \theta_m = \frac{2\pi n}{2M+1}, \quad n \in \mathbb{N}$$

donde $\forall i, U_i \neq 0, \theta_i \in]-\pi, \pi]$ y $\ell, m \in [1, p]$. Entonces la señal $s(k) = C(q)^{-1}u(k)$

$$C(q) = c_0 + c_1 q^{-1} + \cdots + c_{n_c} q^{-n_c}$$

es l.e. de orden $p \forall k \in [-M, M]$.

Demostración. Dado que $u(k)$ está definida $\forall k \in \mathbb{Z}$, entonces

$$M_p(q)s(k) = \frac{M_p(q)}{C(q)}u(k) = \sum_{i=1}^p \frac{M_p(e^{j\theta_i})}{C(e^{j\theta_i})}U_i e^{j\theta_i k}$$

es válido $\forall k \in \mathbb{Z}$, y en particular válido $\forall k \in [-M, M]$. Note ahora que las señales $e^{j\theta_\ell k}$ y $e^{j\theta_m k}$ son linealmente independientes (l.i.) en el intervalo $[-M, M]$. Esto se puede ver teniendo en cuenta el siguiente producto¹, con $\theta_\ell - \theta_m = (2\pi n)/(2M + 1)$

$$\begin{aligned} \langle e^{j\theta_\ell k}, e^{j\theta_m k} \rangle &:= \frac{1}{2M + 1} \sum_{k=-M}^M e^{j\theta_\ell k} (e^{j\theta_m k})^* \\ &= \frac{\text{sen}((\theta_\ell - \theta_m)(2M + 1)/2)}{(2M + 1) \text{sen}((\theta_\ell - \theta_m)/2)} \\ &= \delta_K(n) = 0 \end{aligned}$$

Dado que las señales $e^{j\theta_i k}$, $\theta_\ell - \theta_m = (2\pi n)/(2M + 1)$, son l.i. en el intervalo $[-M, M]$, entonces

$$\sum_{i=1}^p \left(\frac{M_p(e^{j\theta_i})}{C(e^{j\theta_i})} U_i \right) e^{j\theta_i k} = 0 \Leftrightarrow \forall i, \frac{M_p(e^{j\theta_i})}{C(e^{j\theta_i})} U_i = 0$$

pero $\forall i = 1, \dots, p$ $U_i \neq 0$, lo que significa que $\forall i = 1, \dots, p$ $M_p(e^{j\theta_i}) = 0$. Ya que $M_p(q)$ tiene a lo más $p - 1$ ceros sobre la circunferencia unitaria, entonces la única posibilidad de lograr $\forall i, M_p(e^{j\theta_i}) = 0$ es que $M_p(q) = 0$. **QED.**

El teorema anterior indica que las señales $u(k)$ que tienen p componentes espectrales diferentes, separadas entre sí por una distancia que sea múltiplo de $2\pi/(2M + 1)$, son l.e. de orden p en el intervalo $[-M, M]$, y además mantienen esta propiedad aún si son filtradas por un sistema sin ceros, como es el caso de $s(k) = C(q)^{-1}u(k)$. Algunos ejemplos de éstas señales son: secuencias binarias pseudo aleatorias (PRBS), cuyo periodo sea $2M + 1$, o simplemente señales compuestas por sumas de senos y/o cosenos, como por ejemplo

$$u(k) = \sum_{i=\text{par}(p)}^{\text{ent}(p/2)} A_i \cos(\theta_i k - \phi_i), \quad \theta_i = \frac{2\pi i}{2M + 1} \quad (1.22)$$

donde

$$\begin{aligned} \text{par}(p) & \text{ es } 1 \text{ si } p \text{ es par, } 0 \text{ si } p \text{ es impar} \\ \text{ent}(p/2) & \text{ es la parte entera de } p/2 \end{aligned}$$

¹donde $\delta_K(n)$ es la función delta de Kronecker

p	$\det \Pi_u^M(p_0)$	p	$\det \Pi_u^M(p_0)$
1	0	6	2,50e-01
2	-5,68e-49	7	1,83e+00
3	1,82e-32	8	1,55e+01
4	1,25e-19	9	1,76e+01
5	1,10e-03	10	8,58e+01

Tabla 1.1: Singularidad de la matriz $\Pi_u^M(p_0)$, para $u(k)$ con $p = 1, \dots, 10$ componentes espectrales.

la que tiene exactamente p componentes espectrales.

Ejemplo 1.2 Se verificará en forma secuencial si las señales $u(k)$ con $p = 1, \dots, 2p_0$, de la ecuación (1.22), son l.e. de orden $p_0 = 5$ en el intervalo $[-M, M]$.

Para que $u(k)$ pueda tener $2p_0$ componentes espectrales, el periodo fundamental de $u(k)$ debe ser tal que $2M + 1 \geq 2p_0$, es por esto que se elige $M = p_0$. Además, para poder construir la matriz de persistencia $\Pi_u^M(p_0)$ se deben tener datos de $u(k)$ al menos en el intervalo $[-M - p_0 + 1, M]$, por lo que el número total de muestras necesarias es $2M + p_0$.

Con los datos anteriores se construye secuencialmente la matriz $\Pi_u^M(p_0)$, de tamaño $p_0 \times p_0$, para las diferentes señales $u(k)$ con $p = 1, \dots, 10$, cuyas amplitudes se eligen $\forall i, A_i = 1$. De esta forma se obtiene la tabla 1.1, donde se puede ver que si el número de componentes espectrales es $p \geq p_0 = 5$, entonces $\Pi_u^M(p_0)$ es no singular, y por lo tanto las señales son l.e. de orden $p_0 = 5$ en el intervalo $[-5, 5]$. Para más detalles, véase el programa F.2.1.

Puede ocurrir que existan señales que sean l.e. de orden p para el intervalo $[-M, M]$, pero que dejen de serlo cuando $M \rightarrow \infty$, lo que se muestra en el siguiente ejemplo.

Ejemplo 1.3 Considérese la señal $s(k) = \delta_K(k)$, definida $\forall k \in \mathbb{Z}$, la cual es l.e. de orden p en el intervalo $[-M, M]$ si se cumple que $(p - 1) \leq M$. Claramente, si $M \rightarrow \infty$ entonces se tiene que la matriz de excitación resulta

$$\lim_{M \rightarrow \infty} \Pi_s^M(p) = O_p$$

donde O_p es la matriz nula, de dimensión $p \times p$. De esta manera la señal $s(k)$ no es l.e. de orden p en el intervalo $]-\infty, \infty[$.

Este último ejemplo lleva a separar el caso de M finito del caso M infinito, y es por esto que se plantea la siguiente definición

Definición 1.11 La señal $s(k)$ es “persistentemente excitante” (p.e.) de orden p , si la matriz $\Pi_s^\infty(p)$ es definida positiva, o de manera equivalente, si para todo filtro de la forma

$$M_p(q) = m_0 + m_1 q + \dots + m_{p-1} q^{-p+1}$$

se cumple la siguiente implicancia

$$\forall \theta, |M_p(e^{j\theta})|^2 \Phi_{ss}^\infty(e^{j\theta}) = 0 \Rightarrow \forall \theta, M_p(e^{j\theta}) = 0$$

donde $\Phi_{ss}^\infty(e^{j\theta})$ corresponde a la transformada de Fourier de la autocorrelación muestral $\varphi_{ss}^\infty(\tau)$.

Esta última corresponde a la definición de persistencia de excitación ya conocida en la literatura. Para una demostración de esta equivalencia véase [14, lemma 13.1] o la demostración complementaria E.2.

Corolario 1.2 Si $s(k)$ es p.e. de orden p , entonces $s(k)$ es p.e. de orden $n < p$.

Suponga que la autocorrelación muestral $\varphi_{ss}^\infty(\tau)$ es de la siguiente forma

$$\varphi_{ss}^\infty(\tau) = \sum_{i=1}^{\ell} C_i e^{j\theta_i \tau}$$

entonces el concepto anterior puede interpretarse como la mínima cantidad de componentes espectrales que debe tener $\varphi_{ss}^\infty(\tau)$, para que al pasar por el filtro $M_p(z)M_p(z^{-1})$ se tenga salida diferente de cero. Así, $\varphi_{ss}^\infty(\tau)$ debe tener al menos $\ell \geq p$ componentes espectrales, para ser p.e. de orden p , ya que el filtro $M_p(z)M_p(z^{-1})$ tiene a lo más $p - 1$ ceros sobre la circunferencia unitaria.

Otros ejemplos de señales p.e. son las que cumplen $\varphi_{ss}^\infty(\tau) = a^2 \delta_K(\tau)$, con $a > 0$. Este tipo de señales son p.e. de cualquier orden, lo que se puede ver construyendo la matriz de excitación

$$\begin{aligned} \Pi_s^\infty(p) &= \lim_{M \rightarrow \infty} \frac{1}{2M+1} \sum_{k=-M}^M v_s^p(k) v_s^p(k)^T \\ v_s^p(k)^T &= [s(k) \quad \dots \quad s(k-p+1)] \\ \Pi_s^\infty(p) &= \begin{bmatrix} \varphi_{ss}^\infty(0) & \dots & \varphi_{ss}^\infty(p-1) \\ \vdots & \ddots & \vdots \\ \varphi_{ss}^\infty(1-p) & \dots & \varphi_{ss}^\infty(0) \end{bmatrix} \end{aligned}$$

donde $\varphi_{ss}^\infty(\tau)$ fue definida en (1.2), resultando que $\Pi_s^\infty(p) = a^2 I$ para todo p , y por lo tanto es siempre definida positiva.

Teorema 1.3 Sea la señal $s(k) = C(q)^{-1}u(k)$ donde

$$C(q) = c_0 + c_1 q^{-1} + \dots + c_{n_c} q^{-n_c}$$

Si $u(k)$ es p.e. de orden p , entonces $s(k)$ es p.e. de orden p .

Véase la demostración complementaria E.3.

Nótese que el teorema anterior es bastante importante, ya que en el caso asintótico $M \rightarrow \infty$, cualquier señal $u(k)$ que sea p.e. de orden p no pierde esta propiedad al filtrarse por alguna transferencia sin ceros.

1.3.2. Suficiencia de información

Caso sin ruido

Considérese que se toman los siguientes datos, provenientes del sistema en (1.10)

$$Z_{ux}^{L,M} = \begin{bmatrix} u(-M-L) & \dots & u(M) \\ x(-M-L) & \dots & x(M) \end{bmatrix}$$

Este conjunto de datos debería contener la suficiente información, de forma que dos predictores diferentes, asociados a este sistema, entreguen respuestas diferentes. Es por esto que se introduce la siguiente definición

Definición 1.12 El conjunto de datos $Z_{ux}^{L,M}$ es “suficientemente informativo” (s.i.) para la estructura de los predictores en (1.14), si para todo par de predictores de dicha estructura

$$\begin{aligned} \hat{x}_1(k) &= W_1(q) \zeta_{ux}(k) \\ \hat{x}_2(k) &= W_2(q) \zeta_{ux}(k) \end{aligned}$$

se cumple

$$\overline{\langle [\hat{x}_1(k) - \hat{x}_2(k)]^2 \rangle}_{k=-M}^M = 0 \Rightarrow W_1(q) = W_2(q)$$

con $L \geq \max(\hat{n}_b, \hat{n}_a)$.

Esta definición es análoga a [14, definition 8.1], pero restringida a sistemas sin ruido.

Nota 1.4 La definición anterior puede escribirse de forma más compacta como: $Z_{ux}^{L,M}$ es suficientemente informativo, si $\forall \tilde{W}(q)$ de la estructura (1.14) se cumple

$$\overline{\langle \tilde{x}(k)^2 \rangle}_{k=-M}^M = 0 \Rightarrow \tilde{W}(q) = 0$$

donde

$$\tilde{x} = \hat{x}_1 - \hat{x}_2, \quad \tilde{W} = W_1 - W_2$$

con $L \geq \max(\hat{n}_b, \hat{n}_a)$.

El concepto de suficiencia de información es difícil de verificar, considerando la manera en que ha sido definido, pues se debería probar con todos los predictores de una misma estructura, para saber si los datos son suficientemente informativos para la estructura en cuestión. Es por esta razón que, mediante el siguiente teorema, se propone una forma equivalente de verificar esta condición, la cual es simple de implementar computacionalmente.

Lema 1.2 El conjunto de datos $Z_{ux}^{L,M}$ es “suficientemente informativo” (s.i.) para la estructura de los predictores en (1.14) si y sólo si la matriz

$$\Lambda_{xux}^M(\hat{n}_b, \hat{n}_a) = \frac{1}{2M+1} \sum_{k=-M}^M \omega_{ux}^{\hat{n}_b, \hat{n}_a}(k) \omega_{ux}^{\hat{n}_b, \hat{n}_a}(k)^T \quad (1.23)$$

es definida positiva, donde $\omega_{ux}^{\hat{n}_b, \hat{n}_a}(k)$ está definido en (1.15) y $L \geq \max(\hat{n}_b, \hat{n}_a)$.

Para el caso de Z_{ux}^∞ , es decir cuando $M \rightarrow \infty$, el resultado es el mismo. La demostración es idéntica a la del caso finito, mostrada a continuación.

Demostración. Este teorema se demuestra a partir de la siguiente equivalencia

$$\begin{aligned} \overline{\langle \tilde{x}(k)^2 \rangle}_{k=-M}^M &= \tilde{\gamma}^T \overline{\langle \omega_{ux}^{\hat{n}_b, \hat{n}_a}(k) \omega_{ux}^{\hat{n}_b, \hat{n}_a}(k)^T \rangle}_{k=-M}^M \tilde{\gamma} \\ &= \tilde{\gamma}^T \Lambda_{xux}^M(\hat{n}_b, \hat{n}_a) \tilde{\gamma} \geq 0 \end{aligned}$$

donde se ha considerado la representación de regresión lineal en (1.15), con

$$\tilde{x} = \hat{x}_1 - \hat{x}_2, \quad \tilde{\gamma} = \gamma_1 - \gamma_2$$

Además, se debe tener en cuenta que, para todo predictor con la misma estructura que (1.14), se puede establecer una relación 1-1 con cualquier vector de $\mathbb{R}^{\hat{n}_b+1+\hat{n}_a}$

$$W_i(q) \leftrightarrow \gamma_i$$

Con esto se tiene que para todo predictor $\tilde{W}(q)$, de la estructura ya mencionada

$$\begin{aligned} \overline{\langle \tilde{x}(k)^2 \rangle}_{k=-M}^M = 0 &\Leftrightarrow \tilde{\gamma}^T \Lambda_{xux}^M(\hat{n}_b, \hat{n}_a) \tilde{\gamma} = 0 \\ \tilde{W}(q) = 0 &\Leftrightarrow \tilde{\gamma} = 0 \end{aligned}$$

es decir, si $Z_{ux}^{L,M}$ es s.i. para la estructura (1.14), entonces $\Lambda_{xux}^M(\hat{n}_b, \hat{n}_a)$ es definida positiva, y viceversa. **QED.**

Así, la suficiencia de información de los datos $Z_{ux}^{L,M}$, para la estructura en (1.14), está estrechamente ligada a la condición de no singularidad de la matriz $\Lambda_{xux}^M(\hat{n}_b, \hat{n}_a)$, la cual permite una cómoda implementación computacional.

Considérese ahora el conjunto de mediciones $Z_{ux}^{L,M}$ obtenidas al someter la planta en (1.10) a un experimento, cuya señal $u(k)$ ha sido diseñada previamente. Entonces cabe hacerse la siguiente pregunta: ¿Se puede diseñar $u(k)$, de forma que al aplicarla a la planta en (1.10), se obtenga un conjunto de datos $Z_{ux}^{L,M}$ que sea s.i. para la estructura de predictores en (1.14)?

Teorema 1.4 Sea $Z_{ux}^{L,M}$ el conjunto de datos generado por el sistema de la ecuación (1.10), donde $L \geq \max(\hat{n}_b, \hat{n}_a)$. Si la señal $s(k) = A(q)^{-1}u(k)$ es l.e. de orden $\max(\hat{n}_b + n_a, n_b + \hat{n}_a) + 1 \forall k \in [-M, M]$, entonces la matriz $\Lambda_{xux}^M(\hat{n}_b, \hat{n}_a)$, de la ecuación (1.23), es definida positiva para $\min(\hat{n}_b - n_b, \hat{n}_a - n_a) \leq 0$, donde $2M + 1 \geq \max(\hat{n}_b + n_a, n_b + \hat{n}_a) + 1$.

Para el caso de Z_{ux}^∞ , es decir cuando $M \rightarrow \infty$, el resultado es el mismo, pero l.e. debe ser reemplazado por p.e. La demostración para el caso infinito se puede consultar en la demostración complementaria E.4.

Demostración. (Caso finito). La demostración es por contradicción. Supóngase que $\Lambda_{xux}^M(\hat{n}_b, \hat{n}_a)$ no es definida positiva, entonces existe un vector no nulo $\gamma_p = [d_0 \ \cdots \ d_{\hat{n}_b} \ c_1 \ \cdots \ c_{\hat{n}_a}]^T$ tal que

$$\gamma_p^T \Lambda_{xux}^M(\hat{n}) \gamma_p = 0 \Leftrightarrow \forall k \in [-M, M], \omega_{ux}^{\hat{n}}(k)^T \gamma_p = 0$$

donde el lado derecho puede reescribirse como

$$\forall k \in [-M, M], \quad [D(q)A(q) + C(q)B(q)] \frac{u(k)}{A(q)} = 0$$

$$D(q) = \sum_{i=0}^{\hat{n}_b} d_i q^{-i}, \quad C(q) = \sum_{i=1}^{\hat{n}_a} c_i q^{-i}$$

Se puede notar que $[D(q)A(q) + C(q)B(q)] \neq 0$, ya que las siguientes estructuras son siempre diferentes para $\min(\hat{n}_b - n_b, \hat{n}_a - n_a) \leq 0$

$$\frac{d_0 q^{-0} + \dots + d_{\hat{n}_b} q^{-\hat{n}_b}}{c_1 q^{-1} + \dots + c_{\hat{n}_a} q^{-\hat{n}_a}} \neq \frac{b_0 + b_1 q^{-1} + \dots + b_{n_b} q^{-n_b}}{1 + a_1 q^{-1} + \dots + a_{n_a} q^{-n_a}}$$

donde $B(q)$ y $A(q)$ son coprimos. Finalmente se tiene que $[D(q)A(q) + C(q)B(q)] \neq 0$ es una contradicción, ya que $s(k)$ es l.e. de orden $\max(\hat{n}_b + n_a, n_b + \hat{n}_a) + 1 \forall k \in [-M, M]$. **QED.**

Corolario 1.3 Sea $Z_{ux}^{L,M}$ el conjunto de datos generado por el sistema de la ecuación (1.10), donde $L \geq \max(\hat{n}_b, \hat{n}_a)$. Si $s(k) = A(q)^{-1}u(k)$ es l.e. de orden $\max(\hat{n}_b + n_a, n_b + \hat{n}_a) + 1 \forall k \in [-M, M]$, entonces $Z_{ux}^{L,M}$ es s.i. para la estructura en (1.14), con $\min(\hat{n}_b - n_b, \hat{n}_a - n_a) \leq 0$.

Nota 1.5 Obsérvese que a partir del conjunto de datos $Z_{ux}^{L,M}$, con $L \geq \max(\hat{n}_b, \hat{n}_a)$, es posible obtener la señal $s(k) \forall k \in [-M - p + 1, M]$, donde $p = \max(\hat{n}_b + n_a, n_b + \hat{n}_a) + 1$, la cual es una señal intermedia entre la entrada $u(k)$ y la salida $x(k)$

$$x(k) = B(q)s(k), \quad s(k) = A(q)^{-1}u(k)$$

Esta señal $s(k) \forall k \in [-M - p + 1, M]$ se puede calcular a partir del siguiente sistema de ecuaciones

$$\begin{aligned} \forall k \in [-M - \hat{n}_b, M], \quad u(k) &= A(q)s(k) \\ \forall k \in [-M - \hat{n}_a, -M - i - 1], \quad x(k) &= B(q)s(k) \end{aligned}$$

donde $i = \min(\hat{n}_b - n_b, \hat{n}_a - n_a)$, el que se puede reescribir en términos de la matriz de Sylvester de tamaño $(2M + p) \times (2M + p)$, definida en el teorema D.12

$$\begin{aligned} v_{ux} &= \mathcal{S}(A, q^{-2M-i-1}B) v_s \\ v_s &= \begin{bmatrix} s(M) \\ \vdots \\ s(-M - p + 1) \end{bmatrix} \\ v_{ux} &= \begin{bmatrix} u(M) \\ \vdots \\ u(-M - \hat{n}_b) \\ x(-M - i - 1) \\ \vdots \\ x(-M - \hat{n}_a) \end{bmatrix} \end{aligned}$$

Este sistema tiene una única solución, ya que la matriz de Sylvester $\mathcal{S}(A, q^{-2M-i-1}B)$ es de rango completo, por el teorema D.12, debido a que $A(q)$ y $B(q)$ son coprimos.

El corolario 1.3 muestra que si la entrada $u(k)$ se diseña de forma que la señal $s(k)$ sea l.e. de orden $\max(\hat{n}_b + n_a, n_b + \hat{n}_a) + 1 \forall k \in [-M, M]$, entonces al aplicar $u(k)$ a la planta en (1.10) se obtienen datos $Z_{ux}^{L,M}$ que son s.i. para una estructura de predictores, cuyo numerador o denominador tienen menor grado que el numerador o denominador de la planta respectivamente, esto es, para estructuras de predictores “menores” que la planta. Mediante el siguiente teorema y su corolario, se verá qué ocurre en el caso en que la estructura de predictores es “mayor” que la planta.

Teorema 1.5 Considérese la siguiente estructura de predictores, escrita en su forma de regresión lineal

$$\hat{x}_l(k) = w_{ux}^{\hat{n}_b, \hat{n}_a, i}(k)^T \gamma_l \quad (1.24)$$

$$w_{ux}^{\hat{n}_b, \hat{n}_a, i}(k) := \begin{bmatrix} u(k) \\ \vdots \\ u(k - \hat{n}_b) \\ x(k - i - 1) \\ \vdots \\ x(k - \hat{n}_a) \end{bmatrix}, \quad \gamma_l = \begin{bmatrix} b_0^l \\ \vdots \\ b_{\hat{n}_b}^l \\ -a_{i+1}^l \\ \vdots \\ -a_{\hat{n}_a}^l \end{bmatrix}$$

donde $i = \min(\hat{n}_b - n_b, \hat{n}_a - n_a) \geq 0$. Considérese además el conjunto de datos $Z_{ux}^{L,M}$ generado por el sistema (1.10), donde $L \geq \max(\hat{n}_b, \hat{n}_a)$. Entonces la matriz $p \times p$, con $p := \max(\hat{n}_b + n_a, n_b + \hat{n}_a) + 1$

$$L_{ux}^M(\hat{n}_b, \hat{n}_a, i) := \frac{1}{2M + 1} \sum_{k=-M}^M w_{ux}^{\hat{n}_b, \hat{n}_a, i}(k) w_{ux}^{\hat{n}_b, \hat{n}_a, i}(k)^T$$

es definida positiva si y sólo si $s(k) = A(q)^{-1}u(k)$ es l.e. de orden $p \forall k \in [-M, M]$, donde $2M + 1 \geq p$.

Para el caso de Z_{ux}^∞ , es decir cuando $M \rightarrow \infty$, el resultado es el mismo, pero l.p.e. debe ser reemplazado por p.e. La demostración es idéntica a la del caso finito, mostrada a continuación.

Demostración. El vector $w_{ux}^{\hat{n}_b, \hat{n}_a, i}(k)$ puede ser escrito como

$$w_{ux}^{\hat{n}_b, \hat{n}_a, i}(k) = \frac{1}{A(q)} \begin{bmatrix} A(q)u(k) \\ \vdots \\ A(q)u(k - \hat{n}_b) \\ B(q)u(k - i - 1) \\ \vdots \\ B(q)u(k - \hat{n}_a) \end{bmatrix}$$

el cual puede ser representado en términos de la matriz de Sylvester del teorema D.12, para los polinomios $A(q)$ y $q^{-i-1}B(q)$

$$w_{ux}^{\hat{n}_b, \hat{n}_a, i}(k) = \mathcal{S}(A, q^{-i-1}B)v_s^p(k)$$

donde la matriz $\mathcal{S}(A, q^{-i-1}B)$ es $p \times p$ y no singular por el teorema D.12, y además

$$v_s^p(k) := \begin{bmatrix} s(k) \\ \vdots \\ s(k - p + 1) \end{bmatrix}, \quad s(k) := \frac{u(k)}{A(q)}$$

En base a lo anterior se puede obtener el siguiente producto

$$L_{ux}^M(\hat{n}_b, \hat{n}_a, i) = \mathcal{S}(A, q^{-i-1}B)\Pi_s^M(p)\mathcal{S}(A, q^{-i-1}B)^T$$

donde $\Pi_s^M(p)$ es la misma que en la definición 1.11. Así, la demostración sigue a partir del teorema D.12 y de la definición 1.11. **QED.**

Corolario 1.4 Sea $Z_{ux}^{L,M}$ el conjunto de datos generado por el sistema de la ecuación (1.10), donde $L \geq \max(\hat{n}_b, \hat{n}_a)$. La señal $s(k) = A(q)^{-1}u(k)$ es l.e. de orden $\max(\hat{n}_b + n_a, n_b + \hat{n}_a) + 1 \forall k \in [-M, M]$ si y sólo si $Z_{ux}^{L,M}$ es s.i. para la estructura en (1.24), con $i = \min(\hat{n}_b - n_b, \hat{n}_a - n_a) \geq 0$.

Finalmente, este corolario muestra que si la entrada $u(k)$ se diseña de forma que la señal $s(k)$ sea l.e. de orden $\max(\hat{n}_b + n_a, n_b + \hat{n}_a) + 1 \forall k \in [-M, M]$, entonces al aplicar $u(k)$ a la planta en (1.10) se obtienen datos $Z_{ux}^{L,M}$ que son s.i. para la estructura de predictores (1.24), cuyo numerador o denominador tienen mayor grado que el numerador o denominador de la planta respectivamente.

Ejemplo 1.4 Considérese el siguiente sistema generador de datos o planta $x(k) = G(q)u(k)$

$$G(q) = \frac{0,3}{(q-0,8e^{j\pi/4})(q-0,8e^{-j\pi/4})(q-0,7)}$$

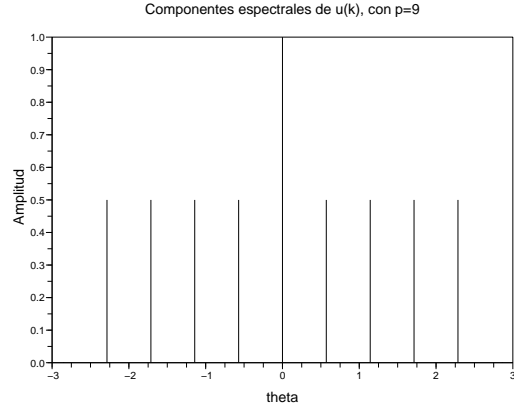


Figura 1.4: Componentes espectrales de la señal $u(k)$, con $p = 9$.

para el cual $n_b = n_a = 3$. Supóngase ahora que a este sistema se aplica la señal (1.22), de periodo $2M + 1$, con $p = 9$ y $\forall i, A_i = 1$, es decir, formada por 9 componentes espectrales, las cuales se muestran en la figura 1.4.

Ahora, se desea verificar para cuáles estructuras de modelos, con $1 \leq (\hat{n}_b, \hat{n}_a) \leq 7$, los datos $Z_{ux}^{L,M}$ son suficientemente informativos. Para ello se necesita primero definir los valores de M y L . Dado que se desea probar hasta $\hat{n}_b = \hat{n}_a = 7$, y que se deben cumplir las siguientes condiciones para los teoremas 1.4 y 1.5

$$\begin{aligned} 2M + 1 &\geq \max(\hat{n}_b + n_a, n_b + \hat{n}_a) + 1 \\ L &\geq \max(\hat{n}_b, \hat{n}_a) \end{aligned}$$

entonces se eligen $M = 5$ y $L = 7$.

De esta manera se mide el conjunto de datos $Z_{ux}^{L,M}$, compuesto de $2M + 1 + L$ pares de datos en estado estacionario.

Haciendo primero un análisis teórico, mediante el teorema 1.4 se tiene que $\Lambda_{xux}^M(\hat{n}_b, \hat{n}_a)$ es no singular para todo (\hat{n}_b, \hat{n}_a) tal que

$$\begin{aligned} \min(\hat{n}_b - n_b, \hat{n}_a - n_a) &\leq 0 \\ \max(\hat{n}_b + n_a, n_b + \hat{n}_a) + 1 &\leq 9 \end{aligned}$$

pues $s(k) = A(q)^{-1}u(k)$ es l.e. de orden 9 en el intervalo $[-M, M]$, mostrándose esta situación con color negro en la tabla 1.2 izquierda, y en blanco, los sectores no cubiertos por este teorema. Ahora, mediante el teorema 1.5 se puede notar que $L_{ux}^M(\hat{n}_b, \hat{n}_a, i)$ es no

		\hat{n}_a									\hat{n}_a						
		1	2	3	4	5	6	7			1	2	3	4	5	6	7
\hat{n}_b	1	Λ	Λ	Λ	Λ	Λ			1	2e00	3e-01	6e-02	2e-02	9e-03	6e-03	4e-03	
	2	Λ	Λ	Λ	Λ	Λ			2	1e00	3e-01	6e-02	2e-02	8e-03	3e-03	2e-07	
	3	Λ	Λ	$\Lambda=L$	$\Lambda=L$	$\Lambda=L$	L	L	3	9e-01	3e-01	6e-02	1e-02	4e-03	7e-09	5e-12	
	4	Λ	Λ	$\Lambda=L$	L	L	L	L	4	6e-01	2e-01	7e-03	3e-02	9e-03	1e-08	7e-12	
	5	Λ	Λ	$\Lambda=L$	L	L	L	L	5	3e-01	8e-02	2e-03	6e-03	3e-02	3e-08	2e-11	
	6			L	L	L	L	L	6	1e-01	4e-02	2e-09	4e-09	1e-08	9e-08	4e-11	
	7			L	L	L	L	L	7	6e-07	6e-09	9e-11	3e-10	1e-10	1e-10	2e-10	

Tabla 1.2: A la izquierda: En negro las combinaciones en que $\Lambda_{xux}^M(\hat{n}_b, \hat{n}_a)$ es no singular, en azul y rojo las combinaciones en que $L_{ux}^M(\hat{n}_b, \hat{n}_a, i)$ es no singular y singular respectivamente. A la derecha: El menor valor propio de $\Lambda_{xux}^M(\hat{n}_b, \hat{n}_a)$ para $\min(\hat{n}_b - n_b, \hat{n}_a - n_a) \leq 0$ y el menor valor propio de $L_{ux}^M(\hat{n}_b, \hat{n}_a, i)$ para $\min(\hat{n}_b - n_b, \hat{n}_a - n_a) \geq 0$.

singular para todo (\hat{n}_b, \hat{n}_a) tal que

$$\begin{aligned} \min(\hat{n}_b - n_b, \hat{n}_a - n_a) &\geq 0 \\ \max(\hat{n}_b + n_a, n_b + \hat{n}_a) + 1 &\leq 9 \end{aligned}$$

mostrándose esta situación en color azul en la tabla 1.2 izquierda; también se puede notar que $L_{ux}^M(\hat{n}_b, \hat{n}_a, i)$ es singular para todo (\hat{n}_b, \hat{n}_a) tal que

$$\begin{aligned} \min(\hat{n}_b - n_b, \hat{n}_a - n_a) &\geq 0 \\ \max(\hat{n}_b + n_a, n_b + \hat{n}_a) + 1 &> 9 \end{aligned}$$

mostrándose esta situación en color rojo en la tabla 1.2 izquierda.

El análisis teórico anterior se puede contrastar con la simulación de la tabla 1.2 derecha, donde se muestra el menor valor propio de $\Lambda_{xux}^M(\hat{n}_b, \hat{n}_a)$ y de $L_{ux}^M(\hat{n}_b, \hat{n}_a, i)$, indicándose en color rojo cuando cada una de las matrices es singular, lo que se puede evidenciar por la fuerte caída que presenta el valor propio más pequeño. De esta manera se puede comprobar que la simulación coincide con el análisis teórico mostrado en la tabla izquierda, indicando que los datos $Z_{ux}^{L,M}$ son s.i. cuando las matrices $\Lambda_{xux}^M(\hat{n}_b, \hat{n}_a)$ y $L_{ux}^M(\hat{n}_b, \hat{n}_a, i)$ son no singulares. Para más detalles de la simulación, ver el programa F.2.2.

Caso con ruido

Considérese por simplicidad sólo el caso $M \rightarrow \infty$, y que para éste se miden los datos Z_{uy}^∞ , provenientes de la planta en (1.16). Al igual que en el caso sin ruido, este conjunto de datos debería contener la suficiente información, de forma que dos predictores diferentes entreguen respuestas diferentes. Es por esto que se introduce la siguiente definición

Definición 1.13 El conjunto de datos Z_{uy}^∞ es “suficientemente informativo” (s.i.) para alguna estructura definida según (1.20), si para todo par de predictores de dicha estructura

$$\begin{aligned} \hat{y}_1(k) &= W_1(q)\zeta_{uy}(k) \\ \hat{y}_2(k) &= W_2(q)\zeta_{uy}(k) \end{aligned}$$

se cumple

$$\overline{\langle [\hat{y}_1(k) - \hat{y}_2(k)]^2 \rangle}_{k \in \mathbb{Z}} = 0 \Rightarrow W_1(e^{j\theta}) = W_2(e^{j\theta})$$

La definición anterior corresponde a [14, definition 8.1].

A semejanza del caso sin ruido, es posible verificar la condición de suficiencia de información mediante el uso de matrices, lo que permitiría una fácil implementación computacional. Sin embargo, esta equivalencia matricial se puede obtener sólo para el caso ARX, tal como se muestra en el siguiente teorema.

Lema 1.3 El conjunto de datos Z_{uy}^∞ es s.i. para la estructura de modelos ARX, definida en (1.21), si y sólo si $\Lambda_{yuy}^\infty(\hat{n}_b, \hat{n}_a)$, en la ecuación (1.23), es definida positiva.

Demostración. Nótese que el predictor en (1.20) define también una regresión lineal, dada en (1.21). Es por esto que, para todo predictor de la estructura ARX, se puede definir una relación 1-1 con el vector de parámetros, de la siguiente manera

$$W_i(q) \leftrightarrow \gamma_i$$

así, para todo filtro ARX según (1.21), se tiene

$$W_1(e^{j\theta}) = W_2(e^{j\theta}) \Leftrightarrow \gamma_1 = \gamma_2$$

También nótese la siguiente igualdad

$$\begin{aligned} \overline{\langle \tilde{y}(k)^2 \rangle}_{k \in \mathbb{Z}} &= \tilde{\gamma}^T \overline{\langle \omega_{uy}^{\hat{n}_b, \hat{n}_a}(k) \omega_{uy}^{\hat{n}_b, \hat{n}_a}(k)^T \rangle}_{k \in \mathbb{Z}} \tilde{\gamma} \\ &= \tilde{\gamma}^T \Lambda_{yuy}^\infty(\hat{n}_b, \hat{n}_a) \tilde{\gamma} \geq 0 \end{aligned}$$

donde $\tilde{y} = \hat{y}_1 - \hat{y}_2$ y $\tilde{\gamma} = \gamma_1 - \gamma_2$. Con esto, se tiene finalmente que, para todo predictor ARX según (1.21)

$$\begin{aligned} \overline{\langle \tilde{y}(k)^2 \rangle}_{k \in \mathbb{Z}} = 0 &\Leftrightarrow \tilde{\gamma}^T \Lambda_{yuy}^\infty(\hat{n}_b, \hat{n}_a) \tilde{\gamma} = 0 \\ W_1(e^{j\theta}) = W_2(e^{j\theta}) &\Leftrightarrow \gamma_1 = \gamma_2 \end{aligned}$$

es decir, si Z_{uy}^∞ es s.i. para la estructura ARX en (1.21), entonces $\Lambda_{yuy}^\infty(\hat{n}_b, \hat{n}_a)$ es definida positiva, y viceversa. **QED.**

Ahora, es necesario saber qué condiciones de diseño debe cumplir la señal de entrada $u(k)$, de manera que, al realizar un experimento sobre la planta (1.16) se obtenga un conjunto de datos Z_{uy}^∞ que sea s.i. para alguna estructura de predictores. Estas condiciones se muestran en el siguiente teorema.

Teorema 1.6 *Considérese la estructura de predictores según (1.20), donde*

$$G_i(q) = \frac{b_0^i + b_1^i q^{-1} + \dots + b_{\hat{n}_b}^i q^{-\hat{n}_b}}{1 + a_1^i q^{-1} + \dots + a_{\hat{n}_a}^i q^{-\hat{n}_a}}$$

$H_i(q)$ es cualquier filtro bipropio y $e(k)$ es ruido blanco estricto de media cero y varianza $a_e^2 > 0$. Si la señal $u(k)$ es p.e. de orden $\hat{n}_a + \hat{n}_b + 1$ y cumple $\varphi_{ue}^\infty(\tau) = 0$, entonces el conjunto de datos Z_{uy}^∞ es s.i. para todo predictor de la estructura ya mencionada.

Este teorema corresponde a [14, theorem 13.1]. Véase también la demostración complementaria E.5.

Nota 1.6 *Nótese que el resultado del teorema anterior es independiente de la estructura del sistema generador de datos (planta), en este caso de $G(q)$ y $H(q)$, y sólo depende de la estructura del modelo $G_i(q)$ y $H_i(q)$.*

Corolario 1.5 *Considérese la estructura ARX de la ecuación (1.21), donde $e(k)$ es ruido blanco estricto de media cero y varianza $a_e^2 > 0$. Si $u(k)$ es p.e. de orden $\hat{n}_b + \hat{n}_a + 1$ y cumple $\varphi_{ue}^\infty(\tau) = 0$, entonces la matriz $\Lambda_{yuy}^\infty(\hat{n}_b, \hat{n}_a)$, definida en (1.23), es definida positiva.*

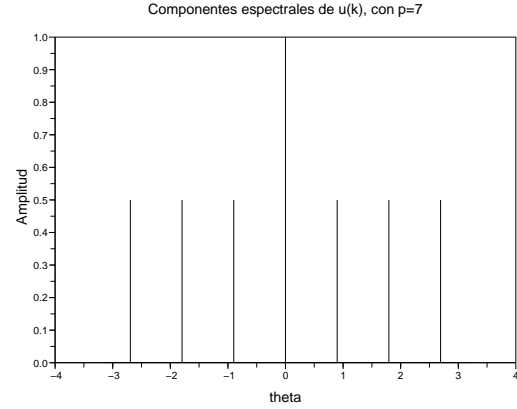


Figura 1.5: Componentes espectrales de la señal $u(k)$, con $p = 7$.

Demostración. Si la señal de entrada $u(k)$ es p.e. de orden $\hat{n}_a + \hat{n}_b + 1$, entonces Z_{uy}^∞ es s.i. para la estructura ARX en (1.21), según el teorema 1.6. Ahora, si Z_{uy}^∞ es s.i. para la estructura ARX, entonces $\Lambda_{yuy}^\infty(\hat{n}_b, \hat{n}_a)$ es definida positiva, según el lema 1.3. **QED.**

Este último teorema y su corolario traen importantes consecuencias, ya que si se utiliza una señal $u(k)$ que sea p.e. de todo orden, entonces los datos Z_{uy}^∞ son s.i. para cualquier estructura ARX, de cualquier orden, es decir, que la matriz $\Lambda_{yuy}^\infty(\hat{n}_b, \hat{n}_a)$ sería definida positiva $\forall(\hat{n}_b, \hat{n}_a)$.

Ejemplo 1.5 *Considérese el siguiente sistema generador de datos o planta $x(k) = G(q)u(k)$*

$$G(q) = \frac{0, 3}{(q - 0, 8e^{j\pi/4})(q - 0, 8e^{-j\pi/4})(q - 0, 7)}$$

donde $n_b = n_a = 3$, cuya salida se ve afectada por ruido coloreado $v(k) = H(q)e(k)$

$$H(q) = \frac{0, 4q}{q - 0, 6}$$

donde $e(k)$ es ruido blanco estricto gaussiano, de media cero y varianza 1, resultando la salida total medible $y(k) = x(k) + v(k)$. Supóngase ahora que a este sistema se le aplica la señal $u(k)$, generada de forma independiente de $e(k)$

$$u(k) = \sum_{i=\text{par}(p)}^{\text{ent}(p/2)} A_i \cos(\theta_i k - \phi_i), \quad \theta_i = \frac{2\pi i}{p} \quad (1.25)$$

$\text{par}(p)$ es 1 si p es par, 0 si p es impar
 $\text{ent}(p/2)$ es la parte entera de $p/2$

		\hat{n}_a									\hat{n}_a						
		1	2	3	4	5	6	7			1	2	3	4	5	6	7
\hat{n}_b	1	Λ	Λ	Λ	Λ	Λ	Λ		1	1e-00	3e-01	2e-01	9e-02	9e-02	8e-02	8e-02	
	2	Λ	Λ	Λ	Λ	Λ			2	1e-00	2e-01	1e-01	9e-02	9e-02	8e-02	8e-02	
	3	Λ	Λ	Λ	Λ				3	1e-00	2e-01	1e-01	8e-02	8e-02	8e-02	8e-02	
	4	Λ	Λ	Λ					\hat{n}_b 4	9e-01	1e-01	8e-02	8e-02	7e-02	7e-02	7e-02	
	5	Λ	Λ						5	5e-01	9e-02	8e-02	8e-02	7e-02	7e-02	7e-02	
	6	Λ							6	1e-01	9e-02	8e-02	8e-02	7e-02	7e-02	7e-02	
	7								7	2e-15	3e-15	3e-15	4e-15	4e-15	4e-15	4e-15	

Tabla 1.3: A la izquierda: Combinaciones en que $\Lambda_{yuy}^\infty(\hat{n}_b, \hat{n}_a)$ es no singular. A la derecha: El menor valor propio de $\Lambda_{yuy}^M(\hat{n}_b, \hat{n}_a)$ con $M = 500$.

con $p = 7$ y $\forall i, A_i = 1$, es decir, formada por 7 componentes espectrales, las cuales se muestran en la figura 1.5.

Supóngase que se desea determinar si el conjunto de mediciones Z_{uy}^∞ es s.i. para las estructuras de modelos ARX con $1 \leq (\hat{n}_b, \hat{n}_a) \leq 7$. Para ello se mide el conjunto de datos $Z_{uy}^{L,M}$, compuesto de $2M + 1 + L$ pares de datos tomados en estado estacionario, usando $M = 500$ y $L = 7$ como una aproximación al caso infinito.

Haciendo primero un análisis teórico, mediante el corolario 1.5, se puede notar que $\Lambda_{yuy}^\infty(\hat{n}_b, \hat{n}_a)$ es no singular para todo par (\hat{n}_b, \hat{n}_a) tal que

$$\hat{n}_b + \hat{n}_a + 1 \leq 7$$

ya que $u(k)$ es p.e. de orden 7, mostrándose esta situación en la tabla 1.3 izquierda, donde los casos en blanco no se encuentran cubiertos por este corolario.

El análisis teórico anterior se puede contrastar con la simulación de la tabla 1.3 derecha, donde se muestra el menor valor propio de $\Lambda_{yuy}^M(\hat{n}_b, \hat{n}_a)$, indicándose en color rojo cuando cada la matriz es singular, lo que se puede evidenciar por la fuerte caída que presenta el valor propio más pequeño. De esta manera se puede comprobar que la simulación coincide con el análisis teórico mostrado en la tabla izquierda, indicando que los datos Z_{uy}^∞ son s.i. para la estructura ARX, cuando la matriz $\Lambda_{yuy}^\infty(\hat{n}_b, \hat{n}_a)$ es no singular.

Para más detalles véase el programa F.2.3.

1.4. Comentarios del capítulo

La eficacia de una señal, para fines de identificación, se puede medir mediante el “orden de excitación

local” (definición 1.10), el que está relacionado con la cantidad de componentes espectrales que presenta la señal en cuestión y con su capacidad de resistir el filtraje en un intervalo de tiempo definido. En general, es suficiente, para que la señal $s(k)$ sea l.e. de orden p en el intervalo $[-M, M]$, que tenga p componentes espectrales no nulas y que cada una de ellas esté separada de las demás por una distancia múltiplo de $2\pi/(2M + 1)$ (teorema 1.2).

Puede ocurrir que algunas señales que sean l.e. de orden p en un intervalo finito $[-M, M]$ pierdan esta condición cuando $M \rightarrow \infty$, es por esto que se define un “orden de excitación persistente” (definición 1.11). Cabe destacar que, a diferencia del caso finito, es suficiente que la señal $s(k)$ tenga p componentes espectrales para que ella sea p.e. de orden p , es decir, no es necesario tener algún tipo de separación entre dichas componentes espectrales.

Al momento de recoger datos de entrada y salida de un sistema, para su posterior modelación, es necesario que este conjunto de datos $Z_{ux}^{L,M}$ tenga la suficiente información, para así poder diferenciar dos modelos candidatos de una estructura particular, por este motivo se define el concepto de “suficiencia de información” en los datos $Z_{ux}^{L,M}$ (definición 1.12).

Cuando la cantidad de datos medidos es finita, el concepto de suficiencia de información está muy relacionado con el concepto de excitación local, ya que es a través de una señal de entrada $u(k)$ l.e. que se puede generar un conjunto de mediciones $Z_{ux}^{L,M}$ s.i. para alguna estructura de modelos. En el caso en que la cantidad de datos medidos es infinita ocurre algo similar, ya que es a través de una señal p.e. que se puede generar un conjunto de mediciones Z_{ux}^∞ que sea s.i. para alguna estructura de modelos (corolarios 1.3, 1.4 y en el caso con ruido el corolario 1.5).

Capítulo 2

Matrices de correlaciones para identificación estructural

A modo de motivación, se introducirá el tema tratando de identificar el orden del siguiente sistema o *planta*

$$\begin{aligned} x(k) + a_1x(k-1) + a_2x(k-2) = \\ b_0u(k) + b_1u(k-1) + b_2u(k-2) \end{aligned} \quad (2.1)$$

a partir del cual se obtienen experimentalmente los siguientes $2M + 1 + L$ pares de datos (considerando una señal $u(k)$ adecuada)

$$Z_{ux}^{L,M} = \begin{bmatrix} u(-M-L) & \cdots & u(M) \\ x(-M-L) & \cdots & x(M) \end{bmatrix}$$

Supóngase que se cree que un buen modelo para el sistema anterior es un predictor de orden 2

$$\begin{aligned} \hat{x}(k) = & b_0^m u(k) + b_1^m u(k-1) + b_2^m u(k-2) \\ & - a_1^m x(k-1) - a_2^m x(k-2) \end{aligned}$$

Al ordenar los datos $Z_{ux}^{L,M}$ para $k = -M, \dots, M$, en la siguiente matriz asociada al predictor de orden 2

$$\Omega_{ux}^M(2) = \begin{bmatrix} u(-M) & \cdots & u(-M-2) & x(-M-1) & x(-M-2) \\ \vdots & & \vdots & \vdots & \vdots \\ u(M) & \cdots & u(M-2) & x(M-1) & x(M-2) \end{bmatrix}$$

se puede notar que ella es de rango completo, considerando $M \geq \hat{n}$, pues todas sus columnas son linealmente independientes (l.i.) entre sí.

Supóngase ahora que se cree que un buen *modelo*, para la *planta* en (2.1), es uno de orden 3

$$\begin{aligned} \hat{x}(k) = & b_0^m u(k) + b_1^m u(k-1) + b_2^m u(k-2) + b_3^m u(k-3) \\ & - a_1^m x(k-1) - a_2^m x(k-2) - a_3^m x(k-3) \end{aligned}$$

Entonces, al ordenar los mismos datos experimentales $Z_{ux}^{L,M}$ en la matriz $\Omega_{ux}^M(3)$, asociada al orden del predictor

$$\Omega_{ux}^M(3) = \begin{bmatrix} u(-M) & \cdots & u(-M-3) & x(-M-1) & \cdots & x(-M-3) \\ \vdots & & \vdots & \vdots & & \vdots \\ u(M) & \cdots & u(M-3) & x(M-1) & \cdots & x(M-3) \end{bmatrix}$$

se obtiene que es deficiente de rango, ya que según la ecuación de la planta (2.1) se tiene

$$\begin{aligned} x(k) = & b_0u(k) + b_1u(k-1) + b_2u(k-2) \\ & - a_1x(k-1) - a_2x(k-2) \end{aligned}$$

o de forma equivalente

$$\begin{aligned} x(k-1) = & b_0u(k-1) + b_1u(k-2) + b_2u(k-3) \\ & - a_1x(k-2) - a_2x(k-3) \end{aligned}$$

Así, la columna asociada a $x(k-1)$, en la matriz del modelo $\Omega_{ux}^M(3)$, depende linealmente de las columnas asociadas a $u(k-1)$, $u(k-2)$, $u(k-3)$, $x(k-2)$ y $x(k-3)$, lo que indica que los datos medidos no pueden provenir de un sistema de orden 3.

De esta manera, si se prueba secuencialmente el rango de la matriz $\Omega_{ux}^M(\hat{n})$, para modelos con órdenes $\hat{n} = 1 \dots 3$, se tiene que

$$\begin{aligned} \Omega_{ux}^M(1) & \text{ es de rango completo} \\ \Omega_{ux}^M(2) & \text{ es de rango completo} \\ \Omega_{ux}^M(3) & \text{ no es de rango completo} \end{aligned}$$

donde el cambio desde rango completo hacia rango incompleto en la matriz $\Omega_{ux}^M(\hat{n})$ sirve como prueba para identificar el orden de la *planta*, obteniéndose de esta manera que el orden más adecuado para el modelo es $\hat{n} = 2$.

Este ejemplo considera sólo la idea básica, ya que no se tiene información sobre qué clase de señal $u(k)$ utilizar para obtener resultados adecuados, ni tampoco se especifica qué se debe hacer cuando hay ruido presente. Estos y otros temas más, relacionados con la determinación de una estructura apropiada para el modelo, se abordarán en este capítulo.

2.1. Matriz de producto

2.1.1. Construcción

Considérese el siguiente sistema generador de datos o *planta*, el que es completamente desconocido

$$\begin{aligned} x(k) &= G(q)u(k), & G(q) &= \frac{B(q)}{A(q)} & (2.2) \\ B(q) &= b_0 + b_1q^{-1} + \dots + b_nq^{-n} \\ A(q) &= 1 + a_1q^{-1} + \dots + a_nq^{-n} \end{aligned}$$

donde

a_n y b_n no son simultáneamente cero
 $B(q)$ y $A(q)$ son coprimos
 $G(q)$ es estable

Este sistema puede ser escrito como regresión lineal, de forma similar a (1.12), según

$$x(k) = \omega_{ux}^n(k)^T \beta \quad (2.3)$$

$$\omega_{ux}^n(k) := \begin{bmatrix} u(k) \\ \vdots \\ u(k-n) \\ x(k-1) \\ \vdots \\ x(k-n) \end{bmatrix}, \quad \beta = \begin{bmatrix} b_0 \\ \vdots \\ b_n \\ -a_1 \\ \vdots \\ -a_n \end{bmatrix}$$

El principal problema ahora consiste en obtener el orden n del sistema (2.2), a partir de los datos $Z_{ux}^{L,M}$ medidos desde él

$$Z_{ux}^{L,M} = \begin{bmatrix} u(-M-L) & \dots & u(M) \\ x(-M-L) & \dots & x(M) \end{bmatrix} \quad (2.4)$$

Esto se puede lograr considerando el siguiente modelo, en su forma de regresión lineal

$$\begin{aligned} \hat{x}_i(k) &= \omega_{ux}^{\hat{n}}(k)^T \gamma_i & (2.5) \\ \gamma_i &= [b_0^i \quad \dots \quad b_{\hat{n}}^i \quad -a_1^i \quad \dots \quad -a_{\hat{n}}^i]^T \end{aligned}$$

donde $\omega_{ux}^{\hat{n}}(k)$ está definido en (2.3). Si los datos medidos $Z_{ux}^{L,M}$ son ordenados utilizando este modelo, para cada instante de tiempo entre $[-M, M]$, entonces se puede definir la siguiente matriz de tamaño $(2M+1) \times (2\hat{n}+1)$

$$\Omega_{ux}^M(\hat{n}) := \begin{bmatrix} \omega_{ux}^{\hat{n}}(-M)^T \\ \vdots \\ \omega_{ux}^{\hat{n}}(M)^T \end{bmatrix} \quad (2.6)$$

donde $L \geq \hat{n}$ y el número total de pares de datos medidos es $N = 2M + 1 + L$.

Supóngase por ahora que M es suficientemente grande. Claramente, para el caso $\hat{n} > n$, la matriz $\Omega_{ux}^M(\hat{n})$ es siempre deficiente de rango, ya que ella tiene al menos $(\hat{n} - n)$ columnas linealmente dependientes (l.d.) debido a la ecuación de la planta

$$x(k-i) = \omega_{ux}^n(k-i)^T \beta, \quad i = 1 \dots (\hat{n} - n)$$

También se puede notar, para el caso $\hat{n} \leq n$, que la matriz $\Omega_{ux}^M(\hat{n})$ puede ser de rango completo, ya que aparentemente no presenta columnas l.d., de manera que su rango satisface

$$\begin{aligned} \text{si } 1 \leq \hat{n} \leq n, & \quad \text{rango } \Omega_{ux}^M(\hat{n}) \leq 2\hat{n} + 1 \\ \text{si } \hat{n} > n, & \quad \text{rango } \Omega_{ux}^M(\hat{n}) \leq \hat{n} + 1 + n \end{aligned}$$

Con el fin de poder transformar las desigualdades anteriores en igualdades, se define el siguiente producto, que resulta ser una matriz de tamaño $(2\hat{n}+1) \times (2\hat{n}+1)$

$$\begin{aligned} \Lambda_{zuy}^M(\hat{n}) &:= \frac{1}{2M+1} \Omega_{uz}^M(\hat{n})^T \Omega_{uy}^M(\hat{n}) & (2.7) \\ &= \frac{1}{2M+1} \sum_{k=-M}^M \omega_{uz}^{\hat{n}}(k) \omega_{uy}^{\hat{n}}(k)^T \end{aligned}$$

donde cada elemento de esta matriz está definido por la correlación muestral en (1.2), para $0 \leq (\tau_1, \tau_2) \leq L$.

Así la *matriz de producto* (PM) $\Lambda_{xux}^M(\hat{n})$, es decir, con $z(k) = y(k) = x(k)$, tiene el mismo rango que $\Omega_{ux}^M(\hat{n})$

$$\text{rango } \Lambda_{xux}^M(\hat{n}) = \text{rango } \Omega_{ux}^M(\hat{n})$$

Por lo tanto, mediante el siguiente lema, se puede obtener de manera exacta el rango de las matrices $\Lambda_{xux}^M(\hat{n})$ y $\Omega_{ux}^M(\hat{n})$

Lema 2.1 *Considere los datos $Z_{ux}^{L,M}$, medidos a partir del sistema de orden n en la ecuación (2.2). Si $s(k) = A(q)^{-1}u(k)$ es localmente excitante¹ (l.e.) de orden $(\hat{n} + 1 + n) \forall k \in [-M, M]$, entonces la matriz $\Lambda_{xux}^M(\hat{n})$ tiene el siguiente rango*

$$\text{rango } \Lambda_{xux}^M(\hat{n}) = \begin{cases} 2\hat{n} + 1, & 1 \leq \hat{n} \leq n \\ \hat{n} + 1 + n, & \hat{n} > n \end{cases}$$

donde $2M + 1 \geq \hat{n} + 1 + n$ y $L \geq \hat{n}$.

Para el caso $M \rightarrow \infty$ el resultado es el mismo, pero l.e. debe ser reemplazado por p.e². La demostración para el caso $M \rightarrow \infty$ es similar a la del caso finito, mostrada a continuación.

Demostración. Para el caso en que $\hat{n} \leq n$, la matriz $\Lambda_{xux}^M(\hat{n})$ es definida positiva por el teorema 1.4, considerando $\hat{n} = \hat{n}_b = \hat{n}_a$ y $n = \max(n_b, n_a)$. Para el caso $\hat{n} > n$, el rango se determina removiendo todas las columnas l.d. de $\Omega_{ux}^M(\hat{n})$ y verificando que las restantes son l.i., por medio del teorema 1.5, considerando $\hat{n} = \hat{n}_b = \hat{n}_a$ y $n = \max(n_b, n_a)$. **QED.**

Un enfoque basado en “total least squares” (TLS) se utilizó en [7] para obtener un resultado similar al de este lema, sin embargo [7] presenta una deficiencia, ya que utiliza señales p.e. para generar el conjunto finito de datos $Z_{ux}^{L,M}$, lo cual es válido si M es muy grande.

Nota 2.1 *Nótese que los requerimientos para el lema 2.1 son bastante accesibles, ya que sólo se necesita tomar un número finito de muestras ($N \geq 2\hat{n} + 1 + n$) y además que la señal $u(k)$ se diseñe de manera que $s(k) = A(q)^{-1}u(k)$ satisfaga el orden de excitación requerido.*

En el caso que el conjunto de datos $Z_{ux}^{L,M}$ esté dado, es decir, que el experimento no se pueda diseñar, el número total de datos y el orden de excitación de la señal $s(k)$ deben ser verificados antes de usar el lema 2.1. Este orden de excitación puede ser verificada usando la definición 1.10 y el teorema 1.2, donde el orden máximo que se puede probar depende de L . En caso que el número de muestras sea muy grande, el orden de excitación de la señal puede ser verificado mediante la definición 1.11 y el teorema 1.3.

2.1.2. Búsqueda secuencial

El lema 2.1 provee un criterio para estimar el orden más apropiado de un modelo, de manera similar a [23], [24], [13] y [8], en los cuales se prueba la no singularidad de la matriz $\Lambda_{xux}^M(\hat{n})$, aumentando secuencialmente el orden del modelo \hat{n} , hasta que la matriz se vuelve singular ($\hat{n} = n + 1$). Esta búsqueda se realiza mediante el uso de algún índice de singularidad $\text{IS}(\hat{n})$ como por ejemplo

$$\text{IS}(\hat{n}) = \lambda_{\min}(\Lambda_{xux}^M(\hat{n})), \quad \text{IS}(\hat{n}) = |\det \Lambda_{xux}^M(\hat{n})|$$

donde $\lambda_{\min}(A)$ es el menor valor propio de A , o como el índice de singularidad usado en [13]

$$\text{IS}(\hat{n}) = \lambda_{\min}(\Lambda_{xux}^M(\hat{n})) \cdot (N^{1/N})^{2\hat{n}}$$

donde N es el número de datos utilizados. Algo similar es usado en [18], definiendo la siguiente razón de determinantes

$$\text{DR}(\hat{n}) := \left| \frac{\det \Lambda_{xux}^M(\hat{n})}{\det \Lambda_{xux}^M(\hat{n} + 1)} \right|$$

Si se aumenta secuencialmente el orden del modelo, la razón se vuelve infinita al momento de alcanzar el orden de la planta (es decir, $\hat{n} = n$).

Supóngase ahora que se quiere buscar el orden de la planta entre los modelos $\hat{n} = 1, \dots, \hat{n}_{\max}$. Dado que el orden n de la planta no se conoce, es suficiente cumplir las siguientes condiciones

$$\begin{aligned} \hat{n}_{\max} &> n \\ p \geq \hat{n}_{\max} + 1 + \hat{n}_{\max} &\geq \hat{n} + 1 + n \\ 2M + 1 \geq p &\geq \hat{n} + 1 + n \\ L \geq \hat{n}_{\max} &\geq \hat{n} \end{aligned}$$

donde p es el orden de excitación a utilizar. Estas condiciones satisfacen el lema 2.1 y, a la vez, permiten realizar una búsqueda secuencial, la cual se resume en el algoritmo 2.1.

Nótese que para el cálculo del índice $\text{DR}(\hat{n})$ se necesita construir la matriz $\Lambda_{xux}^M(\hat{n} + 1)$. Por ello es suficiente cumplir con

$$\begin{aligned} \hat{n}_{\max} &> n \\ p \geq 2(\hat{n}_{\max} + 1) &\geq (\hat{n} + 1) + 1 + n \\ 2M + 1 \geq p &\geq (\hat{n} + 1) + 1 + n \\ L \geq \hat{n}_{\max} + 1 &\geq \hat{n} + 1 \end{aligned}$$

A continuación se muestra un ejemplo para ilustrar el uso del algoritmo 2.1

¹ver definición 1.10

²ver definición 1.11

```

// región de prueba
1  escoja  $\hat{n}_{\text{máx}} > n$ ;
   // diseño de señal de entrada
2  escoja  $p \geq 2\hat{n}_{\text{máx}} + 1$ ;
3  escoja  $M$  tal que  $2M + 1 \geq p$ ;
4  escoja  $L \geq \hat{n}_{\text{máx}}$ ;
5  genere  $u(k)$  l.e./p.e. de orden  $p$ ;
   // medición
6  mida  $Z_{ux}^{L,M}$  en estado estacionario;
   // identificación
7  for  $\hat{n} = 1$  to  $\hat{n}_{\text{máx}}$  do
8    tabla( $\hat{n}$ ) = IS( $\hat{n}$ );
9  end
10 buscar el menor  $\hat{n}$  en tabla( $\hat{n}$ ), tal que  $\Lambda_{xux}^M(\hat{n})$ 
    es singular;
11  $n = \hat{n} - 1$ ;

```

Algoritmo 2.1: Búsqueda secuencial de n a través de la matriz $\Lambda_{xux}^M(\hat{n})$.

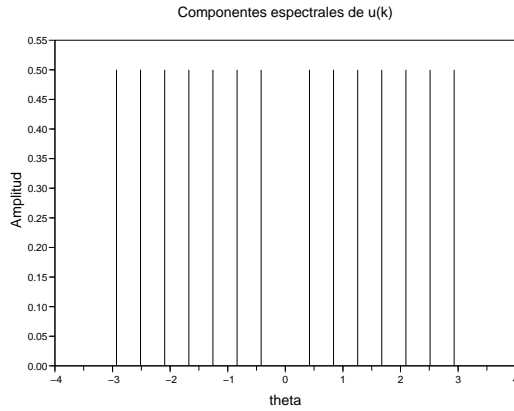


Figura 2.1: Componentes espectrales de la señal $u(k)$, con $p = 14$.

\hat{n}	$\lambda_{\text{mín}}(\Lambda_{xux}^M(\hat{n}))$	DR(\hat{n})
1	2,31e+00	2,92e-01
2	5,72e-01	1,07e+00
3	6,97e-02	9,09e+13
4	3,37e-16	7,34e+14
5	2,98e-16	5,44e+15
6	7,67e-17	1,94e+14

Tabla 2.1: Singularidad de $\Lambda_{xux}^M(\hat{n})$, para $M = 7$ y $L = 7$.

Ejemplo 2.1 Considérese el siguiente sistema generador de datos o planta $x(k) = G(q)u(k)$

$$G(q) = \frac{0,3}{(q - 0,8e^{j\pi/4})(q - 0,8e^{-j\pi/4})(q - 0,7)}$$

que es completamente desconocido y cuyo orden es $n = 3$.

Para poder identificar el orden de este sistema, se prueban secuencialmente varios órdenes de modelos $\hat{n} = 1, \dots, 2n$, utilizando el resultado obtenido en el lema 2.1, a través del algoritmo 2.1 con el indicador de singularidad $\lambda_{\text{mín}}(\Lambda_{xux}^M(\hat{n}))$ y con la razón DR(\hat{n}). Por este motivo se eligen

$$\begin{aligned} \hat{n}_{\text{máx}} &= 2n = 6, & p &= 2(\hat{n}_{\text{máx}} + 1) = 14 \\ M &= \hat{n}_{\text{máx}} + 1 = 7, & L &= \hat{n}_{\text{máx}} + 1 = 7 \end{aligned}$$

Teniendo en cuenta lo expuesto anteriormente, se aplica la señal $u(k)$ de la ecuación (1.22) considerando $p = 14$ y $\forall i, A_i = 1$, es decir, 14 componentes espectrales, espaciadas cada $2\pi/(2M + 1)$, las cuales se muestran en la figura 2.1.

Así, se toma el conjunto de datos $Z_{ux}^{L,M}$ en estado estacionario, obteniéndose como resultado la tabla 2.1, donde se puede ver que a partir de $\hat{n} \geq 4$ la matriz $\Lambda_{xux}^M(\hat{n})$ es singular, ya que $\lambda_{\text{mín}}(\Lambda_{xux}^M(\hat{n}))$ cae fuertemente, indicando que el orden más apropiado para el modelo es $\hat{n} = 3$. De igual forma la razón DR(\hat{n}) confirma este hecho, ya que cuando $\hat{n} = 3$, la razón crece abruptamente.

Para más detalles de la simulación, véase el programa F.3.1.

2.2. Matriz de producto instrumental

2.2.1. Construcción

Supóngase ahora que la salida $x(k)$ del sistema en la ecuación (2.2) se ve afectada por ruido coloreado $v(k)$, de distribución de probabilidad desconocida, obtenido filtrando ruido blanco³ $e(k)$

³definición 1.5

$$y(k) = x(k) + v(k) \quad (2.8)$$

$$v(k) = H(q)e(k), \quad H(q) = \sum_{j=0}^{\infty} h(j)q^{-j}$$

$e(k)$ ruido blanco de media 0 y varianza 1
 $H(q)$ es un filtro estable

de manera que la salida medible total es $y(k)$.

Para estimar el orden n del sistema, considérese el siguiente predictor

$$\hat{y}_i(k) = W_i(q)\zeta_{uy}(k) \quad (2.9)$$

$$W_i(q) = \begin{bmatrix} \bar{H}_i(q)^{-1}G_i(q) & 1 - \bar{H}_i(q)^{-1} \end{bmatrix}$$

$$\bar{H}_i(q) = \frac{H_i(q)}{h_i(0)}, \quad H_i(q) = \sum_{l=0}^{\infty} h_i(l)q^{-l}$$

$$G_i(q) = \frac{B_i(q)}{A_i(q)}, \quad \zeta_{uy}(k) = \begin{bmatrix} u(k) & y(k) \end{bmatrix}^T$$

donde

$G_i(q)$ es un modelo estable para el sistema $G(q)$
 $H_i(q)$ es un modelo bipropio y de fase mínima para el filtro del ruido $H(q)$

Si se utiliza un modelo ARX, es decir $\bar{H}_i(q) = A_i(q)^{-1}$, entonces el predictor define una regresión lineal, de forma similar a (1.21)

$$\hat{y}_i(k) = \omega_{uy}^{\hat{n}}(k)^T \gamma_i \quad (2.10)$$

$$\gamma_i = \begin{bmatrix} b_0^i & \cdots & b_n^i & -a_1^i & \cdots & -a_n^i \end{bmatrix}^T$$

donde $\omega_{uy}^{\hat{n}}(k)$ está definido en (2.3).

Siguiendo el razonamiento de la sección anterior, a partir del conjunto de datos medidos $Z_{uy}^{L,M}$ se pueden construir las matrices $\Omega_{uy}^M(\hat{n})$ y $\Lambda_{yuy}^M(\hat{n})$, pero, a primera vista, no se puede asegurar nada acerca de su rango. De hecho, cuando la *relación señal-ruido* (SNR) es suficientemente baja, estas matrices son en muchos casos de rango completo, dificultando el proceso de estimación de orden. Lo expresado se puede comprobar a través del corolario 1.5, ilustrado en el siguiente ejemplo

Ejemplo 2.2 *Considérese la planta del ejemplo 2.1*

$$G(q) = \frac{0,3}{(q-0,8e^{j\pi/4})(q-0,8e^{-j\pi/4})(q-0,7)}$$

\hat{n}	$\lambda_{\min}(\Lambda_{yuy}^M(\hat{n}))$	DR(\hat{n})
1	1,01e+00	2,12e+00
2	2,61e-01	2,64e+00
3	1,10e-01	3,80e+00
4	7,41e-02	4,75e+00
5	7,15e-02	5,56e+00
6	6,84e-02	5,77e+00

Tabla 2.2: Singularidad de $\Lambda_{yuy}^M(\hat{n})$, para $M = 10^3$ y SNR = 5,7[dB].

\hat{n}	$\lambda_{\min}(\Lambda_{yuy}^M(\hat{n}))$	DR(\hat{n})
1	9,98e-01	2,21e+00
2	2,50e-01	2,75e+00
3	1,05e-01	3,75e+00
4	7,30e-02	4,96e+00
5	7,08e-02	5,73e+00
6	6,86e-02	5,95e+00

Tabla 2.3: Singularidad de $\Lambda_{yuy}^M(\hat{n})$, para $M = 5 \cdot 10^3$ y SNR = 5,92[dB].

\hat{n}	$\lambda_{\min}(\Lambda_{yuy}^M(\hat{n}))$	DR(\hat{n})
1	9,84e-01	2,25e+00
2	2,48e-01	2,81e+00
3	1,04e-01	3,76e+00
4	7,10e-02	4,99e+00
5	6,84e-02	5,79e+00
6	6,57e-02	6,00e+00

Tabla 2.4: Singularidad de $\Lambda_{yuy}^M(\hat{n})$, para $M = 10^4$ y SNR = 5,84[dB].

\hat{n}	$\lambda_{\min}(\Lambda_{yuy}^M(\hat{n}))$	DR(\hat{n})
1	1,00e+00	2,24e+00
2	2,48e-01	2,80e+00
3	1,04e-01	3,77e+00
4	7,08e-02	5,01e+00
5	6,83e-02	5,78e+00
6	6,59e-02	5,97e+00

Tabla 2.5: Singularidad de $\Lambda_{yuy}^M(\hat{n})$, para $M = 5 \cdot 10^4$ y SNR = 5,78[dB].

cuyo orden $n = 3$ es desconocido. Considérese además que la salida de este sistema se ve afectada por ruido coloreado $v(k)$, de forma que la salida total medible es $y(k) = x(k) + v(k)$, donde el ruido de salida es generado mediante

$$v(k) = H(q)e(k), \quad H(q) = \frac{0,4q}{q-0,6}$$

con $e(k)$ ruido blanco estricto gaussiano, de media cero y varianza 1.

Ahora se intentará identificar el orden del sistema $G(q)$, probando secuencialmente varios órdenes de modelos $\hat{n} = 1, \dots, 2n$, utilizando los indicadores de singularidad $\lambda_{\min}(\Lambda_{yuy}^{\infty}(\hat{n}))$ y $\text{DR}(\hat{n})$, ya vistos en el ejemplo 2.1.

Considérese que $u(k)$ es una realización de ruido blanco gaussiano de media cero y varianza 1. De esta manera se miden aproximaciones $Z_{uy}^{L,M}$ (en estado estacionario) del conjunto de datos Z_{uy}^{∞} , considerando $M = \{10^3, 5 \cdot 10^3, 10^4, 5 \cdot 10^4\}$ y $L = 7$, obteniéndose como resultado las tablas 2.2-2.5.

Según el corolario 1.5, dado que $u(k)$ es p.e. de todo orden, entonces para $\hat{n} = 1, \dots, 2n$ las matrices $\Lambda_{yuy}^{\infty}(\hat{n})$ son no singulares, lo que se puede apreciar en las tablas 2.2-2.5, ya que las aproximaciones del indicador $\lambda_{\min}(\Lambda_{yuy}^M(\hat{n}))$ no presentan caídas importantes, y además el indicador $\text{DR}(\hat{n})$ tampoco presenta crecimientos bruscos en ninguno de los casos, lo que impide la identificación del orden del sistema.

Para más detalles sobre la simulación, véase el programa F.3.2.

Para resolver el problema de contaminación con ruido, se define la matriz de producto instrumental⁴ (IPM) $\Lambda_{zuy}^M(\hat{n})$, cuyos elementos corresponden a correlaciones como en la ecuación (1.2). Esta matriz incorpora una secuencia instrumental $z(k)$, que se obtiene al filtrar $u(k)$ según

$$\begin{aligned} z(k) &= F(q)u(k), \quad F(q) = \frac{L(q)}{K(q)} \quad (2.11) \\ L(q) &= l_0 + l_1q^{-1} + \dots + l_mq^{-m} \\ K(q) &= 1 + k_1q^{-1} + \dots + k_mq^{-m} \end{aligned}$$

con

⁴ver ecuación (2.7)

k_m y l_m	no simultáneamente cero
$L(q)$ y $K(q)$	coprimos
$F(q)$	es un filtro estable

Ahora, en vista del siguiente lema, el ruido puede ser eliminado del problema de estimación de orden

Lema 2.2 Considere los datos Z_{uy}^{∞} y Z_{uz}^{∞} , medidos de (2.8) y (2.11) respectivamente, donde $u(k)$ y $e(k)$ satisfacen $\varphi_{ue}^{\infty}(\tau) = 0$, entonces

$$\varphi_{uy}^{\infty} = \varphi_{ux}^{\infty}, \quad \varphi_{zy}^{\infty} = \varphi_{zx}^{\infty}$$

Demostración. Dada la definición en (1.2), se tiene que

$$\varphi_{uy}^{\infty} = \varphi_{ux}^{\infty} + \varphi_{uv}^{\infty}, \quad \varphi_{zy}^{\infty} = \varphi_{zx}^{\infty} + \varphi_{zv}^{\infty}$$

Ahora, según (1.4), se tiene

$$\varphi_{uv}^{\infty} = \varphi_{ue}^{\infty} * \mathcal{J}h^* = 0, \quad \varphi_{zv}^{\infty} = f * \varphi_{ue}^{\infty} * \mathcal{J}h^* = 0$$

QED.

Nota 2.2 Para el lema anterior, nótese que $u(k)$ puede ser estocástica o determinística, pero debe satisfacer la condición $\varphi_{ue}^{\infty}(\tau) = 0$.

Por ejemplo, si $u(k)$ es cualquier señal determinística, mientras $e(k)$ es ruido blanco de media cero, se puede ver que $s(k) = u(k)e(k - \tau)$ tiene media ergódica, lo que significa

$$E \{u(k)e(k - \tau)\} = \varphi_{ue}^{\infty}(\tau) = 0$$

satisfaciendo así el teorema anterior. Ver apéndice C, para temas de ergodicidad.

Si ahora se considera que $u(k)$ es ruido blanco estricto de media cero y varianza a_u^2 , mientras que $e(k)$ es ruido blanco estricto de media cero y varianza a_e^2 , ambos generados de forma independiente, entonces se puede notar que ambos tienen correlación ergódica, es decir

$$r_{ue}(\tau) = \varphi_{ue}^{\infty}(\tau) = 0$$

satisfaciendo igualmente el teorema anterior. Ver apéndice C, para temas de ergodicidad.

Así este último resultado conduce a la siguiente igualdad

$$\Lambda_{zuy}^{\infty}(\hat{n}) = \Lambda_{zux}^{\infty}(\hat{n})$$

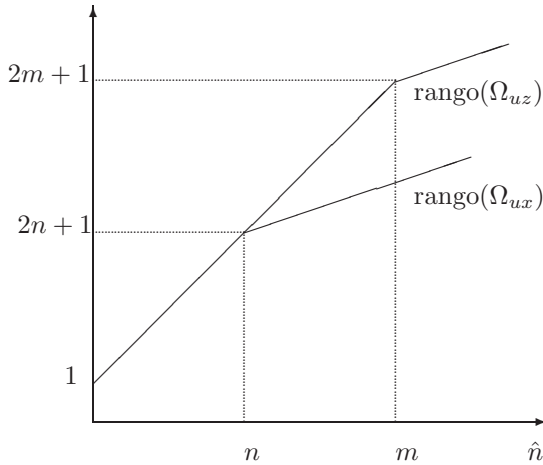


Figura 2.2: Rango de las matrices $\Omega_{uz}^M(\hat{n})$ and $\Omega_{ux}^M(\hat{n})$.

Una vez que el ruido ha sido eliminado del problema, sólo queda conocer las condiciones de singularidad de esta matriz libre de ruido $\Lambda_{zux}^M(\hat{n})$. Sería deseable que esta matriz tuviera las mismas propiedades de rango que $\Omega_{ux}^M(\hat{n})$ o $\Lambda_{xux}^M(\hat{n})$, para así poder utilizarla para estimación de orden.

Teorema 2.1 *Considérese los datos $Z_{ux}^{L,M}$ y $Z_{uz}^{L,M}$ medidos de (2.2) y (2.11) respectivamente. Si la señal $s(k) = A(q)^{-1}u(k)$ es l.e. de orden $\hat{n} + 1 + m \forall k \in [-M, M]$, y el orden de $F(q)$ satisface $m \geq n$, entonces*

$$\text{rango } \Lambda_{zux}^M(\hat{n}) = \text{rango } \Lambda_{xux}^M(\hat{n}), \quad 1 \leq \hat{n} \leq m$$

es decir

$$\text{rango } \Lambda_{zux}^M(\hat{n}) = \begin{cases} 2\hat{n} + 1, & 1 \leq \hat{n} \leq n \\ \hat{n} + 1 + n, & n < \hat{n} \leq m \end{cases}$$

donde $2M + 1 \geq \hat{n} + 1 + m$ y $L \geq \hat{n}$.

Demostración. Debido al corolario D.4, se requiere que $\Omega_{uz}^M(\hat{n})$ sea de rango completo, para que se pueda conseguir

$$\begin{aligned} \text{rango } \Lambda_{zux}^M(\hat{n}) &= \text{rango } \Omega_{uz}^M(\hat{n})^T \Omega_{ux}^M(\hat{n}) \\ &= \text{rango } \Omega_{ux}^M(\hat{n}) \end{aligned}$$

y de acuerdo al lema 2.1, si $s(k) = A(q)^{-1}u(k)$ es l.e. de orden $\hat{n} + 1 + m \forall k \in [-M, M]$, entonces $\Omega_{uz}^M(\hat{n})$ es de rango completo para $1 \leq \hat{n} \leq m$.

Además, se puede ver de la figura 2.2, que se necesita que $m \geq n$ para que $\Omega_{uz}^M(\hat{n})$ sea de rango

completo, mientras $\Omega_{ux}^M(\hat{n})$ pasa de rango completo a deficiente de rango. **QED.**

Nota 2.3 *Si M es muy grande, pero finito, la condición l.e. en el teorema 2.1 puede ser aproximada por la condición p.e., y el resultado debería permanecer sin cambios. Sin embargo, cuando se considera M infinito, el resultado del teorema 2.1 se mantiene “genéricamente”, es decir, pueden aparecer casos extraños en los que no se cumple. Estos casos son bastante raros, y son producidos por combinaciones desafortunadas de $F(q)$, $G(q)$ y $u(k)$, pero desde un punto de vista práctico no hay necesidad de preocuparse, ya que éstos aparecen muy esporádicamente. Para más detalles consúltese la sección 2.4 de este mismo capítulo.*

Nota 2.4 *Nótese que el teorema 2.1 provee condiciones no asintóticas (en el número de mediciones) para el proceso de identificación estructural. De esta forma, las propiedades asintóticas sólo se necesitan para reducción de ruido, según el lema 2.2.*

2.2.2. Indicaciones para el diseño del filtro instrumental

Una pregunta importante que queda, se refiere a cómo escoger el filtro instrumental $F(q)$. Por ahora, la elección de este filtro queda bastante abierta, requiriéndose solamente que $m \geq n$ según el teorema 2.1. También en [14, section 7.6] y [19, example 3.1] se dan algunas indicaciones para la elección del filtro $F(q)$, siendo $F(q) = q^{-m}$ la elección más simple.

Dado que en muchos casos se utilizan señales de entrada $u(k)$ con propiedades de ruido blanco, es decir $\varphi_{uu}^\infty(\tau) = a_u^2 \delta_K(\tau)$, entonces se buscarán filtros instrumentales que potencien la identificación, basándose en estas características de la señal de entrada. Para ello, considere primero la siguiente partición de la matriz de producto instrumental

$$\Lambda_{zuy}^\infty(\hat{n}) = \begin{bmatrix} \Gamma_{uu}^\infty(\hat{n} + 1) & F_{yu}^\infty(\hat{n})^T \\ F_{zu}^\infty(\hat{n}) & \Gamma_{zy}^\infty(\hat{n}) \end{bmatrix} \quad (2.12)$$

donde

$$\Gamma_{\alpha\beta}^\infty(\hat{n}) = \begin{bmatrix} \varphi_{\alpha\beta}^\infty(0) & \cdots & \varphi_{\alpha\beta}^\infty(\hat{n} - 1) \\ \vdots & \ddots & \vdots \\ \varphi_{\alpha\beta}^\infty(1 - \hat{n}) & \cdots & \varphi_{\alpha\beta}^\infty(0) \end{bmatrix} \quad (2.13)$$

$$F_{\alpha\beta}^{\infty}(\hat{n})^T = \begin{bmatrix} \varphi_{\alpha\beta}^{\infty}(-1) & \cdots & \varphi_{\alpha\beta}^{\infty}(-\hat{n}) \\ \varphi_{\alpha\beta}^{\infty}(0) & \cdots & \varphi_{\alpha\beta}^{\infty}(1-\hat{n}) \\ \vdots & \ddots & \vdots \\ \varphi_{\alpha\beta}^{\infty}(\hat{n}-1) & \cdots & \varphi_{\alpha\beta}^{\infty}(0) \end{bmatrix} \quad (2.14)$$

Es claro que al utilizar un número finito de muestras se producen errores en la matriz $\Lambda_{zuy}^M(\hat{n})$, debido a que la reducción de ruido no es completa

$$\Lambda_{zuy}^M = \underbrace{\begin{bmatrix} \Gamma_{uu}^M & (F_{xu}^M)^T \\ F_{zu}^M & \Gamma_{zx}^M \end{bmatrix}}_{\Lambda_{zux}^M} + \underbrace{\begin{bmatrix} 0 & (F_{vu}^M)^T \\ 0 & \Gamma_{zv}^M \end{bmatrix}}_{E_{zuv}^M}$$

La idea ahora es diseñar un filtro $F(q)$ que permita que la matriz de ruido E_{zuv}^M sea lo más pequeña posible frente a la matriz libre de ruido Λ_{zux}^M .

La opción inicial más tentadora es elegir un filtro $F(q)$ con una ganancia alta para así aumentar el valor de la matriz Λ_{zux}^M . Sin embargo, al analizar con mayor detalle, se puede notar que la matriz de ruido E_{zuv}^M también depende del filtro $F(q)$, y por lo tanto, también se vería aumentada en la misma proporción. Dado que no es conveniente aumentar la ganancia del filtro $F(q)$, hay que buscar otra forma de aumentar el valor de los elementos de Λ_{zux}^M . Por esto se analizará qué ocurre con las matrices Γ_{zx}^M y F_{zu}^M .

Para facilitar el análisis, supóngase que $u(k)$ satisface $\varphi_{uu}^{\infty}(\tau) = a_u^2 \delta_K(\tau)$. Entonces, los elementos de las matrices Γ_{zx}^{∞} y F_{zu}^{∞} están formados respectivamente por las siguientes correlaciones debido a (1.4)

$$\begin{aligned} \varphi_{zx}^{\infty}(\tau) &= a_u^2 (f * \mathcal{J}g^*)(\tau) \\ \varphi_{zu}^{\infty}(\tau) &= a_u^2 f(\tau) \end{aligned}$$

donde $f(k)$ y $g(k)$ son las respuestas a impulso de los filtros $F(q)$ y $G(q)$, respectivamente. Supóngase además que ambas respuestas a impulso se encuentran acotadas por

$$\begin{aligned} |f(k)| &\leq A_f \epsilon_f (k - d_f) = A_f \lambda_f^{(k-d_f)} \mu(k - d_f) \\ |g(k)| &\leq A_g \epsilon_g (k - d_g) = A_g \lambda_g^{(k-d_g)} \mu(k - d_g) \end{aligned}$$

donde d_f y d_g son respectivamente los retardos de los filtros $F(q)$ y $G(q)$. Entonces, la correlación $\varphi_{zx}^{\infty}(\tau)$ se ve acotada por

$$\begin{aligned} |\varphi_{zx}^{\infty}(\tau)| &\leq a_u^2 \sum_{k \in \mathbb{Z}} |f(k)| |g(k - \tau)| \\ &\leq a_u^2 A_f A_g \sum_{k \in \mathbb{Z}} \epsilon_f (k - d_f) \epsilon_g (k - d_g - \tau) \\ &\leq a_u^2 A_f A_g (\epsilon_f * \mathcal{J}\epsilon_g)(\tau - d_f + d_g) \end{aligned}$$

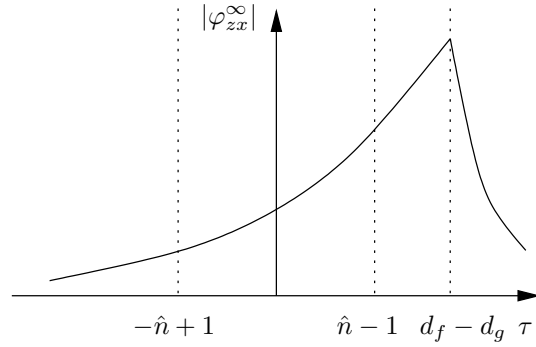


Figura 2.3: Cota para la correlación muestral $\varphi_{zx}^{\infty}(\tau)$, donde la exponencial izquierda está asociada a $G(q)$, y la derecha, asociada a $F(q)$.

y considerando el teorema B.2 esto conduce a

$$|\varphi_{zx}^{\infty}(\tau)| \leq \frac{a_u^2 A_f A_g}{1 - \lambda_f \lambda_g} \lambda_f^{(\tau - d_f + d_g)^+} \lambda_g^{-(\tau - d_f + d_g)^-} \quad (2.15)$$

donde

$$x^+ = \max(0, x), \quad x^- = \min(0, x) \quad (2.16)$$

En la figura 2.3 se muestra la cota para $|\varphi_{zx}^{\infty}(\tau)|$, donde se puede ver que si la distancia

$$(d_f - d_g) - (\hat{n} - 1) \geq 4\kappa_g$$

sobrepasa 4 constantes de tiempo del sistema $G(q)$, donde la constante de tiempo κ_g está dada por

$$\kappa_g = \frac{-1}{\ln \lambda_g} \Leftrightarrow e^{-1/\kappa_g} = \lambda_g$$

entonces los elementos de $\varphi_{zx}^{\infty}(\tau)$ se vuelven muy cercanos a cero, afectando así el tamaño de la matriz $\Gamma_{zx}^{\infty}(\hat{n})$, ya que ella está formada por los elementos

$$\varphi_{zx}^{\infty}(-\hat{n} + 1), \dots, \varphi_{zx}^{\infty}(\hat{n} - 1)$$

Así, para que la matriz $\Gamma_{zx}^{\infty}(\hat{n})$ tenga elementos apreciables, la distancia mostrada anteriormente debe ser no superior a media constante de tiempo

$$(d_f - d_g) - (\hat{n} - 1) \leq \frac{\kappa_g}{2}$$

ya que en el peor caso se tendría

$$|\varphi_{zx}^{\infty}(\hat{n} - 1)| \propto \lambda_g^{\kappa_g/2} = e^{-1/2} \approx 0,606$$

es decir, el 60,6% del valor de $|\varphi_{zx}^{\infty}(d_f - d_g)|$. Ahora, dado que muchos sistemas muestreados tienen retardo

$d_g = 1$, entonces el retardo del filtro instrumental d_f debería satisfacer

$$0 \leq d_f \leq \frac{\kappa_g}{2} + \hat{n} \quad (2.17)$$

Otro criterio que simplificaría la estructura de la matriz $\Lambda_{zux}^\infty(\hat{n})$ es diseñar $F(q)$ para que $F_{zu}^\infty(\hat{n})$ sea lo más cercana posible a la matriz nula. Esto último permitiría también concentrar los valores propios de $\Lambda_{zux}^\infty(\hat{n})$ más cerca de los elementos diagonales (debido al teorema D.6 de los discos de Gershgorin), los cuales deberían ser relativamente grandes, si d_f satisface la ecuación (2.17). Así, considérese los elementos que forman parte de la matriz $F_{zu}^\infty(\hat{n})$, es decir

$$\forall \tau = -\hat{n}, \dots, (\hat{n} - 1), \quad \varphi_{zu}^\infty(\tau) = a_u^2 f(\tau)$$

Dado que $F(q)$ es causal y tiene retardo d_f , entonces

$$\forall \tau = -\hat{n}, \dots, (d_f - 1), \quad \varphi_{zu}^\infty(\tau) = a_u^2 f(\tau) = 0$$

Por este motivo, si se escoge

$$d_f \geq \hat{n} \quad (2.18)$$

entonces la matriz $F_{zu}^\infty(\hat{n}) = 0$, y por lo tanto se lograría concentrar mejor los valores propios de $\Lambda_{zux}^\infty(\hat{n})$.

También en las secciones y capítulos posteriores se darán más indicaciones para poder diseñar este filtro instrumental, las que proveerán propiedades interesantes a la matriz $\Lambda_{zux}^M(\hat{n})$.

2.2.3. Búsqueda secuencial

A partir del lema 2.2 y del teorema 2.1, es posible recuperar el método de estimación de orden, aun si existe ruido de salida de distribución de desconocida, mediante el uso de la matriz $\Lambda_{zuy}^\infty(\hat{n})$ y alguno de los indicadores de singularidad

$$\text{IS}(\hat{n}) = |\lambda|_{\min}(\Lambda_{zuy}^\infty(\hat{n})), \quad \text{IS}(\hat{n}) = \det \Lambda_{zuy}^\infty(\hat{n})$$

donde $|\lambda|_{\min}(A)$ corresponde al valor propio de A más cercano a cero. De la misma manera que en [18], también se puede definir una *razón de determinantes instrumentales* para probar secuencialmente el orden más apropiado para el modelo

$$\text{IDR}(\hat{n}) := \left| \frac{\det \Lambda_{zuy}^\infty(\hat{n})}{\det \Lambda_{zuy}^\infty(\hat{n} + 1)} \right|$$

De esta manera, para buscar el orden de la planta entre los modelos $\hat{n} = 1, \dots, \hat{n}_{\text{máx}}$, es suficiente cumplir las siguientes condiciones

$$\begin{aligned} \hat{n}_{\text{máx}} &> n \\ m \geq \hat{n}_{\text{máx}} &\geq n \\ p \geq \hat{n}_{\text{máx}} + 1 + m &\geq \hat{n} + 1 + m \\ 2M + 1 \geq p &\geq \hat{n} + 1 + m \\ L \geq \hat{n}_{\text{máx}} &\geq \hat{n} \end{aligned}$$

donde p es el orden de excitación a utilizar, satisfaciendo así el teorema 2.1. Además, se pueden considerar los resultados de las ecuaciones (2.17) y (2.18) para el diseño del filtro $F(q)$, los que mejoran el desempeño de la identificación. El resumen de estas condiciones y de la búsqueda secuencial se muestra en el algoritmo 2.2.

```

// región de prueba
1  escoja  $\hat{n}_{\text{máx}} > n$ ;
// filtro instrumental
2  escoja  $m \geq \hat{n}_{\text{máx}}$ ;
3  escoja el retardo del filtro  $\hat{n} \leq d_f \leq \hat{n} + \kappa_g/2$ ;
// diseño de la señal de entrada
4  escoja  $p \geq \hat{n}_{\text{máx}} + 1 + m$ ;
5  escoja  $M$  tal que  $2M + 1 \geq p$ ;
6  escoja  $L \geq \hat{n}_{\text{máx}}$ ;
7  genere  $u(k)$  l.e./p.e. de orden  $p$ ;
// medición
8  mida  $Z_{uy}^{L,M}$  en estado estacionario;
// identificación
9  for  $\hat{n} = 1$  to  $\hat{n}_{\text{máx}}$  do
10   tabla( $\hat{n}$ ) = IS( $\hat{n}$ );
11 end
12 buscar el menor  $\hat{n}$  en tabla( $\hat{n}$ ), tal que  $\Lambda_{zuy}^M(\hat{n})$ 
   es singular;
13  $n = \hat{n} - 1$ ;

```

Algoritmo 2.2: Búsqueda secuencial de n a través de la matriz $\Lambda_{zuy}^M(\hat{n})$.

Nótese que para el cálculo del índice $\text{IDR}(\hat{n})$ se necesita construir la matriz $\Lambda_{zuy}^M(\hat{n} + 1)$. Sin embargo, es suficiente cumplir con las siguientes condiciones

$$\begin{aligned} \hat{n}_{\text{máx}} &> n \\ m \geq \hat{n}_{\text{máx}} + 1 &\geq n \\ p \geq \hat{n}_{\text{máx}} + 2 + m &\geq (\hat{n} + 1) + 1 + m \\ 2M + 1 \geq p &\geq (\hat{n} + 1) + 1 + m \\ L \geq \hat{n}_{\text{máx}} + 1 &\geq \hat{n} + 1 \end{aligned}$$

A continuación se muestra un ejemplo para ilustrar el uso del algoritmo 2.2

Ejemplo 2.3 *Considérese la planta del ejemplo 2.2*

$$G(q) = \frac{0,3}{(q-0,8e^{j\pi/4})(q-0,8e^{-j\pi/4})(q-0,7)}$$

sometida a ruido de salida $v(k)$, generado filtrando ruido blanco estricto gaussiano $e(k)$ de media cero y varianza 1 mediante

$$v(k) = H(q)e(k), \quad H(q) = \frac{0,4q}{q-0,6}$$

Para poder identificar el orden de este sistema, se prueban secuencialmente varios órdenes de modelos $\hat{n} = 1, \dots, 2n$, utilizando los resultados obtenidos en el lema 2.2 y el teorema 2.1, a través del algoritmo 2.2, con los indicadores de singularidad $\text{IDR}(\hat{n})$ y $|\lambda|_{\min}(\Lambda_{zuy}^{\infty}(\hat{n}))$.

Así se diseña $u(k)$ como una realización de ruido blanco estricto gaussiano de media cero y varianza 1, generada independientemente del ruido $v(k)$, la cual satisface $\varphi_{uu}^{\infty}(\tau) = \delta_K(\tau)$ y por lo tanto es p.e. de cualquier orden. Además se eligen

$$\hat{n}_{\max} = 2n = 6, \quad m = \hat{n}_{\max} + 1 = 7$$

$$L = \hat{n}_{\max} + 1, \quad M = \{10^3, 5 \cdot 10^3, 10^4, 5 \cdot 10^4\}$$

y el filtro instrumental se diseña lo más simple posible, satisfaciendo también las condiciones del algoritmo 2.2

$$F(q) = q^{-d_f} + q^{-m}, \quad d_f = \hat{n}, \quad m = \hat{n}_{\max} + 1$$

De esta manera se realizan experimentos para cada valor de M , manteniéndose la misma señal de entrada, y obteniéndose como resultado las tablas 2.6-2.9. En estas tablas se muestra en color rojo cuando $|\lambda|_{\min}(\Lambda_{zuy}^M)$ cae en 2 órdenes de magnitud, detectándose así que $\Lambda_{zuy}^M(\hat{n})$ es “singular” y por lo tanto, el orden más adecuado para el modelo sería $\hat{n} = 3$. De forma similar ocurre con $\text{IDR}(\hat{n})$, mostrándose en azul cuando se produce un aumento de 2 órdenes de magnitud, lo cual también indicaría que el orden más apropiado para el modelo es $\hat{n} = 3$. Nótese que al aumentar el número de datos es posible reducir el efecto del ruido de salida, sin embargo nótese también el alto número de datos requeridos para mermar este efecto.

\hat{n}	$ \lambda _{\min}(\Lambda_{zux}^M)$	$ \lambda _{\min}(\Lambda_{zuy}^M)$	$\text{IDR}(\hat{n})$
1	1,39e-01	1,28e-01	2,69e+00
2	1,77e-01	2,05e-01	1,62e+00
3	8,84e-02	7,39e-02	8,17e+00
4	6,94e-17	2,76e-02	1,64e+01
5	8,14e-17	2,53e-02	5,13e+01
6	1,63e-16	8,08e-03	1,01e+02

Tabla 2.6: Singularidad de $\Lambda_{zuy}^M(\hat{n})$, para $M = 10^3$, $\text{SNR} = 5,70[\text{dB}]$.

\hat{n}	$ \lambda _{\min}(\Lambda_{zux}^M)$	$ \lambda _{\min}(\Lambda_{zuy}^M)$	$\text{IDR}(\hat{n})$
1	1,49e-01	1,45e-01	3,43e+00
2	1,74e-01	1,74e-01	1,82e+00
3	9,05e-02	8,70e-02	3,13e+01
4	2,62e-15	1,69e-03	3,79e+01
5	1,85e-15	4,39e-03	4,71e+02
6	2,93e-15	1,27e-03	7,59e+02

Tabla 2.7: Singularidad de $\Lambda_{zuy}^M(\hat{n})$, para $M = 5 \cdot 10^3$, $\text{SNR} = 5,92[\text{dB}]$.

\hat{n}	$ \lambda _{\min}(\Lambda_{zux}^M)$	$ \lambda _{\min}(\Lambda_{zuy}^M)$	$\text{IDR}(\hat{n})$
1	1,64e-01	1,66e-01	4,63e+00
2	1,70e-01	1,67e-01	1,97e+00
3	9,26e-02	9,08e-02	1,86e+02
4	3,05e-15	1,03e-04	1,97e+01
5	1,12e-15	1,42e-03	1,57e+02
6	6,70e-15	8,44e-04	6,75e+02

Tabla 2.8: Singularidad de $\Lambda_{zuy}^M(\hat{n})$, para $M = 10^4$, $\text{SNR} = 5,84[\text{dB}]$.

\hat{n}	$ \lambda _{\min}(\Lambda_{zux}^M)$	$ \lambda _{\min}(\Lambda_{zuy}^M)$	$\text{IDR}(\hat{n})$
1	1,57e-01	1,53e-01	4,02e+00
2	1,73e-01	1,72e-01	1,88e+00
3	9,17e-02	9,31e-02	2,30e+02
4	8,09e-15	6,16e-04	5,32e+02
5	1,23e-14	5,66e-04	3,42e+03
6	7,97e-15	3,70e-05	8,68e+02

Tabla 2.9: Singularidad de $\Lambda_{zuy}^M(\hat{n})$, para $M = 5 \cdot 10^4$, $\text{SNR} = 5,78[\text{dB}]$.

\hat{n}	$ \lambda _{\min}(\Lambda_{zux}^M)$	$ \lambda _{\min}(\Lambda_{zuy}^M)$	IDR(\hat{n})
1	2,18e-01	2,21e-01	3,98e+00
2	2,31e-01	2,36e-01	2,29e+01
3	7,31e-03	1,05e-02	3,28e+01
4	7,98e-17	3,29e-03	6,79e+01
5	6,30e-16	3,04e-03	2,62e+01
6	2,82e-16	3,16e-03	3,27e+01

Tabla 2.10: Singularidad de $\Lambda_{zuy}^M(\hat{n})$, con $M = 10^4$, SNR = 5,84[dB].

Considérese ahora los mismos valores de los experimentos anteriores, pero con $M = 10^4$ y utilizando el siguiente filtro instrumental

$$m = \hat{n}_{\max} + 1, \quad F(q) = q^{-m} + q^{-2m}$$

Así, se realiza la misma búsqueda secuencial obteniéndose la tabla 2.10. En esta tabla se puede apreciar que al utilizar $F(q)$ con retardo mayor que en los casos anteriores, se obtienen valores propios más cercanos al nivel de ruido, lo cual dificulta la correcta estimación del orden del sistema. Esto se debe a que el retardo de $F(q)$ se encuentra cerca de los límites establecidos en la ecuación (2.17).

Para más detalles de la simulación, véase el programa F.3.3.

2.3. Matriz generalizada de producto instrumental

2.3.1. Construcción

Considérese el siguiente sistema generador de datos o *planta*, el que es completamente desconocido

$$x(k) = G(q)u(k), \quad G(q) = \frac{B(q)}{A(q)} \quad (2.19)$$

$$\begin{aligned} B(q) &= b_0 + b_1q^{-1} + \dots + b_{n_b}q^{-n_b} \\ A(q) &= 1 + a_1q^{-1} + \dots + a_{n_a}q^{-n_a} \end{aligned}$$

donde

$$\begin{aligned} b_{n_b} \text{ y } a_{n_a} &\text{ son diferentes de cero} \\ B(q) \text{ y } A(q) &\text{ son coprimos} \\ G(q) &\text{ es estable} \end{aligned}$$

El sistema puede ser escrito como regresión lineal

según

$$\begin{aligned} x(k) &= \omega_{ux}^{n_b, n_a}(k)^T \beta \quad (2.20) \\ \omega_{ux}^{n_b, n_a}(k) &= \begin{bmatrix} u(k) \\ \vdots \\ u(k - n_b) \\ x(k - 1) \\ \vdots \\ x(k - n_a) \end{bmatrix}, \quad \beta = \begin{bmatrix} b_0 \\ \vdots \\ b_{n_b} \\ -a_1 \\ \vdots \\ -a_{n_a} \end{bmatrix} \end{aligned}$$

El principal problema ahora consiste en estimar los valores n_a y n_b del sistema anterior, ya que éstos definen su estructura, y su conocimiento facilitaría una posterior etapa de identificación paramétrica. Así, estos valores n_a y n_b deben ser estimados a partir de la única información que se tiene acerca de este sistema, que corresponde a un conjunto de datos $Z_{ux}^{L, M}$, definidos en (2.4), los cuales han sido obtenidos experimentalmente.

Al igual que en los casos de la *matriz de producto* (PM) y de la *matriz de producto instrumental* (IPM), la estimación de la estructura del sistema puede llevarse a cabo teniendo en cuenta el siguiente modelo, en su forma de regresión lineal

$$\begin{aligned} \hat{x}_i(k) &= \omega_{ux}^{\hat{n}_b, \hat{n}_a}(k)^T \gamma_i \quad (2.21) \\ \gamma_i &= [b_0^i \quad \dots \quad b_{\hat{n}_b}^i \quad -a_1^i \quad \dots \quad -a_{\hat{n}_a}^i]^T \end{aligned}$$

Si los datos medidos $Z_{ux}^{L, M}$ son ordenados utilizando este modelo, para cada instante de tiempo entre $[-M, M]$, entonces se puede definir la siguiente matriz de tamaño $(2M + 1) \times (\hat{n}_b + 1 + \hat{n}_a)$

$$\Omega_{ux}^M(\hat{n}_b, \hat{n}_a) := \begin{bmatrix} \omega_{ux}^{\hat{n}_b, \hat{n}_a}(-M)^T \\ \vdots \\ \omega_{ux}^{\hat{n}_b, \hat{n}_a}(M)^T \end{bmatrix} \quad (2.22)$$

donde $L \geq \max(\hat{n}_b, \hat{n}_a)$, para que la matriz $\Omega_{ux}^M(\hat{n}_b, \hat{n}_a)$ esté bien construida.

De manera similar al caso de la matriz PM, es el rango de la matriz $\Omega_{ux}^M(\hat{n}_b, \hat{n}_a)$ el que provee de información para poder identificar la estructura del sistema. Así, supóngase por el momento que M es suficientemente grande. Entonces, se puede notar que $\Omega_{ux}^M(\hat{n}_b, \hat{n}_a)$ con $i = \min(\hat{n}_b - n_b, \hat{n}_a - n_a) > 0$ es siempre deficiente de rango, ya que ella tiene al menos i columnas l.d., debido a la ecuación del sistema

$$x(k - j) = \omega_{ux}^{n_b, n_a}(k - j)^T \beta, \quad j = 1, \dots, i$$

En el caso $i = \min(\hat{n}_b - n_b, \hat{n}_a - n_a) \leq 0$ también se puede notar que la matriz $\Omega_{ux}^M(\hat{n}_b, \hat{n}_a)$ puede ser de rango completo, ya que aparentemente no tendría columnas l.d. Así su rango resulta

$$\begin{aligned} \text{si } i \leq 0, \quad & \text{rango } \Omega_{ux}^M(\hat{n}_b, \hat{n}_a) \leq \hat{n}_b + 1 + \hat{n}_a \\ \text{si } i > 0, \quad & \text{rango } \Omega_{ux}^M(\hat{n}_b, \hat{n}_a) \leq \hat{n}_b + 1 + \hat{n}_a - i \end{aligned}$$

donde

$$\hat{n}_b + 1 + \hat{n}_a - i = \max(\hat{n}_b + n_a, n_b + \hat{n}_a) + 1$$

Ahora, para asegurar que los resultados obtenidos anteriormente correspondan efectivamente al rango de la matriz $\Omega_{ux}^M(\hat{n}_b, \hat{n}_a)$, se define la *matriz de producto instrumental generalizada* (GIPM)

$$\begin{aligned} \Lambda_{zuy}^M(\hat{n}_b, \hat{n}_a) &:= \frac{\Omega_{uz}^M(\hat{n}_b, \hat{n}_a)^T \Omega_{uy}^M(\hat{n}_b, \hat{n}_a)}{2M + 1} \quad (2.23) \\ &= \frac{1}{2M + 1} \sum_{k=-M}^M \omega_{uz}^{\hat{n}_b, \hat{n}_a}(k) \omega_{uy}^{\hat{n}_b, \hat{n}_a}(k)^T \end{aligned}$$

donde cada elemento de esta matriz está definido por la correlación muestral en (1.2), para $0 \leq (\tau_1, \tau_2) \leq L$.

Por el momento se analizará el caso básico, en que $z(k) = y(k) = x(k)$, para el cual se cumple la siguiente igualdad

$$\text{rango } \Lambda_{xux}^M(\hat{n}_b, \hat{n}_a) = \text{rango } \Omega_{ux}^M(\hat{n}_b, \hat{n}_a)$$

Así, mediante el siguiente lema, se puede obtener de manera exacta el rango de las matrices $\Lambda_{xux}^M(\hat{n}_b, \hat{n}_a)$ y $\Omega_{ux}^M(\hat{n}_b, \hat{n}_a)$

Lema 2.3 *Considérese los datos $Z_{ux}^{L,M}$, medidos a partir del sistema en la ecuación (2.19). Si la señal $s(k) = A(q)^{-1}u(k)$ es l.e. de orden $p \forall k \in [-M, M]$, entonces la matriz $\Lambda_{xux}^M(\hat{n}_b, \hat{n}_a)$ satisface*

$$\begin{aligned} \text{rango } \Lambda_{xux}^M(\hat{n}_b, \hat{n}_a) &= \\ &\begin{cases} \hat{n}_b + 1 + \hat{n}_a, & i \leq 0 \\ \max(\hat{n}_b + n_a, n_b + \hat{n}_a) + 1, & i > 0 \end{cases} \end{aligned}$$

donde

$$\begin{aligned} p &= \max(\hat{n}_b + n_a, n_b + \hat{n}_a) + 1 \\ i &= \min(\hat{n}_b - n_b, \hat{n}_a - n_a) \\ 2M + 1 &\geq p, \quad L \geq \max(\hat{n}_b, \hat{n}_a) \end{aligned}$$

Para el caso $M \rightarrow \infty$ el resultado es el mismo, pero l.e. debe ser reemplazado por p.e. La demostración

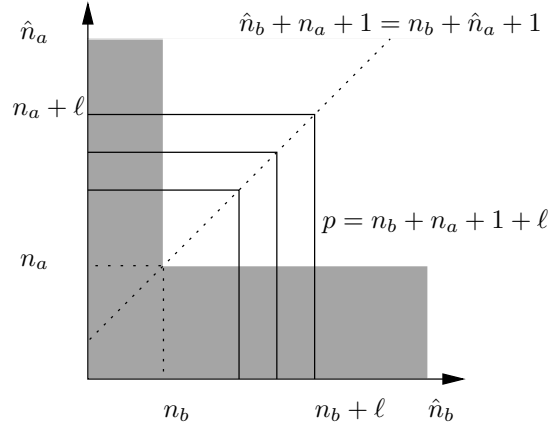


Figura 2.4: En gris: región donde $\Lambda_{xux}^M(\hat{n}_b, \hat{n}_a)$ tiene rango completo.

para el caso $M \rightarrow \infty$ es similar a la del caso finito, mostrada a continuación.

Demostración. Para el caso en que $i \leq 0$, la matriz $\Lambda_{xux}^M(\hat{n}_b, \hat{n}_a)$ es definida positiva por el teorema 1.4. Para el caso $i > 0$, el rango se determina removiendo todas las columnas l.d. de $\Omega_{ux}^M(\hat{n}_b, \hat{n}_a)$ y verificando que las restantes son l.i., por medio del teorema 1.5. **QED.**

Un enfoque basado en *total least squares* (TLS) se utilizó en [7] para obtener un resultado similar al de este lema, sin embargo [7] presenta una deficiencia, ya que utiliza señales p.e. para generar el conjunto finito de datos $Z_{ux}^{L,M}$, lo que es válido sólo cuando M es muy grande.

A diferencia del caso de la matriz PM $\Lambda_{xux}^M(\hat{n})$, para el caso de la matriz $\Lambda_{xux}^M(\hat{n}_b, \hat{n}_a)$ se debe tener mayor cuidado al analizar la región en que ella es no singular. Por ejemplo, la figura 2.4 muestra en color gris la región $\min(\hat{n}_b - n_b, \hat{n}_a - n_a) \leq 0$. De esta manera, si para cada punto (\hat{n}_b, \hat{n}_a) de esa región se satisface el orden de excitación $\max(\hat{n}_b + n_a, n_b + \hat{n}_a) + 1$, entonces la matriz $\Lambda_{xux}^M(\hat{n}_b, \hat{n}_a)$ es no singular en esa región. En la región $\min(\hat{n}_b - n_b, \hat{n}_a - n_a) > 0$, mostrada en color blanco, la matriz $\Lambda_{xux}^M(\hat{n}_b, \hat{n}_a)$ es deficiente de rango y su rango está determinado por el lema 2.3, si se satisface el orden de excitación $\max(\hat{n}_b + n_a, n_b + \hat{n}_a) + 1$. Nótese también que en la figura 2.4 se muestran 3 líneas continuas que son “líneas isoexcitantes”; es decir, todas las matrices $\Lambda_{xux}^M(\hat{n}_b, \hat{n}_a)$, que están ubicadas sobre una de estas líneas, requieren el mismo orden de excitación para satisfacer el lema 2.3, el

que aumenta a medida que \hat{n}_a o \hat{n}_b crecen.

Considérese ahora que la salida $x(k)$ del sistema en (2.19) se ve afectada por ruido coloreado $v(k)$, cuya distribución de probabilidad es desconocida, y es generado al filtrar ruido blanco $e(k)$ según

$$y(k) = x(k) + v(k) \quad (2.24)$$

$$v(k) = H(q)e(k), \quad H(q) = \sum_{j=0}^{\infty} h(j)q^{-j}$$

$e(k)$ ruido blanco de media 0 y varianza 1
 $H(q)$ es un filtro estable

Entonces la salida medible resulta $y(k) = x(k) + v(k)$, y por consiguiente los datos experimentales $Z_{uy}^{L,M}$ se ven igualmente perturbados con ruido.

Tal como se ha visto en el corolario 1.5 y en el ejemplo 2.2, la presencia de ruido en la salida del sistema perjudica el proceso de estimación de estructura, ya que se utilizan matrices contaminadas con ruido $\Lambda_{yuy}^M(\hat{n}_b, \hat{n}_a)$ u $\Omega_{uy}^M(\hat{n}_b, \hat{n}_a)$. Para resolver este problema, se define la *matriz de producto instrumental*⁵ (IPM) $\Lambda_{zuy}^M(\hat{n}_b, \hat{n}_a)$. Esta matriz incorpora una secuencia instrumental $z(k)$, que se obtiene al filtrar $u(k)$ según

$$z(k) = F(q)u(k), \quad F(q) = \frac{L(q)}{K(q)} \quad (2.25)$$

$$L(q) = l_0 + l_1q^{-1} + \dots + l_{m_l}q^{-m_l}$$

$$K(q) = 1 + k_1q^{-1} + \dots + k_{m_k}q^{-m_k}$$

con

k_{m_k} y l_{m_l} diferentes de cero
 $L(q)$ y $K(q)$ coprimos
 $F(q)$ es un filtro estable

Ahora, al igual que en el caso de la matriz IPM, gracias al lema 2.2 es posible eliminar el ruido del proceso de identificación estructural, resultando

$$\Lambda_{zuy}^{\infty}(\hat{n}_b, \hat{n}_a) = \Lambda_{zux}^{\infty}(\hat{n}_b, \hat{n}_a)$$

Una vez que el ruido ha sido eliminado del problema, sólo queda conocer las condiciones de singularidad de esta matriz libre de ruido $\Lambda_{zux}^M(\hat{n}_b, \hat{n}_a)$. Sería deseable que esta matriz tuviera las mismas propiedades de rango que $\Lambda_{xux}^M(\hat{n}_b, \hat{n}_a)$, para así poder utilizarla para estimación estructural.

⁵ver ecuación (2.23)

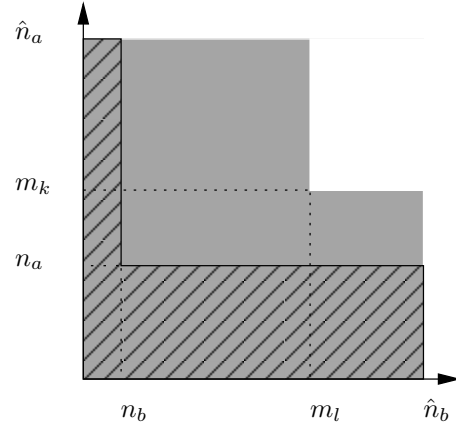


Figura 2.5: En gris: región donde $\Omega_{uz}^M(\hat{n}_b, \hat{n}_a)$ tiene rango completo. Achurada: región donde $\Omega_{ux}^M(\hat{n}_b, \hat{n}_a)$ tiene rango completo.

Teorema 2.2 *Considérese los datos $Z_{ux}^{L,M}$ y $Z_{uz}^{L,M}$ medidos de (2.19) y (2.25) respectivamente. Si la señal $s(k) = A(q)^{-1}u(k)$ es l.e. de orden $p \forall k \in [-M, M]$, y el filtro $F(q)$ satisface $\max(m_l - n_b, m_k - n_a) \geq 0$, entonces*

$$\text{rango } \Lambda_{zux}^M(\hat{n}_b, \hat{n}_a) = \text{rango } \Lambda_{xux}^M(\hat{n}_b, \hat{n}_a), \quad j \leq 0$$

es decir

$$\text{rango } \Lambda_{zux}^M(\hat{n}_b, \hat{n}_a) = \begin{cases} \hat{n}_b + 1 + \hat{n}_a, & i \leq 0 \text{ y } j \leq 0 \\ \max(\hat{n}_b + n_a, n_b + \hat{n}_a) + 1, & i > 0 \text{ y } j \leq 0 \end{cases}$$

donde

$$\begin{aligned} p &= \max(\hat{n}_b, \hat{n}_a) + \max(m_l, m_k, n_b, n_a) + 1 \\ i &= \min(\hat{n}_b - n_b, \hat{n}_a - n_a) \\ j &= \min(\hat{n}_b - m_l, \hat{n}_a - m_k) \\ L &\geq \max(\hat{n}_b, \hat{n}_a), \quad 2M + 1 \geq p \end{aligned}$$

Demostración. Debido al corolario D.4, se requiere que $\Omega_{uz}^M(\hat{n}_b, \hat{n}_a)$ sea de rango completo. Así

$$\begin{aligned} \text{rango } \Lambda_{zux}^M(\hat{n}_b, \hat{n}_a) &= \text{rango } \Omega_{uz}^M(\hat{n}_b, \hat{n}_a)^T \Omega_{ux}^M(\hat{n}_b, \hat{n}_a) \\ &= \text{rango } \Omega_{ux}^M(\hat{n}_b, \hat{n}_a) \end{aligned}$$

Para que $\Omega_{uz}^M(\hat{n}_b, \hat{n}_a)$ sea de rango completo en la región $j \leq 0$, es suficiente cumplir que $s(k) = A(q)^{-1}u(k)$ sea l.e. de orden $p \forall k \in [-M, M]$, según el lema 2.3

$$p \geq \max(\hat{n}_b + m_k, \hat{n}_a + m_l) + 1$$

Ahora, para que el rango $\Omega_{ux}^M(\hat{n}_b, \hat{n}_a)$ coincida con el rango del lema 2.3, es suficiente que $s(k) = A(q)^{-1}u(k)$ sea l.e. de orden p

$$p \geq \max(\hat{n}_b + n_a, \hat{n}_a + n_b) + 1$$

Además, se puede ver de la figura 2.5, que se necesita que $\max(m_l - n_b, m_k - n_a) \geq 0$ para que $\Omega_{uz}^M(\hat{n}_b, \hat{n}_a)$ sea de rango completo, mientras $\Omega_{ux}^M(\hat{n}_b, \hat{n}_a)$ pasa de *rango completo* a *deficiente de rango*, resultando de esta manera que el punto (n_b, n_a) queda cubierto por este teorema. **QED.**

Nota 2.5 Si M es muy grande, pero finito, la condición l.e. en el teorema 2.1 puede ser aproximada por la condición p.e., y el resultado debería permanecer sin cambios. Sin embargo, cuando se considera M infinito, el resultado del teorema 2.2 se mantiene “genéricamente”, es decir, pueden aparecer casos extraños en los que no se cumple. Estos casos son bastante raros, y son producidos por combinaciones desafortunadas de $F(q)$, $G(q)$ y $u(k)$, pero desde un punto de vista práctico no hay necesidad de preocuparse, ya que éstos aparecen muy esporádicamente. Para más detalles consúltese la sección 2.4 de este mismo capítulo.

Nota 2.6 El teorema 2.2 provee de condiciones no asintóticas (en el número de mediciones) para el proceso de identificación estructural mediante matrices instrumentales. De esta forma, las propiedades asintóticas sólo se necesitan para reducción de ruido, según el lema 2.2.

En [19] y [14] ya era conocido que

$$\det \Lambda_{zux}^\infty(\hat{n}_b, \hat{n}_a) \begin{cases} \neq 0, & i \leq 0 \text{ y } j \leq 0 \\ = 0, & i > 0 \text{ y } j \leq 0 \end{cases}$$

es válido “genéricamente”, con

$$\begin{aligned} i &= \min(\hat{n}_b - n_b, \hat{n}_a - n_a) \\ j &= \min(\hat{n}_b - m_l, \hat{n}_a - m_k) \end{aligned}$$

ver [19, example 4.1, theorem 4.1] y al final de [14, sección 8.6]. Sin embargo, el teorema 2.2 aporta nueva información, indicando además el rango de $\Lambda_{zux}^M(\hat{n}_b, \hat{n}_a)$, el que puede ser usado para mejorar algunos métodos de identificación estructural ya existentes (ver capítulo 3).

2.3.2. Indicaciones para el diseño del filtro instrumental

Según el teorema 2.2, es suficiente que el filtro instrumental satisfaga la condición

$$\max(m_l - n_b, m_k - n_a) \geq 0$$

es decir, que el orden del numerador o del denominador del filtro instrumental superen al orden del numerador o denominador de la planta, la cual es una condición muy poco restrictiva. Sin embargo, según lo visto en la sección 2.2.2, el diseño del retardo d_f del filtro instrumental puede contribuir a mejorar la identificación en los casos con ruido.

Ahora se generalizarán, para el caso de la matriz GIPM, los resultados obtenidos en (2.17) y (2.18). Para ello considérese la siguiente partición

$$\Lambda_{zuy}^\infty(\hat{n}_b, \hat{n}_a) = \begin{bmatrix} \Gamma_{uu}^\infty(\hat{n}_b + 1) & F_{yu}^\infty(\hat{n}_b, \hat{n}_a)^T \\ F_{zu}^\infty(\hat{n}_b, \hat{n}_a) & \Gamma_{zy}^\infty(\hat{n}_a) \end{bmatrix} \quad (2.26)$$

donde

$$\Gamma_{\alpha\beta}^\infty(\hat{n}) = \begin{bmatrix} \varphi_{\alpha\beta}^\infty(0) & \cdots & \varphi_{\alpha\beta}^\infty(\hat{n} - 1) \\ \vdots & \ddots & \vdots \\ \varphi_{\alpha\beta}^\infty(1 - \hat{n}) & \cdots & \varphi_{\alpha\beta}^\infty(0) \end{bmatrix} \quad (2.27)$$

$$F_{\alpha\beta}^\infty(\hat{n}_b, \hat{n}_a)^T = \begin{bmatrix} \varphi_{\alpha\beta}^\infty(-1) & \cdots & \varphi_{\alpha\beta}^\infty(-\hat{n}_a) \\ \varphi_{\alpha\beta}^\infty(0) & \cdots & \varphi_{\alpha\beta}^\infty(1 - \hat{n}_a) \\ \vdots & \ddots & \vdots \\ \varphi_{\alpha\beta}^\infty(\hat{n}_b - 1) & \cdots & \varphi_{\alpha\beta}^\infty(\hat{n}_b - \hat{n}_a) \end{bmatrix} \quad (2.28)$$

Es claro que al utilizar un número finito de muestras se producen errores en la matriz Λ_{zuy}^M , debido a que la reducción de ruido no es completa

$$\Lambda_{zuy}^M = \underbrace{\begin{bmatrix} \Gamma_{uu}^M & (F_{xu}^M)^T \\ F_{zu}^M & \Gamma_{zx}^M \end{bmatrix}}_{\Lambda_{zux}^M} + \underbrace{\begin{bmatrix} 0 & (F_{vu}^M)^T \\ 0 & \Gamma_{zv}^M \end{bmatrix}}_{E_{zuv}^M}$$

La idea ahora es diseñar un filtro $F(q)$ que permita que la matriz de ruido E_{zuv}^M sea lo más pequeña posible frente a la matriz libre de ruido Λ_{zux}^M .

Una forma de lograr lo anterior, como ya se vio en la sección 2.2.2, es potenciar el tamaño de la matriz $\Gamma_{zx}^M(\hat{n}_a)$ a través del diseño del retardo d_f del filtro instrumental. Para ello considérese por simplicidad que $u(k)$ satisface $\varphi_{uu}^\infty(\tau) = a_u^2 \delta_K(\tau)$. Entonces, la

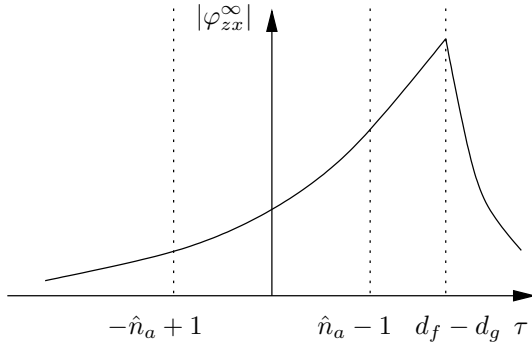


Figura 2.6: Cota para la correlación muestral $\varphi_{zx}^{\infty}(\tau)$, donde la exponencial izquierda está asociada a $G(q)$, y la derecha, asociada a $F(q)$.

matriz $\Gamma_{zx}^{\infty}(\hat{n}_a)$ queda compuesta por los siguientes elementos

$$\forall \tau = -\hat{n}_a + 1, \dots, \hat{n}_a - 1, \varphi_{zx}^{\infty}(\tau) = a_u^2 (f * \mathcal{J}g^*)(\tau)$$

donde $f(k)$ y $g(k)$ son las respuestas a impulso de los filtros $F(q)$ y $G(q)$ respectivamente. Supóngase además que ambas respuestas a impulso se encuentran acotadas por

$$\begin{aligned} |f(k)| &\leq A_f \epsilon_f(k - d_f) = A_f \lambda_f^{(k-d_f)} \mu(k - d_f) \\ |g(k)| &\leq A_g \epsilon_g(k - d_g) = A_g \lambda_g^{(k-d_g)} \mu(k - d_g) \end{aligned}$$

donde d_f es el retardo del filtro instrumental y d_g es el retardo de la planta. Entonces, los elementos $\varphi_{zx}^{\infty}(\tau)$ se ven acotados por el resultado obtenido en la ecuación (2.15), el que se muestra en la figura 2.6.

Así, para que $\varphi_{zx}^{\infty}(\tau)$ concentre valores grandes en el intervalo $|\tau| \leq (\hat{n}_a - 1)$, es suficiente cumplir

$$(d_f - d_g) - (\hat{n}_a - 1) \leq \frac{\kappa_g}{2}$$

donde κ_g es la constante de tiempo de la planta $G(q)$, dada por

$$\kappa_g = \frac{-1}{\ln \lambda_g} \Leftrightarrow e^{-1/\kappa_g} = \lambda_g$$

Ahora, dado que en muchos casos de sistemas muestrados el retardo de la planta es $d_g = 1$, entonces el retardo del filtro instrumental d_f debería satisfacer

$$0 \leq d_f \leq \frac{\kappa_g}{2} + \hat{n}_a \quad (2.29)$$

Para concentrar los valores propios de Λ_{zux}^{∞} cerca de los elementos diagonales, al igual que en la sección 2.2.2, considérese los elementos que conforman la matriz $F_{zu}^{\infty}(\hat{n}_b, \hat{n}_a)$, es decir

$$\forall \tau = -\hat{n}_a, \dots, (\hat{n}_b - 1), \varphi_{zu}^{\infty}(\tau) = a_u^2 f(\tau)$$

Dado que $F(q)$ es causal y tiene retardo d_f , entonces

$$\forall \tau = -\hat{n}_a, \dots, (d_f - 1), \varphi_{zu}^{\infty}(\tau) = a_u^2 f(\tau) = 0$$

Por este motivo, si se escoge

$$d_f \geq \hat{n}_b \quad (2.30)$$

entonces la matriz $F_{zu}^{\infty}(\hat{n}_b, \hat{n}_a) = 0$, y por lo tanto se lograría concentrar mejor los valores propios de Λ_{zux}^{∞} cerca de los elementos diagonales, los cuales deberían ser relativamente grandes, si d_f satisface la ecuación (2.29).

2.3.3. Búsqueda secuencial

En la figura 2.4 se puede ver que el problema de identificación estructural en el caso sin ruido, para el sistema en (2.19), consiste básicamente en encontrar la esquina (n_b, n_a) en el plano (\hat{n}_b, \hat{n}_a) . Este problema se ha abordado en [13], [1] y [2] mediante una búsqueda secuencial, a través de la construcción de tablas cuyos índices son \hat{n}_b y \hat{n}_a . Cada entrada (\hat{n}_b, \hat{n}_a) de esta tabla tiene asociada un índice de singularidad para la matriz Λ_{xux}^M , como por ejemplo:

$$\begin{aligned} \text{IS}(\hat{n}_b, \hat{n}_a) &= \det \Lambda_{xux}^M(\hat{n}_b, \hat{n}_a) \\ \text{IS}(\hat{n}_b, \hat{n}_a) &= \lambda_{\min}(\Lambda_{xux}^M(\hat{n}_b, \hat{n}_a)) \end{aligned}$$

o como el utilizado en [13]

$$\text{IS}(\hat{n}_b, \hat{n}_a) = \lambda_{\min}(\Lambda_{xux}^M(\hat{n}_b, \hat{n}_a)) \cdot (N^{1/N})^{(\hat{n}_b + \hat{n}_a)}$$

donde N es el número de datos utilizados. De esta forma, el orden más apropiado para el modelo correspondería a la esquina inmediatamente anterior a donde ocurre una caída brusca en el índice de singularidad utilizado.

Finalmente, para buscar el orden de la planta entre los modelos $(\hat{n}_b, \hat{n}_a) \in [1, \hat{n}_{\max}] \times [1, \hat{n}_{\max}]$, es suficiente cumplir las siguientes condiciones

$$\begin{aligned} \hat{n}_{\max} &> n = \max(n_b, n_a) \\ p \geq 2\hat{n}_{\max} + 1 &\geq \max(\hat{n}_b + n_a, n_b + \hat{n}_a) + 1 \\ 2M + 1 \geq p &\geq \max(\hat{n}_b + n_a, n_b + \hat{n}_a) + 1 \\ L \geq \hat{n}_{\max} &\geq \max(\hat{n}_b, \hat{n}_a) \end{aligned}$$

		\hat{n}_a					
		1	2	3	4	5	6
\hat{n}_b	1	1,02e+00	8,61e-01	2,74e-01	1,02e-01	5,88e-02	3,94e-02
	2	9,15e-01	7,36e-01	2,74e-01	7,19e-02	2,58e-02	1,14e-02
	3	9,08e-01	6,70e-01	6,62e-02	1,61e-14	2,72e-14	2,05e-14
	4	8,83e-01	4,39e-01	2,11e-02	2,12e-14	1,24e-15	4,63e-15
	5	8,04e-01	2,39e-01	8,85e-03	3,29e-14	3,02e-15	3,78e-16
	6	6,15e-01	1,16e-01	4,64e-03	2,31e-14	1,97e-15	1,45e-15

Tabla 2.11: Tabla de singularidad con el indicador $IS(\hat{n}_b, \hat{n}_a) = \lambda_{\min}(\Lambda_{xux}^M(\hat{n}_b, \hat{n}_a))$.

donde p es el orden de excitación a utilizar, cumpliendo así las condiciones del lema 2.3, las que se resumen en el algoritmo 2.3.

Los parámetros a utilizar en el algoritmo 2.3 se escogen según

$$\hat{n}_{\max} = 6, \quad M = 500, \quad L = \hat{n}_{\max}$$

y la señal de entrada $u(k)$ se diseña como una realización de ruido blanco estricto gaussiano de media cero y varianza 1, la cual es p.e. de cualquier orden.

Así se obtiene el resultado de la tabla 2.11, en la cual se puede encontrar fácilmente la región donde $\Lambda_{xux}^M(\hat{n}_b, \hat{n}_a)$ es singular (delimitada con color rojo), y por consiguiente, la esquina $(\hat{n}_b^o, \hat{n}_a^o) = (3, 4)$, donde se produce una caída brusca en el índice de singularidad. Esto último indica que la estructura identificada resulta $(n_b, n_a) = (\hat{n}_b^o - 1, \hat{n}_a^o - 1) = (2, 3)$, mostrada en color azul, la que corresponde a la estructura de la planta.

Para más detalles de la simulación, consúltese el programa F.3.4.

Al igual que en el caso sin ruido, el problema de identificación estructural para el sistema con ruido en (2.24), consiste básicamente en encontrar la esquina (n_b, n_a) en el plano (\hat{n}_b, \hat{n}_a) , mostrada en la figura 2.5. Este problema se puede abordar mediante una búsqueda secuencial, a través de la construcción de tablas cuyos índices son \hat{n}_b y \hat{n}_a . Cada entrada (\hat{n}_b, \hat{n}_a) de esta tabla tiene asociada un índice de singularidad para la matriz Λ_{zuy}^M , como por ejemplo:

$$\begin{aligned} IS(\hat{n}_b, \hat{n}_a) &= \det(\Lambda_{zuy}^M(\hat{n}_b, \hat{n}_a)) \\ IS(\hat{n}_b, \hat{n}_a) &= |\lambda|_{\min}(\Lambda_{zuy}^M(\hat{n}_b, \hat{n}_a)) \end{aligned}$$

donde $|\lambda|_{\min}$ corresponde al valor propio más cercano a cero. De esta forma, el orden más apropiado para el modelo correspondería a la esquina inmediatamente anterior a donde ocurre una caída brusca en el índice de singularidad utilizado.

```

// región de prueba
1  escoja  $\hat{n}_{\max} > n = \max(n_b, n_a)$ ;
// diseño de señal de entrada
2  escoja  $p \geq 2\hat{n}_{\max} + 1$ ;
3  escoja  $M$  tal que  $2M + 1 \geq p$ ;
4  escoja  $L \geq \hat{n}_{\max}$ ;
5  genere  $u(k)$  l.e./p.e. de orden  $p$ ;
// medición
6  mida  $Z_{ux}^{L,M}$  en estado estacionario;
// identificación
7  for  $\hat{n}_b = 1$  to  $\hat{n}_{\max}$  do
8    for  $\hat{n}_a = 1$  to  $\hat{n}_{\max}$  do
9      tabla( $\hat{n}_b, \hat{n}_a$ ) = IS( $\hat{n}_b, \hat{n}_a$ );
10   end
11 end
12 buscar la menor esquina  $(\hat{n}_b^o, \hat{n}_a^o)$  en
    tabla( $\hat{n}_b, \hat{n}_a$ ), tal que  $\Lambda_{xux}^M(\hat{n}_b^o, \hat{n}_a^o)$  es singular;
13  $(n_b, n_a) = (\hat{n}_b^o - 1, \hat{n}_a^o - 1)$ ;

```

Algoritmo 2.3: Búsqueda secuencial de (n_b, n_a) a través de la matriz $\Lambda_{xux}^M(\hat{n}_b, \hat{n}_a)$.

A continuación se muestra un ejemplo, para ilustrar el uso del algoritmo 2.3

Ejemplo 2.4 Considérese la planta en [2, example 3]

$$G(q) = \frac{1 - 0,45q^{-1} - q^{-2}}{1 - 2,2q^{-1} + 1,77q^{-2} - 0,52q^{-3}}$$

Para identificar la estructura de este sistema, se prueban secuencialmente los modelos $(\hat{n}_b, \hat{n}_a) \in [1, 6] \times [1, 6]$, utilizando el resultado obtenido en el lema 2.3 a través del algoritmo 2.3, con el indicador de singularidad $\lambda_{\min}(\Lambda_{xux}^M(\hat{n}_b, \hat{n}_a))$.

$$\begin{aligned}
\hat{n}_{\text{máx}} &> n = \text{máx}(n_b, n_a) \\
0 \leq \text{máx}(m_l, m_k) - \hat{n}_{\text{máx}} &\leq \text{máx}(m_l - n_b, m_k - n_a) \\
p \geq \hat{n}_{\text{máx}} + \text{máx}(m_l, m_k) + 1 &\geq \text{máx}(\hat{n}_b, \hat{n}_a) + \text{máx}(m_l, m_k, n_b, n_a) + 1 \\
2M + 1 \geq p &\geq \text{máx}(\hat{n}_b, \hat{n}_a) + \text{máx}(m_l, m_k, n_b, n_a) + 1 \\
L \geq \hat{n}_{\text{máx}} &\geq \text{máx}(\hat{n}_b, \hat{n}_a)
\end{aligned} \tag{2.31}$$

```

// region de prueba
1  escoja  $\hat{n}_{\text{máx}} > n = \text{máx}(n_b, n_a)$ ;
// filtro instrumental
2  escoja  $(m_l, m_k)$  tal que  $\text{máx}(m_l, m_k) \geq \hat{n}_{\text{máx}}$ ;
3  escoja el retardo del filtro  $\hat{n}_b \leq d_f \leq \hat{n}_a + \kappa_g/2$ ;
// diseño de señal de entrada
4  escoja  $p \geq \hat{n}_{\text{máx}} + \text{máx}(m_l, m_k) + 1$ ;
5  escoja  $M$  tal que  $2M + 1 \geq p$ ;
6  escoja  $L \geq \hat{n}_{\text{máx}}$ ;
7  genere  $u(k)$  l.e./p.e. de orden  $p$ ;
// medición
8  mida  $Z_{uy}^{L,M}$  en estado estacionario;
// identificación
9  for  $\hat{n}_b = 1$  to  $\hat{n}_{\text{máx}}$  do
10   for  $\hat{n}_a = 1$  to  $\hat{n}_{\text{máx}}$  do
11     tabla $(\hat{n}_b, \hat{n}_a) = \text{IS}(\hat{n}_b, \hat{n}_a)$ ;
12   end
13 end
14 buscar la menor esquina  $(\hat{n}_b^o, \hat{n}_a^o)$  en
    tabla $(\hat{n}_b, \hat{n}_a)$ , tal que  $\Lambda_{zuy}^M(\hat{n}_b^o, \hat{n}_a^o)$  es singular;
15  $(n_b, n_a) = (\hat{n}_b^o - 1, \hat{n}_a^o - 1)$ ;

```

Algoritmo 2.4: Búsqueda secuencial de (n_b, n_a) a través de la matriz $\Lambda_{zuy}^M(\hat{n}_b, \hat{n}_a)$.

Así, para buscar el orden de la planta entre los modelos $(\hat{n}_b, \hat{n}_a) \in [1, \hat{n}_{\text{máx}}] \times [1, \hat{n}_{\text{máx}}]$, es suficiente cumplir el conjunto de condiciones en (2.31), donde p es el orden de excitación a utilizar, satisfaciendo así el teorema 2.2. Además se pueden considerar los resultados de las ecuaciones (2.29) y (2.30) para el diseño del filtro $F(q)$, los cuales mejoran el desempeño de la identificación. Estas condiciones se resumen en el algoritmo 2.4. En caso de existir niveles importantes de ruido de salida, se debe considerar M muy grande, según el lema 2.2, para así atenuarlo.

A continuación se muestran algunos ejemplos, para ilustrar el uso del algoritmo 2.4

Ejemplo 2.5 Considere la planta en [2, example 3]

$$G(q) = \frac{1 - 0,45q^{-1} - q^{-2}}{1 - 2,2q^{-1} + 1,77q^{-2} - 0,52q^{-3}}$$

sometida a ruido de salida $v(k)$, generado filtrando ruido blanco estricto uniforme $e(k)$ de media cero y varianza 1 mediante

$$v(k) = H(q)e(k), \quad H(q) = \frac{0,4q}{q - 0,6}$$

Para identificar la estructura de este sistema, se prueban secuencialmente las estructuras $(\hat{n}_b, \hat{n}_a) \in [1, 6] \times [1, 6]$, utilizando los resultados del teorema 2.2 y del lema 2.2 a través del algoritmo 2.4, con el indicador de singularidad $|\lambda|_{\text{mín}}(\Lambda_{zuy}^M(\hat{n}_b, \hat{n}_a))$.

Los parámetros a utilizar en el algoritmo 2.4 se escogen según

$$\begin{aligned}
\hat{n}_{\text{máx}} &= 6, & (m_l, m_k) &= (\hat{n}_{\text{máx}}, 0) \\
L &= \hat{n}_{\text{máx}}, & M &= \{10^4, 5 \cdot 10^4\}
\end{aligned}$$

La señal de entrada $u(k)$ se diseña como una realización de ruido blanco estricto gaussiano de media cero y varianza 1, la que es p.e. de cualquier orden. Finalmente, el filtro instrumental se diseña de la manera más simple posible, satisfaciendo las condiciones del algoritmo 2.4

$$F(q) = q^{-d_f} + q^{-m_l}, \quad d_f = \hat{n}_b, \quad m_l = \hat{n}_{\text{máx}}$$

Así se obtiene el resultado de la tabla 2.12, con la cual se puede reconstruir la región donde $\Lambda_{zuy}^M(\hat{n}_b, \hat{n}_a)$ es singular, a partir de los valores indicados en color rojo, los que corresponden a caídas de aproximadamente dos órdenes de magnitud. De esta manera, se puede encontrar la esquina $(\hat{n}_b^o, \hat{n}_a^o) = (3, 4)$ donde se produce una caída apreciable en el índice de singularidad. Esto último indica que la estructura identificada resulta $(n_b, n_a) = (\hat{n}_b^o - 1, \hat{n}_a^o - 1) = (2, 3)$, mostrada en color azul, la que corresponde a la estructura de la planta. Para más detalles de la simulación, véase el programa F.3.5.

$M = 10^4$		\hat{n}_a					
		1	2	3	4	5	6
\hat{n}_b	1	7,95e-01	4,47e-01	3,71e-01	8,53e-02	9,43e-02	1,35e-01
	2	9,75e-01	3,27e-01	4,88e-01	5,16e-01	7,87e-02	8,59e-02
	3	9,63e-01	6,32e-01	5,29e-02	3,00e-03	3,59e-03	4,00e-03
	4	9,55e-01	6,17e-01	2,68e-02	6,14e-03	3,46e-03	9,50e-03
	5	9,51e-01	5,70e-01	2,47e-02	3,50e-04	6,95e-04	7,63e-04
	6	9,48e-01	4,88e-01	1,39e-02	7,50e-03	2,46e-03	1,26e-03

$M = 5 \cdot 10^4$		\hat{n}_a					
		1	2	3	4	5	6
\hat{n}_b	1	7,62e-01	4,64e-01	3,77e-01	9,08e-02	7,76e-02	1,30e-01
	2	9,97e-01	3,15e-01	4,85e-01	5,72e-01	8,86e-02	9,02e-02
	3	9,95e-01	6,27e-01	5,84e-02	2,17e-03	4,57e-03	1,13e-02
	4	9,94e-01	6,19e-01	3,13e-02	7,39e-04	2,35e-03	3,49e-03
	5	9,93e-01	5,62e-01	2,52e-02	1,23e-03	1,30e-03	9,25e-04
	6	9,93e-01	4,87e-01	2,00e-02	1,32e-03	3,45e-03	8,46e-04

Tabla 2.12: Tabla de singularidad con el indicador $|\lambda|_{\min}(\Lambda_{zuy}^M(\hat{n}_b, \hat{n}_a))$, considerando $\text{SNR} \approx 20[\text{dB}]$.

Ejemplo 2.6 Considérese ahora la planta expuesta en [13, example 4]

$$G(q) = \frac{B(q)}{A(q)}$$

$$B(q) = 1 + 0,3452q^{-1} + 0,53q^{-2} + 0,3985q^{-3} + 0,8138q^{-4}$$

$$A(q) = 1 + 0,7907q^{-1} + 0,042q^{-2} - 0,5556q^{-3} - 0,0247q^{-4} + 0,3846q^{-5} + 0,3026q^{-6}$$

la que es sometida a ruido de salida $v(k)$, generado filtrando ruido blanco estricto gaussiano $e(k)$ de media cero y varianza 1 mediante

$$v(k) = H(q)e(k), \quad H(q) = \frac{0,4q}{q - 0,6}$$

Para identificar la estructura de este sistema, se prueban secuencialmente las estructuras $(\hat{n}_b, \hat{n}_a) \in [1, 12] \times [1, 12]$, utilizando el algoritmo 2.4, con el indicador de singularidad $|\lambda|_{\min}(\Lambda_{zuy}^M(\hat{n}_b, \hat{n}_a))$.

Los parámetros a utilizar en el algoritmo 2.4 se escogen según

$$\hat{n}_{\max} = 12, \quad (m_l, m_k) = (\hat{n}_{\max}, 0)$$

$$L = \hat{n}_{\max}, \quad M = \{10^4, 5 \cdot 10^4\}$$

La señal de entrada $u(k)$ se diseña como una realización de ruido blanco estricto gaussiano de media cero y varianza 1, la que es p.e. de cualquier orden. Finalmente, el filtro instrumental se diseña de la manera

más simple posible, satisfaciendo las condiciones del algoritmo 2.4

$$F(q) = q^{-d_f} + q^{-m_l}, \quad d_f = \hat{n}_b, \quad m_l = \hat{n}_{\max}$$

Así se obtiene el resultado de la tabla 2.13, con la cual se puede reconstruir la región donde $\Lambda_{zuy}^M(\hat{n}_b, \hat{n}_a)$ es singular, a partir de los valores indicados en color rojo, los que corresponden a caídas de aproximadamente dos órdenes de magnitud. De esta manera, se puede encontrar la esquina $(\hat{n}_b^o, \hat{n}_a^o) = (5, 7)$ donde se produce una caída apreciable en el índice de singularidad. Esto último indica que la estructura identificada resulta $(n_b, n_a) = (\hat{n}_b^o - 1, \hat{n}_a^o - 1) = (4, 6)$, mostrada en color azul, la que corresponde con la estructura de la planta.

Para más detalles de la simulación, vea el programa F.3.5.

Nótese que en los ejemplos anteriores se presentan algunas dificultades al determinar la estructura del sistema, ya que no es fácil decidir cuándo la matriz GIPM $\Lambda_{zuy}^M(\hat{n}_b, \hat{n}_a)$ es realmente singular. Esto se debe a que el ruido de salida perturba los valores propios nulos de $\Lambda_{zuy}^M(\hat{n}_b, \hat{n}_a)$, haciéndolos relativamente comparables con los valores no nulos más pequeños, lo que provoca que la discriminación entre matrices singulares y no singulares sea un poco subjetiva. Afortunadamente, al construir las tablas de singularidad,

$M = 10^4$		\hat{n}_a											
		1	2	3	4	5	6	7	8	9	10	11	12
\hat{n}_b	1	2,86e-01	5,79e-01	4,22e-01	1,83e-02	5,92e-01	3,46e-01	4,54e-02	3,34e-01	3,81e-01	8,25e-02	2,61e-02	1,61e-01
	2	9,83e-01	9,77e-01	8,92e-01	8,17e-01	8,85e-01	8,30e-01	4,81e-01	2,68e-01	2,72e-01	2,83e-01	3,28e-01	2,39e-01
	3	4,40e-01	3,89e-03	3,45e-01	7,80e-01	6,46e-01	2,06e-01	3,54e-01	2,79e-01	2,59e-01	6,07e-02	2,67e-01	2,20e-01
	4	4,56e-01	5,24e-01	3,95e-01	3,75e-01	4,98e-01	5,70e-01	3,95e-01	3,23e-01	2,97e-01	2,99e-01	3,01e-01	2,21e-01
	5	4,14e-02	1,58e-01	6,02e-02	8,98e-02	2,41e-01	4,21e-01	5,88e-03	8,98e-03	1,15e-02	2,05e-02	2,76e-02	2,49e-02
	6	3,42e-01	1,18e-01	3,06e-03	2,29e-01	2,13e-01	3,22e-01	2,43e-03	3,53e-03	3,76e-03	9,53e-04	6,91e-03	6,95e-03
	7	2,56e-02	2,40e-01	2,68e-01	2,35e-01	4,55e-02	1,79e-01	1,94e-03	1,27e-03	1,02e-03	1,17e-03	3,14e-03	1,03e-02
	8	2,17e-01	2,76e-01	2,97e-01	3,25e-01	3,02e-01	2,09e-01	2,92e-04	2,97e-03	1,77e-03	2,31e-03	6,82e-03	5,91e-03
	9	3,18e-01	2,22e-02	2,28e-01	2,03e-01	1,62e-01	9,37e-02	1,06e-03	1,74e-03	3,27e-03	3,42e-03	2,45e-03	1,96e-03
	10	9,25e-02	1,40e-01	1,10e-01	2,57e-01	9,38e-02	1,07e-01	1,36e-04	6,09e-04	6,37e-04	8,75e-04	3,79e-04	2,01e-03
	11	9,36e-02	1,91e-01	8,19e-02	2,59e-01	6,26e-02	1,91e-02	8,45e-04	3,00e-03	2,88e-03	2,96e-03	3,03e-03	2,82e-04
	12	2,75e-01	1,92e-01	1,94e-01	2,21e-01	1,47e-02	1,26e-01	1,60e-03	4,04e-03	6,38e-04	4,33e-03	5,34e-03	7,30e-03

$M = 5 \cdot 10^4$		\hat{n}_a											
		1	2	3	4	5	6	7	8	9	10	11	12
\hat{n}_b	1	3,05e-01	5,74e-01	4,30e-01	8,64e-03	5,73e-01	3,49e-01	4,81e-02	3,43e-01	3,73e-01	8,06e-02	2,41e-02	1,59e-01
	2	9,94e-01	1,00e+00	8,54e-01	7,90e-01	8,53e-01	8,60e-01	4,93e-01	2,72e-01	2,72e-01	2,80e-01	3,28e-01	2,45e-01
	3	4,52e-01	1,13e-02	3,82e-01	8,00e-01	6,61e-01	2,26e-01	3,52e-01	2,71e-01	2,57e-01	6,94e-02	2,65e-01	2,17e-01
	4	4,52e-01	5,19e-01	3,95e-01	3,89e-01	5,08e-01	5,78e-01	4,13e-01	3,24e-01	3,00e-01	2,99e-01	2,96e-01	2,17e-01
	5	4,85e-02	1,57e-01	4,13e-02	8,44e-02	2,43e-01	4,11e-01	4,34e-04	1,07e-02	2,43e-02	2,90e-02	5,10e-02	3,80e-03
	6	3,40e-01	1,25e-01	1,05e-02	2,24e-01	2,01e-01	3,27e-01	2,47e-03	3,38e-03	5,76e-03	7,47e-03	9,86e-03	1,96e-02
	7	1,74e-02	2,48e-01	2,45e-01	2,23e-01	3,95e-02	1,81e-01	4,02e-03	2,64e-04	6,33e-04	7,64e-04	3,12e-03	1,05e-02
	8	2,16e-01	2,72e-01	2,92e-01	3,18e-01	2,96e-01	2,06e-01	4,34e-03	1,11e-04	7,26e-04	8,88e-04	2,51e-03	6,27e-03
	9	3,17e-01	1,48e-02	2,13e-01	1,90e-01	1,64e-01	9,54e-02	2,91e-03	6,01e-04	5,10e-04	6,22e-04	6,46e-04	5,39e-04
	10	9,24e-02	1,43e-01	1,07e-01	2,47e-01	9,98e-02	1,10e-01	6,85e-04	8,57e-04	1,33e-03	2,16e-03	6,28e-04	8,74e-04
	11	8,97e-02	1,84e-01	7,30e-02	2,47e-01	6,43e-02	1,91e-02	1,37e-03	5,84e-04	2,49e-03	5,18e-04	5,21e-04	3,48e-04
	12	2,73e-01	1,82e-01	1,74e-01	2,11e-01	9,15e-03	1,31e-01	1,16e-03	2,51e-03	3,33e-03	3,74e-03	1,09e-03	4,96e-04

Tabla 2.13: Tabla de singularidad con el indicador $|\lambda|_{\min}(\Lambda_{zuy}^M(\hat{n}_b, \hat{n}_a))$, considerando $\text{SNR} \approx 10[\text{dB}]$.

se tienen suficientes casos como para reconstruir la región en la cual la matriz GIPM es singular, permitiendo así la identificación estructural del sistema.

2.4. Rango genérico para la matriz GIPM

En esta sección se mostrará que el resultado para el caso $M \rightarrow \infty$ del teorema 2.2, y por consiguiente del teorema 2.1, son válidos de forma “genérica”, salvo un conjunto pequeño de casos, en el cual ambos teoremas no rigen, aún si $u(k)$, $G(q)$ y $F(q)$ satisfacen todas sus condiciones.

Lema 2.4 *Considérese la matriz $\Lambda_{zux}^\infty(\hat{n}_b, \hat{n}_a)$, definida en la ecuación (2.23), y la matriz $L_{ux}^\infty(\hat{n}_b, \hat{n}_a, i)$, definida en el teorema 1.5, con $i \geq 0$. Entonces es necesario (pero no suficiente) que las matrices*

$$\Lambda_{zuz}^\infty(\hat{n}_b, \hat{n}_a) \text{ y } L_{ux}^\infty(\hat{n}_b, \hat{n}_a, i)$$

sean no singulares, para que la matriz $L_{zux}^\infty(\hat{n}_b, \hat{n}_a, i)$

$$L_{zux}^M(\hat{n}_b, \hat{n}_a, i) := \frac{1}{2M+1} \sum_{k=-M}^M \omega_{uz}^{\hat{n}_b, \hat{n}_a}(k) w_{ux}^{\hat{n}_b, \hat{n}_a, i}(k)^T$$

de tamaño $(\hat{n}_b + 1 + \hat{n}_a) \times (\hat{n}_b + 1 + \hat{n}_a - i)$, tenga rango completo.

Demostración. Supóngase que $\Lambda_{zuz}^\infty(\hat{n}_b, \hat{n}_a)$ es singular. Entonces, existe un vector $\gamma_f \neq 0$ tal que

$$\gamma_f^T \Lambda_{zuz}^\infty(\hat{n}_b, \hat{n}_a) \gamma_f = 0 \Leftrightarrow \forall k, \quad \gamma_f^T \omega_{uz}^{\hat{n}_b, \hat{n}_a}(k) = 0$$

Así, dada esta última condición, se tiene que

$$\gamma_f^T L_{zux}^\infty(\hat{n}_b, \hat{n}_a, i) = 0$$

y por lo tanto $L_{zux}^\infty(\hat{n}_b, \hat{n}_a, i)$ es deficiente de rango. Ahora supóngase que $L_{ux}^\infty(\hat{n}_b, \hat{n}_a, i)$ es singular, entonces existe un vector $\gamma_g \neq 0$ tal que

$$\gamma_g^T L_{ux}^\infty(\hat{n}_b, \hat{n}_a, i) \gamma_g = 0 \Leftrightarrow \forall k, \quad w_{ux}^{\hat{n}_b, \hat{n}_a, i}(k)^T \gamma_g = 0$$

Así, dada esta última condición, se tiene que

$$L_{zux}^\infty(\hat{n}_b, \hat{n}_a, i) \gamma_g = 0$$

y por lo tanto $L_{zux}^\infty(\hat{n}_b, \hat{n}_a, i)$ es deficiente de rango.

QED.

Corolario 2.1 *Considérese la matriz $\Lambda_{zux}^\infty(\hat{n}_b, \hat{n}_a)$, definida en la ecuación (2.23). Entonces es necesario (pero no suficiente) que $\Lambda_{zuz}^\infty(\hat{n}_b, \hat{n}_a)$ y $\Lambda_{ux}^\infty(\hat{n}_b, \hat{n}_a)$ sean no singulares, para que la matriz $\Lambda_{zux}^\infty(\hat{n}_b, \hat{n}_a)$ tenga rango completo.*

Demostración. Utilícese el lema 2.4 considerando

$$\begin{aligned} \Lambda_{xux}^\infty(\hat{n}_b, \hat{n}_a) &= L_{ux}^\infty(\hat{n}_b, \hat{n}_a, 0) \\ \Lambda_{zux}^\infty(\hat{n}_b, \hat{n}_a) &= L_{zux}^\infty(\hat{n}_b, \hat{n}_a, 0) \end{aligned}$$

QED.

Ahora se definen dos conceptos, que son de utilidad para los próximos teoremas, ya que ayudan a entender el comportamiento “genérico” del teorema 2.2 para el caso $M \rightarrow \infty$, y también dan una medida del conjunto de casos en los cuales este teorema no rige.

Definición 2.1 Analiticidad. *Se dice que una función $f : D \subseteq \mathbb{C}^n \rightarrow \mathbb{C}$ es “analítica en una vecindad de $x_o \in D$ ”, si tiene representación en serie de Taylor en dicha vecindad.*

Definición 2.2 Medida de Lebesgue. *Sea $D \subset \mathbb{R}^\ell$, entonces su “medida de Lebesgue” es*

$$m(D) = \int_D 1 \, dx_1 \cdots dx_\ell$$

Ver [19, appendix 2].

Así, si $\ell = 2$ entonces $m(D)$ puede interpretarse como el área de D , y si $\ell = 3$ entonces $m(D)$ puede interpretarse como el volumen de D . Ahora, note que si D es un subconjunto con menor dimensión que \mathbb{R}^ℓ , como una recta en el caso $\ell = 2$ o un plano en el caso $\ell = 3$, entonces la medida de Lebesgue resulta $m(D) = 0$.

Lema 2.5 *Considérese una matriz $L(\rho)$, cuyo tamaño es $p \times q$ ($p \geq q$), la cual depende del vector $\rho \in D$, donde $D \subseteq \mathbb{R}^\ell$ un es conjunto abierto y conexo. Si se cumplen las siguientes condiciones*

1. cada elemento de $L(\rho)$ es una función analítica en los puntos $\rho_i \in D$, $\forall i = 1, \dots, \ell$
2. existe un $\rho_o \in D$ tal que $\text{rango } L(\rho_o) = q$

Entonces, el conjunto de casos en los cuales $L(\rho)$ es deficiente de rango

$$R_d = \{\rho \in D \mid \text{rango } L(\rho) < q\}$$

tiene medida de Lebesgue cero, es decir $m(R_d) = 0$.

Una demostración para este lema se puede encontrar en [19, corollary of lemma A2.3].

Este último lema es bastante potente, ya que una vez probada la analiticidad de los elementos de $L(\rho)$, entonces sólo es necesario encontrar un caso $L(\rho_o)$ que sea de rango completo, para que todos los demás casos $\forall \rho \in D$, $L(\rho)$ sean de rango completo. Así, éste es el teorema fundamental para demostrar que el rango de la matriz $\Lambda_{zux}^\infty(\hat{n}_b, \hat{n}_a)$ está dado “genéricamente” por el siguiente teorema, salvo un conjunto pequeño de casos.

Teorema 2.3 *Considere los datos Z_{ux}^∞ y Z_{uz}^∞ medidos de (2.19) y (2.25) respectivamente. Si se cumplen las siguientes condiciones*

1. la señal $s(k) = A(q)^{-1}u(k)$ es p.e. de orden p
2. el filtro $F(q)$ satisface $\max(m_l - n_b, m_k - n_a) \geq 0$
3. los elementos de $\Lambda_{zux}^\infty(\hat{n}_b, \hat{n}_a)(\rho_{\hat{n}_b, \hat{n}_a})$, son funciones analíticas $\forall \rho_{\hat{n}_b, \hat{n}_a} \in D_{\hat{n}_b, \hat{n}_a} \subset \mathbb{R}^\ell$
4. existe $\rho_{\hat{n}_b, \hat{n}_a}^o \in D_{\hat{n}_b, \hat{n}_a}$ tal que

$$\text{rango } \Lambda_{zux}^\infty(\hat{n}_b, \hat{n}_a)(\rho_{\hat{n}_b, \hat{n}_a}^o) = \begin{cases} \hat{n}_b + 1 + \hat{n}_a, & i \leq 0 \text{ y } j \leq 0 \\ \max(\hat{n}_b + n_a, n_b + \hat{n}_a) + 1, & i > 0 \text{ y } j \leq 0 \end{cases}$$

entonces

$$\text{rango } \Lambda_{zux}^\infty(\hat{n}_b, \hat{n}_a) = \begin{cases} \hat{n}_b + 1 + \hat{n}_a, & i \leq 0 \text{ y } j \leq 0 \\ \max(\hat{n}_b + n_a, n_b + \hat{n}_a) + 1, & i > 0 \text{ y } j \leq 0 \end{cases}$$

donde

$$\begin{aligned} p &= \max(\hat{n}_b, \hat{n}_a) + \max(m_l, m_k, n_b, n_a) + 1 \\ i &= \min(\hat{n}_b - n_b, \hat{n}_a - n_a) \\ j &= \min(\hat{n}_b - m_l, \hat{n}_a - m_k) \\ L &\geq \max(\hat{n}_b, \hat{n}_a) \end{aligned}$$

Este rango es válido “genéricamente”, salvo en un conjunto de casos $R_d^{\hat{n}_b, \hat{n}_a} \subset D_{\hat{n}_b, \hat{n}_a}$, cuya medida de Lebesgue es cero.

Demostración. (Caso $i \leq 0$ y $j \leq 0$). Dado que el orden de persistencia de $s(k)$ satisface

$$p \geq \max(\hat{n}_b + m_k, m_l + \hat{n}_a) + 1$$

entonces $\Lambda_{zuz}^\infty(\hat{n}_b, \hat{n}_a)$ es no singular para $j \leq 0$, por el lema 2.3. Además, como el orden de persistencia de $s(k)$ también satisface

$$p \geq \max(\hat{n}_b + n_a, n_b + \hat{n}_a) + 1$$

entonces $\Lambda_{xux}^\infty(\hat{n}_b, \hat{n}_a)$ es no singular para $i \leq 0$. Según el corolario 2.1, ambos resultados son necesarios (pero no suficientes) para que

$$\text{rango } \Lambda_{zux}^\infty(\hat{n}_b, \hat{n}_a) = \hat{n}_b + 1 + \hat{n}_a, \quad i \leq 0 \text{ y } j \leq 0$$

(Caso $i > 0$ y $j \leq 0$). Nótese que en este caso $\Lambda_{zuz}^\infty(\hat{n}_b, \hat{n}_a)$ aún es no singular. Ahora, dado que el orden de persistencia de $s(k)$ satisface

$$p \geq \max(\hat{n}_b + n_a, n_b + \hat{n}_a) + 1$$

entonces $L_{ux}^\infty(\hat{n}_b, \hat{n}_a, i)$ es no singular, por el teorema 1.5. Según el lema 2.4, estos resultados son necesarios (pero no suficientes) para que

$$\text{rango } L_{zux}^\infty(\hat{n}_b, \hat{n}_a, i) = \hat{n}_b + 1 + \hat{n}_a - i, \quad i > 0 \text{ y } j \leq 0$$

donde

$$\hat{n}_b + 1 + \hat{n}_a - i = \max(\hat{n}_b + n_a, n_b + \hat{n}_a) + 1$$

Note además que para $i > 0$ y $j \leq 0$ se tiene

$$\text{rango } \Lambda_{zux}^\infty(\hat{n}_b, \hat{n}_a) = \text{rango } L_{zux}^\infty(\hat{n}_b, \hat{n}_a, i)$$

ya que $L_{zux}^\infty(\hat{n}_b, \hat{n}_a, i)$ equivale a eliminar i columnas de la matriz $\Lambda_{zux}^\infty(\hat{n}_b, \hat{n}_a)$, las cuales son l.d. producto de la ecuación de la planta.

Una vez examinadas las condiciones necesarias para los casos $(i \leq 0, j \leq 0)$ y $(i > 0, j \leq 0)$ solamente resta: parametrizar con respecto a algún vector $\rho_{\hat{n}_b, \hat{n}_a} \in D_{\hat{n}_b, \hat{n}_a}$, luego comprobar que cada elemento de la matriz $\Lambda_{zux}^\infty(\hat{n}_b, \hat{n}_a)(\rho_{\hat{n}_b, \hat{n}_a})$ es analítico para todo $\rho_{\hat{n}_b, \hat{n}_a} \in D$ y finalmente encontrar $\rho_{\hat{n}_b, \hat{n}_a}^o \in D_{\hat{n}_b, \hat{n}_a}$ tal que

$$\text{rango } \Lambda_{zux}^\infty(\hat{n}_b, \hat{n}_a)(\rho_{\hat{n}_b, \hat{n}_a}^o) = \hat{n}_b + 1 + \hat{n}_a, \quad \text{si } i \leq 0 \text{ y } j \leq 0$$

$$\text{rango } L_{zux}^\infty(\hat{n}_b, \hat{n}_a, i)(\rho_{\hat{n}_b, \hat{n}_a}^o) = \max(\hat{n}_b + n_a, n_b + \hat{n}_a) + 1, \quad \text{si } i > 0 \text{ y } j \leq 0$$

$$\rho_{\hat{n}_b, \hat{n}_a} = [l_0 \quad \cdots \quad l_{m_l} \quad k_1 \quad \cdots \quad k_{m_k} \mid b_0 \quad \cdots \quad b_{n_b} \quad a_1 \quad \cdots \quad a_{n_a} \mid \beta_0 \quad \cdots \quad \beta_{n_\beta} \quad \alpha_0 \quad \cdots \quad \alpha_{n_\alpha}] \quad (2.32)$$

es decir

$$\text{rango } \Lambda_{zux}^\infty(\hat{n}_b, \hat{n}_a)(\rho_{\hat{n}_b, \hat{n}_a}^\circ) = \begin{cases} \hat{n}_b + 1 + \hat{n}_a, & i \leq 0 \text{ y } j \leq 0 \\ \text{máx}(\hat{n}_b + n_a, n_b + \hat{n}_a) + 1, & i > 0 \text{ y } j \leq 0 \end{cases}$$

para así utilizar el lema 2.5. **QED.**

Este teorema entrega, para el caso $M \rightarrow \infty$, básicamente el mismo resultado que el teorema 2.2. Sin embargo, se exigen dos condiciones adicionales, especificadas en los puntos 3 y 4. Estas condiciones pueden, a simple vista, parecer un poco rebuscadas, pero en realidad son condiciones relativamente fáciles de satisfacer. En seguida se discutirá una forma posible de parametrizar la matriz $\Lambda_{zux}^\infty(\hat{n}_b, \hat{n}_a)$ y verificar que sus elementos son analíticos, satisfaciendo así el punto 3. Posteriormente se encontrará un vector $\rho_{\hat{n}_b, \hat{n}_a}^\circ$, satisfaciéndose así el punto 4.

Considérese la planta $G(q)$ de la ecuación (2.19) sometida a una excitación $u(k) = N(q)\bar{u}(k)$, donde

$$N(q) = \frac{\beta_0 + \beta_1 q^{-1} + \cdots + \beta_{n_\beta} q^{-n_\beta}}{1 + \alpha_1 q^{-1} + \cdots + \alpha_{n_\alpha}}$$

y $\bar{u}(k)$ es tal que su autocorrelación muestral satisficiera $\varphi_{\bar{u}\bar{u}}^\infty(\tau) = \delta_K(k)$, por lo que $u(k)$ es p.e. de cualquier orden. Considérese también el filtro instrumental $F(q)$ de la ecuación (2.25).

De esta manera, los elementos de $\Lambda_{zux}^\infty(\hat{n}_b, \hat{n}_a)$ van a depender de los coeficientes de $F(q)$, $G(q)$ y $N(q)$, por lo que se puede tomar la parametrización $\rho_{\hat{n}_b, \hat{n}_a}$ de la ecuación (2.32), con

$$\rho_{\hat{n}_b, \hat{n}_a} \in D_{\hat{n}_b, \hat{n}_a} \subset \mathbb{R}^{(m_l+1+m_k)+(n_b+1+n_a)+(n_\beta+1+n_\alpha)}$$

donde $D_{\hat{n}_b, \hat{n}_a}$ debe cumplir las siguientes restricciones

- $F(q)$ debe tener numerador y denominador coprimos, debe ser estable y cumplir $\text{máx}(m_l - n_b, m_k - n_a) \geq 0$
- $G(q)$ debe tener numerador y denominador coprimos y debe ser estable
- $N(q)$ debe ser estable

Así, con estas condiciones, los elementos de la matriz $\Lambda_{zux}^\infty(\hat{n}_b, \hat{n}_a)(\rho_{\hat{n}_b, \hat{n}_a}^\circ)$ son funciones analíticas del vector $\rho_{\hat{n}_b, \hat{n}_a}^\circ \in D_{\hat{n}_b, \hat{n}_a}$, satisfaciendo así el punto 3 del teorema 2.3.

Note que la parametrización dada por $\rho_{\hat{n}_b, \hat{n}_a}$ depende en general del tamaño de la matriz en cuestión, dado por el par (\hat{n}_b, \hat{n}_a) . Por ejemplo, para el caso

$$F(q) = \sum_{i=0}^{\hat{n}_b} l_i q^{-i}, \quad G(q) = \frac{1}{1 - a_1 q^{-1}}, \quad N(q) = \beta_0$$

con $\hat{n}_b \geq 1$, cada matriz $\Lambda_{zux}^\infty(\hat{n}_b, \hat{n}_a)$ se puede parametrizar como

$$\rho_{\hat{n}_b, \hat{n}_a} = [l_0 \quad \cdots \quad l_{\hat{n}_b} \quad a_1 \quad \beta_0] \in \mathbb{R}^{\hat{n}_b+3}$$

Sin embargo, en ciertos casos, es posible encontrar la misma parametrización para algún conjunto de matrices. Por ejemplo, para el caso

$$F(q) = l_{\hat{n}_b} q^{-\hat{n}_b}, \quad G(q) = \frac{1}{1 - a_1 q^{-1}}, \quad N(q) = \beta_0$$

con $\hat{n}_b \geq 1$, cada matriz $\Lambda_{zux}^\infty(\hat{n}_b, \hat{n}_a)$ se puede parametrizar como

$$\rho_{\hat{n}_b, \hat{n}_a} = [l_{\hat{n}_b} \quad a_1 \quad \beta_0] \in \mathbb{R}^3$$

Ahora es necesario encontrar $\rho_{\hat{n}_b, \hat{n}_a}^\circ$, es decir, un caso particular de $F(q)$, $G(q)$ y $N(q)$, que satisfaga el punto 4 del teorema 2.3. Por ejemplo, para el caso en que

$$\begin{aligned} i &= \text{mín}(\hat{n}_b - n_b, \hat{n}_a - n_a) \leq 0 \\ j &= \text{mín}(\hat{n}_b - m_l, \hat{n}_a - m_k) \leq 0 \end{aligned}$$

se puede utilizar la siguiente combinación para $\rho_{\hat{n}_b, \hat{n}_a}^\circ$

$$F(q) = G(q)$$

Así, mediante el lema 2.3 se tiene que

$$\text{rango } \Lambda_{zux}^\infty(\hat{n}_b, \hat{n}_a)(\rho_{\hat{n}_b, \hat{n}_a}^\circ) = \hat{n}_b + 1 + \hat{n}_a, \quad i \leq 0 \text{ y } j \leq 0$$

Ahora, para el caso en que $i > 0$ y $j \leq 0$, se puede escoger la siguiente combinación para $\rho_{\hat{n}_b, \hat{n}_a}^\circ$

$$F(q) = q^{-\hat{n}_b}, \quad N(q) = 1$$

con la que se tiene

$$\text{rango } \Lambda_{zux}^\infty(\hat{n}_b, \hat{n}_a)(\rho_{\hat{n}_b, \hat{n}_a}^\circ) = \text{máx}(\hat{n}_b + n_a, n_b + \hat{n}_a) + 1, \quad i > 0 \text{ y } j \leq 0$$

debido al siguiente teorema

Teorema 2.4 Si la señal de entrada $u(k)$ satisface $\varphi_{uu}^\infty(\tau) = \delta_K(\tau)$ y el filtro instrumental se diseña $F(q) = q^{-\hat{n}_b}$, entonces

$$\text{rango } \Lambda_{zux}^\infty(\hat{n}_b, \hat{n}_a) = \max(\hat{n}_b + n_a, n_b + \hat{n}_a) + 1$$

para $\min(\hat{n}_b - n_b, \hat{n}_a - n_a) \geq 0$.

Véase la demostración complementaria E.6.

Como se acaba de ver, no es una tarea difícil encontrar una parametrización $\rho_{\hat{n}_b, \hat{n}_a}$ para la matriz $\Lambda_{zux}^\infty(\hat{n}_b, \hat{n}_a)$, de manera que sus elementos sean analíticos. Además, como ya se pudo comprobar, es posible encontrar $\rho_{\hat{n}_b, \hat{n}_a}^o$ que satisfaga el punto 4 del teorema 2.3. Por este motivo, y dado que son condiciones bastante generales, en la práctica se supondrán ciertas.

En resumen, gracias al teorema 2.3, se puede conocer el rango de $\Lambda_{zux}^\infty(\hat{n}_b, \hat{n}_a)$, que es “genéricamente” cierto bajo condiciones poco restrictivas. Sin embargo, pueden aparecer casos particulares en los cuales el rango no es cierto, los que son producidos por combinaciones desafortunadas de $F(q)$, $G(q)$ y $u(k)$. Pero desde un punto de vista práctico no hay necesidad de preocuparse por estos casos, ya que aparecen muy esporádicamente. Para más información sobre este tema véase [20, “generic consistency” en section 8.2], [19, theorem 4.1], [19, example 4.4] y [20, example 8.6].

Ahora, desde un punto de vista teórico, podría resultar interesante encontrar algún ejemplo, en el que se puedan visualizar los casos $R_d^{\hat{n}_b, \hat{n}_a} \subset D_{\hat{n}_b, \hat{n}_a}$, para los que el teorema 2.3 no es válido.

Ejemplo 2.7 Considérese las siguientes combinaciones de señal de entrada, planta y filtro instrumental, las que dependen de un parámetro $\rho \in D \subseteq \mathbb{R}$

$$\begin{aligned} u(k) &= (1 - 2\rho^2 q^{-2} + \rho^4 q^{-4}) \bar{u}(k) \\ G(q) &= \frac{1}{1 - 2\rho q^{-1} + \rho^2 q^{-2}} \\ F(q) &= \frac{1}{1 + 2\rho q^{-1} + \rho^2 q^{-2}} \end{aligned}$$

donde

$$D = \{\rho \in \mathbb{R} : |\rho| < 1\}$$

Estas satisfacen al teorema 2.3, por lo que el rango de la siguiente matriz

$$\Lambda_{zux}^\infty(0, 2) = \begin{bmatrix} \varphi_{uu}^\infty(0) & \varphi_{ux}^\infty(1) & \varphi_{ux}^\infty(2) \\ \varphi_{zu}^\infty(-1) & \varphi_{zx}^\infty(0) & \varphi_{zx}^\infty(1) \\ \varphi_{zu}^\infty(-2) & \varphi_{zx}^\infty(-1) & \varphi_{zx}^\infty(0) \end{bmatrix}$$

es conocido e igual a $\text{rango } \Lambda_{zux}^\infty(0, 2) = 3$, salvo en un conjunto de casos $R_d^{0,2}$ para los cuales no es cierto.

Para obtener dicho conjunto de casos $R_d^{0,2}$, se calculan los elementos de la matriz $\Lambda_{zux}^\infty(0, 2)$, teniendo en cuenta que las correlaciones se pueden obtener de

$$\varphi_{zx}^\infty(\tau) = \frac{1}{j2\pi} \oint_{\zeta=e^{j\theta}} \Phi_{zx}^\infty(\zeta) \zeta^{\tau-1} d\zeta$$

donde

$$\begin{aligned} \Phi_{uu}^\infty(\zeta) &= (1 - \rho\zeta^{-1})^2 (1 + \rho\zeta^{-1})^2 (1 - \rho\zeta)^2 (1 + \rho\zeta)^2 \\ \Phi_{ux}^\infty(\zeta) &= \Phi_{uu}^\infty(\zeta) G(\zeta^{-1}) = (1 - \rho\zeta^{-1})^2 (1 + \rho\zeta^{-1})^2 (1 + \rho\zeta)^2 \\ \Phi_{zu}^\infty(\zeta) &= F(\zeta) \Phi_{uu}^\infty(\zeta) = (1 - \rho\zeta^{-1})^2 (1 - \rho\zeta)^2 (1 + \rho\zeta)^2 \\ \Phi_{zx}^\infty(\zeta) &= F(\zeta) \Phi_{ux}^\infty(\zeta) G(\zeta^{-1}) = (1 - \rho\zeta^{-1})^2 (1 + \rho\zeta)^2 \end{aligned}$$

son las transformaciones de Fourier de todas las correlaciones, considerando $\zeta = e^{j\theta}$. De esta manera la matriz $\Lambda_{zux}^\infty(0, 2)$ resulta

$$\Lambda_{zux}^\infty(0, 2)(\rho) = \begin{bmatrix} \rho^8 + 4\rho^4 + 1 & -4\rho^3 & \rho^6 - 2\rho^2 \\ 4\rho^3 & \rho^4 - 4\rho^2 + 1 & 2\rho^3 - 2\rho \\ \rho^6 - 2\rho^2 & 2\rho - 2\rho^3 & \rho^4 - 4\rho^2 + 1 \end{bmatrix}$$

Entonces, para comprobar en qué casos esta matriz es de rango completo, se calcula su determinante

$$\begin{aligned} \det \Lambda_{zux}^\infty(0, 2)(\rho) &= (\rho^2 - \rho - 1)(\rho^2 + \rho - 1)(\rho^8 - \rho^6 + 6\rho^4 - \rho^2 + 1) \end{aligned}$$

obteniéndose que $\det \Lambda_{zux}^\infty(0, 2)(\rho) = 0$ para

$$\rho = \frac{-1 \pm \sqrt{5}}{2}, \quad \rho = \frac{1 \pm \sqrt{5}}{2}$$

por lo que $\text{rango } \Lambda_{zux}^\infty(0, 2) \neq 3$ para $\rho \in R_d^{0,2} \subset D$

$$R_d^{0,2} = \left\{ \rho = \frac{-1 + \sqrt{5}}{2}, \quad \rho = \frac{1 - \sqrt{5}}{2} \right\}$$

el cual efectivamente cumple

$$m(R_d^{0,2}) = \int_{R_d^{0,2}} 1 d\rho = 0$$

es decir, tiene medida de Lebesgue cero.

2.5. Comentarios del capítulo

se satisface un conjunto de condiciones poco restrictivas. Para más detalles consúltese la sección 2.4.

En un ambiente libre de ruido, y cumpliendo las condiciones de los lemas 2.1 y 2.3, el rango de las matrices $\Lambda_{xux}^M(\hat{n})$ y $\Lambda_{xux}^M(\hat{n}_b, \hat{n}_a)$ se puede conocer de manera exacta, cuando se utiliza una señal de entrada cuyo orden de excitación es suficientemente grande, incluso si se utiliza un número finito de mediciones. El conocimiento de este rango permite crear pruebas para identificar el orden de un sistema, las cuales se basan en determinar si Λ_{xux}^M es singular. Al calcular los valores propios de estas matrices, para determinar si son singulares, se puede notar que los valores propios nulos son del orden de magnitud del ϵ del computador, lo que permite discriminar fácilmente entre matrices singulares y no singulares.

El rango de las matrices instrumentales $\Lambda_{zux}^M(\hat{n})$ y $\Lambda_{zux}^M(\hat{n}_b, \hat{n}_a)$, en un ambiente libre de ruido y bajo las condiciones de los teoremas 2.1 y 2.2, también es conocido de manera exacta, aun con un número finito de mediciones. Sin embargo, cuando existe ruido de salida, se necesita un número infinito de mediciones para removerlo totalmente. En la práctica, el número de mediciones puede ser muy grande, pero es siempre finito, lo que provoca que el ruido no se pueda eliminar completamente. Este último hecho hace que los valores propios nulos de Λ_{zuy}^M se vean perturbados, lo que dificulta discriminar cuándo la matriz es realmente singular, ya que no se dispone de criterios objetivos para saber cuándo un valor propio perturbado podría ser cero.

Las pruebas para identificar la estructura del sistema, mostradas en este capítulo, son búsquedas secuenciales exhaustivas, es decir, si se sospecha que la estructura del sistema $(n_b, n_a) \in [1, \hat{n}_{\text{máx}}] \times [1, \hat{n}_{\text{máx}}]$, entonces se busca la estructura correcta en cada uno de los puntos de esta región. Este tipo de búsqueda, existente ya en la literatura, requiere el análisis de muchos casos cuando $\hat{n}_{\text{máx}}$ es alto, es decir, cuando el conjunto de estructuras candidatas es grande.

Para el caso de las matrices instrumentales, cuando el número de mediciones tiende a infinito, pueden ocurrir situaciones extrañas, en las cuales los resultados de los teoremas 2.1 y 2.2 no se cumplen, incluso si se satisfacen todas sus condiciones. Estos casos son bastante raros, y son producidos por combinaciones desafortunadas de $F(q)$, $G(q)$ y $u(k)$, pero desde un punto de vista práctico no hay necesidad de preocuparse, ya que éstos aparecen muy esporádicamente, si

Capítulo 3

Perfiles de valores singulares para matrices de correlación

EN este capítulo se intenta solucionar los dos problemas principales que presentan los métodos de búsqueda secuencial, vistos en el capítulo anterior

1. Reducir el costo computacional de la búsqueda
2. Encontrar un criterio objetivo, para poder determinar cuándo las matrices son singulares

En cuanto a la notación general para este capítulo, se utiliza $\lambda(A)$ para referirse a los *valores propios* de la matriz A , y $\sigma(A)$ para sus *valores singulares*.

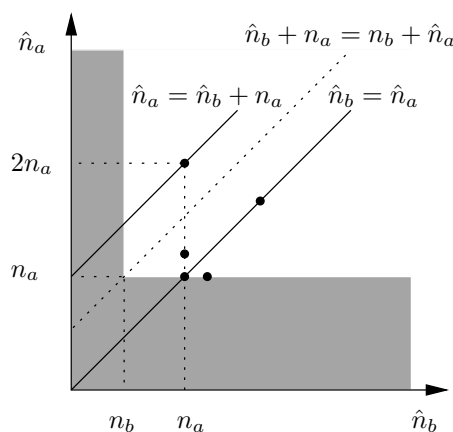


Figura 3.1: Búsqueda de (n_b, n_a) en el caso $n_b < n_a$. En gris: región donde Λ_{xux}^M es de rango completo. En blanco: región donde Λ_{xux}^M es deficiente de rango.

3.1. Caso sin ruido

3.1.1. Construcción del método

Considérese nuevamente el sistema libre de ruido de la ecuación (2.19), cuya estructura (n_b, n_a) se quiere identificar.

Los métodos de búsqueda secuencial, del capítulo anterior, están basados solamente en la determinación de la singularidad de la matriz $\Lambda_{xux}^M(\hat{n}_b, \hat{n}_a)$, sin embargo, no hacen uso de la información adicional que entrega el lema 2.3, es decir, del rango de la matriz $\Lambda_{xux}^M(\hat{n}_b, \hat{n}_a)$. Esta información adicional puede ayudar a reducir la cantidad de matrices analizadas, para determinar la estructura de la planta.

Por ejemplo, considérese que se escoge el modelo (\hat{n}, \hat{n}) , con $\hat{n} > \max(n_b, n_a)$, y que para él se construye la matriz Λ_{xux}^M . En este caso, y satisfaciendo las

condiciones del lema 2.3, se sabe que

$$\text{rango } \Lambda_{xux}^M(\hat{n}, \hat{n}) = \hat{n} + \max(n_b, n_a) + 1$$

es decir, es posible obtener inmediatamente el orden n de la planta, una vez determinado experimentalmente el rango de $\Lambda_{xux}^M(\hat{n}, \hat{n})$

$$n := \max(n_b, n_a) = \text{rango } \Lambda_{xux}^M(\hat{n}, \hat{n}) - (\hat{n} + 1)$$

Una vez que el orden del sistema $n = \max(n_b, n_a)$ es conocido, hay que determinar si este orden corresponde a n_a o n_b , presentándose tres casos posibles

1. En la figura 3.1 se muestra la situación $n_b < n_a$, entonces, si ella es cierta, los rangos de las siguientes matrices satisfacen

$$\text{rango } \Lambda_{xux}^M(n+1, n) = \text{rango completo}$$

rango $\Lambda_{xux}^M(n, n+1)$ = deficiente de rango

Así, esta combinación de rangos permite determinar que $n_a = n$. Una vez que n_a es conocido, sólo falta obtener n_b , el cual puede ser determinado utilizando nuevamente el lema 2.3, cuando $\hat{n}_b + n_a \leq n_b + \hat{n}_a$

$$\text{rango } \Lambda_{xux}^M(\hat{n}_b, \hat{n}_a) = n_b + \hat{n}_a + 1$$

Así, el punto $(\hat{n}_b, \hat{n}_a) = (n_a, 2n_a)$ satisface esta última desigualdad, por lo que

$$n_b = \text{rango } \Lambda_{xux}^M(n_a, 2n_a) - (2n_a + 1)$$

permitiendo obtener finalmente n_b .

- En la figura 3.2 se muestra la situación $n_b > n_a$, entonces, si ella es cierta, los rangos de las siguientes matrices satisfacen

$$\begin{aligned} \text{rango } \Lambda_{xux}^M(n+1, n) &= \text{deficiente de rango} \\ \text{rango } \Lambda_{xux}^M(n, n+1) &= \text{rango completo} \end{aligned}$$

así, esta combinación de rangos permite determinar que $n_b = n$. Una vez que n_b es conocido, sólo falta obtener n_a , el cual puede ser determinado utilizando nuevamente el lema 2.3, cuando $\hat{n}_b + n_a \geq n_b + \hat{n}_a$

$$\text{rango } \Lambda_{xux}^M(\hat{n}_b, \hat{n}_a) = \hat{n}_b + n_a + 1$$

así, el punto $(\hat{n}_b, \hat{n}_a) = (2n_b, n_b)$ satisface esta última desigualdad, por lo que

$$n_a = \text{rango } \Lambda_{xux}^M(2n_b, n_b) - (2n_b + 1)$$

permitiendo obtener finalmente n_a

- Si la situación $n_b = n_a$ es cierta, los rangos de las siguientes matrices cumplen

$$\begin{aligned} \text{rango } \Lambda_{xux}^M(n+1, n) &= \text{rango completo} \\ \text{rango } \Lambda_{xux}^M(n, n+1) &= \text{rango completo} \end{aligned}$$

los cuales sirven para determinar que n_b y n_a son iguales al orden del sistema

$$n_b = n_a = n$$

Nótese que las matrices $\Lambda_{xux}^M(n+1, n)$ y $\Lambda_{xux}^M(n, n+1)$ no pueden ser simultáneamente deficientes de rango, ya que esto significaría que el orden del sistema no fue correctamente determinado, o que alguno de los rangos anteriores no fue bien calculado.

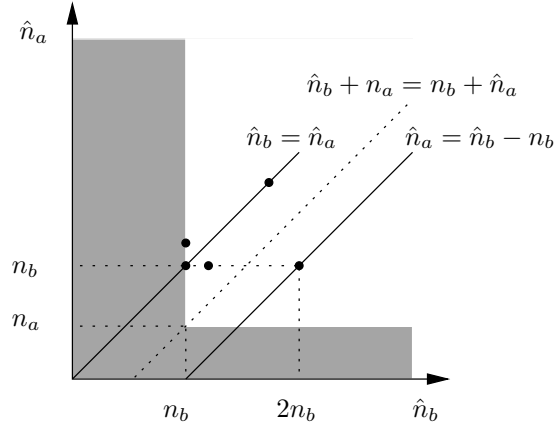


Figura 3.2: Búsqueda de (n_b, n_a) en el caso $n_b > n_a$. En gris: región donde Λ_{xux}^M es de rango completo. En blanco: región donde Λ_{xux}^M es deficiente de rango.

Nótese además que, para la búsqueda de la estructura correcta del sistema, se han utilizado modelos que pertenecen a la región en que $\Lambda_{xux}^M(\hat{n}_b, \hat{n}_a)$ es deficiente de rango, es decir

$$\begin{aligned} &\Lambda_{xux}^M(\hat{n}, \hat{n}), \quad \hat{n} > n \\ &\Lambda_{xux}^M(n+1, n) \\ &\Lambda_{xux}^M(n, n+1) \\ &\Lambda_{xux}^M(n_a, 2n_a) \text{ o } \Lambda_{xux}^M(2n_b, n_b) \end{aligned}$$

Ahora, dado que los modelos

$$\begin{aligned} (\hat{n}_b, \hat{n}_a) &= (n_a, 2n_a) = (n, 2n) \\ (\hat{n}_b, \hat{n}_a) &= (2n_b, n_b) = (2n, n) \end{aligned}$$

son los modelos más grandes a utilizar en el procedimiento de identificación desarrollado anteriormente, entonces es suficiente que la región, en la cual se va a realizar búsqueda, sea

$$(\hat{n}_b, \hat{n}_a) \in [n_b, \hat{n}_{\text{máx}}] \times [n_a, \hat{n}_{\text{máx}}]$$

donde $\hat{n}_{\text{máx}}$ debe ser escogido tal que

$$\hat{n}_{\text{máx}} \geq 2n = 2 \text{ máx}(n_b, n_a)$$

Además, para satisfacer las condiciones del lema 2.3, es suficiente cumplir con

$$\begin{aligned} \hat{n}_{\text{máx}} &\geq 2n \\ p \geq 2\hat{n}_{\text{máx}} + 1 &\geq \text{máx}(\hat{n}_b + n_a, n_b + \hat{n}_a) + 1 \\ 2M + 1 \geq p &\geq \text{máx}(\hat{n}_b + n_a, n_b + \hat{n}_a) + 1 \\ L \geq \hat{n}_{\text{máx}} &\geq \text{máx}(\hat{n}_b, \hat{n}_a) \end{aligned}$$

donde p es el orden de excitación a utilizar. Estas condiciones y el procedimiento de identificación, desarrollado anteriormente, se resumen en el algoritmo 3.1.

```

// región de prueba
1  escoja  $\hat{n}_{\text{máx}} \geq 2n = 2 \text{máx}(n_b, n_a)$ ;
// diseño de señal de entrada
2  escoja  $p \geq 2\hat{n}_{\text{máx}} + 1$ ;
3  escoja  $M$  tal que  $2M + 1 \geq p$ ;
4  escoja  $L \geq \hat{n}_{\text{máx}}$ ;
5  genere  $u(k)$  l.e./p.e. de orden  $p$ ;
// medición
6  mida  $Z_{ux}^{L,M}$  en estado estacionario;
// identificación
7  escoger  $\hat{n}$  tal que  $n < \hat{n} \leq \hat{n}_{\text{máx}}$ ;
8   $n = \text{rango } \Lambda_{xux}^M(\hat{n}, \hat{n}) - (\hat{n} + 1)$ ;
9   $\Lambda_b = \Lambda_{xux}^M(n + 1, n)$ ;
10  $\Lambda_a = \Lambda_{xux}^M(n, n + 1)$ ;
11 if ( $\Lambda_b$  no singular y  $\Lambda_a$  singular) then
12    $n_a = n$ ;
13    $n_b = \text{rango } \Lambda_{xux}^M(n_a, 2n_a) - (2n_a + 1)$ ;
14 else if ( $\Lambda_b$  singular y  $\Lambda_a$  no singular) then
15    $n_b = n$ ;
16    $n_a = \text{rango } \Lambda_{xux}^M(2n_b, n_b) - (2n_b + 1)$ ;
17 else if ( $\Lambda_b$  no singular y  $\Lambda_a$  no singular) then
18    $n_b = n$ ;
19    $n_a = n$ ;
20 else
21   Advertencia: Esta situación no puede
    ocurrir. Revisar los rangos de las matrices;
22 end

```

Algoritmo 3.1: Búsqueda del par (n_b, n_a) con el método de perfiles de valores singulares. Caso sin ruido.

3.1.2. Determinación del rango

Para poder determinar experimentalmente el rango de $\Lambda_{xux}^M(\hat{n}_b, \hat{n}_a)$, se puede utilizar su descomposición en valores singulares, debido a que el rango de una matriz A cualquiera corresponde al número de valores singulares no nulos, es decir

$$\text{rango } A = \# \{ \sigma(A) \mid \sigma(A) > 0 \}$$

Sin embargo, al calcular computacionalmente los valores singulares de A , se incurre en errores numéricos, que perturban los valores singulares, afectando especialmente a los valores singulares nulos. Este último hecho dificulta la determinación del rango de la matriz A .

El problema de determinar el rango de A , cuando sus valores singulares nulos se ven perturbados,

se llama *rango efectivo* de A , y puede ser abordado mediante la siguiente estrategia

$$\text{rango } A = \# \{ \sigma(A) \mid \sigma(A) > \epsilon_\sigma \} \quad (3.1)$$

donde ϵ_σ da una medida del nivel de perturbación que presentarían los valores singulares de A . Así, cualquier valor singular que sea menor que la medida de perturbación ϵ_σ podría considerarse como un “valor singular nulo”, y no debería aportar al rango de A .

Dado que al calcular $\Lambda_{xux}^M(\hat{n}_b, \hat{n}_a)$ se incurre sólo en errores numéricos, los valores singulares de esta matriz se ven perturbados proporcionalmente a la precisión del computador ϵ_{ps} . Es por este motivo que puede utilizarse el siguiente criterio, para una matriz A de tamaño $p \times q$

$$\epsilon_\sigma = \text{máx}(p, q) \|A\|_2 \epsilon_{\text{ps}} \quad (3.2)$$

que es usado por omisión en la mayoría de los programas de cálculo numérico.

Nótese que si se considera la siguiente partición

$$\Lambda_{xux}^M = (\tilde{\Omega}_{ux}^M)^T \tilde{\Omega}_{ux}^M, \quad \tilde{\Omega}_{ux}^M = \frac{\Omega_{ux}^M}{\sqrt{2M+1}}$$

se puede evitar el cálculo de las correlaciones de la matriz $\Lambda_{xux}^M(\hat{n}_b, \hat{n}_a)$, ya que

$$\text{rango } \Lambda_{xux}^M = \text{rango } \tilde{\Omega}_{ux}^M$$

obteniéndose igualmente el rango de Λ_{xux}^M a partir del rango de $\tilde{\Omega}_{ux}^M$. Además, el cálculo de los valores singulares de $\tilde{\Omega}_{ux}^M$ da una medida de independencia lineal de cada una de sus columnas, indicándose así las columnas l.d. a través de valores singulares cero. Nótese también que esta medida de independencia de las columnas de $\tilde{\Omega}_{ux}^M$ puede ser obtenida a través de la matriz Λ_{xux}^M , ya que

$$\sigma(\Lambda_{xux}^M) = \lambda(\Lambda_{xux}^M) = \sigma^2(\tilde{\Omega}_{ux}^M) \quad (3.3)$$

A continuación se muestran algunos ejemplos, que hacen uso del algoritmo 3.1, calculando los rangos a través de representaciones gráficas de perfiles de valores singulares

Ejemplo 3.1 *Considérese la siguiente planta*

$$G(q) = \frac{0,3}{q^2(q-0,8e^{j\pi/4})(q-0,8e^{-j\pi/4})(q-0,7)}$$

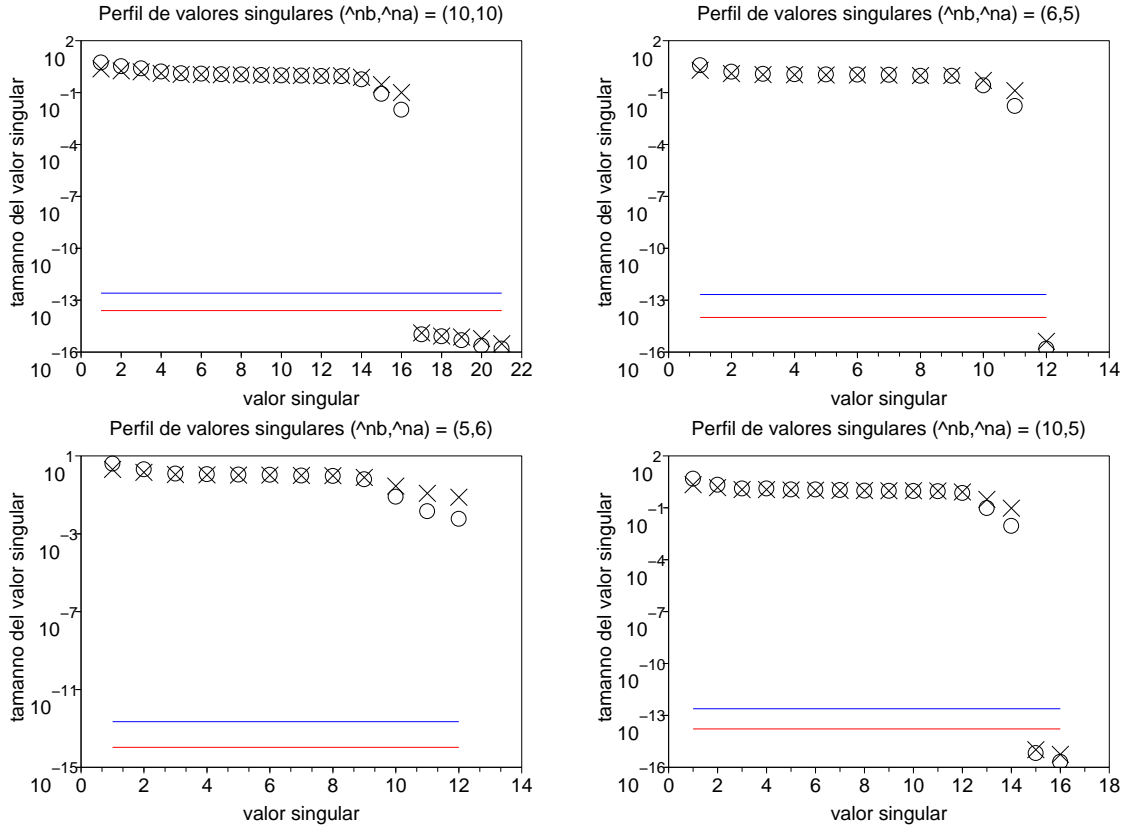


Figura 3.3: Con “o” los valores singulares de $\Lambda_{xux}^M(\hat{n}_b, \hat{n}_a)$ y en rojo su ϵ_σ . Con “x” los valores singulares de $\tilde{\Omega}_{ux}^M(\hat{n}_b, \hat{n}_a)$ y en azul su ϵ_σ .

cuya estructura $(n_b, n_a) = (5, 3)$ se desea identificar. del sistema

Los parámetros a utilizar en el algoritmo 3.1 se escogen según

$$\hat{n}_{\text{máx}} = 15, \quad M = 250, \quad L = \hat{n}_{\text{máx}}$$

y la señal $u(k)$ se diseña como una realización de ruido blanco gaussiano de media cero y varianza 1, la que es p.e. de todo orden.

De esta forma se miden los datos $Z_{ux}^{L,M}$ en estado estacionario, con los que se construyen los perfiles de valores singulares para las matrices $\Lambda_{xux}^M(\hat{n}_b, \hat{n}_a)$ y $\tilde{\Omega}_{ux}^M(\hat{n}_b, \hat{n}_a)$, mostrados en la figura 3.3. Para el cálculo de los rangos se usa la estrategia de la ecuación (3.1), considerando ϵ_σ dado por la ecuación (3.2).

Siguiendo el algoritmo 3.1, se escoge inicialmente el modelo $(\hat{n}_b, \hat{n}_a) = (10, 10)$, para el que se obtiene $\text{rango } \Lambda_{xux}^M(10, 10) = 16$, así se puede obtener el orden

$$n = \text{rango } \Lambda_{xux}^M(10, 10) - (10 + 1) = 5$$

Una vez que el orden del sistema es conocido $n = 5$, se computan los rangos de las siguientes matrices

$$\text{rango } \Lambda_{xux}^M(6, 5) = \text{deficiente (singular)}$$

$$\text{rango } \Lambda_{xux}^M(5, 6) = \text{completo (no singular)}$$

Esto indica, según el algoritmo 3.1, que el orden del sistema corresponde a n_b , es decir

$$n_b = n = 5$$

Finalmente se obtiene el valor de n_a , mediante la matriz asociada al modelo $(\hat{n}_b, \hat{n}_a) = (2n_b, n_b)$, teniendo en cuenta que $n_b = n = 5$. De esta forma, $\text{rango } \Lambda_{xux}^M(10, 5) = 14$, y por lo tanto

$$n_a = \text{rango } \Lambda_{xux}^M(10, 5) - (2 \cdot 5 + 1) = 3$$

con lo que la estructura identificada del sistema resulta $(n_b, n_a) = (5, 3)$.

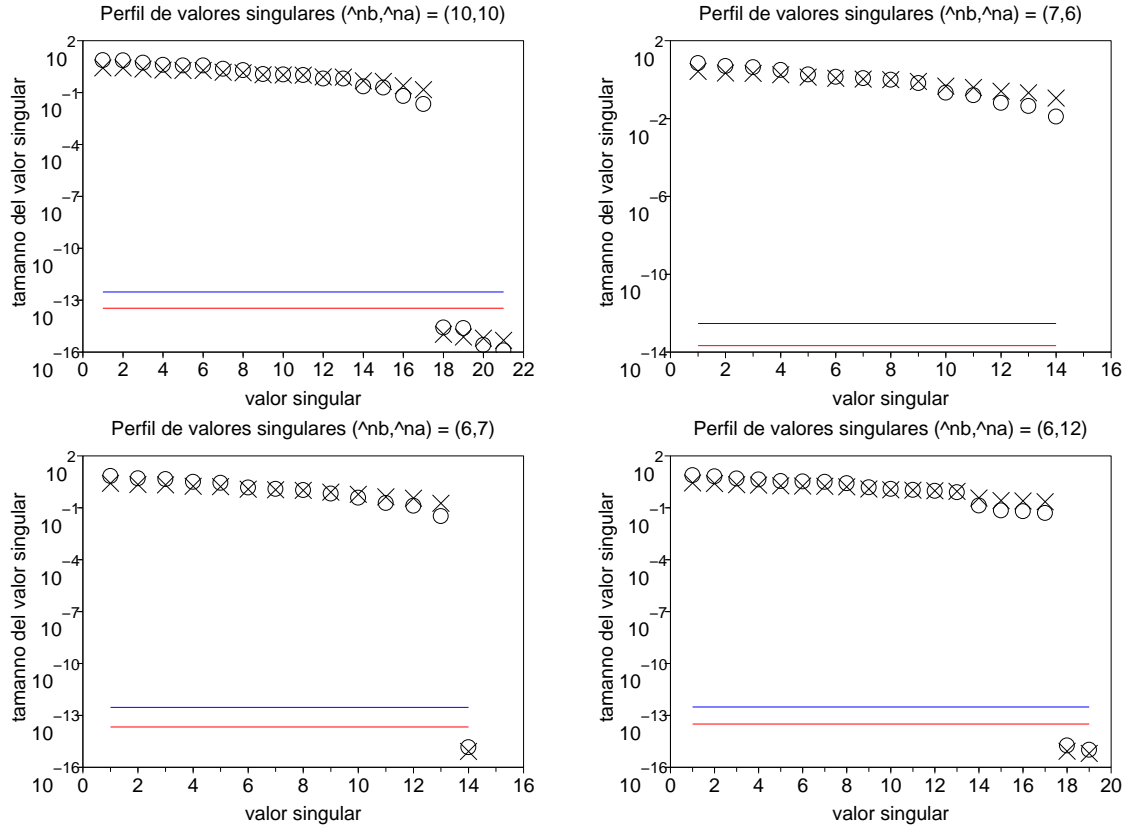


Figura 3.4: Con “o” los valores singulares de $\Lambda_{xux}^M(\hat{n}_b, \hat{n}_a)$ y en rojo su ϵ_σ . Con “x” los valores singulares de $\tilde{\Omega}_{ux}^M(\hat{n}_b, \hat{n}_a)$ y en azul su ϵ_σ .

Para más detalles de la simulación véase el programa *F.4.1* cual es p.e. de todo orden.

Ejemplo 3.2 Considérese ahora el mismo sistema que en [13, example 4]

$$G(q) = \frac{B(q)}{A(q)}$$

$$B(q) = 1 + 0,3452q^{-1} + 0,53q^{-2} + 0,3985q^{-3} + 0,8138q^{-4}$$

$$A(q) = 1 + 0,7907q^{-1} + 0,042q^{-2} - 0,5556q^{-3} - 0,0247q^{-4} + 0,3846q^{-5} + 0,3026q^{-6}$$

cuya estructura $(n_b, n_a) = (4, 6)$ se desea identificar.

Los parámetros a utilizar en el algoritmo 3.1 se escogen según

$$\hat{n}_{\text{máx}} = 15, \quad M = 250, \quad L = \hat{n}_{\text{máx}}$$

y la señal $u(k)$ se diseña como una realización de ruido blanco gaussiano de media cero y varianza 1, la

De esta forma se miden los datos $Z_{ux}^{L,M}$ en estado estacionario, con los cuales se construyen los perfiles de valores singulares para las matrices $\Lambda_{xux}^M(\hat{n}_b, \hat{n}_a)$ y $\tilde{\Omega}_{ux}^M(\hat{n}_b, \hat{n}_a)$, mostrados en la figura 3.3. Para el cálculo de los rangos se usa la estrategia de la ecuación (3.1), considerando ϵ_σ dado por la ecuación (3.2).

Siguiendo el algoritmo 3.1, se escoge inicialmente el modelo $(\hat{n}_b, \hat{n}_a) = (10, 10)$, para el cual se obtiene $\text{rango } \Lambda_{xux}^M(10, 10) = 17$, así se puede determinar el orden del sistema

$$n = \text{rango } \Lambda_{xux}^M(10, 10) - (10 + 1) = 6$$

Una vez que el orden del sistema es conocido $n = 6$, se computan los rangos de las siguientes matrices

$$\begin{aligned} \text{rango } \Lambda_{xux}^M(7, 6) &= \text{completo (no singular)} \\ \text{rango } \Lambda_{xux}^M(6, 7) &= \text{deficiente (singular)} \end{aligned}$$

Esto indica, según el algoritmo 3.1, que el orden del

sistema corresponde a n_a , es decir

$$n_a = n = 6$$

Finalmente se obtiene el valor de n_b , mediante la matriz asociada al modelo $(\hat{n}_b, \hat{n}_a) = (n_a, 2n_a)$, teniendo en cuenta que $n_a = n = 6$. De esta forma, $\text{rango } \Lambda_{xux}^M(6, 12) = 17$, y por lo tanto

$$n_b = \text{rango } \Lambda_{xux}^M(6, 12) - (2 \cdot 6 + 1) = 4$$

con lo que la estructura identificada del sistema resulta $(n_b, n_a) = (4, 6)$.

Para más detalles de la simulación vea el programa F.4.1.

Nótese que, cualquiera sea el caso, como máximo se requiere calcular el rango de 4 matrices, para identificar la estructura del sistema, tal como se mostró en los ejemplos anteriores. Esto sin duda trae ventajas frente a los métodos de búsqueda exhaustiva del capítulo anterior, ya que la cantidad de matrices analizadas se reduce considerablemente.

3.2. Caso con ruido

3.2.1. Construcción del método

Considérese ahora que el sistema (2.19) se ve perturbado con ruido de salida, tal como muestra la ecuación (2.24). Se sabe, del capítulo anterior, que para eliminar el efecto del ruido se puede utilizar la matriz instrumental Λ_{zuy}^M , ya que gracias al lema 2.2 se tiene

$$\Lambda_{zuy}^\infty = \Lambda_{zux}^\infty$$

donde $z(k)$ es la secuencia instrumental, definida por (2.25). Además, debido al teorema 2.2 se puede concluir que

$$\text{rango } \Lambda_{zux}^M(\hat{n}_b, \hat{n}_a) = \text{rango } \Lambda_{xux}^M(\hat{n}_b, \hat{n}_a)$$

Estos dos resultados indican que se puede recuperar el procedimiento del algoritmo 3.1, para el caso con ruido, pero ahora a través la matriz de Λ_{zuy}^M , es decir

1. determinar el orden del sistema $n = \text{máx}(n_b, n_a)$ a partir de

$$n = \text{rango } \Lambda_{zuy}^M(\hat{n}, \hat{n}) - (\hat{n} + 1)$$

donde $\hat{n} > n$.

2. determinar si el orden del sistema corresponde a n_b o n_a , mediante la verificación de la singularidad de las siguientes matrices

$$\text{rango } \Lambda_{zuy}^M(n+1, n) = \text{¿singular?}$$

$$\text{rango } \Lambda_{zuy}^M(n, n+1) = \text{¿singular?}$$

3. dependiendo del resultado del punto anterior, calcular el parámetro restante según

$$n_b = \text{rango } \Lambda_{zuy}^M(n_a, 2n_a) - (2n_a + 1)$$

$$n_a = \text{rango } \Lambda_{zuy}^M(2n_b, n_b) - (2n_b + 1)$$

En todos los puntos anteriores la estructura del filtro instrumental (m_l, m_k) debe cumplir

$$\text{máx}(m_l - n_b, m_k - n_a) \geq 0$$

para así satisfacer el teorema 2.2.

Dado que los modelos

$$(\hat{n}_b, \hat{n}_a) = (n_a, 2n_a) = (n, 2n)$$

$$(\hat{n}_b, \hat{n}_a) = (2n_b, n_b) = (2n, n)$$

son los modelos más grandes a utilizar, entonces es suficiente que la región, en la cual se va a realizar búsqueda, sea

$$(\hat{n}_b, \hat{n}_a) \in [n_b, \hat{n}_{\text{máx}}] \times [n_a, \hat{n}_{\text{máx}}]$$

donde $\hat{n}_{\text{máx}}$ debe ser escogido tal que

$$\hat{n}_{\text{máx}} \geq 2n = 2 \text{máx}(n_b, n_a)$$

Además, para satisfacer las condiciones del teorema 2.2, es suficiente cumplir con el conjunto de ecuaciones en (3.4), donde p es el orden de excitación a utilizar. También se pueden usar los criterios vistos en (2.29) y (2.30) para el diseño del filtro instrumental. Estas condiciones y el procedimiento de identificación se resumen en el algoritmo 3.2.

3.2.2. Determinación del rango

Como ya se vio anteriormente, el rango de una matriz A corresponde al número de valores singulares no nulos, sin embargo, discriminar si un valor singular es nulo, es una tarea complicada cuando la matriz en cuestión se ve perturbada. En una matriz perturbada $\tilde{A} = A + E$, los valores singulares “nulos” son en realidad pequeños, y pueden interferir en la determinación

$$\begin{aligned}
\hat{n}_{\text{máx}} &> 2n = 2 \text{máx}(n_b, n_a) \\
0 \leq \text{máx}(m_l - \hat{n}_b, m_k - \hat{n}_a) &\leq \text{máx}(m_l - n_b, m_k - n_a) \\
p \geq \hat{n}_{\text{máx}} + \text{máx}(m_l, m_k, \hat{n}_{\text{máx}}) + 1 &\geq \text{máx}(\hat{n}_b, \hat{n}_a) + \text{máx}(m_l, m_k, n_b, n_a) + 1 \\
2M + 1 \geq p &\geq \text{máx}(\hat{n}_b, \hat{n}_a) + \text{máx}(m_l, m_k, n_b, n_a) + 1 \\
L \geq \hat{n}_{\text{máx}} &\geq \text{máx}(\hat{n}_b, \hat{n}_a)
\end{aligned} \tag{3.4}$$

```

// región de prueba
1  escoja  $\hat{n}_{\text{máx}} \geq 2n = 2 \text{máx}(n_b, n_a)$ ;
   // filtro instrumental
2  escoja  $(m_l, m_k)$  tal que
    $\text{máx}(m_l - \hat{n}_b, m_k - \hat{n}_a) \geq 0$ ;
3  escoja el retardo  $\hat{n}_b \leq d_f \leq \hat{n}_a + \kappa_g/2$ ;
   // diseño de señal de entrada
4  escoja  $p \geq \hat{n}_{\text{máx}} + \text{máx}(m_l, m_k, \hat{n}_{\text{máx}}) + 1$ ;
5  escoja  $M$  tal que  $2M + 1 \geq p$ ;
6  escoja  $L \geq \hat{n}_{\text{máx}}$ ;
7  genere  $u(k)$  l.e./p.e. de orden  $p$ ;
   // medición
8  mida  $Z_{uy}^{L,M}$  en estado estacionario;
   // identificación
9  escoger  $\hat{n}$  tal que  $n < \hat{n} \leq \hat{n}_{\text{máx}}$ ;
10  $n = \text{rango } \Lambda_{zuy}^M(\hat{n}, \hat{n}) - (\hat{n} + 1)$ ;
11  $\Lambda_b = \Lambda_{zuy}^M(n + 1, n)$ ;
12  $\Lambda_a = \Lambda_{zuy}^M(n, n + 1)$ ;
13 if ( $\Lambda_b$  no singular y  $\Lambda_a$  singular) then
14    $n_a = n$ ;
15    $n_b = \text{rango } \Lambda_{zuy}^M(n_a, 2n_a) - (2n_a + 1)$ ;
16 else if ( $\Lambda_b$  singular y  $\Lambda_a$  no singular) then
17    $n_b = n$ ;
18    $n_a = \text{rango } \Lambda_{zuy}^M(2n_b, n_b) - (2n_b + 1)$ ;
19 else if ( $\Lambda_b$  no singular y  $\Lambda_a$  no singular) then
20    $n_b = n$ ;
21    $n_a = n$ ;
22 else
23   Advertencia: Esta situación no puede
   ocurrir. Revisar los rangos de las matrices;
24 end

```

Algoritmo 3.2: Búsqueda del par (n_b, n_a) con el método de perfiles de valores singulares. Caso con ruido.

correcta del rango, si la perturbación E es suficientemente grande. El procedimiento para determinar el rango de la matriz $\tilde{A} = A + E$, descartando estos valores singulares pequeños, se llama *rango efectivo de la matriz* y se puede abordar mediante la ecuación (3.1), es decir

$$\text{rango } \tilde{A} = \# \left\{ \sigma(\tilde{A}) \mid \sigma(\tilde{A}) > \epsilon_\sigma \right\}$$

donde ϵ_σ da una medida del nivel de perturbación que presentarían los valores singulares de \tilde{A} .

El *nivel de discriminación* o *umbral* ϵ_σ no es fácil de obtener para la matriz $\Lambda_{zuy}^M(\hat{n}_b, \hat{n}_a)$, ya que ella no sólo tiene errores numéricos, sino que también tiene errores debido a las correlaciones instrumentales, ya que ellas no eliminan completamente el ruido cuando se utiliza un número finito de muestras. De hecho, en estos casos, los errores numéricos son despreciables frente a los errores en las correlaciones instrumentales, resultando así que

$$\Lambda_{zuy}^M = \underbrace{\begin{bmatrix} \Gamma_{uu}^M & (F_{xu}^M)^T \\ F_{zu}^M & \Gamma_{zx}^M \end{bmatrix}}_{\Lambda_{zux}^M} + \underbrace{\begin{bmatrix} 0 & (F_{vu}^M)^T \\ 0 & \Gamma_{zv}^M \end{bmatrix}}_{E_{zuv}^M}$$

donde las matrices $\Gamma_{\alpha\beta}^M$ y $F_{\alpha\beta}^M$ se encuentran definidas en (2.27) y (2.28) respectivamente, y la matriz E_{zuv}^M corresponde a los errores producidos por la eliminación incompleta del ruido, la cual tiende a cero cuando M tiende a infinito, bajo las condiciones del lema 2.2.

Según el corolario D.1, es posible saber cuán desviados están los valores singulares de la matriz perturbada Λ_{zuy}^M , con respecto a los valores singulares de la matriz sin perturbar Λ_{zux}^M

$$|\sigma(\Lambda_{zuy}^M) - \sigma(\Lambda_{zux}^M)| \leq \|E_{zuv}^M\|_2 = \|\bar{E}_{zuv}^M\|_2$$

donde

$$\bar{E}_{zuv}^M = \begin{bmatrix} (F_{vu}^M)^T \\ \Gamma_{zv}^M \end{bmatrix}$$

Así, para conocer el nivel de perturbación de cada valor singular, es necesario saber qué tan grande es el error \bar{E}_{zuv}^M .

Supóngase ahora que las correlaciones que conforman la matriz \bar{E}_{zuv}^M se encuentran acotadas

$$|\varphi_{uv}^M(-\tau_1, \tau_2)| \leq \epsilon_{uv}, \quad |\varphi_{zv}^M(-\tau_1, \tau_2)| \leq \epsilon_{zv}$$

entonces según el teorema D.7, la medida del error $\|\bar{E}_{zuv}^M\|_2$ se encuentra acotada por

$$\|\bar{E}_{zuv}^M\|_2 \leq \sqrt{[\epsilon_{uv}^2(\hat{n}_b + 1) + \epsilon_{zv}^2\hat{n}_a]\hat{n}_a} \quad (3.5)$$

donde \bar{E}_{zuv}^M es de tamaño $[(\hat{n}_b + 1) + \hat{n}_a] \times \hat{n}_a$.

Estas cotas ϵ_{uv} y ϵ_{zv} existen bajo ciertas condiciones, pero en general no son fáciles de obtener. Su existencia depende principalmente de la convergencia de las correlaciones φ_{uv}^M y φ_{zv}^M cuando $M \rightarrow \infty$, la que se obtiene, por ejemplo, cuando las correlaciones son ergódicas

$$\begin{aligned} \lim_{M \rightarrow \infty} \varphi_{uv}^M(-\tau_1, \tau_2) &= r_{uv}(\tau_2 - \tau_1) \\ \lim_{M \rightarrow \infty} \varphi_{zv}^M(-\tau_1, \tau_2) &= r_{zv}(\tau_2 - \tau_1) \end{aligned}$$

las cuales son iguales a cero si $u(k)$ y $e(k)$ tienen media cero y no están correlacionadas¹

$$\forall \tau, r_{uv}(\tau) = c_{uv}(\tau) = 0 \text{ y } r_{zv}(\tau) = c_{zv}(\tau) = 0$$

De esta manera, y bajo las condiciones del teorema C.2, esta convergencia se logra con probabilidad 1

$$\begin{aligned} (\forall \epsilon_{uv} > 0)(\exists M_{uv}^o)(\forall M > M_{uv}^o) \\ P(|\varphi_{uv}^M - r_{uv}| \leq \epsilon_{uv}) = 1 \\ (\forall \epsilon_{zv} > 0)(\exists M_{zv}^o)(\forall M > M_{zv}^o) \\ P(|\varphi_{zv}^M - r_{zv}| \leq \epsilon_{zv}) = 1 \end{aligned}$$

Es decir, existe una cantidad de muestras M_{uv}^o y M_{zv}^o que permite que las correlaciones se encuentren acotadas por ϵ_{uv} y ϵ_{zv} con probabilidad 1

$$\begin{aligned} \forall M > M_{uv}^o, \quad P(|\varphi_{uv}^M| \leq \epsilon_{uv}) = 1 \\ \forall M > M_{zv}^o, \quad P(|\varphi_{zv}^M| \leq \epsilon_{zv}) = 1 \end{aligned}$$

lo que demuestra la existencia de dichas cotas.

A pesar de la existencia de las cotas anteriores, no es una tarea fácil poder determinarlas, por lo que

¹ $\rho_{ue}(\tau) = 0$, ver definición 1.4

para estimar aproximadamente el nivel de estas cotas puede utilizarse el otro resultado de convergencia obtenido del teorema C.2, donde la convergencia se logra en sentido cuadrático medio

$$\begin{aligned} (\forall \epsilon_{uv} > 0)(\exists M_{uv})(\forall M > M_{uv}) \\ E \left\{ |\varphi_{uv}^M - r_{uv}|^2 \right\} \leq \epsilon_{uv}^2 \\ (\forall \epsilon_{zv} > 0)(\exists M_{zv})(\forall M > M_{zv}) \\ E \left\{ |\varphi_{zv}^M - r_{zv}|^2 \right\} \leq \epsilon_{zv}^2 \end{aligned}$$

Es decir, existe una cantidad de muestras M_{uv} y M_{zv} que permite que las correlaciones se encuentren “acotadas” por ϵ_{uv} y ϵ_{zv} en sentido cuadrático medio

$$\begin{aligned} \forall M > M_{uv}, \quad E \left\{ |\varphi_{uv}^M|^2 \right\} \leq \epsilon_{uv}^2 \\ \forall M > M_{zv}, \quad E \left\{ |\varphi_{zv}^M|^2 \right\} \leq \epsilon_{zv}^2 \end{aligned}$$

donde ϵ_{uv} y ϵ_{zv} corresponderían a una suerte de cotas RMS², las cuales entregan un nivel cercano a ϵ_{uv} y ϵ_{zv}

$$\epsilon_{uv} \sim \epsilon_{uv}, \quad \epsilon_{zv} \sim \epsilon_{zv}$$

y además son mucho más fáciles de obtener, ya que se puede utilizar la varianza de los estimadores φ_{uv}^M y φ_{zv}^M para calcularlas.

Considérese por el momento sólo el estimador

$$\varphi_{zv}^M(-\tau_1, \tau_2) = \frac{1}{2M+1} \sum_{k=-M}^M z(k-\tau_1)v(k-\tau_2)$$

que puede ser escrito, de manera conveniente, como

$$\tilde{\varphi}_{zv}^M(\tau, \tau_1) = \frac{1}{2M+1} \sum_{k=-M-\tau_1}^{M-\tau_1} z(k)v_\tau(k)$$

donde

$$\begin{aligned} \varphi_{zv}^M(-\tau_1, \tau_2) &= \tilde{\varphi}_{zv}^M(\tau, \tau_1), \quad \tau = \tau_2 - \tau_1 \\ v_\tau(k) &= H_\tau(q)e(k), \quad H_\tau(q) = q^{-\tau}H(q) \end{aligned}$$

Ahora, teniendo en cuenta las siguientes asociaciones

$$p_M(-\tau_1) = \tilde{\varphi}_{zv}^M(\tau, \tau_1), \quad s(k) = z(k)v_\tau(k)$$

y según el lema C.1, se tiene que la varianza del estimador φ_{zv}^M es

$$\begin{aligned} E \left\{ |\varphi_{zv}^M(-\tau_1, \tau_2) - r_{zv}(\tau_2 - \tau_1)|^2 \right\} = \\ \frac{1}{2M+1} \sum_{l=-2M}^{2M} \left(1 - \frac{|l|}{2M+1} \right) c_{ss}(l) \end{aligned}$$

²root mean square o valor eficaz

donde la covarianza $c_{ss}(l)$ está dada por el lema C.2. Si para el cálculo de $c_{ss}(l)$ se considera que el filtro instrumental tiene la siguiente forma

$$F(q) = q^{-m_l} \Leftrightarrow f(k) = \delta_K(k - m_l)$$

y que $u(k)$ y $e(k)$ no están correlacionadas, entonces la covarianza $c_{ss}(l)$ resulta

$$c_{ss}(l) = \delta_K(l) \left(\sum_{j \in \mathbb{Z}} h_\tau(l+j) h_\tau(j) \right) a_e^2 a_u^2$$

donde a_e^2 y a_u^2 son respectivamente las varianzas de $e(k)$ y $u(k)$, lo que equivale a

$$c_{ss}(l) = \delta_K(l) a_v^2 a_u^2$$

donde a_v^2 es la varianza del ruido de salida $v(k)$

$$a_v^2 = a_e^2 \sum_{j \in \mathbb{Z}} h_\tau(j)^2 = a_e^2 \sum_{k \in \mathbb{Z}} h(k)^2$$

De esta manera, la varianza del estimador φ_{zv}^M resulta

$$E \left\{ |\varphi_{zv}^M(-\tau_1, \tau_2)|^2 \right\} = \frac{(a_v a_u)^2}{2M+1} =: \varepsilon_{zv}^2$$

Ahora, la varianza del estimador φ_{uv}^M se obtiene a partir del resultado anterior, considerando que el filtro instrumental es igual a $f(k) = \delta_K(k)$, así

$$E \left\{ |\varphi_{uv}^M(-\tau_1, \tau_2)|^2 \right\} = \frac{(a_v a_u)^2}{2M+1} =: \varepsilon_{uv}^2$$

es decir

$$\varepsilon^2 := \varepsilon_{uv}^2 = \varepsilon_{zv}^2 = \frac{(a_v a_u)^2}{2M+1}$$

Finalmente, estos valores RMS $\varepsilon = \varepsilon_{uv} = \varepsilon_{zv}$ entregan niveles cercanos a ε_{uv} y ε_{zv}

$$\varepsilon_{uv} \sim \varepsilon_{uv} = \varepsilon, \quad \varepsilon_{zv} \sim \varepsilon_{zv} = \varepsilon$$

Entonces el nivel de perturbación que tienen los valores singulares de Λ_{zuy}^M , dado por (3.5), es aproximadamente

$$\sqrt{[\varepsilon_{uv}^2(\hat{n}_b + 1) + \varepsilon_{zv}^2 \hat{n}_a] \hat{n}_a} \sim \varepsilon \sqrt{(\hat{n}_b + 1 + \hat{n}_a) \hat{n}_a}$$

lo que se puede resumir en el siguiente teorema

Teorema 3.1 *Considérese los siguientes procesos estocásticos reales*

$$z(k) = F(q)u(k), \quad v(k) = H(q)e(k)$$

donde $F(q)$ y $H(q)$ son filtros tales que su respuestas a impulso satisfacen

$$f(k) = \delta_K(k - m_l), \quad \sum_{k=0}^{\infty} |h(k)|^2 < \infty$$

Considérese además que $\xi(k) = [u(k) \ e(k)]^T$ es un proceso de ruido blanco estricto³ de media cero, tal que

$$c_{\xi\xi}(\tau) = \begin{bmatrix} a_u^2 & 0 \\ 0 & a_e^2 \end{bmatrix} \delta_K(\tau)$$

Entonces, los valores singulares de $\Lambda_{zuy}^M(\hat{n}_b, \hat{n}_a)$ se ven perturbados según

$$|\sigma(\Lambda_{zuy}^M) - \sigma(\Lambda_{zux}^M)| \leq \|\bar{E}_{zuv}^M\|_2 \sim \varepsilon_\sigma$$

donde

$$\begin{aligned} \varepsilon_\sigma &:= \varepsilon \sqrt{(\hat{n}_b + 1 + \hat{n}_a) \hat{n}_a} \\ \varepsilon^2 &:= E \left\{ |\varphi_{uv}^M|^2 \right\} = E \left\{ |\varphi_{zv}^M|^2 \right\} = \frac{(a_v a_u)^2}{2M+1} \\ a_v^2 &= E \left\{ v(k)^2 \right\} = a_e^2 \sum_{\ell \in \mathbb{Z}} h(\ell)^2 \end{aligned}$$

Este teorema permite abordar el cálculo del rango de $\Lambda_{zuy}^M(\hat{n}_b, \hat{n}_a)$, ya que los valores singulares nulos de $\Lambda_{zux}^M(\hat{n}_b, \hat{n}_a)$ corresponderían aproximadamente a los valores singulares de $\Lambda_{zuy}^M(\hat{n}_b, \hat{n}_a)$ que cumplen con

$$|\sigma(\Lambda_{zuy}^M)| \leq \varepsilon_\sigma$$

y por lo tanto, su rango efectivo estaría dado por

$$\text{rango } \Lambda_{zuy}^M = \# \{ \sigma(\Lambda_{zuy}^M) \mid \sigma(\Lambda_{zuy}^M) > \varepsilon_\sigma \} \quad (3.6)$$

siempre que el filtro instrumental sea de la forma $F(q) = q^{-m_l}$, que la potencia del ruido de salida a_v^2 sea conocida, y que $u(k)$ y $e(k)$ sean señales de ruido blanco estricto e independientes entre sí.

El teorema anterior entrega también un criterio de diseño para el filtro instrumental, y además, proporciona una herramienta para poder diseñar el número de mediciones a tomar, el cual está dado por la siguiente relación

$$\varepsilon_\sigma^2 = \frac{(a_v a_u)^2}{2M+1} (\hat{n}_b + 1 + \hat{n}_a) \hat{n}_a \quad (3.7)$$

donde M está relacionado con el número de pares de datos del conjunto de mediciones Z_{uy}^M , definido en (2.4).

³ $\xi(k_i)$ es independiente de $\xi(k_j) \forall k_i \neq k_j$, ver definición 1.5

De esta manera, si se quiere tener una precisión de d decimales en los valores singulares de $\Lambda_{zuy}^M(\hat{n}_b, \hat{n}_a)$, entonces el número de datos debe ser proporcional a

$$2M + 1 \propto \varepsilon_\sigma^{-2} = (0,5 \cdot 10^{-d})^{-2} = 4 \cdot 10^{2d}$$

Es decir, por cada decimal de precisión que se quiere en los valores singulares de $\Lambda_{zuy}^M(\hat{n}_b, \hat{n}_a)$, se debe aumentar el número de datos en dos órdenes de magnitud.

A continuación se presentan algunos ejemplos, que ilustran el algoritmo 3.2 para identificación estructural, en conjunto con los principales resultados obtenidos para el caso con ruido

Ejemplo 3.3 *Considérese la planta en [2, example 3], la cual tiene ceros ubicados en $\{-0,8; 1,25\}$ y polos ubicados en $\{0,8; 0,7 \pm j0,4\}$*

$$G(q) = \frac{1 - 0,45q^{-1} - q^{-2}}{1 - 2,2q^{-1} + 1,77q^{-2} - 0,52q^{-3}}$$

cuya estructura $(n_b, n_a) = (2, 3)$ se desea identificar. Esta planta se encuentra sometida a ruido de salida $v(k)$, generado filtrando ruido blanco estricto uniforme $e(k)$ de media cero y varianza 1 mediante

$$v(k) = H(q)e(k), \quad H(q) = \frac{0,4q}{q - 0,6}$$

resultando así que el ruido de salida $v(k)$ tiene varianza conocida $a_v^2 = 0,25$.

En general, los sistemas de órdenes pequeños son relativamente fáciles de identificar, sin embargo, este sistema presenta un factor pasa-todo

$$\frac{q - 1,25}{q - 0,8}$$

el cual dificulta el proceso de identificación estructural.

Para diseñar el número de datos a medir, se utiliza la ecuación (3.7), suponiendo que el orden del sistema es cercano a $\hat{n}_b = \hat{n}_a = 5$, que se desea precisión de un decimal $\varepsilon_\sigma = 0,5 \cdot 10^{-1}$ en los valores singulares y que la señal de entrada tiene varianza $a_u^2 = 1$

$$2M + 1 = \frac{(a_v a_u)^2}{\varepsilon_\sigma^2} (\hat{n}_b + 1 + \hat{n}_a) \hat{n}_a = 5500$$

Por este motivo se escoge $M = 3000$.

Así, los parámetros a utilizar en el algoritmo 3.2 se escogen según

$$\hat{n}_{\text{máx}} = 15, \quad M = 3000, \quad L = \hat{n}_{\text{máx}}$$

y la señal $u(k)$ se diseña como una realización de ruido blanco gaussiano de media cero y varianza $a_u^2 = 1$, generada independientemente de $e(k)$, la cual es p.e. de cualquier orden. Finalmente, el filtro instrumental se diseña de la forma más simple posible

$$F(q) = q^{-\hat{n}_b}$$

satisfaciendo las condiciones del algoritmo 3.2 y del teorema 3.1.

De esta forma se miden los datos $Z_{uy}^{L,M}$ en estado estacionario, con los cuales se construyen los perfiles de valores singulares para la matriz $\Lambda_{zuy}^M(\hat{n}_b, \hat{n}_a)$, mostrados en la figura 3.5. Para el cálculo de los rangos se usa la estrategia de la ecuación (3.6), considerando ε_σ dado por el teorema 3.1. Además, para verificar el comportamiento de los valores singulares "nulos", la figura muestra 50 realizaciones de los valores singulares para cada matriz.

Siguiendo el algoritmo 3.2, se escoge inicialmente el modelo $(\hat{n}_b, \hat{n}_a) = (10, 10)$, para el cual se obtiene rango $\Lambda_{zuy}^M(10, 10) = 14$, y así se puede determinar el orden del sistema

$$n = \text{rango } \Lambda_{zuy}^M(10, 10) - (10 + 1) = 3$$

Una vez que el orden del sistema es conocido $n = 3$, se computan los rangos de las siguientes matrices

$$\begin{aligned} \text{rango } \Lambda_{zuy}^M(4, 3) &= \text{deficiente (singular)} \\ \text{rango } \Lambda_{zuy}^M(3, 4) &= \text{deficiente (singular)} \end{aligned}$$

Esta situación, según el algoritmo 3.2, no puede ocurrir, lo que indica que alguno de los rangos anteriores fue mal calculado. De hecho la matriz $\Lambda_{zuy}^M(4, 3)$ se encuentra justo en el límite de discriminación, mientras que la matriz $\Lambda_{zuy}^M(3, 4)$ es claramente singular. Es por este motivo que la situación debe ser

$$\begin{aligned} \text{rango } \Lambda_{zuy}^M(4, 3) &= \text{completo (no singular)} \\ \text{rango } \Lambda_{zuy}^M(3, 4) &= \text{deficiente (singular)} \end{aligned}$$

lo que indica que el orden del sistema corresponde a n_a , es decir

$$n_a = n = 3$$

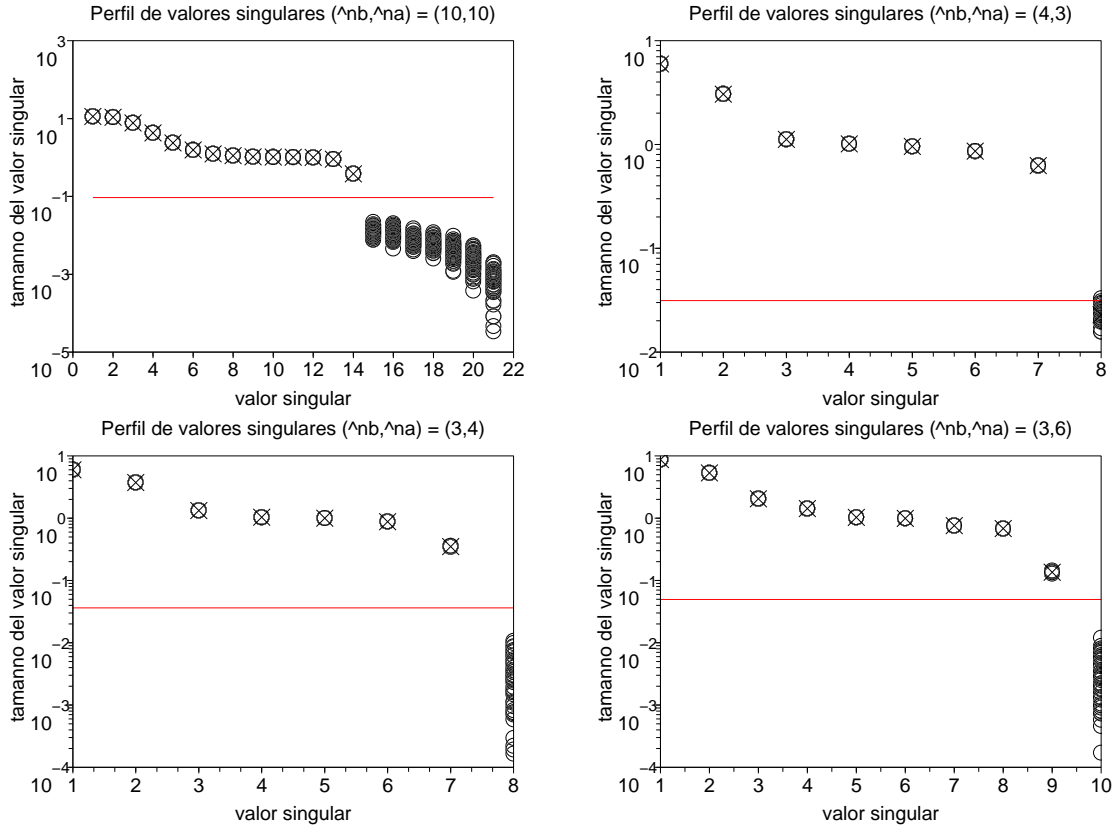


Figura 3.5: Con “o” los valores singulares de $\Lambda_{zuy}^M(\hat{n}_b, \hat{n}_a)$, con “x” los valores singulares de $\Lambda_{zux}^M(\hat{n}_b, \hat{n}_a)$ y en rojo ε_σ . Cada gráfico muestra 50 realizaciones con $M = 3 \cdot 10^3$ y SNR de salida ≈ 20 [dB].

Finalmente se obtiene el valor de n_b , mediante la matriz asociada al modelo $(\hat{n}_b, \hat{n}_a) = (n_a, 2n_a)$, teniendo en cuenta que $n_a = n = 3$. De esta forma, $\text{rango } \Lambda_{zuy}^M(3, 6) = 9$, y por lo tanto

$$n_b = \text{rango } \Lambda_{zuy}^M(3, 6) - (2 \cdot 3 + 1) = 2$$

con lo que la estructura identificada del sistema resulta $(n_b, n_a) = (2, 3)$.

Para más detalles de la simulación vea el programa F.4.2.

Ejemplo 3.4 Considérese ahora el mismo sistema que en [13, example 4], el que tiene ceros ubicados en

$$\{0, 48 \pm j0, 84; -0, 65 \pm j0, 65\}$$

y polos ubicados en

$$\{0, 71 \pm j0, 41; -0, 67 \pm j0, 39; -0, 43 \pm j0, 74\}$$

$$G(q) = \frac{B(q)}{A(q)}$$

$$B(q) = 1 + 0,3452q^{-1} + 0,53q^{-2} + 0,3985q^{-3} + 0,8138q^{-4}$$

$$A(q) = 1 + 0,7907q^{-1} + 0,042q^{-2} - 0,5556q^{-3} - 0,0247q^{-4} + 0,3846q^{-5} + 0,3026q^{-6}$$

cuya estructura $(n_b, n_a) = (4, 6)$ se desea identificar. Esta planta se encuentra sometida a ruido de salida $v(k)$, generado filtrando ruido blanco estricto gaussiano $e(k)$ de media cero y varianza 1 mediante

$$v(k) = H(q)e(k), \quad H(q) = \frac{0, 4q}{q - 0, 6}$$

resultando así que el ruido de salida $v(k)$ tiene varianza conocida $a_v^2 = 0, 25$.

Los parámetros a utilizar en el algoritmo 3.2 se escogen iguales al ejemplo anterior, es decir

$$\hat{n}_{\text{máx}} = 15, \quad M = 3000, \quad L = \hat{n}_{\text{máx}}$$

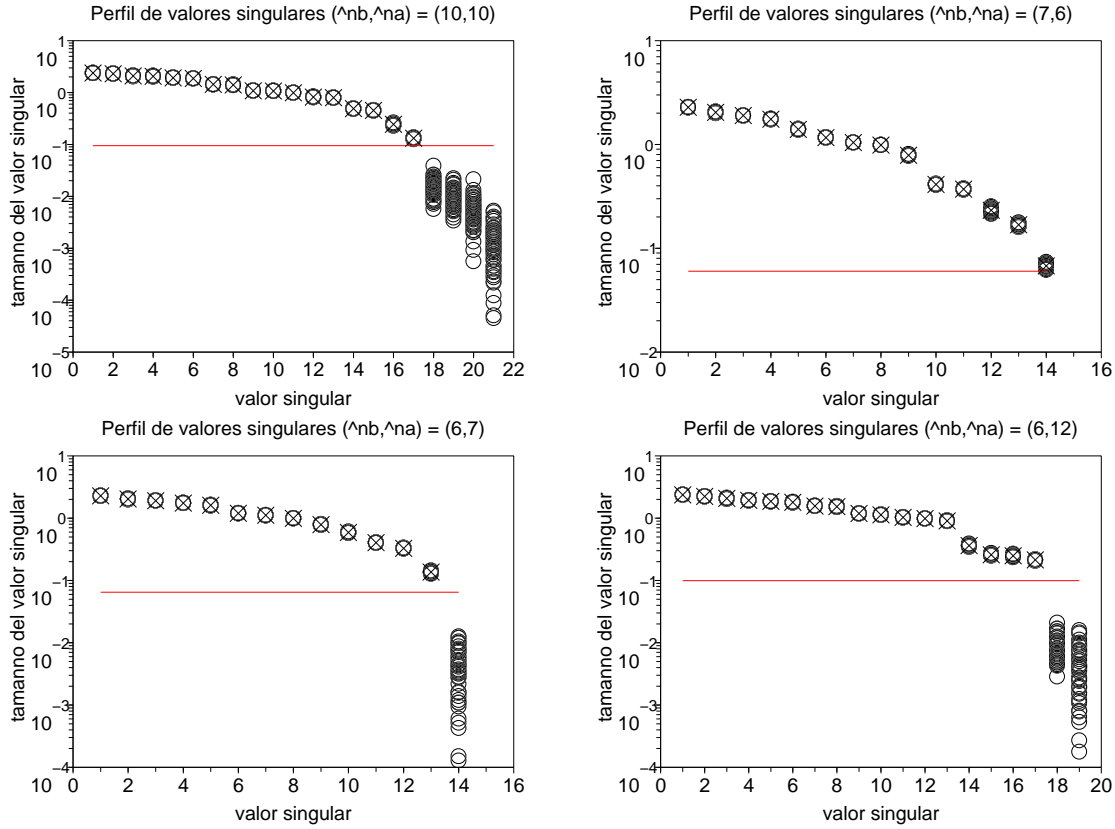


Figura 3.6: Con “o” los valores singulares de $\Lambda_{zuy}^M(\hat{n}_b, \hat{n}_a)$, con “x” los valores singulares de $\Lambda_{zux}^M(\hat{n}_b, \hat{n}_a)$ y en rojo ε_σ . Cada gráfico muestra 50 realizaciones con $M = 3 \cdot 10^3$ y SNR de salida $\approx 10[dB]$.

La señal $u(k)$ se diseña como una realización de ruido blanco gaussiano de media cero y varianza 1, generada independientemente de $e(k)$, la cual es p.e. de todo orden. Finalmente, el filtro instrumental se diseña de la forma más simple posible

$$F(q) = q^{-\hat{n}_b}$$

satisfaciendo las condiciones del algoritmo 3.2 y del teorema 3.1.

De esta forma se miden los datos $Z_{uy}^{L,M}$ en estado estacionario, con los cuales se construyen los perfiles de valores singulares para la matriz $\Lambda_{zuy}^M(\hat{n}_b, \hat{n}_a)$, mostrados en la figura 3.6. Para el cálculo de los rangos se usa la estrategia de la ecuación (3.6), considerando ε_σ dado por el teorema 3.1. Además, para verificar el comportamiento de los valores singulares “nulos”, la figura muestra 50 realizaciones de los valores singulares para cada matriz.

Siguiendo el algoritmo 3.2, se escoge inicialmente el modelo $(\hat{n}_b, \hat{n}_a) = (10, 10)$, para el cual se obtiene

$\text{rango } \Lambda_{zuy}^M(10, 10) = 17$, así se puede determinar el orden del sistema

$$n = \text{rango } \Lambda_{zuy}^M(10, 10) - (10 + 1) = 6$$

Una vez que el orden del sistema es conocido $n = 6$, se computan los rangos de las siguientes matrices

$$\text{rango } \Lambda_{zuy}^M(7, 6) = \text{completo (no singular)}$$

$$\text{rango } \Lambda_{zuy}^M(6, 7) = \text{deficiente (singular)}$$

lo que indica que el orden del sistema corresponde a n_a , es decir

$$n_a = n = 6$$

Finalmente se obtiene el valor de n_b , mediante la matriz asociada al modelo $(\hat{n}_b, \hat{n}_a) = (n_a, 2n_a)$, teniendo en cuenta que $n_a = n = 6$. De esta forma, $\text{rango } \Lambda_{zuy}^M(6, 12) = 17$, y por lo tanto

$$n_b = \text{rango } \Lambda_{zuy}^M(6, 12) - (2 \cdot 6 + 1) = 4$$

con lo que la estructura identificada del sistema resulta $(n_b, n_a) = (4, 6)$.

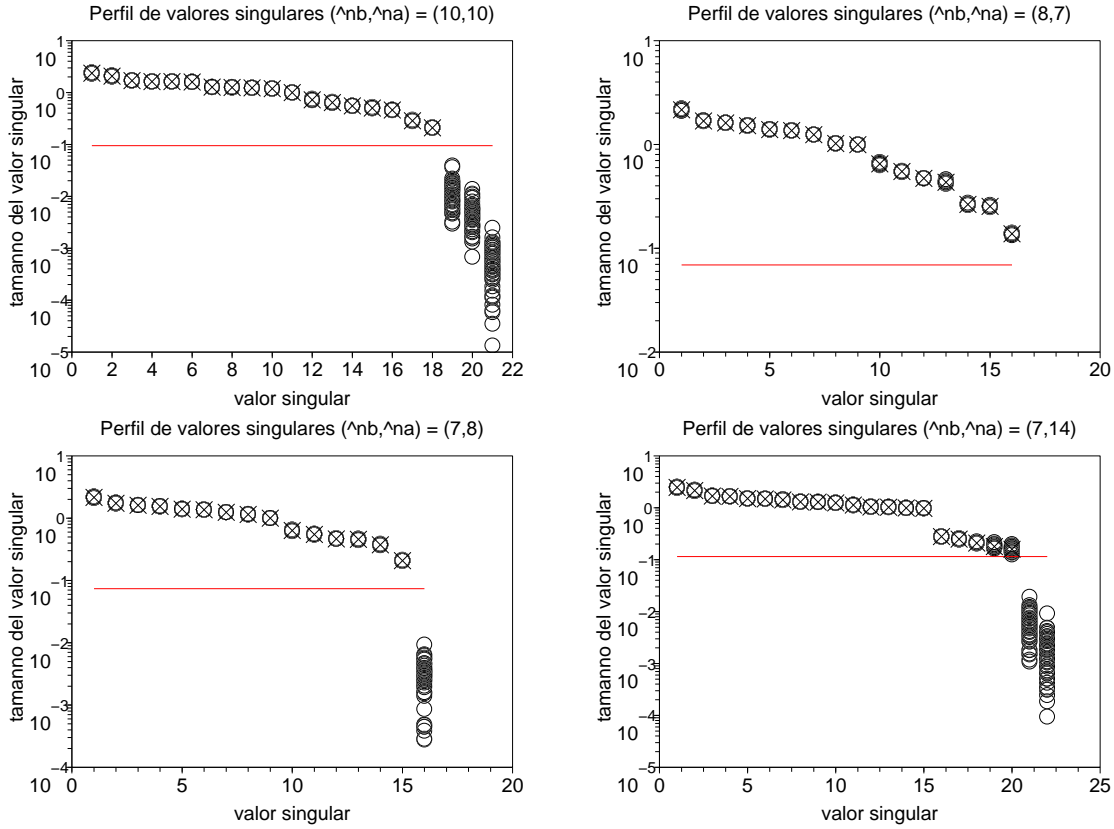


Figura 3.7: Con “o” los valores singulares de $\Lambda_{zuy}^M(\hat{n}_b, \hat{n}_a)$, con “x” los valores singulares de $\Lambda_{zux}^M(\hat{n}_b, \hat{n}_a)$ y en rojo ε_σ . Cada gráfico muestra 50 realizaciones con $M = 3 \cdot 10^3$ y SNR de salida ≈ 9 [dB].

Para más detalles de la simulación vea el programa F.4.2.

Ejemplo 3.5 Considérese ahora el mismo sistema que en [1, example 1]

$$\begin{aligned}
 G(q) &= \frac{B(q)}{A(q)} \\
 B(q) &= 1 - 0,1292q^{-1} - 0,3893q^{-2} \\
 &\quad + 0,3163q^{-3} + 0,2549q^{-4} - 0,5823q^{-5} \\
 A(q) &= 1 - 0,2459q^{-1} \\
 &\quad - 0,3786q^{-2} - 0,0707q^{-3} - 0,0944q^{-4} \\
 &\quad - 0,2881q^{-5} + 0,107q^{-6} + 0,4695q^{-7}
 \end{aligned}$$

cuya estructura $(n_b, n_a) = (5, 7)$ se desea identificar. Esta planta se encuentra sometida a ruido de salida $v(k)$, generado filtrando ruido blanco estricto gaussiano $e(k)$ de media cero y varianza 1 mediante

$$v(k) = H(q)e(k)$$

$$H(q) = \frac{0,0976q^2 + 0,1953q + 0,0976}{q^2 - 0,9428q + 0,3333}$$

resultando así que el ruido de salida $v(k)$ tiene varianza conocida $a_v^2 = 0,25$.

Los parámetros a utilizar en el algoritmo 3.2 se escogen iguales al ejemplo anterior, es decir

$$\hat{n}_{\text{máx}} = 15, \quad M = 3000, \quad L = \hat{n}_{\text{máx}}$$

La señal $u(k)$ se diseña como una realización de ruido blanco gaussiano de media cero y varianza 1, generada independientemente de $e(k)$, la cual es p.e. de todo orden. Finalmente, el filtro instrumental se diseña de la forma más simple posible

$$F(q) = q^{-\hat{n}_b}$$

satisfaciendo las condiciones del algoritmo 3.2 y del teorema 3.1.

De esta forma se miden los datos $Z_{uy}^{L,M}$ en estado estacionario, con los cuales se construyen los per-

files de valores singulares para la matriz $\Lambda_{zuy}^M(\hat{n}_b, \hat{n}_a)$, mostrados en la figura 3.7. Para el cálculo de los rangos se usa la estrategia de la ecuación (3.6), considerando ε_σ dado por el teorema 3.1. Además, para verificar el comportamiento de los valores singulares “nulos”, la figura muestra 50 realizaciones de los valores singulares para cada matriz.

Siguiendo el algoritmo 3.2, se escoge inicialmente el modelo $(\hat{n}_b, \hat{n}_a) = (10, 10)$, para el cual se obtiene $\text{rango } \Lambda_{zuy}^M(10, 10) = 18$, así se puede determinar el orden del sistema

$$n = \text{rango } \Lambda_{zuy}^M(10, 10) - (10 + 1) = 7$$

Una vez que el orden del sistema es conocido $n = 7$, se computan los rangos de las siguientes matrices

$$\begin{aligned} \text{rango } \Lambda_{zuy}^M(8, 7) &= \text{completo (no singular)} \\ \text{rango } \Lambda_{zuy}^M(7, 8) &= \text{deficiente (singular)} \end{aligned}$$

lo que indica que el orden del sistema corresponde a n_a , es decir

$$n_a = n = 7$$

Finalmente se obtiene el valor de n_b , mediante la matriz asociada al modelo $(\hat{n}_b, \hat{n}_a) = (n_a, 2n_a)$, teniendo en cuenta que $n_a = n = 7$. De esta forma, $\text{rango } \Lambda_{zuy}^M(7, 14) = 20$, y por lo tanto

$$n_b = \text{rango } \Lambda_{zuy}^M(7, 14) - (2 \cdot 7 + 1) = 5$$

con lo que la estructura identificada del sistema resulta $(n_b, n_a) = (5, 7)$.

Para más detalles de la simulación vea el programa F.4.2.

3.3. Comentarios del capítulo

En un ambiente libre de ruido, y bajo las condiciones del lema 2.3, el rango de la matriz de correlaciones muestrales $\Lambda_{xux}^M(\hat{n}_b, \hat{n}_a)$ queda completamente determinado, si se utiliza una señal de entrada cuyo orden de excitación sea suficientemente alto. Este resultado conduce a métodos *no exhaustivos* de identificación estructural, como el propuesto en el algoritmo 3.1, el que aprovecha el rango obtenido en el lema 2.3, para reducir el número total de matrices analizadas (a un máximo de 4), y así obtener la estructura del sistema. Esto presenta una evidente ventaja, comparado con los métodos exhaustivos del capítulo 2, los

que requieren analizar un conjunto mayor matrices, para poder encontrar la estructura correcta dentro de la región

$$(\hat{n}_b, \hat{n}_a) \in [1, \hat{n}_{\text{máx}}] \times [1, \hat{n}_{\text{máx}}]$$

Si la salida del sistema se ve perturbada por ruido de distribución de probabilidad desconocida, se puede utilizar la matriz de correlaciones instrumentales $\Lambda_{zuy}^M(\hat{n}_b, \hat{n}_a)$, la cual reduce el efecto del ruido de salida, debido al lema 2.2. Además, bajo las condiciones del teorema 2.2 el rango de la matriz instrumental $\Lambda_{zux}^M(\hat{n}_b, \hat{n}_a)$ también queda completamente determinado, si se utiliza una señal cuyo orden de excitación es suficientemente alto. Así, a través del lema 2.2 y del teorema 2.2, es posible recuperar el método *no exhaustivo* de identificación estructural, el cual se resume en el algoritmo 3.2, aún cuando el ruido de salida tiene distribución de probabilidad desconocida.

En los casos sin y con ruido de salida, se puede utilizar la descomposición en valores singulares de las diferentes matrices para poder determinar su rango. En el caso sin ruido los valores singulares nulos se pueden descartar utilizando el criterio de la ecuación (3.1), cuyo umbral de discriminación ε_σ está dado por la ecuación (3.2), ya que estas matrices tienen sólo errores de tipo numérico. En el caso en que existe ruido de salida, la perturbación producida por el ruido es mucho mayor que las perturbaciones numéricas, por lo que se utiliza la estrategia de la ecuación (3.6) para descartar los valores singulares nulos, considerando el umbral de discriminación ε_σ , obtenido en el teorema 3.1.

La elección del filtro instrumental de la forma $F(q) = q^{-m_i}$, además de ser simple, permite obtener el nivel de discriminación ε_σ , útil para calcular el rango efectivo de Λ_{zuy}^M . Gracias a la elección de este filtro, dicho umbral depende sólo de la potencia del ruido de salida, de la potencia de la señal de entrada y del número de muestras.

Capítulo 4

Relación con la matriz de Yule-Walker y con la matriz de Hankel del sistema

EN este capítulo se muestran algunas propiedades de valores singulares y valores propios para las matrices GPM y GIPM, las que dependen básicamente del retardo de la planta y del retardo de la secuencia instrumental. Posteriormente se verificará que, satisfaciendo ciertas condiciones, la matriz GIPM posee valores singulares que son redundantes al identificar el orden de la planta. También se obtendrá que la submatriz Γ_{zx}^∞ , de la matriz GIPM, es suficiente para determinar la estructura del sistema. Finalmente, se establecerán relaciones de esta matriz Γ_{zx}^∞ con la matriz de ecuaciones de Yule-Walker y con la matriz de Hankel del sistema, permitiendo estimar los valores singulares de Hankel, a partir de mediciones de entrada y salida.

4.1. Propiedades de valores singulares y valores propios

4.1.1. Caso de la matriz GPM

Considérese el sistema en (2.19) y la siguiente partición para la *matriz de producto generalizada* (GPM) $\Lambda_{xux}^\infty(\hat{n}_b, \hat{n}_a)$, definida en (2.23)

$$\Lambda_{xux}^\infty(\hat{n}_b, \hat{n}_a) = \begin{bmatrix} \Gamma_{uu}^\infty(\hat{n}_b + 1) & F_{xu}^\infty(\hat{n}_b, \hat{n}_a)^T \\ F_{xu}^\infty(\hat{n}_b, \hat{n}_a) & \Gamma_{xx}^\infty(\hat{n}_a) \end{bmatrix}$$

donde $\Gamma_{\alpha\beta}^\infty(\hat{n})$ y $F_{\alpha\beta}^\infty(\hat{n}_b, \hat{n}_a)$ se encuentran definidas en (2.27) y (2.28) respectivamente.

Al utilizar una actuación con características de ruido blanco, es decir, $\varphi_{uu}^\infty(\tau) = a_u^2 \delta_K(\tau)$, se obtiene

que $\Gamma_{uu}^\infty(\hat{n}_b + 1) = a_u^2 I$ y además que

$$F_{xu}^\infty(\hat{n}_b, \hat{n}_a)^T = a_u^2 \begin{bmatrix} g(-1) & \cdots & g(-\hat{n}_a) \\ g(0) & \cdots & g(1 - \hat{n}_a) \\ \vdots & \ddots & \vdots \\ g(\hat{n}_b - 1) & \cdots & g(\hat{n}_b - \hat{n}_a) \end{bmatrix}$$

la que tiene siempre la primera fila nula, ya que la planta es evidentemente causal.

Teorema 4.1 *Considérese que $u(k)$ es tal que $\varphi_{uu}^\infty(\tau) = a_u^2 \delta_K(\tau)$. Si la planta tiene la siguiente forma*

$$G(q) = \frac{b_{d_g} q^{-d_g} + \cdots + b_{n_b} q^{-n_b}}{1 + a_1 q^{-1} + \cdots + a_{n_a} q^{-n_a}}$$

entonces la matriz $\Lambda_{xux}^\infty(\hat{n}_b, \hat{n}_a)$ tiene por lo menos $1 + d_g$ valores singulares iguales a a_u^2 y al menos $1 + d_g$ valores propios iguales a a_u^2 .

Demostración. Lo anterior se debe a que la matriz $\Gamma_{uu}^\infty(\hat{n}_b + 1) = a_u^2 I$, y también a que las primeras $1 + d_g$ filas de $F_{xu}^\infty(\hat{n}_b, \hat{n}_a)^T$ y las primeras $1 + d_g$ columnas de $F_{xu}^\infty(\hat{n}_b, \hat{n}_a)$ son nulas. **QED.**

Notar que todos los valores singulares y valores propios iguales a a_u^2 , mencionados en el teorema anterior, tienen su origen gracias a que la matriz $\Gamma_{uu}^\infty(\hat{n}_b + 1)$ aporta tantos a_u^2 como retardos tiene la planta.

4.1.2. Caso de la matriz GIPM

Considere el sistema de la ecuación (2.19) y la secuencia instrumental definida en (2.25). De manera similar a la sección anterior, puede establecerse la

siguiente partición para la *matriz generalizada de producto instrumental* (GIPM) $\Lambda_{zux}^\infty(\hat{n}_b, \hat{n}_a)$

$$\Lambda_{zux}^\infty(\hat{n}_b, \hat{n}_a) = \begin{bmatrix} \Gamma_{uu}^\infty(\hat{n}_b + 1) & F_{xu}^\infty(\hat{n}_b, \hat{n}_a)^T \\ F_{zu}^\infty(\hat{n}_b, \hat{n}_a) & \Gamma_{zx}^\infty(\hat{n}_a) \end{bmatrix}$$

donde $\Gamma_{\alpha\beta}^\infty(\hat{n})$ y $F_{\alpha\beta}^\infty(\hat{n}_b, \hat{n}_a)$ se encuentran definidas en (2.27) y (2.28) respectivamente.

En el caso GIPM es mucho más fácil controlar la cantidad de valores propios iguales a a_u^2 , ya que esto se puede lograr diseñando de manera adecuada la secuencia instrumental; sin embargo, no ocurre lo mismo con los valores singulares iguales a a_u^2 , que siguen sujetos al retardo de la planta.

Lema 4.1 *Considérese que la actuación $u(k)$ es tal que $\varphi_{uu}^\infty(\tau) = a_u^2 \delta_K(\tau)$, que la planta tiene la siguiente forma*

$$G(q) = \frac{b_{d_g} q^{-d_g} + \dots + b_{n_b} q^{-n_b}}{1 + a_1 q^{-1} + \dots + a_{n_a} q^{-n_a}}$$

y que la secuencia instrumental se diseña de la siguiente manera

$$F(q) = \frac{l_{d_f} q^{-d_f} + \dots + l_{m_l} q^{-m_l}}{1 + k_1 q^{-1} + \dots + k_{m_k} q^{-m_k}}$$

donde $d_f \geq \hat{n}_b$. Entonces, $\Lambda_{zux}^\infty(\hat{n}_b, \hat{n}_a)$:

1. es una matriz triangular por bloques
2. tiene al menos $\hat{n}_b + 1$ valores propios iguales a a_u^2 , provenientes de la submatriz $\Gamma_{uu}^\infty(\hat{n}_b + 1)$
3. $\text{rango } \Lambda_{zux}^\infty(\hat{n}_b, \hat{n}_a) = (\hat{n}_b + 1) + \text{rango } \Gamma_{zx}^\infty(\hat{n}_a)$
4. tiene al menos $1 + d_g$ valores singulares iguales a a_u^2 , provenientes de la submatriz $\Gamma_{uu}^\infty(\hat{n}_b + 1)$

Demostración. La forma triangular por bloques de $\Lambda_{zux}^\infty(\hat{n}_b, \hat{n}_a)$ se debe exclusivamente a que

$$F_{zu}^\infty(\hat{n}_b, \hat{n}_a) = a_u^2 \begin{bmatrix} f(-1) & f(0) & \dots & f(\hat{n}_b - 1) \\ \vdots & \vdots & \ddots & \vdots \\ f(-\hat{n}_a) & f(1 - \hat{n}_a) & \dots & f(\hat{n}_b - \hat{n}_a) \end{bmatrix}$$

es una matriz nula para la secuencia instrumental escogida. Los valores propios se obtienen directamente, ya que $\Lambda_{zux}^\infty(\hat{n}_b, \hat{n}_a)$ es triangular por bloques, entonces se cumple que

$$\begin{aligned} \det(\Lambda_{zux}^\infty - \lambda I) &= \det(\Gamma_{uu}^\infty - \lambda I) \det(\Gamma_{zx}^\infty - \lambda I) \\ &= (a_u^2 - \lambda)^{\hat{n}_b + 1} \det(\Gamma_{zx}^\infty - \lambda I) \end{aligned}$$

donde el rango corresponde al número de valores propios que son diferentes de 0. Ahora, $\Lambda_{zux}^\infty(\hat{n}_b, \hat{n}_a)$ tiene $1 + d_g$ valores singulares iguales a a_u^2 , debido a que $F_{zu}(\hat{n}_b, \hat{n}_a)$ es una matriz nula y a que $F_{xu}(\hat{n}_b, \hat{n}_a)^T$ tiene sus primeras $1 + d_g$ filas nulas. **QED.**

Con las condiciones del resultado anterior, se puede ver que hay $\hat{n}_b + 1$ valores propios que no aportan información para determinar si $\Lambda_{zux}^\infty(\hat{n}_b, \hat{n}_a)$ es de rango completo, y que por lo tanto se considerarían redundantes. Esto hace pensar que podría existir una forma de prescindir de la matriz $\Lambda_{zux}^\infty(\hat{n}_b, \hat{n}_a)$ para identificar la estructura de la planta, lo que se podría lograr sólo mediante su submatriz $\Gamma_{zx}^\infty(\hat{n}_a)$.

Considérese por ejemplo que la matriz $\Lambda_{zux}^\infty(\hat{n}_b, \hat{n}_a)$ se construye utilizando $\hat{n}_b = d_f$, donde d_f es el retardo del filtro instrumental

$$F(q) = \frac{l_{d_f} q^{-d_f} + \dots + l_{m_l} q^{-m_l}}{1 + k_1 q^{-1} + \dots + k_{m_k} q^{-m_k}}$$

que la actuación $u(k)$ se diseña de manera que $\varphi_{uu}^\infty(\tau) = a_u^2 \delta_K(\tau)$, y que se satisfacen las condiciones del teorema 2.2, entonces el rango de $\Gamma_{zx}^\infty(\hat{n}_a)$

$$\text{rango } \Gamma_{zx}^\infty(\hat{n}_a) = \text{rango } \Lambda_{zux}^\infty(d_f, \hat{n}_a) - (d_f + 1)$$

queda completamente determinado, ya que el rango de $\Lambda_{zux}^\infty(d_f, \hat{n}_a)$ es conocido, gracias al teorema 2.2, lo que conduce al siguiente teorema

Teorema 4.2 *Considere los datos Z_{ux}^∞ y Z_{uz}^∞ medidos de (2.19) y (2.25) respectivamente. Si $u(k)$ es tal que $\varphi_{uu}^\infty(\tau) = a_u^2 \delta_K(\tau)$, y el filtro instrumental*

$$F(q) = \frac{l_{d_f} q^{-d_f} + \dots + l_{m_l} q^{-m_l}}{1 + k_1 q^{-1} + \dots + k_{m_k} q^{-m_k}}$$

satisface $\text{máx}(m_l - n_b, m_k - n_a) \geq 0$, entonces

$$\text{rango } \Gamma_{zx}^\infty(\hat{n}_a) = \begin{cases} \hat{n}_a, & i \leq 0 \\ \text{máx}(\hat{n}_a - d_f + n_b, n_a), & i > 0 \end{cases}$$

donde $i = \text{mín}(d_f - n_b, \hat{n}_a - n_a)$.

Nota 4.1 *Este último resultado se cumple “genéricamente”, es decir, pueden aparecer casos extraños en los cuales no es válido. Estos casos son bastante raros, y son producidos por combinaciones desafortunadas de $F(q)$, $G(q)$ y $u(k)$, pero desde un punto de vista práctico, no hay necesidad de preocuparse, ya que este conjunto de casos aparece de manera muy esporádica.*

Este resultado “genérico” para $\Gamma_{zx}^\infty(\hat{n}_a)$ se debe a que el rango de la matriz $\Lambda_{zux}^\infty(\hat{n}_b, \hat{n}_a)$ también se cumple “genéricamente” cuando $M \rightarrow \infty$. Para más detalles consulte la sección 2.4 del capítulo 2.

Nótese que la construcción de la submatriz GIPM $\Gamma_{zx}^\infty(\hat{n}_a)$ depende explícitamente de \hat{n}_a e implícitamente de d_f , es decir, del par (d_f, \hat{n}_a) . Así, la región en la cual $\Gamma_{zx}^\infty(\hat{n}_a)$ es no singular ($i \leq 0$) corresponde exactamente a la región gris de la figura 2.4, reemplazando \hat{n}_b por d_f . Análogamente, en la misma figura, se muestra en color blanco la región en la cual $\Gamma_{zx}^\infty(\hat{n}_a)$ es singular ($i > 0$).

4.2. Método de perfiles para la submatriz GIPM

Con el resultado del teorema 4.2, se evidencia que no es necesaria la matriz GIPM $\Lambda_{zux}^\infty(\hat{n}_b, \hat{n}_a)$ para estimar la estructura de la planta, sino que es suficiente el uso de la submatriz $\Gamma_{zx}^\infty(\hat{n}_a)$. Esta tarea se puede lograr a través de pruebas exhaustivas de singularidad sobre $\Gamma_{zx}^\infty(\hat{n}_a)$, de manera similar a los ejemplos 2.4, 2.5 y 2.6, mostrados en el capítulo 2, ya que la región en que $\Gamma_{zx}^\infty(\hat{n}_a)$ es singular es conocida.

Como se ha visto, estos métodos exhaustivos buscan la estructura correcta (n_b, n_a) en cada punto de la región $(d_f, \hat{n}_a) \in [1, \hat{n}_{\text{máx}}] \times [1, \hat{n}_{\text{máx}}]$, lo que no es muy eficiente, ya que se requiere analizar una gran cantidad de matrices cuando la región es grande.

En analogía al capítulo 3, se puede aprovechar la información adicional que entrega el teorema 4.2 sobre el rango de $\Gamma_{zx}^\infty(\hat{n}_a)$, para obtener un método no exhaustivo de búsqueda de la estructura del sistema, y así reducir el número de matrices utilizadas. Este método no exhaustivo podría resumirse como

1. Determinar el orden del sistema $n = \text{máx}(n_b, n_a)$ a partir de

$$n = \text{rango } \Gamma_{zx}^\infty(\hat{n}), \text{ con } d_f = \hat{n}$$

donde $\hat{n} > n$.

2. Determinar si el orden del sistema corresponde a n_b o n_a , mediante la verificación de la singularidad de las siguientes matrices

$$\Gamma_{zx}^\infty(n) = \text{¿singular?}, \text{ con } d_f = n + 1$$

$$\Gamma_{zx}^\infty(n + 1) = \text{¿singular?}, \text{ con } d_f = n$$

de esta manera, si

- a) $\Gamma_{zx}^\infty(n)$ es no singular y $\Gamma_{zx}^\infty(n + 1)$ es singular, entonces $n_a = n$
- b) $\Gamma_{zx}^\infty(n)$ es singular y $\Gamma_{zx}^\infty(n + 1)$ es no singular, entonces $n_b = n$
- c) $\Gamma_{zx}^\infty(n)$ y $\Gamma_{zx}^\infty(n + 1)$ son no singulares, entonces $n_b = n_a = n$

3. Dependiendo del resultado del punto anterior, calcular el parámetro restante según

$$n_b = \text{rango } \Gamma_{zx}^\infty(2n_a) - n_a, \text{ con } d_f = n_a$$

$$n_a = \text{rango } \Gamma_{zx}^\infty(n_b), \text{ con } d_f = 2n_b$$

donde el cálculo de n_b se muestra en la figura 3.1 y el cálculo de n_a se muestra en la figura 3.2, reemplazando \hat{n}_b por d_f .

En todos los puntos anteriores la estructura del filtro instrumental (m_l, m_k) debe cumplir

$$\text{máx}(m_l - n_b, m_k - n_a) \geq 0$$

para así satisfacer el teorema 4.2.

Dado que las combinaciones

$$(d_f, \hat{n}_a) = (n_a, 2n_a) = (n, 2n)$$

$$(d_f, \hat{n}_a) = (2n_b, n_b) = (2n, n)$$

son las más grandes a utilizar en el procedimiento anterior, es suficiente que la región en que se realiza la búsqueda no exhaustiva sea

$$(d_f, \hat{n}_a) \in [n_b, \hat{n}_{\text{máx}}] \times [n_a, \hat{n}_{\text{máx}}]$$

donde $\hat{n}_{\text{máx}}$ debe ser escogido tal que

$$\hat{n}_{\text{máx}} \geq 2n = 2 \text{máx}(n_b, n_a)$$

Nótese además que, al trabajar en esta región, la condición

$$\text{máx}(m_l - n_b, m_k - n_a) \geq 0$$

siempre se cumple, ya que

$$\text{máx}(m_l - n_b, m_k - n_a) \geq \text{máx}(m_l - d_f, m_k - \hat{n}_a)$$

donde $m_l - d_f \geq 0$.

En la práctica, para implementar este método no exhaustivo, los conjuntos de datos Z_{ux}^∞ y Z_{uz}^∞ deben

```

// región de prueba
1  escoja  $\hat{n}_{\text{máx}} \geq 2n = 2 \text{máx}(n_b, n_a)$ ;
2  escoja  $M$  muy grande y  $L \geq \hat{n}_{\text{máx}}$ ;
   // diseño de señal de entrada
3  diseñe  $u(k)$  tal que  $\varphi_{uu}^\infty(\tau) = a_u^2 \delta_K(\tau)$ ;
   // medición
4  mida  $Z_{ux}^{L,M}$  en estado estacionario;
   // identificación
5  escoja  $\hat{n}$  tal que  $n < \hat{n} \leq \hat{n}_{\text{máx}}$ ;
6   $n = \text{rango } \Gamma_{zx}^M(\hat{n})$ , con  $d_f = \hat{n}$ ;
7   $\Gamma_b = \Gamma_{zx}^M(n)$ , con  $d_f = n + 1$ ;
8   $\Gamma_a = \Gamma_{zx}^M(n + 1)$ , con  $d_f = n$ ;
9  if ( $\Gamma_b$  no singular y  $\Gamma_a$  singular) then
10    $n_a = n$ ;
11    $n_b = \text{rango } \Gamma_{zx}^M(2n_a) - n_a$ , con  $d_f = n_a$ ;
12 else if ( $\Gamma_b$  singular y  $\Gamma_a$  no singular) then
13    $n_b = n$ ;
14    $n_a = \text{rango } \Gamma_{zx}^M(n_b)$ , con  $d_f = 2n_b$ ;
15 else if ( $\Gamma_b$  no singular y  $\Gamma_a$  no singular) then
16    $n_b = n$ ;
17    $n_a = n$ ;
18 else
19   Advertencia: Esta situación no puede
   ocurrir. Revisar los rangos de las matrices;
20 end

```

Algoritmo 4.1: Búsqueda del par (n_b, n_a) con el método de perfiles para $\Gamma_{zx}^M(\hat{n}_a)$.

ser aproximados por las mediciones $Z_{ux}^{L,M}$ y $Z_{uz}^{L,M}$, definidas en la ecuación (2.4), por lo que la submatriz debe construirse según

$$\Gamma_{zx}^M(\hat{n}_a) = \frac{1}{2M+1} \sum_{k=-M}^M \begin{bmatrix} z(k-1) \\ \vdots \\ z(k-\hat{n}_a) \end{bmatrix} \begin{bmatrix} x(k-1) \\ \vdots \\ x(k-\hat{n}_a) \end{bmatrix}^T$$

considerando M muy grande y $L \geq \hat{n}_{\text{máx}}$.

De esta forma, el procedimiento no exhaustivo puede resumirse en el algoritmo 4.1, donde el cálculo de los rangos puede abordarse mediante la siguiente estrategia, basada en la obtención de valores singulares

$$\text{rango } \Gamma_{zx}^M = \# \{ \sigma(\Gamma_{zx}^M) \mid \sigma(\Gamma_{zx}^M) > \epsilon_\sigma \}$$

la cual permite descartar los valores singulares de Γ_{zx}^M que son comparativamente despreciables.

Ejemplo 4.1 Considérese la siguiente planta

$$G(q) = \frac{0,3}{q^2(q-0,8e^{j\pi/4})(q-0,8e^{-j\pi/4})(q-0,7)}$$

cuya estructura $(n_b, n_a) = (5, 3)$ se desea identificar.

Los parámetros a utilizar en el algoritmo 4.1 se escogen según

$$\hat{n}_{\text{máx}} = 15, \quad M = \{10^3, 10^5\}, \quad L = \hat{n}_{\text{máx}}$$

$u(k)$ se diseña como una realización de ruido blanco estricto gaussiano de media cero y varianza 1, y el filtro instrumental se diseña de la forma más simple posible

$$F(q) = q^{-d_f}$$

De esta manera, se miden los datos $Z_{ux}^{L,M}$ en estado estacionario, con los cuales se construyen los perfiles de valores singulares para $\Gamma_{zx}^M(\hat{n}_a)$, mostrados en las figuras 4.1 y 4.2. Para verificar el comportamiento de los valores singulares “nulos”, estas figuras muestran 50 realizaciones de los valores singulares para cada matriz.

Siguiendo el algoritmo 4.1, se escoge inicialmente el modelo $(d_f, \hat{n}_a) = (10, 10)$, para el cual se obtiene el orden del sistema

$$n = \text{rango } \Gamma_{zx}^M(10) = 5$$

Una vez que el orden del sistema es conocido $n = 5$, se computan los rangos de las siguientes matrices

$$\begin{aligned} \text{rango } \Gamma_{zx}^M(5) &= \text{deficiente, con } d_f = 6 \\ \text{rango } \Gamma_{zx}^M(6) &= \text{completo, con } d_f = 5 \end{aligned}$$

Lo anterior indica, según el algoritmo 4.1, que el orden del sistema corresponde a n_b , es decir

$$n_b = n = 5$$

Finalmente se obtiene el valor de n_a , mediante la matriz asociada a $(d_f, \hat{n}_a) = (2n_b, n_b)$, teniendo en cuenta que $n_b = n = 5$

$$n_a = \text{rango } \Gamma_{zx}^M(5) = 3$$

con lo que la estructura identificada del sistema resulta $(n_b, n_a) = (5, 3)$.

Para más detalles véase el programa F.5.1.

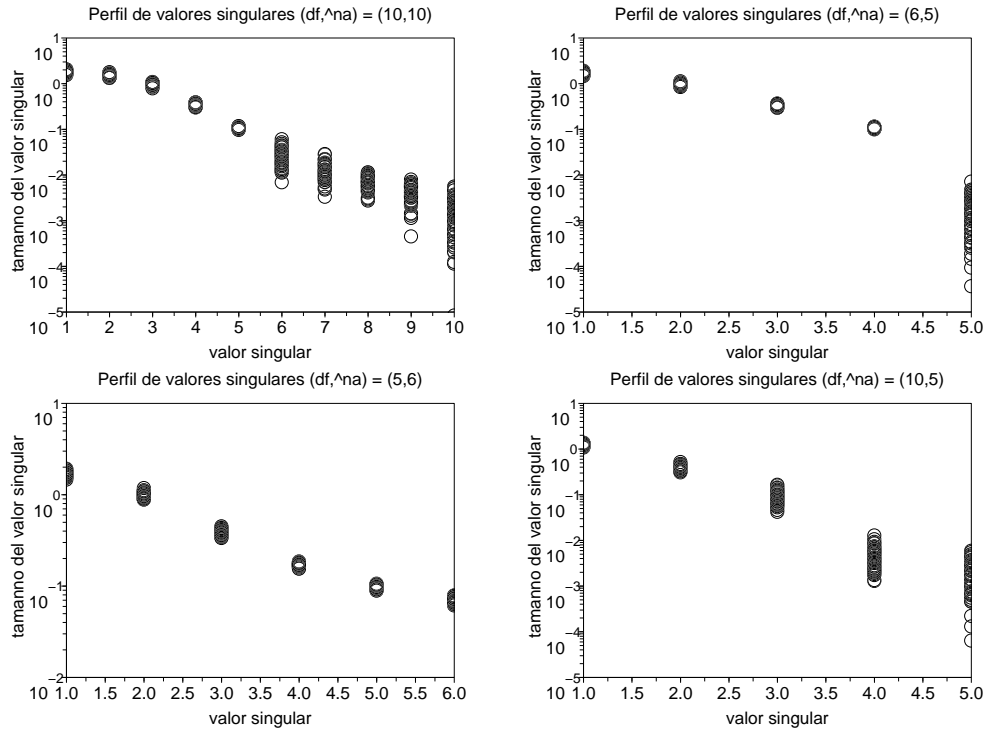


Figura 4.1: Búsqueda de $(n_b, n_a) = (5, 3)$. Cada gráfico muestra 50 realizaciones de los valores singulares de $\Gamma_{zx}^M(\hat{n}_a)$ con $M = 10^3$.

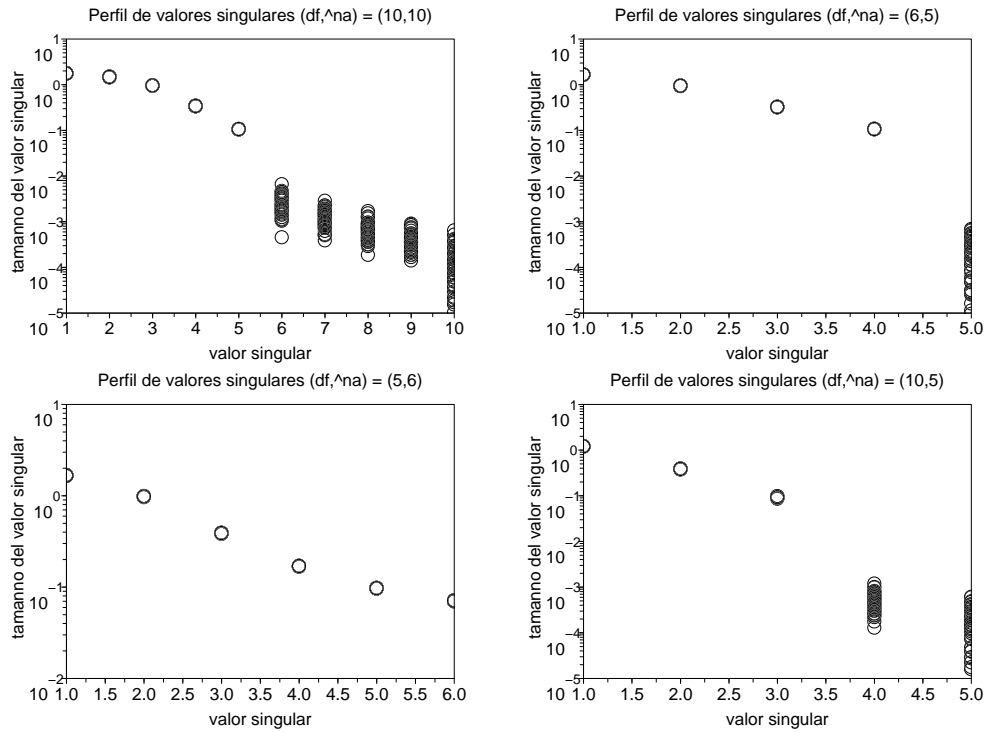


Figura 4.2: Búsqueda de $(n_b, n_a) = (5, 3)$. Cada gráfico muestra 50 realizaciones de los valores singulares de $\Gamma_{zx}^M(\hat{n}_a)$ con $M = 10^5$.

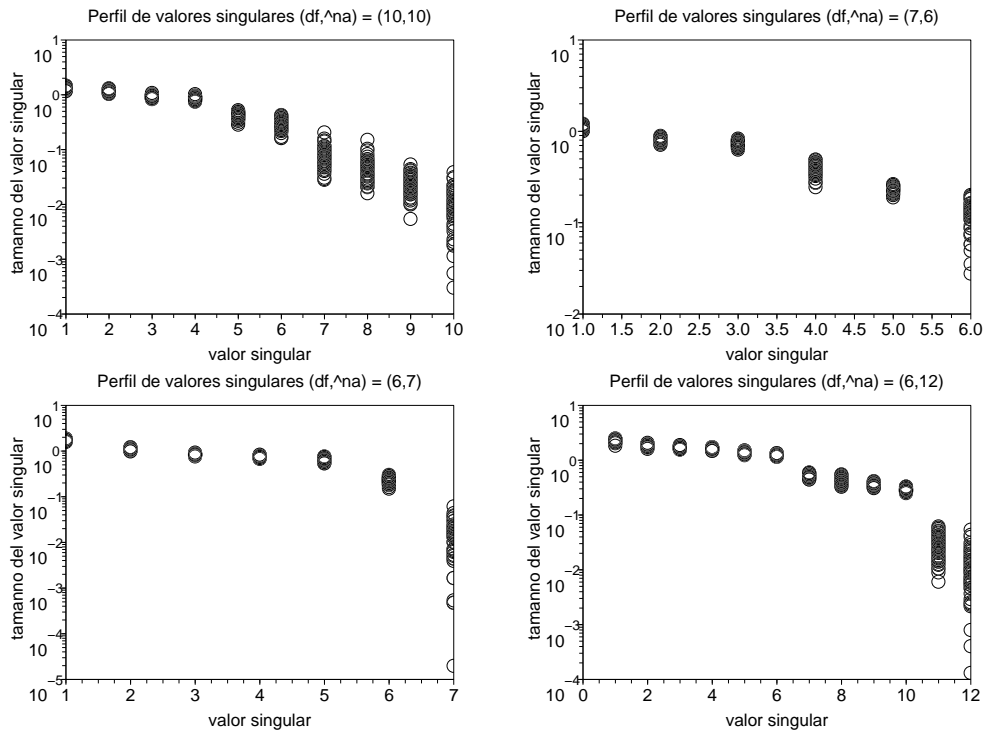


Figura 4.3: Búsqueda de $(n_b, n_a) = (4, 6)$. Cada gráfico muestra 50 realizaciones de los valores singulares de $\Gamma_{zx}^M(\hat{n}_a)$ con $M = 10^3$.

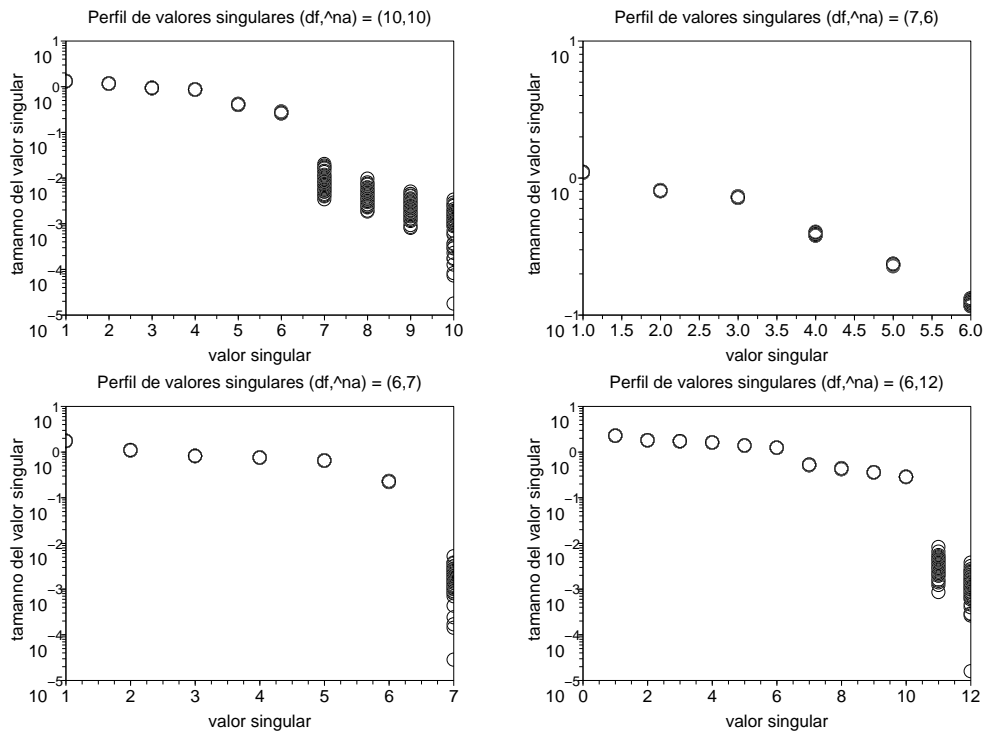


Figura 4.4: Búsqueda de $(n_b, n_a) = (4, 6)$. Cada gráfico muestra 50 realizaciones de los valores singulares de $\Gamma_{zx}^M(\hat{n}_a)$ con $M = 10^5$.

Ejemplo 4.2 *Considérese ahora el mismo sistema que en [13, example 4]*

$$\begin{aligned} G(q) &= \frac{B(q)}{A(q)} \\ B(q) &= 1 + 0,3452q^{-1} + 0,53q^{-2} \\ &\quad + 0,3985q^{-3} + 0,8138q^{-4} \\ A(q) &= 1 + 0,7907q^{-1} + 0,042q^{-2} - 0,5556q^{-3} \\ &\quad - 0,0247q^{-4} + 0,3846q^{-5} + 0,3026q^{-6} \end{aligned}$$

cuya estructura $(n_b, n_a) = (4, 6)$ se desea identificar.

Los parámetros a utilizar en el algoritmo 4.1 se escogen según

$$\hat{n}_{\text{máx}} = 15, \quad M = \{10^3, 10^5\}, \quad L = \hat{n}_{\text{máx}}$$

$u(k)$ se diseña como una realización de ruido blanco estricto uniforme de media cero y varianza 1, y el filtro instrumental se diseña de la forma más simple posible

$$F(q) = q^{-d_f}$$

De esta forma se miden los datos $Z_{ux}^{L,M}$ en estado estacionario, con los cuales se construyen los perfiles de valores singulares para $\Gamma_{zx}^M(\hat{n}_a)$, mostrados en las figuras 4.3 y 4.4. Para verificar el comportamiento de los valores singulares “nulos”, estas figuras muestran 50 realizaciones de los valores singulares para cada matriz.

Siguiendo el algoritmo 4.1, se escoge inicialmente el modelo $(d_f, \hat{n}_a) = (10, 10)$, para el cual se obtiene

$$n = \text{rango } \Gamma_{zx}^M(10) = 6$$

Una vez que el orden del sistema es conocido $n = 6$, se computan los rangos de las siguientes matrices

$$\begin{aligned} \text{rango } \Gamma_{zx}^M(6) &= \text{completo, con } d_f = 7 \\ \text{rango } \Gamma_{zx}^M(7) &= \text{deficiente, con } d_f = 6 \end{aligned}$$

Lo anterior indica, según el algoritmo 4.1, que el orden del sistema corresponde a n_a , es decir

$$n_a = n = 6$$

Finalmente, se obtiene el valor de n_b , mediante la matriz asociada al modelo $(d_f, \hat{n}_a) = (n_a, 2n_a)$, teniendo en cuenta que $n_a = n = 6$. De esta forma, $\text{rango } \Gamma_{zx}^M(12) = 10$, y por lo tanto

$$n_b = \text{rango } \Gamma_{zx}^M(12) - 6 = 4$$

con lo que la estructura identificada del sistema resulta $(n_b, n_a) = (4, 6)$.

Para más detalles véase el programa F.5.1.

Nótese que, al usar un número finito de muestras, se producen errores en Γ_{zx}^M con respecto a la matriz ideal Γ_{zx}^∞ , debido a que el cálculo de las correlaciones es incompleto

$$\Gamma_{zx}^M = \Gamma_{zx}^\infty + E_{zx}^M, \quad \lim_{M \rightarrow \infty} E_{zx}^M = 0$$

Así, los valores singulares nulos de Γ_{zx}^∞ se ven perturbados al utilizar la matriz aproximada Γ_{zx}^M , resultando notablemente mayores que la precisión **eps** del computador. Esto muestra una clara desventaja de Γ_{zx}^M con respecto a la matriz Λ_{zux}^M , ya que el rango de esta última matriz es exacto para un número finito de muestras, según el teorema 2.2, por lo que sus valores singulares “nulos” son del orden de magnitud de **eps**.

4.3. Relación con la matriz de Yule-Walker

A pesar de la desventaja mostrada en la sección anterior, la matriz Γ_{zx}^∞ presenta una notable ventaja sobre las matrices Λ_{zux}^∞ y Λ_{uxx}^∞ , ya que, con una elección adecuada del filtro instrumental $F(q)$, se puede prescindir de la medición de $u(k)$ para la construcción de Γ_{zx}^∞ , resultando así particularmente útil para la identificación estructural de sistemas ARMA.

Supóngase, por ejemplo, que sólo se tienen mediciones de la salida $x(k)$ del sistema en (2.19)

$$Z_x^{L,M} = [x(-M-L) \quad \cdots \quad x(M)] \quad (4.1)$$

Supóngase además que $u(k)$ satisface $\varphi_{uu}^\infty(\tau) = a_u^2 \delta_K(\tau)$ y que el filtro instrumental se diseña tal que

$$F(q) = q^{-\hat{d}_f} G(q)$$

Es decir, $z(k) = x(k - \hat{d}_f)$, donde

$$G(q) = \frac{b_{d_g} q^{-d_g} + \cdots + b_{n_b} q^{-n_b}}{1 + a_1 q^{-1} + \cdots + a_{n_a} q^{-n_a}}$$

Así el retardo real del filtro instrumental sería

$$d_f = \hat{d}_f + d_g$$

Al aplicar estas condiciones al teorema 4.2, se puede obtener el siguiente corolario

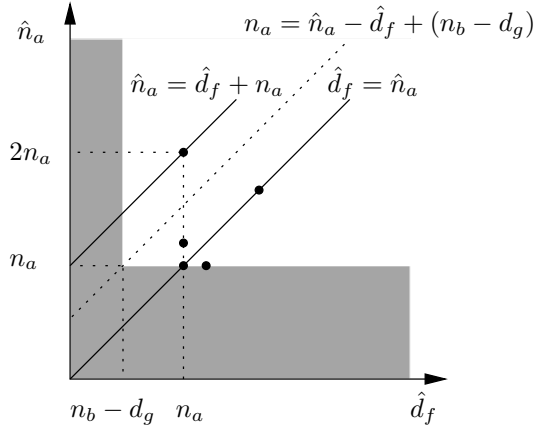


Figura 4.5: Búsqueda de $(n_b - d_g, n_a)$ en el caso $n_b - d_g < n_a$. En gris: región donde Γ_{zx}^∞ es de rango completo. En blanco: región donde Γ_{zx}^∞ es deficiente de rango.

Corolario 4.1 *Considérese los datos Z_x^∞ medidos de (2.19). Si la señal $u(k)$ es tal que $\varphi_{uu}^\infty(\tau) = a_u^2 \delta_K(\tau)$ y se utiliza la secuencia instrumental $z(k) = x(k - \hat{d}_f)$, entonces*

$$\text{rango } \Gamma_{zx}^\infty(\hat{n}_a) = \begin{cases} \hat{n}_a, & i \leq 0 \\ \max(\hat{n}_a - \hat{d}_f - d_g + n_b, n_a), & i > 0 \end{cases}$$

donde $i = \min(\hat{d}_f + d_g - n_b, \hat{n}_a - n_a)$.

Nótese que la condición $\max(m_l - n_b, m_k - n_a) \geq 0$, establecida en el teorema 4.2, siempre se satisface con las condiciones de este corolario, ya que la estructura del filtro instrumental es $(m_l, m_k) = (\hat{d}_f + n_b, n_a)$, donde $\hat{d}_f \geq 0$.

Las correlaciones que componen la matriz

$$\Gamma_{zx}^\infty(\hat{n}_a) = \begin{bmatrix} \varphi_{zx}^\infty(0) & \cdots & \varphi_{zx}^\infty(\hat{n}_a - 1) \\ \vdots & \ddots & \vdots \\ \varphi_{zx}^\infty(1 - \hat{n}_a) & \cdots & \varphi_{zx}^\infty(0) \end{bmatrix}$$

considerando las condiciones del corolario anterior, se ven reducidas a

$$\varphi_{zx}^\infty(\tau) = \varphi_{xx}^\infty(\tau - \hat{d}_f) = \varphi_{xx}^\infty(\hat{d}_f - \tau)$$

por lo que la matriz $\Gamma_{zx}^\infty(\hat{n}_a)$ corresponde a la matriz de Yule-Walker [20, ecuación (C8.1.4)]

$$\Gamma_{zx}^\infty(\hat{n}_a) = \begin{bmatrix} \varphi_{xx}^\infty(\hat{d}_f) & \cdots & \varphi_{xx}^\infty(\hat{d}_f - \hat{n}_a + 1) \\ \vdots & \ddots & \vdots \\ \varphi_{xx}^\infty(\hat{d}_f + \hat{n}_a - 1) & \cdots & \varphi_{xx}^\infty(\hat{d}_f) \end{bmatrix}$$

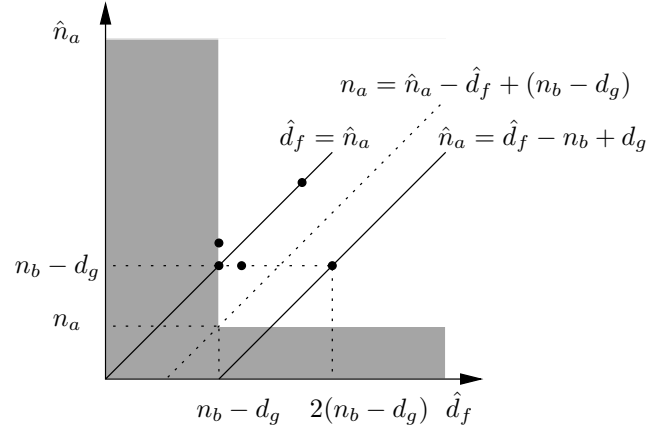


Figura 4.6: Búsqueda de $(n_b - d_g, n_a)$ en el caso $n_b - d_g > n_a$. En gris: región donde Γ_{zx}^∞ es de rango completo. En blanco: región donde Γ_{zx}^∞ es deficiente de rango.

En [21] y [9] era conocido que la matriz de Yule-Walker cumple

$$\det \Gamma_{zx}^\infty(\hat{n}_a) \begin{cases} \neq 0, & i \leq 0 \\ = 0, & i > 0 \end{cases}$$

con $i = \min(\hat{d}_f - n_b, \hat{n}_a - n_a)$ y $d_g = 0$. Además en [4] se conocía el rango de esta matriz para algunos casos

$$\text{rango } \Gamma_{zx}^\infty(\hat{n}_a) = n_a, \quad \text{si } \hat{d}_f - n_b \geq \hat{n}_a - n_a \geq 0$$

al igual que en [25]

$$\text{rango } \Gamma_{zx}^\infty(n_a + 1) = \begin{cases} n_a + 1, & \text{si } \hat{d}_f = n_b \\ n_a, & \text{si } \hat{d}_f > n_b \end{cases}$$

ambos resultados considerando $d_g = 0$. Sin embargo, el corolario 4.1 aporta nueva información, indicando el rango de esta matriz para cualquier valor de \hat{d}_f , \hat{n}_a y d_g , el que puede ser usado para implementar métodos no exhaustivos, en forma similar al algoritmo 4.1.

El siguiente método no exhaustivo, basado en el resultado del corolario 4.1, es capaz de identificar el par $(n_b - d_g, n_a)$ y no directamente la estructura de la planta (n_b, n_a) . Esto se debe a que el retardo real del filtro instrumental $d_f = \hat{d}_f + d_g$ es desconocido, pues el retardo de la planta d_g no se conoce, pudiéndose manejar solamente \hat{d}_f .

1. Determinar $n_o = \max(n_b - d_g, n_a)$ a partir de

$$n_o = \text{rango } \Gamma_{zx}^\infty(\hat{n}), \quad \text{con } \hat{d}_f = \hat{n}$$

donde $\hat{n} > n_o$.

2. Determinar si n_o corresponde a $(n_b - d_g)$ o n_a , mediante la verificación de la singularidad de las siguientes matrices

$$\begin{aligned}\Gamma_{zx}^\infty(n_o) &= \text{¿singular?}, \text{ con } \hat{d}_f = n_o + 1 \\ \Gamma_{zx}^\infty(n_o + 1) &= \text{¿singular?}, \text{ con } \hat{d}_f = n_o\end{aligned}$$

De esta manera, si

- a) $\Gamma_{zx}^\infty(n_o)$ es no singular y $\Gamma_{zx}^\infty(n_o + 1)$ es singular, entonces $n_a = n_o$
b) $\Gamma_{zx}^\infty(n_o)$ es singular y $\Gamma_{zx}^\infty(n_o + 1)$ es no singular, entonces $(n_b - d_g) = n_o$
c) $\Gamma_{zx}^\infty(n_o)$ y $\Gamma_{zx}^\infty(n_o + 1)$ son no singulares, entonces $(n_b - d_g) = n_a = n_o$
3. Dependiendo del resultado del punto anterior, calcular el parámetro restante según

$$\begin{aligned}(n_b - d_g) &= \text{rango } \Gamma_{zx}^\infty(2n_a) - n_a, \\ &\quad \text{con } \hat{d}_f = n_a \\ n_a &= \text{rango } \Gamma_{zx}^\infty(n_b - d_g), \\ &\quad \text{con } \hat{d}_f = 2(n_b - d_g)\end{aligned}$$

donde el cálculo de $(n_b - d_g)$ se muestra en la figura 4.5 y el cálculo de n_a se muestra en la figura 4.6.

Dado que las combinaciones más grandes a utilizar en este método son

$$\begin{aligned}(\hat{d}_f, \hat{n}_a) &= (n_a, 2n_a) = (n_o, 2n_o) \\ (\hat{d}_f, \hat{n}_a) &= (2[n_b - d_g], n_b - d_g) = (2n_o, n_o)\end{aligned}$$

es suficiente que la región de análisis sea

$$(\hat{d}_f, \hat{n}_a) \in [n_b - d_g, \hat{n}_{\text{máx}}] \times [n_a, \hat{n}_{\text{máx}}]$$

donde $\hat{n}_{\text{máx}}$ debe escogerse tal que

$$\hat{n}_{\text{máx}} \geq 2n_o = 2 \text{máx}(n_b - d_g, n_a)$$

Finalmente, dado que en la práctica no se puede medir Z_x^∞ , entonces debe aproximarse como $Z_x^{L,M}$ con M muy grande y $L \geq 2\hat{n}_{\text{máx}}$ de forma que la matriz de Yule-Walker pueda construirse para cualquier punto de la región de análisis. Así, el método y las condiciones mencionadas anteriormente pueden resumirse en el algoritmo 4.2.

Los rangos del algoritmo 4.2 pueden obtenerse a través de valores singulares

$$\text{rango } \Gamma_{zx}^M = \# \{ \sigma(\Gamma_{zx}^M) \mid \sigma(\Gamma_{zx}^M) > \epsilon_\sigma \}$$

```
// región de prueba
1  escoja  $\hat{n}_{\text{máx}} \geq 2n_o = 2 \text{máx}(n_b - d_g, n_a)$ ;
2  escoja  $M$  muy grande y  $L \geq 2\hat{n}_{\text{máx}}$ ;
// medición
3  mida  $Z_x^{L,M}$  en estado estacionario;
// identificación
4  escoja  $\hat{n}$  tal que  $n_o < \hat{n} \leq \hat{n}_{\text{máx}}$ ;
5   $n_o = \text{rango } \Gamma_{zx}^M(\hat{n})$ , con  $\hat{d}_f = \hat{n}$ ;
6   $\Gamma_b = \Gamma_{zx}^M(n_o)$ , con  $\hat{d}_f = n_o + 1$ ;
7   $\Gamma_a = \Gamma_{zx}^M(n_o + 1)$ , con  $\hat{d}_f = n_o$ ;
8  if ( $\Gamma_b$  no singular y  $\Gamma_a$  singular) then
9     $n_a = n_o$ ;
10    $(n_b - d_g) = \text{rango } \Gamma_{zx}^M(2n_a) - n_a$ , con
       $\hat{d}_f = n_a$ ;
11 else if ( $\Gamma_b$  singular y  $\Gamma_a$  no singular) then
12    $(n_b - d_g) = n_o$ ;
13    $n_a = \text{rango } \Gamma_{zx}^M(n_b - d_g)$ , con
       $\hat{d}_f = 2(n_b - d_g)$ ;
14 else if ( $\Gamma_b$  no singular y  $\Gamma_a$  no singular) then
15    $(n_b - d_g) = n_o$ ;
16    $n_a = n_o$ ;
17 else
18   Advertencia: Esta situación no puede
      ocurrir. Revisar los rangos de las matrices;
19 end
```

Algoritmo 4.2: Búsqueda del par $(n_b - d_g, n_a)$ con el método de perfiles para $\Gamma_{zx}^M(\hat{n}_a)$ en el caso ARMA.

lo que permite descartar los valores singulares de Γ_{zx}^M que son comparativamente despreciables. En [9] se propone un criterio para determinación de rangos, basados en valores propios

$$\text{rango } \Gamma_{zx}^M = \# \{ \lambda(\Gamma_{zx}^M) \mid |\lambda(\Gamma_{zx}^M)| > \epsilon_\lambda \}$$

Este criterio permite descartar valores propios de matrices perturbadas que poseen sólo un valor propio “nulo”, y en él se considera que $u(k)$ tiene distribución gaussiana.

Para ilustrar el algoritmo 4.2, se computarán los rangos de las matrices a través de valores singulares, utilizando las mismas plantas de los ejemplos 4.1 y 4.2, pero ahora considerando que no se puede medir $u(k)$.

Ejemplo 4.3 Considérese la planta $x(k) = G(q)u(k)$,

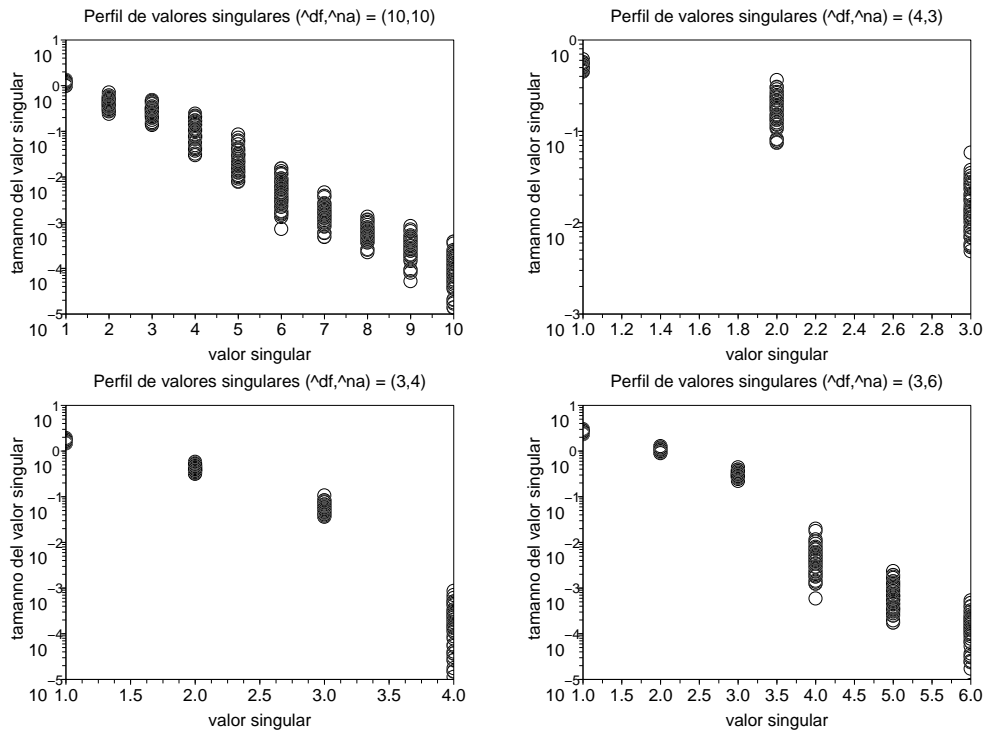


Figura 4.7: Búsqueda de $(n_b - d_g, n_a) = (0, 3)$ caso ARMA. Cada gráfico muestra 50 realizaciones de los valores singulares de $\Gamma_{zx}^M(\hat{n}_a)$ con $M = 10^3$.

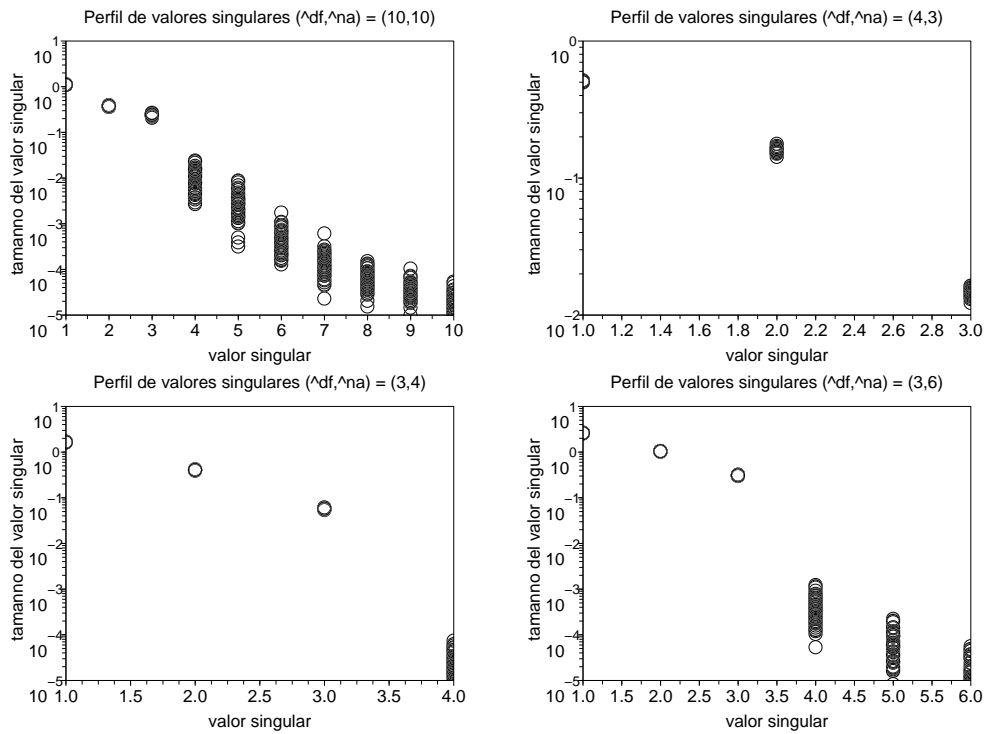


Figura 4.8: Búsqueda de $(n_b - d_g, n_a) = (0, 3)$ caso ARMA. Cada gráfico muestra 50 realizaciones de los valores singulares de $\Gamma_{zx}^M(\hat{n}_a)$ con $M = 10^5$.

cuya estructura es $(n_b, n_a) = (5, 3)$ con $d_g = 5$

$$G(q) = \frac{0,3}{q^2(q-0,8e^{j\pi/4})(q-0,8e^{-j\pi/4})(q-0,7)}$$

donde $u(k)$ es ruido blanco estricto gaussiano de media cero y varianza 1.

Los parámetros a utilizar en el algoritmo 4.2 se escogen según

$$\hat{n}_{\text{máx}} = 15, \quad M = \{10^3, 10^5\}, \quad L = 2\hat{n}_{\text{máx}}$$

De esta manera, se miden los datos $Z_x^{L,M}$ en estado estacionario, con los cuales se construyen los perfiles de valores singulares para $\Gamma_{zx}^M(\hat{n}_a)$, con $z(k) = x(k - \hat{d}_f)$, mostrados en las figuras 4.7 y 4.8. Para verificar el comportamiento de los valores singulares “nulos”, estas figuras muestran 50 realizaciones de los valores singulares para cada matriz.

Siguiendo el algoritmo 4.2, se escoge inicialmente $(\hat{d}_f, \hat{n}_a) = (10, 10)$, para el cual se obtiene

$$n_o = \text{rango } \Gamma_{zx}^M(10) = 3$$

Una vez que $n_o = 3$ es conocido, se computan los rangos de las siguientes matrices

$$\begin{aligned} \text{rango } \Gamma_{zx}^M(3) &= \text{completo, con } \hat{d}_f = 4 \\ \text{rango } \Gamma_{zx}^M(4) &= \text{deficiente, con } \hat{d}_f = 3 \end{aligned}$$

Esto indica, según el algoritmo 4.2, que

$$n_a = n_o = 3$$

Finalmente, se obtiene el valor de $(n_b - d_g)$, mediante la matriz asociada a $(\hat{d}_f, \hat{n}_a) = (n_a, 2n_a)$, teniendo en cuenta que $n_a = n_o = 3$, con lo que se obtiene $\text{rango } \Gamma_{zx}^M(6) = 3$, y por lo tanto

$$n_b - d_g = \text{rango } \Gamma_{zx}^M(6) - 3 = 0$$

con lo que la estructura identificada del sistema resulta $(n_b - d_g, n_a) = (0, 3)$.

Para más detalles véase el programa F.5.1.

Ejemplo 4.4 Considérese la planta $x(k) = G(q)u(k)$, cuya estructura es $(n_b, n_a) = (4, 6)$ con $d_g = 0$

$$\begin{aligned} G(q) &= \frac{B(q)}{A(q)} \\ B(q) &= 1 + 0,3452q^{-1} + 0,53q^{-2} \\ &\quad + 0,3985q^{-3} + 0,8138q^{-4} \\ A(q) &= 1 + 0,7907q^{-1} + 0,042q^{-2} - 0,5556q^{-3} \\ &\quad - 0,0247q^{-4} + 0,3846q^{-5} + 0,3026q^{-6} \end{aligned}$$

donde $u(k)$ es ruido blanco estricto uniforme de media cero y varianza 1.

Los parámetros a utilizar en el algoritmo 4.2 se escogen según

$$\hat{n}_{\text{máx}} = 15, \quad M = \{10^3, 10^5\}, \quad L = 2\hat{n}_{\text{máx}}$$

De esta forma, se miden los datos $Z_x^{L,M}$ en estado estacionario, con los que se construyen los perfiles de valores singulares para $\Gamma_{zx}^M(\hat{n}_a)$, con $z(k) = x(k - \hat{d}_f)$, mostrados en las figuras 4.9 y 4.10. Para verificar el comportamiento de los valores singulares “nulos”, estas figuras muestran 50 realizaciones de los valores singulares para cada matriz.

Siguiendo el algoritmo 4.2, se escoge inicialmente $(\hat{d}_f, \hat{n}_a) = (10, 10)$, para el cual se obtiene

$$n_o = \text{rango } \Gamma_{zx}^M(10) = 6$$

Una vez que $n_o = 6$ es conocido, se computan los rangos de las siguientes matrices

$$\begin{aligned} \text{rango } \Gamma_{zx}^M(6) &= \text{completo, con } \hat{d}_f = 7 \\ \text{rango } \Gamma_{zx}^M(7) &= \text{deficiente, con } \hat{d}_f = 6 \end{aligned}$$

Esto indica, según el algoritmo 4.2, que

$$n_a = n_o = 6$$

Finalmente se obtiene el valor de $(n_b - d_g)$, mediante la matriz asociada a $(\hat{d}_f, \hat{n}_a) = (n_a, 2n_a)$, teniendo en cuenta que $n_a = n_o = 6$, con lo que se obtiene $\text{rango } \Gamma_{zx}^M(12) = 10$, y por lo tanto

$$n_b - d_g = \text{rango } \Gamma_{zx}^M(12) - 6 = 4$$

con lo que la estructura identificada del sistema resulta $(n_b - d_g, n_a) = (4, 6)$.

Para más detalles véase el programa F.5.1.

Con estos dos ejemplos se puede evidenciar que la matriz de Yule-Walker, utilizada para identificar estructura de sistemas ARMA, es “ciega” al retardo de la planta d_g y al orden del numerador n_b , entregando así sólo información acerca de su diferencia, es decir, del par $(n_b - d_g, n_a)$. Sin embargo, este resultado coincide con la estructura real del sistema si el retardo de la planta es $d_g = 0$, es decir, cuando la planta es bipropia.

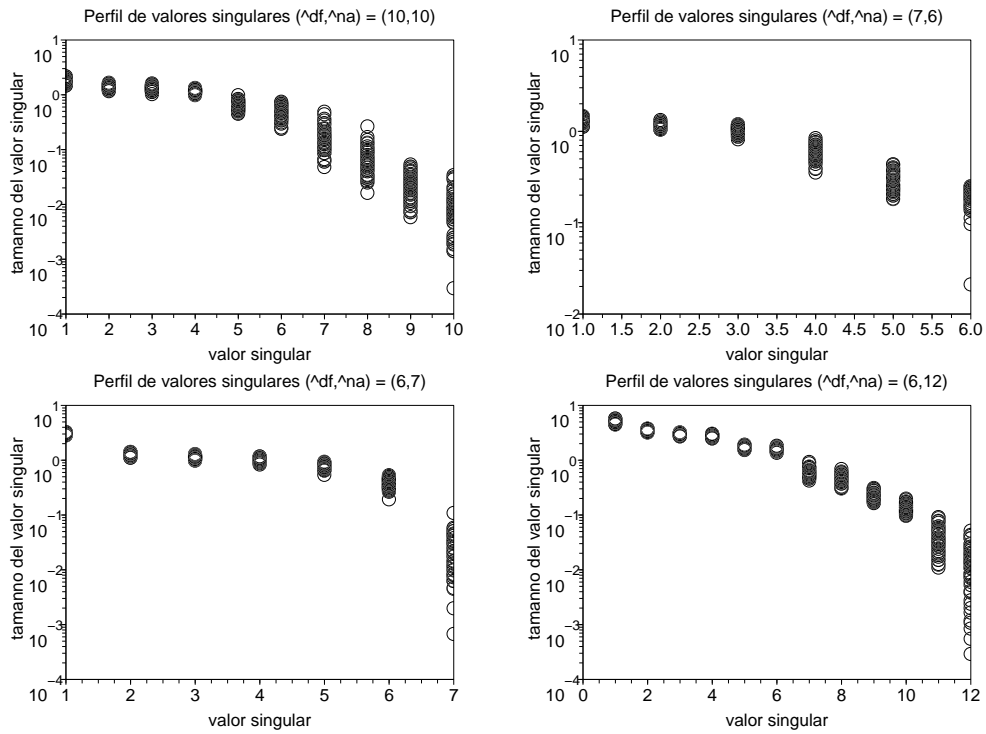


Figura 4.9: Búsqueda de $(n_b - d_g, n_a) = (4, 6)$ caso ARMA. Cada gráfico muestra 50 realizaciones de los valores singulares de $\Gamma_{zx}^M(\hat{n}_a)$ con $M = 10^3$.

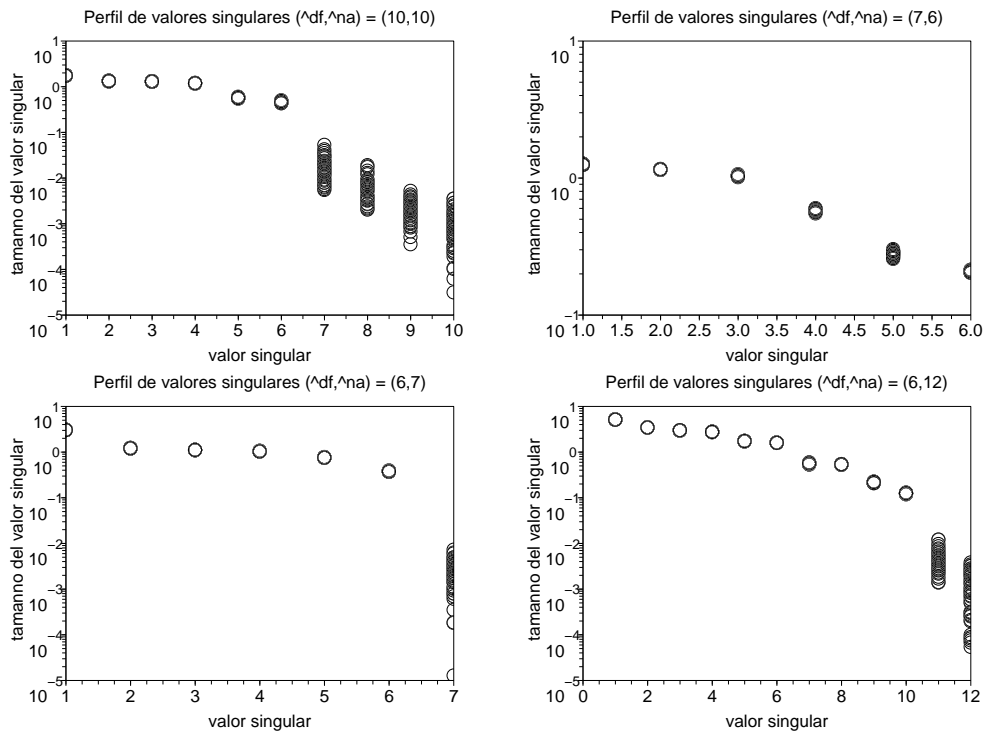


Figura 4.10: Búsqueda de $(n_b - d_g, n_a) = (4, 6)$ caso ARMA. Cada gráfico muestra 50 realizaciones de los valores singulares de $\Gamma_{zx}^M(\hat{n}_a)$ con $M = 10^5$.

4.4. Relación con la matriz de Hankel del sistema

Como se mencionó, y con las condiciones del teorema 4.2, la matriz $\Gamma_{zx}^\infty(\hat{n}_a)$ es suficiente para estimar la estructura la planta, y en particular su orden, tal como lo muestra el siguiente corolario del teorema 4.2

Corolario 4.2 *Considérese los datos Z_{ux}^∞ y Z_{uz}^∞ medidos de (2.19) y (2.25) respectivamente. Si $u(k)$ es tal que $\varphi_{uu}^\infty(\tau) = a_u^2 \delta_K(\tau)$, y el filtro instrumental*

$$F(q) = \frac{l_{\hat{n}}q^{-\hat{n}} + \dots + l_{m_l}q^{-m_l}}{1 + k_1q^{-1} + \dots + k_{m_k}q^{-m_k}}$$

satisface $\max(m_l - n_b, m_k - n_a) \geq 0$, entonces

$$\text{rango } \Gamma_{zx}^\infty(\hat{n}) = \begin{cases} \hat{n}, & \hat{n} \leq n \\ n, & \hat{n} > n \end{cases}$$

donde $n := \max(n_b, n_a)$.

La identificación del orden del sistema corresponde, de hecho, al primer paso de identificación a realizar en el algoritmo 4.1.

Como ya es sabido [12], [10, sección 2.3], la matriz de Hankel $\mathcal{H}(\hat{n})$ asociada al sistema en (2.19) es

$$\mathcal{H}(\hat{n}) = \begin{bmatrix} g(1) & g(2) & \dots & g(\hat{n}) \\ g(2) & g(3) & \dots & g(\hat{n} + 1) \\ \vdots & \vdots & & \vdots \\ g(\hat{n}) & g(\hat{n} + 1) & \dots & g(2\hat{n} - 1) \end{bmatrix}$$

donde $g(k)$ es la respuesta a impulso del sistema, y cuyo rango está dado por

$$\text{rango } \mathcal{H}(\hat{n}) = \begin{cases} \hat{n}, & \hat{n} \leq n \\ n, & \hat{n} > n \end{cases}$$

Esto es válido cuando el sistema es de realización mínima, es decir, cuando es completamente controlable y completamente observable, y coincide con el rango obtenido en el corolario 4.2.

Naturalmente uno puede preguntarse si existe alguna relación entre la matriz $\mathcal{H}(\hat{n})$ y la matriz $\Gamma_{zx}^\infty(\hat{n})$, o preguntarse si comparten alguna otra propiedad además de tener rangos iguales y bajo qué condiciones. El siguiente teorema muestra la existencia de tales relaciones

Teorema 4.3 *Sean Z_{ux}^∞ and Z_{uz}^∞ medidos de (2.19) y (2.25) respectivamente. Si la señal $u(k)$ es tal que $\varphi_{uu}^\infty(\tau) = a_u^2 \delta_K(\tau)$, y el filtro instrumental es diseñado como $F(q) = q^{-\hat{n}}$, entonces $\Gamma_{zx}^\infty(\hat{n})$ tiene la siguiente forma*

$$\Gamma_{zx}^\infty(\hat{n}) = a_u^2 \begin{bmatrix} g(\hat{n}) & \dots & g(2) & g(1) \\ g(1 + \hat{n}) & \dots & g(3) & g(2) \\ \vdots & & \vdots & \vdots \\ g(2\hat{n} - 1) & \dots & g(1 + \hat{n}) & g(\hat{n}) \end{bmatrix}$$

es decir

$$\Gamma_{zx}^\infty(\hat{n}) = a_u^2 \mathcal{H}(\hat{n})J$$

donde

$$J = \begin{bmatrix} 0 & \dots & 0 & 1 \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ddots & \ddots & \vdots \\ 1 & 0 & \dots & 0 \end{bmatrix}$$

Además, si $\mathcal{H}(\hat{n})$ tiene la siguiente descomposición en valores singulares

$$\mathcal{H}(\hat{n}) = U\Sigma V^T$$

entonces la descomposición en valores singulares para $\Gamma_{zx}^\infty(\hat{n})$ es

$$\Gamma_{zx}^\infty(\hat{n}) = U(a_u^2 \Sigma)(V^T J)$$

y así, los valores singulares de $\Gamma_{zx}^\infty(\hat{n})$ son proporcionales a los valores singulares de Hankel del sistema

$$\sigma(\Gamma_{zx}^\infty) = a_u^2 \cdot \sigma(\mathcal{H})$$

Demostración. Dado que cada elemento de $\Gamma_{zx}^\infty(\hat{n})$

$$\Gamma_{zx}^\infty(\hat{n}) = \begin{bmatrix} \varphi_{zx}^\infty(0) & \dots & \varphi_{zx}^\infty(\hat{n} - 1) \\ \vdots & \ddots & \vdots \\ \varphi_{zx}^\infty(1 - \hat{n}) & \dots & \varphi_{zx}^\infty(0) \end{bmatrix}$$

tiene la forma $\varphi_{zx}^\infty(\tau)$, y que $\varphi_{uu}^\infty(\tau) = a_u^2 \delta_K(\tau)$, entonces

$$\varphi_{zx}^\infty = f * \mathcal{J}g * \varphi_{uu}^\infty = a_u^2 \cdot (q^{-\hat{n}} \delta_K) * \mathcal{J}g$$

es decir

$$\varphi_{zx}^\infty(\tau) = a_u^2 (\mathcal{J}g)(\tau - \hat{n}) = a_u^2 g(\hat{n} - \tau)$$

por lo que $\Gamma_{zx}^\infty(\hat{n}) = a_u^2 \mathcal{H}(\hat{n})J$. El resto de la demostración sigue directamente de la estructura de $\Gamma_{zx}^\infty(\hat{n})$, la que es similar a la matriz de Hankel del sistema, salvo que sus columnas están escritas de derecha a

	$M = 10^2$	$M = 10^3$	$M = 10^4$	$M = 10^5$
σ_1	13,7730	9,4405	10,1963	9,9396
σ_2	7,6742	7,2300	7,5370	7,2629
σ_3	3,2199	2,0115	1,8338	1,7908
σ_4	0,7231	0,1148	0,0904	0,0116
σ_5	0,3509	0,0901	0,0345	0,0107
σ_6	0,2200	0,0615	0,0186	0,0075
σ_7	0,1507	0,0556	0,0114	0,0034
σ_8	0,0312	0,0063	0,0060	0,0014
σ_9	0,0134	0,0054	0,0023	0,0009
σ_{10}	0,0042	0,0015	0,0005	0,0008

Tabla 4.1: Valores singulares de $\Gamma_{zx}^M(\hat{n})$ con $\hat{n} = 10$, para planta de orden 3.

	$M = 10^2$	$M = 10^3$	$M = 10^4$	$M = 10^5$
σ_1	1,9020	1,3534	1,2777	1,3110
σ_2	1,3728	1,1248	1,2013	1,1574
σ_3	1,3157	0,9916	0,9237	0,9270
σ_4	1,2793	0,9165	0,8585	0,8561
σ_5	0,9169	0,4043	0,4316	0,4020
σ_6	0,5090	0,3302	0,2714	0,2642
σ_7	0,3946	0,0813	0,0399	0,0136
σ_8	0,1016	0,0564	0,0090	0,0034
σ_9	0,0908	0,0142	0,0089	0,0050
σ_{10}	0,0330	0,0098	0,0050	0,0022

Tabla 4.2: Valores singulares de $\Gamma_{zx}^M(\hat{n})$ con $\hat{n} = 10$, para planta de orden 6.

izquierda, lo que significa que $\Gamma_{zx}^\infty(\hat{n})$ sigue un patrón Toeplitz. **QED.**

Este último teorema es bastante importante, ya que mediante la submatriz $\Gamma_{zx}^\infty(\hat{n})$ es posible identificar la matriz de Hankel del sistema $\mathcal{H}(\hat{n})$, y así conectar la submatriz $\Gamma_{zx}^\infty(\hat{n})$ con una medida de controlabilidad-observabilidad, provista por los valores singulares de Hankel. Estos valores singulares de Hankel son además bastante utilizados para reducción de modelos, lo que confirma el hecho de que $\Gamma_{zx}^\infty(\hat{n})$ puede ser usada, con antelación, para estimar el orden de un sistema.

Ejemplo 4.5 *Considérese el mismo sistema que en [2, example 3]*

$$G(q) = \frac{1 - 0,45q^{-1} - q^{-2}}{1 - 2,2q^{-1} + 1,77q^{-2} - 0,52q^{-3}}$$

Para este sistema se utiliza $u(k)$ como ruido blanco estricto gaussiano de media cero y varianza unitaria, y además se utiliza el filtro instrumental $F(q) = q^{-\hat{n}}$, para de satisfacer las condiciones del teorema 4.3.

Con estas condiciones se utilizan los conjuntos de datos $Z_{ux}^{L,M}$ y $Z_{uz}^{L,M}$ en estado estacionario, considerando

$$L = \hat{n}, \quad M = \{10^2, 10^3, 10^4, 10^5\}$$

para así aproximar los conjuntos ideales Z_{ux}^∞ y Z_{uz}^∞ . De esta manera se construye la matriz $\Gamma_{zx}^M(\hat{n})$ con $\hat{n} = 10$, cuyos valores singulares se muestran en la tabla 4.1, donde se puede ver que la precisión de los valores singulares aumenta cuando M aumenta. Con el fin de poder realizar comparaciones, se muestran

los valores singulares no nulos de $\mathcal{H}(\hat{n})$

$$\sigma_1(\mathcal{H}) = 9,9035; \quad \sigma_2(\mathcal{H}) = 7,2875; \quad \sigma_3(\mathcal{H}) = 1,8156$$

Para más detalles ver el programa F.5.2.

Ejemplo 4.6 *Considérese el mismo sistema que en [13, example 4]*

$$\begin{aligned} G(q) &= \frac{B(q)}{A(q)} \\ B(q) &= 1 + 0,3452q^{-1} + 0,53q^{-2} \\ &\quad + 0,3985q^{-3} + 0,8138q^{-4} \\ A(q) &= 1 + 0,7907q^{-1} + 0,042q^{-2} - 0,5556q^{-3} \\ &\quad - 0,0247q^{-4} + 0,3846q^{-5} + 0,3026q^{-6} \end{aligned}$$

Para este sistema se utiliza $u(k)$ como ruido blanco estricto gaussiano de media cero y varianza unitaria, y además se utiliza el filtro instrumental $F(q) = q^{-\hat{n}}$, para satisfacer las condiciones del teorema 4.3.

Con estas condiciones se utilizan los conjuntos de datos $Z_{ux}^{L,M}$ y $Z_{uz}^{L,M}$ en estado estacionario, considerando

$$L = \hat{n}, \quad M = \{10^2, 10^3, 10^4, 10^5\}$$

para así aproximar los conjuntos ideales Z_{ux}^∞ y Z_{uz}^∞ . De esta manera se construye la matriz $\Gamma_{zx}^M(\hat{n})$ con $\hat{n} = 10$, cuyos valores singulares se muestran en la tabla 4.2, donde se puede ver que la precisión de los valores singulares aumenta cuando M aumenta. Con el fin de poder realizar comparaciones, se muestran los valores singulares no nulos de $\mathcal{H}(\hat{n})$

$$\begin{aligned} \sigma_1(\mathcal{H}) &= 1,3158 & \sigma_2(\mathcal{H}) &= 1,1590 & \sigma_3(\mathcal{H}) &= 0,9309 \\ \sigma_4(\mathcal{H}) &= 0,8598 & \sigma_5(\mathcal{H}) &= 0,4061 & \sigma_6(\mathcal{H}) &= 0,2709 \end{aligned}$$

	$M = 10^2$	$M = 10^3$	$M = 10^4$	$M = 10^5$
σ_1	1,1585	1,0953	1,3924	1,3141
σ_2	0,9384	1,0908	1,3219	1,2711
σ_3	0,6385	0,7831	0,7961	0,7938
σ_4	0,5760	0,7576	0,7954	0,7796
σ_5	0,5445	0,5959	0,6237	0,6518
σ_6	0,4347	0,4387	0,4582	0,4541
σ_7	0,3410	0,3018	0,3398	0,3518
σ_8	0,1732	0,0704	0,0147	0,0036
σ_9	0,0627	0,0694	0,0134	0,0035
σ_{10}	0,0249	0,0077	0,0064	0,0028

Tabla 4.3: Valores singulares de $\Gamma_{zx}^M(\hat{n})$ con $\hat{n} = 10$, para planta de orden 7.

Para más detalles ver el programa F.5.2.

Ejemplo 4.7 *Considérese el mismo sistema que en [1, example 1]*

$$\begin{aligned}
 G(q) &= \frac{B(q)}{A(q)} \\
 B(q) &= 1 - 0,1292q^{-1} - 0,3893q^{-2} \\
 &\quad + 0,3163q^{-3} + 0,2549q^{-4} - 0,5823q^{-5} \\
 A(q) &= 1 - 0,2459q^{-1} \\
 &\quad - 0,3786q^{-2} - 0,0707q^{-3} - 0,0944q^{-4} \\
 &\quad - 0,2881q^{-5} + 0,107q^{-6} + 0,4695q^{-7}
 \end{aligned}$$

Para este sistema se utiliza $u(k)$ como ruido blanco estricto gaussiano de media cero y varianza unitaria, y además se utiliza el filtro instrumental $F(q) = q^{-\hat{n}}$, para satisfacer las condiciones del teorema 4.3.

Con estas condiciones se utilizan los conjuntos de datos $Z_{ux}^{L,M}$ y $Z_{uz}^{L,M}$ en estado estacionario, considerando

$$L = \hat{n}, \quad M = \{10^2, 10^3, 10^4, 10^5\}$$

para así aproximar los conjuntos ideales Z_{ux}^∞ y Z_{uz}^∞ . De esta manera se construye la matriz $\Gamma_{zx}^M(\hat{n})$ con $\hat{n} = 10$, cuyos valores singulares se muestran en la tabla 4.3, donde se puede ver que la precisión de los valores singulares aumenta cuando M aumenta. Con el fin de poder realizar comparaciones, se muestran los valores singulares no nulos de $\mathcal{H}(\hat{n})$

$$\begin{aligned}
 \sigma_1(\mathcal{H}) &= 1,3065 & \sigma_2(\mathcal{H}) &= 1,2604 & \sigma_3(\mathcal{H}) &= 0,7871 \\
 \sigma_4(\mathcal{H}) &= 0,7712 & \sigma_5(\mathcal{H}) &= 0,6474 & \sigma_6(\mathcal{H}) &= 0,4473 \\
 \sigma_7(\mathcal{H}) &= 0,3441 & & & &
 \end{aligned}$$

Para más detalles ver el programa F.5.2.

4.5. Comentarios del capítulo

Al utilizar actuaciones $u(k)$ con características de ruido blanco $\varphi_{uu}^\infty(\tau) = a_u^2 \delta_K(\tau)$, para construir las matrices $\Lambda_{xux}^\infty(\hat{n}_b, \hat{n}_a)$ y $\Lambda_{zux}^\infty(\hat{n}_b, \hat{n}_a)$, aparece un conjunto de valores singulares y valores propios iguales a a_u^2 . En el caso de la matriz $\Lambda_{zux}^\infty(\hat{n}_b, \hat{n}_a)$, es posible controlar la cantidad de valores propios iguales a a_u^2 , mediante una elección adecuada del filtro instrumental (ver lema 4.1). Este resultado permite prescindir de la matriz $\Lambda_{zux}^\infty(\hat{n}_b, \hat{n}_a)$ para el proceso de identificación estructural, siendo suficiente el uso de la submatriz $\Gamma_{zx}^\infty(\hat{n}_a)$ (ver teorema 4.2).

En la práctica, la submatriz $\Gamma_{zx}^\infty(\hat{n}_a)$ no se puede construir, ya que para ella se necesitan infinitos datos, por lo que debe ser aproximada por $\Gamma_{zx}^M(\hat{n}_a)$, con M muy grande. Esto trae desventajas, ya que construir la submatriz con M finito equivale a perturbar la submatriz $\Gamma_{zx}^\infty(\hat{n}_a)$, aun cuando los datos no están contaminados con ruido. Esta perturbación, debida a errores de aproximación en las correlaciones, dificulta el cálculo de los rangos (ejemplos 4.1 y 4.2), situación que no ocurre con las matrices $\Lambda_{xux}^M(\hat{n}_b, \hat{n}_a)$ y $\Lambda_{zux}^M(\hat{n}_b, \hat{n}_a)$, las que no poseen errores para M finito, y por lo tanto no se ven perturbadas.

A pesar de la desventaja mencionada, la submatriz $\Gamma_{zx}^\infty(\hat{n}_a)$ tiene una importante ventaja sobre $\Lambda_{xux}^M(\hat{n}_b, \hat{n}_a)$ y $\Lambda_{zux}^M(\hat{n}_b, \hat{n}_a)$, ya que mediante la elección de un filtro instrumental adecuado, se puede prescindir de la medición de $u(k)$ para la construcción de $\Gamma_{zx}^\infty(\hat{n}_a)$, lo que es útil para identificar la estructura de sistemas ARMA (ver corolario 4.1 y ejemplos 4.3 y 4.4). Nótese que con las condiciones del corolario 4.1 sólo es posible determinar parcialmente la estructura de sistemas ARMA, ya que al aplicar algún sistema de búsqueda de estructura, solamente se puede identificar el par $(n_b - d_g, n_a)$, donde d_g es el retardo del sistema ARMA. Sin embargo, cuando el sistema ARMA tiene retardo $d_g = 0$, puede identificarse directamente la estructura (n_b, n_a) .

La submatriz $\Gamma_{zx}^\infty(\hat{n}_a)$ también es capaz de identificar la matriz de Hankel asociada al sistema, mediante la elección apropiada del filtro instrumental. Esto a su vez, permite que los valores singulares de $\Gamma_{zx}^\infty(\hat{n}_a)$ sean proporcionales a los valores singulares de Hankel, conectando así la submatriz $\Gamma_{zx}^\infty(\hat{n}_a)$ con una medida clásica de controlabilidad-observabilidad (ver teorema 4.3), utilizada comúnmente para reducción de modelos.

Capítulo 5

Conclusiones generales

LA principal motivación de este trabajo consistió en establecer claramente condiciones suficientes, que permitan conocer el rango de matrices de correlaciones, ya que estas condiciones no se encuentran bien establecidas en la literatura, especialmente cuando se utiliza un número finito de mediciones, y más aún, cuando la salida se ve perturbada con ruido de medición.

El desarrollo de condiciones suficientes, para conocer el rango de estas matrices, está íntimamente relacionado con la medición de datos *suficientemente informativos* (s.i.) (definiciones 1.12 y 1.13). Este tipo de datos puede ser generado, en el caso de tomar infinitas mediciones, al aplicar al sistema una señal de actuación que sea *persistentemente excitante* (p.e.) de algún orden, tal como se conoce en la literatura (corolarios 1.3, 1.4 y 1.5). Sin embargo, al utilizar un número finito de mediciones, una señal de actuación p.e. no es suficiente para generar datos s.i., especialmente si el número de mediciones es pequeño. Es por esto que se debió desarrollar otro concepto para señales de actuación, de manera que permita generar mediciones s.i. y finitas. (corolarios 1.3 y 1.4); este concepto corresponde al de señales *localmente excitantes* (l.e.). Para hacer una comparación entre señales p.e. (definición 1.11) y señales l.e. (definición 1.10), se puede decir que las señales p.e. poseen cierta cantidad mínima de componentes espectrales, sin embargo, las señales l.e. deben tener esta cantidad mínima de componentes espectrales, separadas además por un espaciado frecuencial determinado.

Así, gracias a la utilización del nuevo concepto l.e., para el diseño de señales de actuación, y a la medición de datos s.i., es posible conocer de manera exacta el rango de la *matriz de producto* $\Lambda_{zux}^M(\hat{n}_b, \hat{n}_a)$, definida en (2.23) y construida en base a un conjunto finito mediciones de entrada y salida del sistema,

las cuales se encuentran libres de ruido (lema 2.3). De manera análoga, gracias a la utilización de una señal de actuación l.e. y a la medición de datos s.i., es posible conocer de manera exacta el rango de la *matriz de producto instrumental* $\Lambda_{zux}^M(\hat{n}_b, \hat{n}_a)$, definida en (2.23), cuando se utiliza un número finito de mediciones y en un ambiente libre de ruido (teorema 2.2).

De esta manera, el lema 2.3 y el teorema 2.2, obtenidos en este trabajo, permiten el correcto funcionamiento de procedimientos existentes en la literatura, para identificación estructural, los que realizan pruebas de singularidad sobre la *matriz de producto* o la *matriz de producto instrumental*, las que son hechas de forma exhaustiva hasta encontrar la estructura más apropiada, tal como se muestra en los ejemplos 2.4, 2.5 y 2.6 y en los artículos [13], [1] y [2], requiriéndose analizar muchas matrices, si la cantidad de estructuras candidatas es alta. En el caso de que el sistema tenga ruido de salida, de distribución de probabilidad desconocida, la *matriz de producto instrumental* es capaz de eliminar completamente el ruido, si se utiliza un número infinito de mediciones (lema 2.2), sin embargo, en la práctica el número de mediciones puede ser muy grande, pero siempre es finito, lo que provoca que el ruido no se pueda eliminar completamente. Este último hecho hace que los valores propios nulos de la *matriz de producto instrumental* se vean perturbados, lo que dificulta la discriminación de cuándo la matriz es realmente singular, ya que no se sabría cuándo un valor propio perturbado podría ser cero.

Con el fin de reducir la cantidad de matrices a analizar en los métodos exhaustivos de búsqueda de estructura, se puede utilizar la información adicional que provee el rango de la *matriz de producto* y de la *matriz de producto instrumental*, proponiéndose

se así un nuevo método basado en análisis de rangos, el cual reduce el número total de matrices analizadas a un máximo de 4, para identificar la estructura del sistema (algoritmos 3.1 y 3.2), donde el cálculo de los rangos se realiza descartando valores singulares nulos. En caso de que los datos se encuentren perturbados con ruido de salida, puede utilizarse la *matriz de producto instrumental* para atenuar parcialmente su efecto, sin embargo, el remanente de ruido perturba los valores singulares nulos de la *matriz de producto instrumental*, dificultando así la determinación del rango. Este último problema puede ser resuelto gracias al teorema 3.1, obtenido en esta tesis, ya que si se utiliza como señal de entrada ruido blanco no correlacionado con el ruido de medición, y si además se utiliza un retardo como filtro instrumental, es posible obtener un criterio objetivo para discriminar si un valor singular perturbado de la *matriz de producto instrumental* podría ser cero.

Al utilizar actuaciones con características de ruido blanco y un filtro instrumental adecuado para construir la *matriz de producto instrumental*, aparecen una serie de valores propios que no contribuyen a determinar si esta matriz es singular (lema 4.1), este resultado permite prescindir de la *matriz de producto instrumental* para el proceso de identificación estructural, siendo suficiente el uso de su submatriz $\Gamma_{zx}^\infty(\hat{n}_a)$, definida en (2.27), y cuyo rango queda totalmente determinado, gracias al resultado obtenido en el teorema 4.2. En la práctica, esta submatriz no se puede construir, ya que para ella se necesitan infinitas mediciones, por lo que debe ser aproximada por $\Gamma_{zx}^M(\hat{n}_a)$, con M muy grande. Esto equivale a perturbar la submatriz $\Gamma_{zx}^\infty(\hat{n}_a)$, aun cuando los datos no están contaminados con ruido, dificultando así el cálculo de los rangos (ejemplos 4.1 y 4.2).

A pesar de la desventaja anterior, la submatriz $\Gamma_{zx}^\infty(\hat{n}_a)$ tiene una importante ventaja sobre la *matriz de producto (instrumental)*, ya que mediante la elección de un filtro instrumental adecuado se puede obtener la *matriz de Yule-Walker*, cuyo rango queda totalmente determinado, gracias al resultado obtenido en el corolario 4.1, permitiendo así la identificación estructural de sistemas ARMA (algoritmo 4.2). La submatriz $\Gamma_{zx}^\infty(\hat{n}_a)$ presenta además otra ventaja, ya que es capaz de identificar la matriz de Hankel asociada al sistema, considerando una elección adecuada del filtro instrumental. Esto a su vez, permite obtener otro importante resultado, el que indica que los valores singulares de $\Gamma_{zx}^\infty(\hat{n}_a)$ son proporcionales

a los valores singulares de Hankel, conectando así la submatriz $\Gamma_{zx}^\infty(\hat{n}_a)$ con una medida clásica de controlabilidad-observabilidad (ver teorema 4.3), utilizada comúnmente para reducción de modelos.

Trabajo futuro

A continuación se detallan algunos aspectos que no han sido cubiertos por este trabajo de tesis, pero que podrían dar lugar a líneas de investigación y desarrollo en el futuro

- Encontrar un criterio para discriminar valores singulares nulos de la submatriz $\Gamma_{zx}^M(\hat{n}_a)$, debido a que ella se ve perturbada al utilizar un número finito de mediciones $\Gamma_{zx}^M = \Gamma_{zx}^\infty + E_{zx}^M$
- Extensión de los resultados exactos para los rangos de matrices de correlación al caso multivariable, considerando un número finito de mediciones y en el caso libre de ruido
- Extensión al caso multivariable de un criterio objetivo para discriminar valores singulares nulos de la *matriz de producto instrumental*, cuando la salida del sistema se ve perturbada con ruido de medición $\Lambda_{zuy}^M = \Lambda_{zux}^M + E_{zuy}^M$
- Analizar la conveniencia de la utilización de una versión *en línea/recursiva* de estos métodos de identificación estructural en aplicaciones prácticas
- Verificar qué ocurre en caso de que estos métodos de identificación estructural se apliquen a una planta no lineal. ¿Son estos métodos capaces de encontrar la estructura del sistema linealizado en torno a un punto de operación específico? ¿Bajo qué condiciones se puede lograr esto?. En caso de que no se encuentre la estructura del sistema linealizado en un punto de operación ¿qué es lo que estos métodos obtendrían?

Apéndice A

Notación

\xrightarrow{cm}	convergencia cuadrática media	definición 1.1
$\xrightarrow{p1}$	convergencia con probabilidad 1	definición 1.1
$:=$	definición	$x := y$, x se define como y
$\langle v_1, \dots, v_k \rangle$	espacio vectorial generado por v_1, \dots, v_k	
$(\cdot)^+$	parte positiva	$x^+ = \text{máx}(0, x)$, ecuación (2.16)
$(\cdot)^-$	parte negativa	$x^- = \text{mín}(0, x)$, ecuación (2.16)
$(\cdot)^*$	conjugación compleja	
$(\cdot)^H$	operador hermitiano para matrices	
$(\cdot)^T$	operador traspuesto	
$\overline{(\cdot)}_{k=a}^b$	operador promedio	ecuación (1.1)
$(\cdot) * (\cdot)$	convolución discreta	ecuación (1.3)
$\delta_K(k)$	delta de Kronecker	
Δt	tiempo de muestreo	
θ	frecuencia normalizada para señales discretas	$\theta = \omega \Delta t$
$\mu(k)$	escalón unitario	
$\varphi_{xy}^M(-\tau_1, \tau_2)$	correlación muestral	ecuación (1.2)
$\Phi_{xy}^\infty(e^{j\theta})$	transformada de Fourier de $\varphi_{xy}^\infty(\tau)$	$\Phi_{xy}^\infty(e^{j\theta}) = \mathcal{F}_d \{ \varphi_{xy}^\infty \} (\theta)$
$c_{xy}(k_1, k_2)$	covarianza discreta	definición 1.3
$c_{xy}(\tau)$	covarianza estacionaria discreta	definición 1.6
$C_{xy}(e^{j\theta})$	transformada de Fourier de $c_{xy}(\tau)$	$C_{xy}(e^{j\theta}) = \mathcal{F}_d \{ c_{xy} \} (\theta)$
c.e.s.a.	conjuntamente estacionario en sentido amplio	definición 1.7
$E \{ \cdot \}$	operador esperanza	$E \{ g(X) \} := \int_{\mathbb{R}} g(x) f_X(x) dx$
e.s.a.	estacionario en sentido amplio	definición 1.6
GIPM	generalized instrumental product matrix, matriz de producto instrumental generalizada	ecuación (2.23)
	$\Lambda_{zux}^M(\hat{n}_b, \hat{n}_a)$	
IPM	instrumental product matrix, matriz de producto instrumental	ecuación (2.7)
	$\Lambda_{zux}^M(\hat{n})$	

$\mathcal{J}(\cdot)$	operador de inversión temporal	$(\mathcal{J}x)(k) := x(-k)$
l.d.	linealmente dependiente	
l.i.	linealmente independiente	
l.e.	localmente excitante	definición 1.10
LS	método de cuadrados mínimos	
$\mathcal{F}_d\{\cdot\}$	transformada de Fourier discreta	ecuación (1.7)
p.e.	persistentemente excitante	definición 1.11
PEM	método del error de predicción	
PM	product matrix, matriz de producto $\Lambda_{xux}^M(\hat{n})$	ecuación (2.7)
$q^a(\cdot)$	operador de adelanto	$(q^a x)(k) = x(k + a)$
QED	quod erat demonstrandum	lo que se estaba demostrando
$r_{xy}(k_1, k_2)$	correlación discreta	definición 1.3
$r_{xy}(\tau)$	correlación estacionaria discreta	definición 1.6
$R_{xy}(e^{j\theta})$	transformada de Fourier de $r_{xy}(\tau)$	$R_{xy}(e^{j\theta}) = \mathcal{F}_d\{r_{xy}\}(\theta)$
s.i.	suficientemente informativo	definiciones 1.12 y 1.13
SNR	signal to noise ratio	relación señal a ruido
$\text{var}\{\cdot\}$	varianza	$\text{var}\{X\} := E\{(X - EX)^2\}$
$Z_{ux}^{L,M}$	conjunto de mediciones	ecuación (2.4)

Apéndice B

Propiedades de convolución y correlación

ANTES de revisar las propiedades de las operaciones de convolución y correlación, se recordarán las propiedades espectrales del operador de inversión temporal \mathcal{J} y del operador hermitico para matrices $(\cdot)^H$. Para obtener resultados generales se trabajará con señales matriciales del tipo $s(k) \in \mathbb{C}^{n \times m}$.

Sea $S(e^{j\theta}) \in \mathbb{C}^{n \times m}$ la transformación de Fourier de tiempo discreto de $s(k) \in \mathbb{C}^{n \times m}$. Entonces se tienen los siguientes pares

$$\begin{aligned} (\mathcal{J}s)(k) &\leftrightarrow S(e^{-j\theta}) \\ s^H(k) &\leftrightarrow S(e^{-j\theta})^H \end{aligned}$$

Si ahora se aplican ambos operadores simultáneamente se tiene

$$(\mathcal{J}s^H)(k) \leftrightarrow S(e^{j\theta})^H \quad (\text{B.1})$$

Nótese que el operador \mathcal{J} conmuta con el operador $(\cdot)^H$, por lo que no influye el orden en que se apliquen.

B.1. Convolución

Sean $x(k) \in \mathbb{C}^{n \times m}$, $y(k) \in \mathbb{C}^{m \times l}$ y $z(k) \in \mathbb{C}^{l \times p}$. Es claro que la operación de convolución, definida como

$$(x * y)(k) := \sum_{l \in \mathbb{Z}} x(l)y(k-l)$$

hereda la propiedad de asociatividad proveniente de la multiplicación de matrices

$$(x * y) * z = x * (y * z)$$

además de la propiedad de conmutatividad en el caso escalar, es decir $n = m = l = 1$

$$x * y = y * x$$

Así, el comportamiento espectral de la convolución es el siguiente

$$(x * y)(k) \leftrightarrow X(e^{j\theta})Y(e^{j\theta})$$

donde las versiones en mayúscula corresponden a las transformadas de Fourier de las letras en minúscula.

Además los operadores $(\cdot)^H$ y \mathcal{J} presentan el siguiente comportamiento bajo la operación de convolución

$$\begin{aligned} (x * y)^H &= y^H * x^H \\ \mathcal{J}(x * y) &= \mathcal{J}x * \mathcal{J}y \end{aligned} \quad (\text{B.2})$$

esta última propiedad se puede demostrar considerando

$$\begin{aligned} \mathcal{J}(x * y)(k) &= (x * y)(-k) = \sum_{l \in \mathbb{Z}} x(l)y(-k-l) \\ &= \sum_{n \in \mathbb{Z}} x(-n)y(-k+n) \\ &= \sum_{n \in \mathbb{Z}} \mathcal{J}x(n)\mathcal{J}y(k-n) \\ &= (\mathcal{J}x * \mathcal{J}y)(k) \end{aligned}$$

B.2. Correlación

Sean las señales $x(k) \in \mathbb{C}^{n \times m}$ e $y(k) \in \mathbb{C}^{l \times m}$. Como ya se ha visto en la ecuación (1.2), para ellas se puede definir la siguiente correlación muestral

$$\varphi_{xy}^\infty(\tau) := \overline{\left\langle x(k)y(k-\tau)^H \right\rangle_{k \in \mathbb{Z}}}$$

Dicha función puede verse, al igual que la convolución, como una operación binaria entre dos señales x e y

$$\varphi_{(\cdot, \cdot)}^\infty : (x, y) \rightarrow \varphi_{xy}^\infty$$

esta *operación binaria* posee una clase de “conmutatividad” muy particular, ya que involucra al operador \mathcal{J} y al operador $(\cdot)^H$

$$\mathcal{J}\varphi_{xy}^\infty = (\varphi_{yx}^\infty)^H \quad (\text{B.3})$$

la que se puede demostrar considerando

$$\begin{aligned} (\mathcal{J}\varphi_{xy}^\infty)(\tau) &= \varphi_{xy}^\infty(-\tau) \\ &= \left\langle \overline{x(k)y(k+\tau)^H} \right\rangle_{k \in \mathbb{Z}} \\ &= \left(\left\langle \overline{y(k+\tau)x(k)^H} \right\rangle_{k \in \mathbb{Z}} \right)^H \\ &= \varphi_{yx}^\infty(\tau)^H \end{aligned}$$

y cuya transformación de Fourier resulta

$$\begin{aligned} \Phi_{xy}^\infty(e^{-j\theta}) &= \Phi_{yx}^\infty(e^{-j\theta})^H \\ \Phi_{xy}^\infty(e^{j\theta}) &= \Phi_{yx}^\infty(e^{j\theta})^H \end{aligned}$$

Además, si $x(k)$ e $y(k)$ son procesos c.e.s.a., entonces la correlación y la covarianza

$$\begin{aligned} r_{xy}(\tau) &:= E \{ x(k)y(k-\tau)^H \} \\ c_{xy}(\tau) &:= E \{ [x(k) - m_x][y(k-\tau) - m_y]^H \} \end{aligned}$$

presentan la misma propiedad de “conmutatividad”

$$\mathcal{J}r_{xy} = r_{yx}^H \quad (\text{B.4})$$

$$\mathcal{J}c_{xy} = c_{yx}^H \quad (\text{B.5})$$

las que se pueden demostrar de manera análoga.

B.3. Correlación en sistemas

B.3.1. Caso estacionario

Sea $h(k) \in \mathbb{C}^{m \times i}$ la respuesta a impulso de un sistema lineal, $u(k) \in \mathbb{C}^{i \times n}$ la señal de entrada de dicho sistema e $y(k) \in \mathbb{C}^{l \times n}$ una señal arbitraria. ¿Cómo se comporta la correlación muestral entre la salida del sistema $(h * u)(k)$ y la señal $y(k)$? Esto se

puede ver a través de su definición

$$\begin{aligned} \varphi_{(h*u)y}^\infty(\tau) &= \left\langle \overline{(h * u)(k)y(k-\tau)^H} \right\rangle_{k \in \mathbb{Z}} \\ &= \left\langle \overline{\sum_{l \in \mathbb{Z}} h(l)u(k-l)y(k-\tau)^H} \right\rangle_{k \in \mathbb{Z}} \\ &= \sum_{l \in \mathbb{Z}} h(l) \left\langle \overline{u(k-l)y(k-\tau)^H} \right\rangle_{k \in \mathbb{Z}} \end{aligned}$$

$$\begin{aligned} \varphi_{(h*u)y}^\infty(\tau) &= \sum_{l \in \mathbb{Z}} h(l)\varphi_{uy}^\infty(\tau-l) = (h * \varphi_{uy}^\infty)(\tau) \\ \varphi_{(h*u)y}^\infty &= h * \varphi_{uy}^\infty \end{aligned}$$

Es decir, la correlación $\varphi_{(h*u)y}^\infty$ se puede interpretar como la salida obtenida al aplicar la señal φ_{uy}^∞ al sistema lineal $h(k)$. Esta propiedad también se puede obtener al convolucionar φ_{yu}^∞ por la izquierda, según

$$\begin{aligned} \varphi_{(h*u)y}^\infty &= h * \varphi_{uy}^\infty \\ \mathcal{J}\varphi_{(h*u)y}^\infty &= \mathcal{J}h * \mathcal{J}\varphi_{uy}^\infty \\ (\varphi_{y(h*u)}^\infty)^H &= \mathcal{J}h * (\varphi_{yu}^\infty)^H \\ \varphi_{y(h*u)}^\infty &= \varphi_{yu}^\infty * \mathcal{J}h^H \end{aligned}$$

Es decir, la correlación $\varphi_{y(h*u)}^\infty$ se puede interpretar como la salida obtenida al aplicar la señal φ_{yu}^∞ (por la izquierda) al sistema lineal $h(-k)^H$. En resumen

$$\varphi_{(h*u)y}^\infty = h * \varphi_{uy}^\infty \quad (\text{B.6})$$

$$\varphi_{y(h*u)}^\infty = \varphi_{yu}^\infty * \mathcal{J}h^H \quad (\text{B.7})$$

cuyas transformaciones de Fourier son

$$\Phi_{(h*u)y}^\infty(e^{j\theta}) = H(e^{j\theta})\Phi_{uy}^\infty(e^{j\theta})$$

$$\Phi_{y(h*u)}^\infty(e^{j\theta}) = \Phi_{yu}^\infty(e^{j\theta})H(e^{j\theta})^H$$

Si se considera ahora que $u(k)$ e $y(k)$ son procesos c.e.s.a.¹, entonces se tiene que la correlación estadística y la covarianza cumplen las mismas propiedades, las cuales se pueden demostrar de forma análoga

$$r_{(h*u)y} = h * r_{uy} \quad (\text{B.8})$$

$$r_{y(h*u)} = r_{uy} * \mathcal{J}h^H \quad (\text{B.9})$$

$$c_{(h*u)y} = h * c_{uy} \quad (\text{B.10})$$

$$c_{y(h*u)} = c_{uy} * \mathcal{J}h^H \quad (\text{B.11})$$

¹ver definición 1.7

B.3.2. Caso no estacionario

Considérese ahora que las señales $z(k) \in \mathbb{C}^{m \times n}$ y $x(k) \in \mathbb{C}^{l \times n}$ son procesos estocásticos, no necesariamente estacionarios, de manera que $z = f * u$, donde $f(k) \in \mathbb{C}^{m \times i}$ es determinística y $u(k) \in \mathbb{C}^{i \times n}$ es un proceso estocástico. Para estas señales su correlación satisface

$$\begin{aligned} r_{zx}(k_1, k_2) &= E \{ z(k_1)x(k_2)^H \} \\ &= \sum_{k \in \mathbb{Z}} f(k) E \{ u(k_1 - k)x(k_2)^H \} \\ &= \sum_{k \in \mathbb{Z}} f(k) r_{ux}(k_1 - k, k_2) \end{aligned}$$

es decir

$$r_{zx}(\cdot, k_2) = f * r_{ux}(\cdot, k_2) \quad (\text{B.12})$$

Note además, que dado $r_{zx}(k_1, k_2) = r_{xz}(k_2, k_1)^H$ y $r_{ux}(k_1, k_2) = r_{xu}(k_2, k_1)^H$ se tiene

$$r_{xz}(k_2, \cdot) = r_{xu}(k_2, \cdot) * f^H \quad (\text{B.13})$$

Las dos propiedades anteriores se cumplen igualmente para la covarianza, teniendo en cuenta que

$$\begin{aligned} \tilde{z} &= z - Ez & c_{zx} &= r_{z\tilde{x}} \\ \tilde{x} &= x - Ex & \tilde{z} &= f * \tilde{u} \\ \tilde{u} &= u - Eu \end{aligned}$$

así resulta

$$c_{zx}(\cdot, k_2) = f * c_{ux}(\cdot, k_2) \quad (\text{B.14})$$

$$c_{xz}(k_2, \cdot) = c_{xu}(k_2, \cdot) * f^H \quad (\text{B.15})$$

B.3.3. Caso de ruido blanco

Se sabe que el ruido blanco, de media y varianza constantes, es un proceso e.s.a.². Sin embargo, en el caso general, y según la definición 1.5, el ruido blanco no es un proceso e.s.a.

A pesar de lo anterior, una señal $v(k) \in \mathbb{C}^{m \times n}$, generada al filtrar ruido blanco $e(k) \in \mathbb{C}^{i \times n}$ mediante $v = h * e$, continúa cumpliendo las propiedades (B.14) y (B.15)

$$\begin{aligned} c_{ve}(\cdot, k_2) &= h * c_{ee}(\cdot, k_2) \\ c_{ev}(k_1, \cdot) &= c_{ee}(k_1, \cdot) * h^H \end{aligned}$$

²ver definición 1.6

donde $h(k) \in \mathbb{C}^{m \times i}$ es determinística y $c_{ee}(k_1, k_2) = \alpha(k_1)\delta_K(k_1 - k_2)$. Además se cumple una propiedad especial para el cálculo de autocovarianza de $v(k)$, la cual se puede determinar a través de

$$\begin{aligned} c_{vv}(k_1, k_2) &= \sum_{k \in \mathbb{Z}} h(k_1 - k)c_{ee}(k, k_2) \\ c_{ev}(k, k_2) &= \alpha(k)h(k_2 - k)^H \end{aligned}$$

así resulta

$$\begin{aligned} c_{vv}(k_1, k_2) &= \sum_{l \in \mathbb{Z}} [h(l)\alpha(k_1 - l)]h(-k_1 + k_2 + l)^H \\ &= \sum_{l \in \mathbb{Z}} \bar{h}_{k_1}(l)\mathcal{J}h(k_1 - k_2 - l)^H \end{aligned}$$

donde $\bar{h}_{k_1}(l) = h(l)\alpha(k_1 - l)$, lo que se puede escribir de forma más compacta como

$$c_{vv}(k_1, k_2) = (\bar{h}_{k_1} * \mathcal{J}h^H)(k_1 - k_2) \quad (\text{B.16})$$

Nótese, que apesar de que ni $v(k)$ ni $e(k)$ son procesos e.s.a., la ecuación anterior se asemeja bastante al resultado (1.8) obtenido para casos donde las señales son c.e.s.a.

B.4. Algunas correlaciones de señales conocidas

Considérese por ahora señales escalares, las cuales pueden clasificarse dentro de dos categorías

1. Señales de potencia finita
2. Señales de energía finita

Durante la mayor parte de esta tesis se trabaja explícitamente con señales de potencia finita, como por ejemplo sinusoides y procesos estocásticos, es decir señales que se mantienen activas por un periodo muy largo de tiempo. Así las correlaciones para este tipo de señales están definidas por

$$\begin{aligned} \varphi_{zv}^\infty(\tau) &= \lim_{M \rightarrow \infty} \frac{1}{2M+1} \sum_{k=-M}^M z(k)v(k-\tau)^* \\ r_{zv}(\tau) &= Ez(k)v(k-\tau)^* \end{aligned}$$

En el caso de las señales de energía finita, como por ejemplo las respuestas a impulso de sistemas estrictamente estables, es decir, señales que decaen con

el tiempo o que están confinadas a algún intervalo de tiempo finito, también se puede definir una operación de correlación

$$(f \star h)(\tau) := \sum_{k \in \mathbb{Z}} f(k)h(k - \tau)^*$$

lo que puede interpretarse como una convolución

$$f \star h = f * \mathcal{J}h^*$$

donde \mathcal{J} es el operador de inversión temporal.

Estas correlaciones de señales de energía también se pueden encontrar implícitamente relacionadas con las correlaciones de señales de potencia, como por ejemplo en (1.4) y (1.9)

$$\varphi_{zv}^\infty = (f * \mathcal{J}h^*) * \varphi_{ue}^\infty, \quad r_{zv} = (f * \mathcal{J}h^*) * r_{ue}$$

donde $z = f * u$ y $v = h * e$. Es por este motivo que a continuación se muestran dos resultados de correlaciones para señales de energía finita

Teorema B.1 Sea $\chi_{[a,b]}(k)$ la función característica del intervalo $[a, b] \subset \mathbb{Z}$ definida por

$$\chi_{[a,b]}(k) = \begin{cases} 1, & k \in [a, b] \\ 0, & k \notin [a, b] \end{cases}$$

entonces

$$(\chi_{[a,b]} * \mathcal{J}\chi_{[a,b]})(k) = (b - a + 1 - |k|)\chi_{[-b+a, b-a]}(k)$$

Demostración. Según la definición de convolución se tiene

$$\begin{aligned} (\chi_{[a,b]} * \mathcal{J}\chi_{[a,b]})(k) &= \sum_{l \in \mathbb{Z}} \chi_{[a,b]}(l)\chi_{[a,b]}(l - k) \\ &= \sum_{l \in \mathbb{Z}} \chi_{[a,b]}(l)\chi_{[a+k, b+k]}(l) \end{aligned}$$

de la cual se desprenden tres casos

$$\begin{aligned} &= \begin{cases} 0, & k \notin [-b+a, b-a] \\ \sum_{l=a}^{b+k} 1, & k \in [-b+a, 0] \\ \sum_{l=a+k}^b 1, & k \in [0, b-a] \end{cases} \\ &= \begin{cases} 0, & k \notin [-b+a, b-a] \\ b - a + 1 + k, & k \in [-b+a, 0] \\ b - a + 1 - k, & k \in [0, b-a] \end{cases} \end{aligned}$$

lo que conduce al resultado. **QED.**

Teorema B.2 Considere las siguientes exponenciales

$$\begin{aligned} \epsilon_a(k) &= \lambda_a^k \mu(k), \quad |\lambda_a| < 1 \\ \epsilon_b(k) &= \lambda_b^k \mu(k), \quad |\lambda_b| < 1 \end{aligned}$$

con $\lambda_a, \lambda_b \in \mathbb{R}$, entonces

$$(\epsilon_a * \mathcal{J}\epsilon_b)(k) = \frac{\lambda_a^{k^+} \lambda_b^{-k^-}}{1 - \lambda_a \lambda_b}$$

donde

$$x^+ = \max(0, x), \quad x^- = \min(0, x)$$

Demostración. Sean $E_a(e^{j\theta})$ y $E_b(e^{j\theta})$ las transformaciones de Fourier de $\epsilon_a(k)$ y $\epsilon_b(k)$ respectivamente

$$E_a(e^{j\theta}) = \frac{e^{j\theta}}{e^{j\theta} - \lambda_a}, \quad E_b(e^{j\theta}) = \frac{e^{j\theta}}{e^{j\theta} - \lambda_b}$$

según la propiedad espectral de la convolución y del operador \mathcal{J} en la ecuación (B.1), se tiene que $\epsilon_a * \mathcal{J}\epsilon_b$ equivale a

$$\begin{aligned} E_a(e^{j\theta})E_b(e^{j\theta})^* &= \frac{1}{(e^{j\theta} - \lambda_a)(e^{-j\theta} - \lambda_b)} \\ &= \frac{1}{1 - \lambda_a \lambda_b} \left[\frac{\lambda_a}{e^{j\theta} - \lambda_a} + \frac{e^{-j\theta}}{e^{-j\theta} - \lambda_b} \right] \end{aligned}$$

volviendo al dominio del tiempo

$$\begin{aligned} (\epsilon_a * \mathcal{J}\epsilon_b)(k) &= \frac{\lambda_a^k \mu(k-1) + \lambda_b^{-k} \mu(-k)}{1 - \lambda_a \lambda_b} \\ &= \frac{\lambda_a^k \mu(k) + \lambda_b^{-k} \mu(-k-1)}{1 - \lambda_a \lambda_b} \\ &= \frac{1}{1 - \lambda_a \lambda_b} \begin{cases} \lambda_a^k, & k \geq 0 \\ \lambda_b^{-k}, & k < 0 \end{cases} \end{aligned}$$

con lo que finalmente se obtiene el resultado deseado. **QED.**

Apéndice C

Ergodicidad

PARA poder determinar algunas propiedades estadísticas como la media, varianza o algún momento de un proceso estocástico, es necesario conocer la distribución de probabilidad de dicho proceso, por ejemplo si se desea conocer la media del proceso estocástico $s(k) \in \mathbb{R}$, entonces se necesita calcular

$$m_s(k) = E\{s(k)\} = \int_{\mathbb{R}} s f_s(s, k) ds$$

La pregunta ahora es ¿cómo poder estimar esta media, si no se conoce la distribución de probabilidad del proceso estocástico?. Una alternativa para poder resolver este problema es a través del siguiente *estimador*

$$p_M = \frac{1}{2M+1} \sum_{l=-M}^M s(l)$$

Para que éste no tenga *sesgo* en su estimación, es necesario que se cumpla

$$E\{p_M\} = E\{s(k)\}$$

al verificar si se cumple esta condición se tiene

$$E\{p_M\} = \frac{1}{2M+1} \sum_{l=-M}^M E\{s(l)\}$$

lo que indica, que el estimador p_M podría ser sesgado, ya que la siguiente igualdad podría no cumplirse

$$\frac{1}{2M+1} \sum_{l=-M}^M E\{s(l)\} = E\{s(k)\}$$

En muchos casos se tiene que el proceso estocástico $s(k)$ es e.s.a.¹, en este caso se tiene que su media es una constante

$$m_s = E\{s(k)\}$$

¹Estacionario en sentido amplio, er definición 1.6

por lo que el estimador p_M definido anteriormente se vuelve ahora un estimador *no sesgado*, ya que se cumple que

$$E\{p_M\} = E\{s(k)\} = m_s$$

es decir, el estimador p_M obtiene, en promedio, la media m_s que se quiere estimar. Idealmente, se podría pensar que esta situación mejora si se aumenta el número de muestras para el cálculo de p_M , es decir

$$\lim_{M \rightarrow \infty} p = \mu_s$$

La ergodicidad se preocupa en general de este tema, es decir, de poder calcular propiedades estadísticas como media, varianza o algún momento de algún proceso estocástico, a través de promedios temporales.

C.1. Media ergódica

Supóngase que se tiene un proceso e.s.a $s(k) \in \mathbb{C}$ de media

$$m_s = E\{s(k)\}$$

y que se intenta estimar la media m_s , a través de un promedio “móvil” de $2M+1$ muestras

$$p_M(k) = \frac{1}{2M+1} \sum_{l=-M+k}^{M+k} s(l)$$

Note que este promedio móvil incluye al promedio convencional, el cual se obtiene a partir de $p_M(0)$.

Se puede ver que este estimador $p_M(k)$ no tiene sesgo, ya que se cumple que

$$E\{p_M(k)\} = E\{s(l)\} = m_s$$

es decir, es capaz de estimar (en promedio) la media del proceso $s(k)$. La pregunta que queda ahora es la siguiente ¿puede $p_M(k)$ estimar la media m_s cuando aumenta el número de muestras?, es decir

$$\lim_{M \rightarrow \infty} p_M(k) \rightarrow m_s$$

Para responder a la pregunta anterior, es conveniente analizar las propiedades estadísticas del estimador $p_M(k)$.

La señal $p_M(k)$ puede ser vista como la señal $s(k)$ filtrada, ya que el promedio móvil es un sistema lineal, así

$$p_M(k) = \sum_{l \in \mathbb{Z}} s(l)w(k-l) = (w * s)(k)$$

donde

$$w(k) = \frac{1}{2M+1} \chi_{[-M, M]}(k)$$

y $\chi_{[-M, M]}(k)$ está definida en el teorema B.1. Ahora, dado que $s(k)$ es un proceso e.s.a y que $w(k)$ es un filtro lineal estable, entonces $p_M(k)$ también es un proceso e.s.a, y por lo tanto, su autocovarianza cumple con la ecuación (1.8)

$$c_{pp} = (w * \mathcal{J}w) * c_{ss} = \gamma * c_{ss}$$

así la varianza de $p_M(k)$ resulta

$$c_{pp}(0) = (\gamma * c_{ss})(0) = \sum_{l \in \mathbb{Z}} \gamma(l)c_{ss}(l)$$

donde la correlación

$$\gamma = w * \mathcal{J}w = \frac{\chi_{[-M, M]} * \mathcal{J}\chi_{[-M, M]}}{(2M+1)^2}$$

se puede obtener a partir del teorema B.1, resultando

$$\gamma(l) = \frac{1}{2M+1} \left(1 - \frac{|l|}{2M+1}\right) \chi_{[-2M, 2M]}(l)$$

con lo que ya se puede tener una expresión explícita para la varianza de $p_M(k)$, la que se resume en el siguiente lema

Lema C.1 Si $s(k) \in \mathbb{C}$ es un proceso e.s.a de media $Es(k) = m_s$, entonces el estimador de m_s

$$p_M(k) = \frac{1}{2M+1} \sum_{l=-M+k}^{M+k} s(l)$$

tiene varianza

$$\text{var} \{p_M(k)\} = \frac{1}{2M+1} \sum_{l=-2M}^{2M} \left(1 - \frac{|l|}{2M+1}\right) c_{ss}(l)$$

Ver [16, ecuación (13-12)].

De este lema se puede ver que si

$$\lim_{|l| \rightarrow \infty} c_{ss}(l) = 0$$

entonces la varianza del estimador $p_M(k)$

$$\lim_{M \rightarrow \infty} \text{var} \{p_M(k)\} = 0$$

es decir, el estimador $p_M(k)$ converge a $Es(k)$

$$\lim_{M \rightarrow \infty} p_M(k) \stackrel{cm}{=} Es(k) = m_s$$

en sentido *cuadrático medio* (ver definición 1.1). El siguiente teorema muestra básicamente el mismo resultado de ergodicidad, pero con un tipo de convergencia más fuerte

Teorema C.1 Sea $s(k)$ un proceso e.s.a., si su autocovarianza satisface

$$\lim_{|l| \rightarrow \infty} c_{ss}(l) = 0$$

entonces

$$\lim_{M \rightarrow \infty} p_M(k) = Es(k) = m_s$$

con probabilidad 1 y en sentido *cuadrático medio*.

Ver [20, lemma B.1].

Corolario C.1 Si $s(k)$ obtenido a través de

$$s(k) = H(q)e(k)$$

donde $h(k)$ es la respuesta a impulso de $H(q)$ tal que

$$\sum_{k=0}^{\infty} |h(k)|^2 < \infty$$

y $e(k)$ es ruido blanco de varianza a_e^2 y media constante, entonces

$$\lim_{M \rightarrow \infty} p_M(k) = Es(k)$$

con probabilidad 1 y en sentido *cuadrático medio*.

Demostración. Dado que $c_{ee}(\tau) = a_e^2 \delta_K(\tau)$ y por la ecuación (1.8) se tiene

$$\begin{aligned} c_{ss}(\tau) &= a_e^2 (h * \mathcal{J}h^*)(\tau) = a_e^2 \sum_{k \in \mathbb{Z}} h(k)h(k-\tau)^* \\ &= a_e^2 \sum_{k=\tau}^{\infty} h(k)h(k-\tau)^* \end{aligned}$$

donde $\tau \geq 0$, entonces utilizando la desigualdad de Cauchy-Schwartz

$$\begin{aligned} |c_{ss}(\tau)|^2 &\leq \sum_{k=\tau}^{\infty} |h(k)|^2 \sum_{k=\tau}^{\infty} |h(k-\tau)|^2 \\ &\leq \underbrace{\sum_{k=\tau}^{\infty} |h(k)|^2}_{\rightarrow 0} \underbrace{\sum_{k=0}^{\infty} |h(k)|^2}_{< \infty} \end{aligned}$$

lo que indica que

$$\lim_{\tau \rightarrow \infty} |c_{ss}(\tau)| = 0$$

ahora, dado que $|c_{ss}(\tau)|$ es par, entonces

$$\lim_{|\tau| \rightarrow \infty} |c_{ss}(\tau)| = 0$$

ahora la demostración sigue a partir del teorema C.1.

QED.

Este último corolario es bastante importante, ya que cualquier proceso que se obtenga de filtrar ruido blanco, de media y varianza constantes, tiene media ergódica, la que converge en sentido cuadrático medio y con probabilidad 1.

C.2. Correlación ergódica

Supóngase ahora que se tienen los siguientes procesos estocásticos reales

$$z(k) = F(q)u(k), \quad v(k) = H(q)e(k)$$

para los cuales se desea obtener

$$r_{zv}(\tau) = E \{z(k)v(k-\tau)\}$$

a través del estimador de la ecuación (1.2)

$$\varphi_{zv}^M(-\tau_1, \tau_2) = \frac{1}{2M+1} \sum_{k=-M}^M z(k-\tau_1)v(k-\tau_2)$$

Sin pérdida de generalidad, este estimador puede ser reescrito en función de la diferencia de tiempo $\tau = \tau_2 - \tau_1$ y de τ_1 , resultando

$$\tilde{\varphi}_{zv}^M(\tau, \tau_1) = \frac{1}{2M+1} \sum_{k=-M-\tau_1}^{M-\tau_1} z(k)v(k-\tau)$$

donde

$$\varphi_{zv}^M(-\tau_1, \tau_2) = \tilde{\varphi}_{zv}^M(\tau, \tau_1), \quad \tau = \tau_2 - \tau_1$$

Si ahora se considera que

$$v_\tau(k) = H_\tau(q)e(k), \quad H_\tau(q) = q^{-\tau}H(q)$$

entonces el estimador queda reducido a

$$\tilde{\varphi}_{zv}^M(\tau, \tau_1) = \frac{1}{2M+1} \sum_{k=-M-\tau_1}^{M-\tau_1} z(k)v_\tau(k)$$

Teniendo en cuenta las siguientes asociaciones

$$p_M(-\tau_1) = \tilde{\varphi}_{zv}^M(\tau, \tau_1), \quad s(k) = z(k)v_\tau(k)$$

se puede hacer uso del resultado obtenido en el teorema C.1, ya que si $s(k)$ tiene media ergódica, calculable a través de $p_M(-\tau_1)$, entonces significa que $z(k)$ y $v(k)$ tienen correlación ergódica, calculable a través de

$$\varphi_{zv}^M(-\tau_1, \tau_2) = \tilde{\varphi}_{zv}^M(\tau, \tau_1), \quad \tau = \tau_2 - \tau_1$$

Por esto, para analizar la ergodicidad de $r_{zv}(\tau)$ sólo basta hacer un análisis de la varianza de su estimador, es decir

$$\text{var} \{ \varphi_{zv}^M(-\tau_1, \tau_2) \} = \tag{C.1}$$

$$\frac{1}{2M+1} \sum_{l=-2M}^{2M} \left(1 - \frac{|l|}{2M+1}\right) c_{ss}(l)$$

para lo cual se necesita calcular primero $c_{ss}(l)$.

Lema C.2 *Considérese los siguientes procesos estocásticos reales*

$$z(k) = F(q)u(k), \quad v(k) = H(q)e(k)$$

donde $\xi(k) = [u(k) \quad e(k)]^T$ es un proceso de ruido blanco estricto² de media cero, tal que

$$c_{\xi\xi}(\tau) = \begin{bmatrix} a_u^2 & \rho a_u a_e \\ \rho a_u a_e & a_e^2 \end{bmatrix} \delta_K(\tau)$$

y tiene cuarto momento finito $E \{u(k)^2 e(k)^2\} = \eta_{ue}$, entonces la covarianza de $s(k) = z(k)v_\tau(k)$ es

$$\begin{aligned} c_{ss}(l) = & \left(\sum_{p \in \mathbb{Z}} f(l+p)f(p) \right) \left(\sum_{q \in \mathbb{Z}} h_\tau(l+q)h_\tau(q) \right) (a_u a_e)^2 \\ & + \left(\sum_{p \in \mathbb{Z}} h_\tau(l+p)f(p) \right) \left(\sum_{q \in \mathbb{Z}} f(l+q)h_\tau(q) \right) (\rho a_u a_e)^2 \\ & + \left(\sum_{p \in \mathbb{Z}} f(l+p)h_\tau(l+p)f(p)h_\tau(p) \right) [\eta_{ue} - (2\rho^2 + 1)(a_u a_e)^2] \end{aligned}$$

² $\xi(k_i)$ es independiente de $\xi(k_j) \forall k_i \neq k_j$, ver definición 1.5

donde

$$v_\tau(k) = H_\tau(q)e(k), \quad H_\tau(q) = q^{-\tau}H(q)$$

y $f(k)$ y $h_\tau(k)$ son las respuestas a impulso de $F(q)$ y $H_\tau(q)$ respectivamente.

Demostración. Nótese que el cálculo de $c_{ss}(l)$ se reduce al cálculo de la correlación $r_{ss}(l)$, ya que

$$c_{ss}(l) = r_{ss}(l) - [Es(k)]^2$$

donde, debido a la ecuación (1.9) se tiene

$$\begin{aligned} Es(k) &= Ez(k)v_\tau(k) = (\rho a_u a_e)(f * \mathcal{J}h_\tau)(0) \\ &= (\rho a_u a_e) \sum_{k \in \mathbb{Z}} f(k)h_\tau(k) \end{aligned}$$

Ahora sólo falta calcular $r_{ss}(l)$

$$\begin{aligned} r_{ss}(l) &= Es(k)s(k-l) \\ &= Ez(k)v_\tau(k)z(k-l)v_\tau(k-l) \\ &= \sum_{i \in \mathbb{Z}} \sum_{j \in \mathbb{Z}} \sum_{p \in \mathbb{Z}} \sum_{q \in \mathbb{Z}} f(i)h_\tau(j)f(p)h_\tau(q)\epsilon(i, j, p, q) \end{aligned}$$

donde

$$\epsilon(i, j, p, q) = Eu(k-i)e(k-j)u(k-l-p)e(k-l-q)$$

El cálculo de $\epsilon(i, j, p, q)$ se puede realizar considerando que $\xi(k_i)$ y $\xi(k_j)$ son *independientes* $\forall k_i \neq k_j$, de esta manera, si hay un solo tiempo diferente del resto, por ejemplo $(k-i)$ es diferente de los demás, entonces se tiene que

$$Eu(k-i) \cdot Ee(k-j)u(k-l-p)e(k-l-q) = 0$$

ya que es ruido blanco estricto de media cero. Así, las únicas posibilidades de tener resultados diferentes de cero son

1. $(k-i = k-j) \neq (k-l-p = k-l-q)$
2. $(k-i = k-l-p) \neq (k-j = k-l-q)$
3. $(k-i = k-l-q) \neq (k-j = k-l-p)$
4. $k-i = k-j = k-l-p = k-l-q$

Así, para cada uno de los casos anteriores se obtienen los siguientes resultados

1. $\epsilon(i, j, p, q) = E \{u(k)e(k)\}^2 = (\rho a_u a_e)^2$
2. $\epsilon(i, j, p, q) = E \{u(k)^2\}E \{e(k)^2\} = (a_u a_e)^2$
3. $\epsilon(i, j, p, q) = E \{u(k)e(k)\}^2 = (\rho a_u a_e)^2$
4. $\epsilon(i, j, p, q) = E \{u(k)^2 e(k)^2\} = \eta_{ue}$

Entonces, $\epsilon(i, j, p, q)$ puede ser escrito de forma compacta como

$$\begin{aligned} \epsilon(i, j, p, q) &= \\ &(\rho a_u a_e)^2 \delta_K(i-j) \delta_K(p-q) [1 - \delta_K(l+p-i)] \\ &+ (a_u a_e)^2 \delta_K(l+p-i) \delta_K(l+q-j) [1 - \delta_K(i-j)] \\ &+ (\rho a_u a_e)^2 \delta_K(l+q-i) \delta_K(l+p-j) [1 - \delta_K(i-j)] \\ &+ \eta_{ue} \delta_K(i-j) \delta_K(p-q) \delta_K(l+p-i) \end{aligned}$$

Ahora, considerando que

$$\begin{aligned} \delta_K(l+p-i) \delta_K(l+q-j) \delta_K(i-j) &= \\ \delta_K(p-q) \delta_K(l+q-j) \delta_K(i-j) &= \\ \delta_K(l+q-i) \delta_K(l+p-j) \delta_K(i-j) &= \\ \delta_K(l+q-j) \delta_K(p-q) \delta_K(i-j) &= \end{aligned}$$

se tiene finalmente

$$\begin{aligned} \epsilon(i, j, p, q) &= \\ &(\rho a_u a_e)^2 \delta_K(i-j) \delta_K(p-q) \\ &+ (a_u a_e)^2 \delta_K(l+p-i) \delta_K(l+q-j) \\ &+ (\rho a_u a_e)^2 \delta_K(l+q-i) \delta_K(l+p-j) \\ &+ [\eta_{ue} - (2\rho^2 + 1)(a_u a_e)^2] \delta_K(i-j) \delta_K(p-q) \delta_K(l+p-i) \end{aligned}$$

Una vez obtenido $\epsilon(i, j, p, q)$ se reemplaza en $r_{ss}(l)$

$$\begin{aligned} r_{ss}(l) &= \\ &(\rho a_u a_e)^2 \sum_{i \in \mathbb{Z}} \sum_{p \in \mathbb{Z}} f(i)h_\tau(i)f(p)h_\tau(p) \\ &+ (a_u a_e)^2 \sum_{p \in \mathbb{Z}} \sum_{q \in \mathbb{Z}} f(l+p)h_\tau(l+q)f(p)h_\tau(q) \\ &+ (\rho a_u a_e)^2 \sum_{p \in \mathbb{Z}} \sum_{q \in \mathbb{Z}} f(l+q)h_\tau(l+p)f(p)h_\tau(q) \\ &+ [\eta_{ue} - (2\rho^2 + 1)(a_u a_e)^2] \sum_{p \in \mathbb{Z}} f(l+p)h_\tau(l+p)f(p)h_\tau(p) \end{aligned}$$

Ahora, restando

$$c_{ss}(l) = r_{ss}(l) - [Es(k)]^2$$

se obtiene el resultado deseado. **QED.**

Teorema C.2 *Considérese los siguientes procesos estocásticos reales*

$$z(k) = F(q)u(k), \quad v(k) = H(q)e(k)$$

donde $F(q)$ y $H(q)$ son filtros tales que su respuestas a impulso satisfacen

$$\sum_{k=0}^{\infty} |f(k)|^2 < \infty, \quad \sum_{k=0}^{\infty} |h(k)|^2 < \infty$$

Considérese además que $\xi(k) = [u(k) \quad e(k)]^T$ es un proceso de ruido blanco estricto³ de media cero, tal que

$$c_{\xi\xi}(\tau) = \begin{bmatrix} a_u^2 & \rho a_u a_e \\ \rho a_u a_e & a_e^2 \end{bmatrix} \delta_K(\tau)$$

y tiene cuarto momento finito $E\{u(k)^2 e(k)^2\} = \eta_{ue}$, entonces

$$\lim_{M \rightarrow \infty} \varphi_{zv}^M(-\tau_1, \tau_2) = r_{zv}(\tau_2 - \tau_1)$$

con probabilidad 1 y en sentido cuadrático medio.

Un resultado similar se obtiene en [20, lemma B.2].

Demostración. Considerando

$$\varphi_{zv}^M(-\tau_1, \tau_2) = \tilde{\varphi}_{zv}^M(\tau, \tau_1), \quad \tau = \tau_2 - \tau_1$$

y además

$$v_\tau(k) = H_\tau(q)e(k), \quad H_\tau(q) = q^{-\tau} H(q)$$

entonces el estimador queda reducido a

$$\tilde{\varphi}_{zv}^M(\tau, \tau_1) = \frac{1}{2M+1} \sum_{k=-M-\tau_1}^{M-\tau_1} z(k)v_\tau(k)$$

Ahora, teniendo en cuenta las siguientes asociaciones

$$p_M(-\tau_1) = \tilde{\varphi}_{zv}^M(\tau, \tau_1), \quad s(k) = z(k)v_\tau(k)$$

y según el lema C.1, se tiene que la varianza del estimador es

$$\begin{aligned} \text{var} \{ \varphi_{zv}^M(-\tau_1, \tau_2) \} &= \\ &= \frac{1}{2M+1} \sum_{l=-2M}^{2M} \left(1 - \frac{|l|}{2M+1} \right) c_{ss}(l) \end{aligned}$$

³ $\xi(k_i)$ es independiente de $\xi(k_j) \forall k_i \neq k_j$, ver definición 1.5

donde la covarianza $c_{ss}(l)$ está dada por el lema C.2. Así, se puede hacer uso del resultado obtenido en el teorema C.1, por lo cual sólo hay que verificar que

$$\lim_{|l| \rightarrow \infty} c_{ss}(l) = 0$$

Se puede comprobar que los tres términos en $c_{ss}(l)$ tienden a cero, cuando $l \rightarrow \infty$. Esto se puede verificar a través de la desigualdad de Cauchy-Schwartz, por ejemplo

$$\begin{aligned} \left| \sum_{p \in \mathbb{Z}} h_\tau(l+p)f(p) \right|^2 &= \left| \sum_{p=0}^{\infty} h_\tau(l+p)f(p) \right|^2 \\ &\leq \sum_{p=0}^{\infty} h_\tau(l+p)^2 \sum_{p=0}^{\infty} f(p)^2 \\ &\leq \underbrace{\sum_{j=l}^{\infty} h_\tau(j)^2}_{\rightarrow 0} \underbrace{\sum_{p=0}^{\infty} f(p)^2}_{< \infty} \rightarrow 0 \end{aligned}$$

cuando $l \rightarrow \infty$. De manera similar puede probarse para los demás casos, resultando de esta manera que

$$\lim_{l \rightarrow \infty} c_{ss}(l) = 0$$

ahora, dado que $c_{ss}(l)$ es una función par, entonces

$$\lim_{|l| \rightarrow \infty} c_{ss}(l) = 0$$

QED.

C.3. Caso de ruido blanco

Se sabe que una señal de ruido blanco $s(k)$, de media y varianza constantes, es un proceso e.s.a.⁴, y debido al teorema C.1, tiene media ergódica

$$\lim_{M \rightarrow \infty} p_M(k) = \lim_{M \rightarrow \infty} \frac{1}{2M+1} \sum_{j=-M+k}^{M+k} s(j) = Es(l)$$

en sentido cuadrático medio y con probabilidad 1, ya que su autocovarianza cumple

$$c_{ss}(l) = a_s^2 \delta_K(l) \Rightarrow \lim_{|l| \rightarrow \infty} c_{ss}(l) = 0$$

Sin embargo, en el caso general, y según la definición 1.5, el ruido blanco no es un proceso e.s.a., por lo

⁴Estacionario en sentido amplio, ver definición 1.6

que es natural hacerse la siguiente pregunta: Si $s(k)$ es ruido blanco de media constante y autocovarianza

$$c_{ss}(k_1, k_2) = \alpha(k_1)\delta_K(k_1 - k_2), \quad \alpha(k_1) > 0$$

¿tiene media ergódica, al igual que en el caso e.s.a?

Como se ha visto, la señal $p_M(k)$ puede interpretarse como la señal $s(k)$ filtrada, ya que el promedio móvil es un sistema lineal, así

$$p_M(k) = \sum_{l \in \mathbb{Z}} s(l)w(k-l) = (w * s)(k)$$

donde

$$w(k) = \frac{1}{2M+1}\chi_{[-M, M]}(k)$$

y $\chi_{[-M, M]}(k)$ está definida según el teorema B.1. Ahora, dado que $s(k)$ es ruido blanco, su autocovarianza cumple con la ecuación (B.16), es decir

$$c_{pp}(k_1, k_2) = (\bar{w}_{k_1} * \mathcal{J}w)(k_1 - k_2)$$

En otras palabras

$$c_{pp}(k_1, k_2) = \sum_{l \in \mathbb{Z}} w(l)\alpha(k_1 - l)w(-k_1 + k_2 + l) > 0$$

Si se considera $A = \max_{k \in \mathbb{Z}} \alpha(k)$, esta covarianza puede ser acotada según

$$\begin{aligned} c_{pp}(k_1, k_2) &\leq A \sum_{l \in \mathbb{Z}} w(l)w(-k_1 + k_2 + l) \\ &\leq A(w * \mathcal{J}w)(k_1 - k_2) \\ &\leq A \frac{(\chi_{[-M, M]} * \mathcal{J}\chi_{[-M, M]})(k_1 - k_2)}{(2M+1)^2} \end{aligned}$$

Aplicando el resultado del teorema B.1

$$c_{pp}(k_1, k_2) \leq \frac{A}{2M+1} \left(1 - \frac{|k_1 - k_2|}{2M+1}\right) \chi_{[-2M, 2M]}(k_1 - k_2)$$

Así la varianza de $p_M(k)$ resulta

$$\text{var} \{p_M(k)\} = c_{pp}(j, j) \leq \frac{A}{2M+1}$$

la que tiende a cero cuando el número de muestras aumenta

$$\lim_{M \rightarrow \infty} \text{var} \{p_M(k)\} = 0$$

con lo que se puede obtener el siguiente teorema

Teorema C.3 Si $s(k)$ es ruido blanco de media constante y autocovarianza

$$c_{ss}(k_1, k_2) = \alpha(k_1)\delta_K(k_1 - k_2), \quad 0 < \alpha(k_1) < \infty$$

entonces

$$\lim_{M \rightarrow \infty} \frac{1}{2M+1} \sum_{l=-M}^M s(l) = Es(k)$$

en sentido cuadrático medio.

Apéndice D

Matrices

D.1. Descomposiciones

EN [11] se muestra una serie de descomposiciones y factorizaciones útiles, las que son presentadas a continuación.

Teorema D.1 Descomposición de Jordan. *Considérese que $A \in \mathbb{C}^{n \times n}$, entonces existe una matriz $V \in \mathbb{C}^{n \times n}$ no singular, de manera que*

$$V^{-1}AV = \text{diag}(J_1, \dots, J_t)$$

donde $J_i \in \mathbb{C}^{m_i \times m_i}$

$$J_i = \begin{bmatrix} \lambda_i & 1 & \dots & 0 \\ 0 & \lambda_i & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 0 & \dots & 0 & \lambda_i \end{bmatrix}$$

con

$$\sum_{i=1}^t m_i = n$$

donde λ_i con $i = 1, \dots, t$ son los valores propios de A .

En el caso de que los valores propios de A sean todos diferentes, entonces se tiene que la matriz V es capaz de diagonalizar a la matriz A , en cuyo caso se dice que A es diagonalizable.

Teorema D.2 Descomposición en valores singulares. *Si $A \in \mathbb{C}^{m \times n}$, entonces existen matrices unitarias $U \in \mathbb{C}^{m \times m}$ y $V \in \mathbb{C}^{n \times n}$ de manera que*

$$U^H A V = \text{diag}(\sigma_1, \dots, \sigma_p) \in \mathbb{R}^{m \times n}$$

donde $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_p \geq 0$ y $p = \min(m, n)$

En el caso de que $A \in \mathbb{R}^{m \times n}$ se tendrá que $U \in \mathbb{R}^{m \times m}$ y $V \in \mathbb{R}^{n \times n}$, donde U y V son ortogonales.

Teorema D.3 Descomposición “económica” en valores singulares. *Si $A = U \Sigma V^H \in \mathbb{C}^{m \times n}$ es la descomposición en valores singulares de A con $m \geq n$, entonces*

$$A = U_n \Sigma_n V^H$$

es la descomposición económica en valores singulares, donde

$$U \Sigma V^H = \begin{bmatrix} U_n & U_{m-n} \end{bmatrix} \begin{bmatrix} \Sigma_n \\ O_{m-n} \end{bmatrix} V^H$$

con $U_n \in \mathbb{C}^{m \times n}$ unitaria, $U_{m-n} \in \mathbb{C}^{m \times (m-n)}$ unitaria, $\Sigma_n \in \mathbb{R}^{n \times n}$ y $O_{m-n} \in \mathbb{R}^{(m-n) \times n}$ es una matriz nula.

Teorema D.4 Descomposición de Schur. *Si se tiene $A \in \mathbb{C}^{n \times n}$, entonces existe una matriz $Q \in \mathbb{C}^{n \times n}$ unitaria, de manera que $Q^H A Q = T$, donde T es una matriz triangular superior, cuyas entradas diagonales son exactamente los valores propios de A .*

Para la descomposición anterior siempre se puede escoger una matriz Q de manera que los valores propios de A aparezcan en cualquier orden sobre la diagonal.

Teorema D.5 Descomposición de Schur (simétrica). *Si $A \in \mathbb{R}^{n \times n}$ es simétrica, entonces existe una matriz $Q \in \mathbb{R}^{n \times n}$ ortogonal de manera que*

$$Q^T A Q = \text{diag}(\lambda_1, \dots, \lambda_n) \in \mathbb{R}^{n \times n}$$

Para la descomposición anterior, con A simétrica, se pueden ordenar los valores propios de manera que

$$\lambda_n \leq \dots \leq \lambda_2 \leq \lambda_1$$

además los valores singulares de A están dados por

$$\{|\lambda_1|, \dots, |\lambda_n|\}$$

Los valores propios de una matriz real simétrica tienen la particularidad de que pueden ser descritos por el *teorema minimax de Courant-Fischer* (teorema D.8).

D.2. Valores propios y valores singulares

Teorema D.6 Discos de Gershgorin. *Sea la matriz $M = [m_{ij}]$ de tamaño $n \times n$, entonces cada valor propio de M pertenece a al menos un disco D_i , para $i = 1, \dots, n$*

$$D_i = \left\{ z \in \mathbb{C} \mid |z - m_{ii}| \leq \sum_{j \neq i} |m_{ij}| \right\}$$

Teorema D.7 *Sea la matriz de tamaño $(p+q) \times r$*

$$E = \begin{bmatrix} a_{11} & \cdots & a_{1r} \\ \vdots & & \vdots \\ a_{p1} & \cdots & a_{pr} \\ b_{11} & \cdots & b_{1r} \\ \vdots & & \vdots \\ b_{q1} & \cdots & b_{qr} \end{bmatrix}$$

donde cada elemento de esta matriz se encuentra acotado por

$$|a_{ij}| \leq \epsilon_a, \quad |b_{ij}| \leq \epsilon_b$$

entonces todos los valores singulares de E están acotados por

$$\sigma(E) \leq \sqrt{(\epsilon_a^2 p + \epsilon_b^2 q)r}$$

por lo que en particular se cumple

$$\|E\|_2 \leq \sqrt{(\epsilon_a^2 p + \epsilon_b^2 q)r}$$

Demostración. Tómesese la matriz $M = E^T E$, cuyos elementos están dados por m_{ij} , entonces debido al teorema de Gershgorin, un valor propio $\lambda(M) = \sigma^2(E)$ (cualquiera) debe pertenecer a al menos un disco D_i , es decir

$$|\sigma^2(E) - m_{ii}| \leq \sum_{j \neq i} |m_{ij}|$$

para $i \in \{1, \dots, r\}$. Ahora, dado que

$$|\sigma^2(E) - m_{ii}| \leq |\sigma^2(E) - m_{ii}|$$

entonces $\sigma^2(E)$ debe pertenecer a al menos uno de los discos

$$|\sigma^2(E)| \leq \sum_{j=1}^r |m_{ij}|$$

con $i \in \{1, \dots, r\}$. Nótese también que el elemento m_{ij} esta formado por

$$m_{ij} = \sum_{k=1}^p a_{ki} a_{kj} + \sum_{l=1}^q b_{li} b_{lj}$$

Entonces

$$|\sigma^2(E)| \leq \sum_{j=1}^r \left(\sum_{k=1}^p |a_{ki}| |a_{kj}| + \sum_{l=1}^q |b_{li}| |b_{lj}| \right)$$

para al menos un $i \in \{1, \dots, r\}$. Finalmente, como $|a_{ij}| \leq \epsilon_a$ y $|b_{ij}| \leq \epsilon_b$, entonces

$$|\sigma^2(E)| \leq (\epsilon_a^2 p + \epsilon_b^2 q)r$$

para cualquier valor singular de E . **QED.**

Teorema D.8 Representación minimax de Courant-Fischer. *Si $A \in \mathbb{R}^{n \times n}$ es simétrica, entonces se tiene que su k -ésimo valor propio más grande está dado por*

$$\lambda_k = \max_{\dim(S)=k} \min_{0 \neq y \in S} \frac{y^T A y}{y^T y} \quad \forall k = 1, \dots, n$$

Ver [11, theorem 8.1.2].

Demostración. Como A es simétrica, entonces tiene descomposición simétrica de Schur

$$Q^T A Q = \text{diag}(\lambda_1, \dots, \lambda_n)$$

donde $Q = [q_1 \ \cdots \ q_n]$ y $\{q_1, \dots, q_n\}$ son l.i.

Considérese ahora el subespacio de dimensión k , asociado a los valores propios $\lambda_{p_1}, \dots, \lambda_{p_k}$, es decir el subespacio $\langle q_{p_1}, \dots, q_{p_k} \rangle$ y considérese también el vector

$$y = \sum_{l=1}^k \alpha_l q_{p_l} \in \langle q_{p_1}, \dots, q_{p_k} \rangle$$

entonces

$$\frac{y^T A y}{y^T y} = \frac{\sum_{l=1}^k \sum_{m=1}^k \alpha_l \alpha_m q_{p_l}^T A q_{p_m}}{\sum_{l=1}^k \sum_{m=1}^k \alpha_l \alpha_m q_{p_l}^T q_{p_m}} = \frac{\sum_{j=1}^k \alpha_j^2 \lambda_{p_j}}{\sum_{j=1}^k \alpha_j^2}$$

así

$$\lambda_{\hat{p}} \geq \frac{y^T A y}{y^T y} \geq \lambda_{\check{p}}$$

donde

$$\hat{p} = \max_{j=1, \dots, k} p_j, \quad \check{p} = \min_{j=1, \dots, k} p_j$$

de esta manera, la cota superior $\lambda_{\hat{p}}$ y la cota inferior $\lambda_{\check{p}}$ se alcanzan si $y = \beta_1 q_{\hat{p}}$ e $y = \beta_2 q_{\check{p}}$ respectivamente, por lo que se tiene que

$$\lambda_{\hat{p}} = \max_{y \in \langle q_{p_1}, \dots, q_{p_k} \rangle} \frac{y^T A y}{y^T y}, \quad \lambda_{\check{p}} = \min_{y \in \langle q_{p_1}, \dots, q_{p_k} \rangle} \frac{y^T A y}{y^T y}$$

Si ahora se busca en todos los espacios $S \subseteq \mathbb{R}^n$ tal que $\dim(S) = k$ se tiene

$$\max_{\dim(S)=k} \min_{y \in S} \frac{y^T A y}{y^T y} \geq \min_{y \in \langle q_1, \dots, q_k \rangle} \frac{y^T A y}{y^T y} = \lambda_k$$

donde $y \neq 0$.

Ahora sólo basta obtener la desigualdad en sentido inverso, para así poder terminar la demostración, para lograr esto tómesese un espacio $S \subseteq \mathbb{R}^n$ cualquiera tal que $\dim(S) = k$, dicho espacio siempre interseca a $\langle q_k, \dots, q_n \rangle$. Supóngase ahora que el vector x pertenece simultáneamente a S y a $\langle q_k, \dots, q_n \rangle$, entonces se tiene

$$\min_{0 \neq y \in S} \frac{y^T A y}{y^T y} \leq \frac{x^T A x}{x^T x}, \quad x \in S$$

como también $x \in \langle q_k, \dots, q_n \rangle$ entonces se puede escribir

$$x = \sum_{l=k}^n \alpha_l q_l \in \langle q_k, \dots, q_n \rangle$$

y se tiene que

$$\frac{x^T A x}{x^T x} \leq \max_{z \in \langle q_k, \dots, q_n \rangle} \frac{z^T A z}{z^T z} = \lambda_k, \quad x \in \langle q_k, \dots, q_n \rangle$$

por lo tanto

$$\min_{0 \neq y \in S} \frac{y^T A y}{y^T y} \leq \max_{z \in \langle q_k, \dots, q_n \rangle} \frac{z^T A z}{z^T z} = \lambda_k$$

como lo anterior es válido para cualquier espacio S con $\dim(S) = k$, entonces

$$\max_{\dim(S)=k} \min_{0 \neq y \in S} \frac{y^T A y}{y^T y} \leq \max_{z \in \langle q_k, \dots, q_n \rangle} \frac{z^T A z}{z^T z} = \lambda_k$$

así se tiene que

$$\max_{\dim(S)=k} \min_{0 \neq y \in S} \frac{y^T A y}{y^T y} = \lambda_k$$

QED.

D.3. Teoría de perturbaciones

En [11] se analiza cómo se ven afectados los valores propios y los valores singulares de una matriz, si ella se encuentra sometida a perturbaciones aditivas. En cuanto a la notación, se utiliza $\lambda_k(A)$ para enfatizar que es el k -ésimo valor propio más grande de A , ídem para el valor singular $\sigma_k(A)$.

A continuación se muestran dos resultados importantes en perturbaciones de valores propios y valores singulares.

Teorema D.9 Si $A \in \mathbb{R}^{n \times n}$ y $(A + E) \in \mathbb{R}^{n \times n}$ son matrices simétricas, entonces

$$\lambda_k(A) + \lambda_n(E) \leq \lambda_k(A + E) \leq \lambda_k(A) + \lambda_1(E)$$

para todo $k = 1, \dots, n$.

Ver [11, theorem 8.1.5].

Demostración. Ésta se basa principalmente en la representación *minimax* de los valores propios de una matriz simétrica, teorema D.8. Como A y $A + E$ son simétricas, entonces E también es simétrica y por lo tanto tiene descomposición simétrica de Schur D.5

$$Q^T E Q = \text{diag}(\lambda_1(E), \dots, \lambda_n(E))$$

donde $Q = [q_1 \ \dots \ q_n]$ y $\{q_1, \dots, q_n\}$ son vectores l.i. Sea el vector $x \in S \subseteq \mathbb{R}^n$, entonces x se puede escribir como

$$x = \sum_{l=1}^n \alpha_l q_l$$

Así

$$\frac{x^T E x}{x^T x} = \frac{\sum_{l=1}^n \sum_{m=1}^n \alpha_l \alpha_m q_l^T E q_m}{\sum_{l=1}^n \sum_{m=1}^n \alpha_l \alpha_m q_l^T q_m} = \frac{\sum_{j=1}^n \alpha_j^2 \lambda_j(E)}{\sum_{j=1}^n \alpha_j^2}$$

y por lo tanto se tiene que

$$\lambda_n(E) \leq \frac{x^T E x}{x^T x} \leq \lambda_1(E)$$

lo que implica

$$\frac{x^T A x}{x^T x} + \lambda_n(E) \leq \frac{x^T (A + E) x}{x^T x} \leq \frac{x^T A x}{x^T x} + \lambda_1(E) \quad (\text{D.1})$$

válido $\forall x \in S$ y $\forall S \subseteq \mathbb{R}^n$.

Sea $\tilde{x}_1 \in S \subseteq \mathbb{R}^n$

$$\tilde{x}_1 = \arg \min_{0 \neq x \in S} \frac{x^T A x}{x^T x}$$

entonces evaluando en la desigualdad derecha de (D.1) se tiene

$$\min_{0 \neq x \in S} \frac{x^T(A+E)x}{x^T x} \leq \frac{\tilde{x}_1^T(A+E)\tilde{x}_1}{\tilde{x}_1^T \tilde{x}_1} \leq \frac{\tilde{x}_1^T A \tilde{x}_1}{\tilde{x}_1^T \tilde{x}_1} + \lambda_1(E) \quad (\text{D.2})$$

Sea $\tilde{x}_2 \in S \subseteq \mathbb{R}^n$

$$\tilde{x}_2 = \arg \min_{0 \neq x \in S} \frac{x^T(A+E)x}{x^T x}$$

entonces evaluando en la desigualdad izquierda de (D.1) se tiene

$$\min_{0 \neq x \in S} \frac{x^T A x}{x^T x} + \lambda_n(E) \leq \frac{\tilde{x}_2^T A \tilde{x}_2}{\tilde{x}_2^T \tilde{x}_2} + \lambda_n(E) \leq \frac{\tilde{x}_2^T(A+E)\tilde{x}_2}{\tilde{x}_2^T \tilde{x}_2} \quad (\text{D.3})$$

Juntando (D.3) y (D.2)

$$\begin{aligned} \min_{0 \neq x \in S} \frac{x^T A x}{x^T x} + \lambda_n(E) &\leq \\ \min_{0 \neq x \in S} \frac{x^T(A+E)x}{x^T x} &\leq \min_{0 \neq x \in S} \frac{x^T A x}{x^T x} + \lambda_1(E) \end{aligned} \quad (\text{D.4})$$

válido $\forall S \subseteq \mathbb{R}^n$.

Sea $\hat{S}_1 \subseteq \mathbb{R}^n$ un subespacio de dimensión k

$$\hat{S}_1 = \arg \max_{\dim(S)=k} \min_{0 \neq x \in S} \frac{x^T A x}{x^T x}$$

entonces evaluando en la desigualdad izquierda de (D.4)

$$\begin{aligned} \min_{0 \neq x \in \hat{S}_1} \frac{x^T A x}{x^T x} + \lambda_n(E) &\leq \min_{0 \neq x \in \hat{S}_1} \frac{x^T(A+E)x}{x^T x} \\ &\leq \max_{\dim(S)=k} \min_{0 \neq x \in S} \frac{x^T(A+E)x}{x^T x} \end{aligned} \quad (\text{D.5})$$

Sea $\hat{S}_2 \subseteq \mathbb{R}^n$ un subespacio de dimensión k

$$\hat{S}_2 = \arg \max_{\dim(S)=k} \min_{0 \neq x \in S} \frac{x^T(A+E)x}{x^T x}$$

entonces evaluando en la desigualdad derecha de la ecuación (D.4)

$$\begin{aligned} \min_{0 \neq x \in \hat{S}_2} \frac{x^T(A+E)x}{x^T x} &\leq \min_{0 \neq x \in \hat{S}_2} \frac{x^T A x}{x^T x} + \lambda_1(E) \\ &\leq \max_{\dim(S)=k} \min_{0 \neq x \in S} \frac{x^T A x}{x^T x} + \lambda_1(E) \end{aligned} \quad (\text{D.6})$$

Juntando (D.5) y (D.6) se tiene

$$\max_{\dim(S)=k} \min_{0 \neq x \in S} \frac{x^T A x}{x^T x} + \lambda_n(E) \leq$$

$$\begin{aligned} \max_{\dim(S)=k} \min_{0 \neq x \in S} \frac{x^T(A+E)x}{x^T x} &\leq \\ \max_{\dim(S)=k} \min_{0 \neq x \in S} \frac{x^T A x}{x^T x} &+ \lambda_1(E) \end{aligned}$$

Considerando el teorema D.8 se completa la demostración. **QED.**

Para matrices reales, el problema de perturbaciones en valores singulares se puede tratar a partir del problema perturbacional de valores propios, ya que el problema de valores singulares es equivalente a un problema de valores propios de una matriz simétrica.

Sea $A \in \mathbb{R}^{m \times n}$ con $m \geq n$ y valores singulares $\sigma_1(A), \dots, \sigma_n(A)$, que tiene la siguiente descomposición en valores singulares según D.3

$$A = \begin{bmatrix} U_n & U_{m-n} \end{bmatrix} \begin{bmatrix} \Sigma_n \\ O_{m-n} \end{bmatrix} V^T$$

considérese también la matriz $E \in \mathbb{R}^{m \times n}$ con $m \geq n$ y valores singulares $\sigma_1(E), \dots, \sigma_n(E)$, la cual tiene la siguiente descomposición en valores singulares según D.3

$$E = \begin{bmatrix} W_n & W_{m-n} \end{bmatrix} \begin{bmatrix} \Xi_n \\ O_{m-n} \end{bmatrix} Y^T$$

Claramente, las matrices A y E no son simétricas, sin embargo la matriz $B \in \mathbb{R}^{(m+n) \times (m+n)}$ y la matriz $D \in \mathbb{R}^{(m+n) \times (m+n)}$ sí lo son

$$B = \begin{bmatrix} O_{n \times m} & A^T \\ A & O_{m \times n} \end{bmatrix} \quad D = \begin{bmatrix} O_{n \times m} & E^T \\ E & O_{m \times n} \end{bmatrix}$$

donde $O_{n \times m}$ y $O_{m \times n}$ son matrices nulas. Las matrices B y D tienen la particularidad de que tienen la siguiente descomposición simétrica de Schur D.5

$$Q^T B Q = \text{diag}(\sigma_1(A), \dots, \sigma_n(A), -\sigma_1(A), \dots, -\sigma_n(A), \underbrace{0, \dots, 0}_{m-n})$$

$$R^T D R = \text{diag}(\sigma_1(E), \dots, \sigma_n(E), -\sigma_1(E), \dots, -\sigma_n(E), \underbrace{0, \dots, 0}_{m-n})$$

$$\begin{aligned} Q &= \frac{1}{\sqrt{2}} \begin{bmatrix} V & V & O_{m-n}^T \\ U_n & -U_n & \sqrt{2}U_{m-n} \end{bmatrix} \\ R &= \frac{1}{\sqrt{2}} \begin{bmatrix} Y & Y & O_{m-n}^T \\ W_n & -W_n & \sqrt{2}W_{m-n} \end{bmatrix} \end{aligned}$$

de aquí se puede ver que los valores propios de las matrices simétricas B y D están íntimamente relacionados con los valores singulares de las matrices A y

E , es por esto que se puede obtener el siguiente corolario a partir del resultado obtenido en el teorema D.9

Corolario D.1 Sean $A \in \mathbb{R}^{m \times n}$ y $(A + E) \in \mathbb{R}^{m \times n}$ con $m \geq n$, entonces se tiene

$$|\sigma_k(A + E) - \sigma_k(A)| \leq \sigma_1(E) = \|E\|_2, \quad \forall k = 1, \dots, n$$

Ver [11, corollary 8.6.2].

Demostración. Según el teorema D.9 se tiene

$$\lambda_k(B) + \lambda_n(D) \leq \lambda_k(B + D) \leq \lambda_k(B) + \lambda_1(D)$$

$$\lambda_n(D) \leq \lambda_k(B + D) - \lambda_k(B) \leq \lambda_1(D)$$

entonces

$$|\lambda_k(B + D) - \lambda_k(B)| \leq \max\{|\lambda_1(D)|, |\lambda_n(D)|\}$$

para todo $k = 1, \dots, m + n$, lo que equivale a

$$|\sigma_l(A + E) - \sigma_l(E)| \leq \sigma_1(E) = \|E\|_2$$

para todo $l = 1, \dots, n$. **QED.**

D.4. Propiedades de rango

Las demostraciones de la mayoría de los teoremas y corolarios citados en esta sección se pueden encontrar en [15, sec.3.4].

Definición D.1 Sea A una matriz de tamaño $m \times n$, entonces se definen

1. **rango fila** de A $\text{rango}_f(A)$ como el número de filas linealmente independientes de la matriz A
2. **rango columna** de A $\text{rango}_c(A)$ como el número de columnas linealmente independientes de la matriz A
3. **rango** de A

$$\text{rango}(A) := \text{rango}_f(A) = \text{rango}_c(A)$$

Nota D.1 Si A es una matriz $m \times n$, entonces el rango de A puede ser visto también como el mayor entero positivo r , de manera que existe una submatriz $r \times r$ de A , formada al eliminar $(m - r)$ filas y $(n - r)$ columnas, la cual es no singular. Ver [3, apéndice A, sección 7].

Teorema D.10 Sean las matrices A ($m \times n$) y B ($n \times p$), entonces se tienen las siguientes cotas

1. $\text{rango}(A) \leq \min(m, n)$
2. $\text{rango}(AB) \leq \min\{\text{rango}(A), \text{rango}(B)\}$

Teorema D.11 Sea A una matriz ($m \times n$), entonces se tiene que

1. El rango de A es invariante al aplicar operaciones elementales fila a la matriz
2. El rango de A es invariante al aplicar operaciones elementales columna a la matriz
3. $\text{rango}(A) = \text{rango}(A^T)$

Corolario D.2 Sea A una matriz ($m \times n$), y además P ($m \times m$) y Q ($n \times n$) matrices invertibles, entonces

1. $\text{rango}(A) = \text{rango}(PAQ)$
2. A es invertible sólo si $n = m$ y $\text{rango}(A) = n$

Corolario D.3 Sea A una matriz $m \times n$, entonces se tiene que

$$\text{rango}(A) = \text{rango}(A^H A) = \text{rango}(AA^H)$$

Corolario D.4 Sean las matrices A y B , ambas de tamaño ($m \times n$), con $n \leq m$ y $\text{rango}(B) = n$, entonces

$$\text{rango}(B^H A) = \text{rango}(A)$$

Demostración. Considere A y B en sus descomposiciones de valores singulares

$$A = U \Sigma V^H = \begin{bmatrix} U_n & U_{m-n} \end{bmatrix} \begin{bmatrix} \Sigma_n \\ O \end{bmatrix} V^H$$

$$B = W \Xi Y^H = \begin{bmatrix} W_n & W_{m-n} \end{bmatrix} \begin{bmatrix} \Xi_n \\ O \end{bmatrix} Y^H$$

donde $U, W \in \mathbb{C}^{m \times m}$, $V, Y \in \mathbb{C}^{n \times n}$ son matrices unitarias y $U_n, W_n \in \mathbb{C}^{m \times n}$, $U_{m-n}, W_{m-n} \in \mathbb{C}^{m \times (m-n)}$, $\Sigma, \Xi \in \mathbb{R}^{m \times n}$, $\Sigma_n, \Xi_n \in \mathbb{R}^{n \times n}$ y $O \in \mathbb{R}^{(m-n) \times n}$.

Al multiplicar $B^H A = (Y \Xi_n^H W_n^H U_n) \Sigma_n (V^H)$ se puede notar que como V^H es invertible, entonces por el

corolario D.2 sólo basta demostrar que el producto $Y \Xi_n^H W_n^H U_n \in \mathbb{C}^{n \times n}$ es invertible. Dado que Ξ_n es invertible, ya que $\text{rango}(B) = n$, además Y es invertible por definición y $W_n^H U_n$ es unitaria, entonces el producto $Y \Xi_n^H W_n^H U_n \in \mathbb{C}^{n \times n}$ también es invertible. **QED.**

D.5. Matrices especiales

Definición D.2 La matriz $A \in \mathbb{C}^{n \times n}$ es “definida positiva” si se cumplen las siguientes condiciones

1. $\forall x \in \mathbb{C}^n, x^H A x \geq 0$
2. $x_0^H A x_0 = 0 \Rightarrow x_0 = 0$, donde $x_0 \in \mathbb{C}^n$

Teorema D.12 Teorema de Sylvester. Sean los polinomios $p(x)$ y $q(x)$ definidos por

$$\begin{aligned} p(x) &= p_0 + p_1 x + \cdots + p_m x^m \\ q(x) &= q_0 + q_1 x + \cdots + q_n x^n \end{aligned}$$

donde p_m y q_n no son simultáneamente cero. Los polinomios $p(x)$ y $q(x)$ son relativamente primos (coprimos) si y sólo si $\det \mathcal{S}(p, q) \neq 0$, donde la matriz de Sylvester $\mathcal{S}(p, q)$ está definida por

$$\begin{bmatrix} p_0 & p_1 & \cdots & p_{m-1} & p_m & 0 & \cdots & 0 \\ 0 & p_0 & p_1 & \cdots & p_{m-1} & p_m & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & p_0 & p_1 & \cdots & p_{m-1} & p_m \\ \hline q_0 & q_1 & \cdots & q_{n-1} & q_n & 0 & \cdots & 0 \\ 0 & q_0 & q_1 & \cdots & q_{n-1} & q_n & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & q_0 & q_1 & \cdots & q_{n-1} & q_n \end{bmatrix}$$

La submatriz relacionada al polinomio $p(x)$ es $\bar{n} \times (\bar{m} + \bar{n})$ y la submatriz relacionada al polinomio $q(x)$ es $\bar{m} \times (\bar{m} + \bar{n})$, así la matriz $\mathcal{S}(p, q)$ es $(\bar{m} + \bar{n}) \times (\bar{m} + \bar{n})$, donde $\text{mín}(\bar{m} - m, \bar{n} - n) = 0$.

Para una demostración de este teorema véase [19, lemma A3.1].

Apéndice E

Demostraciones complementarias

E.1. Capítulo 1

Demostración E.1 (teorema 1.1) Sólo basta tener en cuenta las siguientes definiciones

$$P(|X - \mu| \geq \varepsilon) = \int_{|x-\mu| \geq \varepsilon} f_X(x) dx$$

$$\sigma^2 = \int_{\mathbb{R}} (x - \mu)^2 f_X(x) dx$$

Así se puede acotar σ^2

$$\sigma^2 \geq \int_{|x-\mu| \geq \varepsilon} (x - \mu)^2 f_X(x) dx$$

$$\geq \int_{|x-\mu| \geq \varepsilon} \varepsilon^2 f_X(x) dx$$

Reemplazando

$$\sigma^2 \geq \varepsilon^2 P(|X - \mu| \geq \varepsilon)$$

y despejando se obtiene

$$P(|X - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}$$

QED.

Demostración E.2 (definición 1.11) Nótese que

$$m^T \Pi_s^\infty(p) m = \overline{\langle (v_s^p(k)^T m)^2 \rangle_{k \in \mathbb{Z}}}$$

donde

$$v_s^p(k)^T = [s(k) \quad \cdots \quad s(k-p+1)]$$

$$m^T = [m_0 \quad \cdots \quad m_{p-1}]$$

Nótese además que el producto $v_s^p(k)^T m$ puede ser escrito como

$$v_s^p(k)^T m = M_p(q) s(k)$$

donde

$$M_p(q) = m_0 + m_1 q^{-1} + \cdots + m_{p-1} q^{-p+1}$$

Ahora, mediante la transformada de Fourier inversa, se puede establecer una relación entre el promedio $\overline{\langle [M_p(q) s(k)]^2 \rangle_{k \in \mathbb{Z}}}$ y el espectro de $s(k)$

$$\overline{\langle [M_p(q) s(k)]^2 \rangle_{k \in \mathbb{Z}}} = \frac{1}{2\pi} \int_{-\pi}^{\pi} |M_p(e^{j\theta})|^2 \Phi_{ss}^\infty(e^{j\theta}) d\theta$$

Así se tienen finalmente las siguientes equivalencias

$$m^T \Pi_s^\infty(p) m = 0 \Leftrightarrow \forall \theta, |M_p(e^{j\theta})|^2 \Phi_{ss}^\infty(e^{j\theta}) = 0$$

$$m = 0 \Leftrightarrow \forall \theta, M_p(e^{j\theta}) = 0$$

QED.

Demostración E.3 (teorema 1.3) Nótese que La demostración es por contradicción. Supóngase que $s(k)$ no es p.e. de orden p , entonces existe un filtro $M_p(q) \neq 0$ que satisface

$$\forall \theta, |M_p(e^{j\theta})|^2 \Phi_{ss}^\infty(e^{j\theta}) = |M_p(e^{j\theta})|^2 \frac{\Phi_{uu}^\infty(e^{j\theta})}{|C(e^{j\theta})|^2} = 0$$

donde $C(q) \neq 0$, entonces

$$\forall \theta, |M_p(e^{j\theta})|^2 \Phi_{uu}^\infty(e^{j\theta}) = 0$$

lo que es una contradicción, pues $u(k)$ es a.p.e. de orden p . **QED.**

Demostración E.4 (teorema 1.4, caso infinito)

Considérese el siguiente vector

$$\gamma_p = [d_0 \quad \cdots \quad d_{\hat{n}_b} \quad c_1 \quad \cdots \quad c_{\hat{n}_a}]^T$$

Ahora, si $u(k)$ es p.e. de orden $\hat{n}_a + \hat{n}_b + 1$, entonces según la definición 1.11 $\forall M_{\hat{n}_a + \hat{n}_b + 1}(q)$ se cumple

$$\left| M_{\hat{n}_a + \hat{n}_b + 1}(e^{j\theta}) \right|^2 \Phi_{uu}^\infty(e^{j\theta}) = 0 \Rightarrow M_{\hat{n}_a + \hat{n}_b + 1}(e^{j\theta}) = 0$$

y por lo tanto, se cumple en particular para

$$\left| \tilde{G}(e^{j\theta}) \right|^2 \Phi_{uu}^\infty(e^{j\theta}) = 0 \Rightarrow \tilde{G}(e^{j\theta}) = 0$$

ya que

$$\tilde{G}(q) = G_1(q) - G_2(q) = \frac{A_2(q)B_1(q) - A_1(q)B_2(q)}{A_1(q)A_2(q)}$$

tiene a lo más $\hat{n}_a + \hat{n}_b + 1$ ceros sobre la circunferencia unitaria. De esta manera la demostración queda completa, resultando que Z_{uy}^∞ es s.i. para la estructura. **QED.**

E.2. Capítulo 2

Demostración E.6 (teorema 2.4) Considérese la planta de la ecuación (2.19). Considérese también que $u(k) = \delta_K(k)$, que es l.e.¹ de orden

$$p = \hat{n}_b + \hat{n}_a + 1 - \min(\hat{n}_b - n_b, \hat{n}_a - n_b)$$

en el intervalo $[-M, M]$, si se escoge

$$M \geq (p - 1)$$

Así, tomando los datos $Z_{ux}^{L,M}$, con $L = \max(\hat{n}_b, \hat{n}_a)$, se puede construir la matriz $\Omega_{ux}^M(\hat{n}_b, \hat{n}_a)$, cuyo rango está dado por el lema 2.3.

Ahora, se verificará la estructura de la matriz $\Omega_{ux}^M(\hat{n}_b, \hat{n}_a)$, considerando que la respuesta al experimento anterior es exactamente $x(k) = g(k)$, es decir, la respuesta a impulso del sistema.

$$\Omega_{ux}^M(\hat{n}_b, \hat{n}_a) = \begin{bmatrix} 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 \\ \hline 1 & \cdots & 0 & g(-1) & \cdots & g(-\hat{n}_a) \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 1 & g(\hat{n}_b - 1) & \cdots & g(\hat{n}_b - \hat{n}_a) \\ 0 & \cdots & 0 & g(\hat{n}_b) & \cdots & g(\hat{n}_b + 1 - \hat{n}_a) \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & g(M - 1) & \cdots & g(M - \hat{n}_a) \end{bmatrix}$$

¹definición 1.10

Dado que las primeras M filas de $\Omega_{ux}^M(\hat{n}_b, \hat{n}_a)$ no contribuyen a determinar su rango, entonces se eliminan dichas filas construyéndose así la matriz

$$\tilde{\Omega}_g^M(\hat{n}_b, \hat{n}_a) = \begin{bmatrix} 1 & \cdots & 0 & g(-1) & \cdots & g(-\hat{n}_a) \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 1 & g(\hat{n}_b - 1) & \cdots & g(\hat{n}_b - \hat{n}_a) \\ 0 & \cdots & 0 & g(\hat{n}_b) & \cdots & g(\hat{n}_b + 1 - \hat{n}_a) \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & g(M - 1) & \cdots & g(M - \hat{n}_a) \end{bmatrix}$$

donde $M \geq (p - 1)$, resultando así que

$$\text{rango } \Omega_{ux}^M(\hat{n}_b, \hat{n}_a) = \text{rango } \tilde{\Omega}_g^M(\hat{n}_b, \hat{n}_a)$$

el cual está dado por el lema 2.3.

Considere ahora la misma planta en (2.19), cuya nueva señal de entrada se diseña de manera que $\varphi_{uu}^\infty(\tau) = \delta_K(\tau)$. Considere también que se utiliza el filtro instrumental $F(q) = q^{-\hat{n}_b}$. De esta manera se miden los datos Z_{uz}^∞ y Z_{ux}^∞ , con los que se construye la matriz $\Lambda_{zux}^\infty(\hat{n}_b, \hat{n}_a)$. Teniendo en cuenta la partición de las ecuaciones (2.26), (2.27) y (2.28), y además

$\varphi_{zu}^\infty(\tau) = f(\tau)$, $\varphi_{ux}^\infty(\tau) = g(-\tau)$, $\varphi_{zx}^\infty(\tau) = g(\hat{n}_b - \tau)$ entonces estructura de la matriz $\Lambda_{zux}^\infty(\hat{n}_b, \hat{n}_a)$ resulta

$$\Lambda_{zux}^\infty(\hat{n}_b, \hat{n}_a) = \begin{bmatrix} 1 & \cdots & 0 & g(-1) & \cdots & g(-\hat{n}_a) \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 1 & g(\hat{n}_b - 1) & \cdots & g(\hat{n}_b - \hat{n}_a) \\ \hline 0 & \cdots & 0 & g(\hat{n}_b) & \cdots & g(\hat{n}_b + 1 - \hat{n}_a) \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & g(\hat{n}_b + \hat{n}_a - 1) & \cdots & g(\hat{n}_b) \end{bmatrix}$$

A partir de este último resultado se puede ver que

$$\tilde{\Omega}_g^{\hat{n}_b + \hat{n}_a}(\hat{n}_b, \hat{n}_a) = \Lambda_{zux}^\infty(\hat{n}_b, \hat{n}_a)$$

donde $M = \hat{n}_b + \hat{n}_a$, y por consiguiente se debe cumplir que

$$\hat{n}_b + \hat{n}_a \geq (p - 1) \Leftrightarrow \min(\hat{n}_b - n_b, \hat{n}_a - n_b) \geq 0$$

En resumen, si $\varphi_{uu}^\infty(\tau) = \delta_K(\tau)$, $F(q) = q^{-\hat{n}_b}$ y $\min(\hat{n}_b - n_b, \hat{n}_a - n_b) \geq 0$, entonces

$$\text{rango } \Lambda_{zux}^\infty(\hat{n}_b, \hat{n}_a) = \max(\hat{n}_b + n_a, n_b + \hat{n}_a) + 1$$

QED.

Apéndice F

Programas

LAS simulaciones de los ejemplos presentados en esta tesis fueron desarrolladas con el programa de cálculo numérico SCILAB 4.0, desarrollado por el INRIA (Institut National de Recherche en Informatique et Automatique) y la ENPC (École Nationale des Ponts et Chaussées). Este software es de libre uso y distribución, bajo los términos de la licencia <http://www.scilab.org/legal/license.html>. Para más información, véase <http://www.scilab.org/>.

F.1. Funciones comunes

F.1.1. funciones.sci

```
// funciones.sci
// version 2008.08.19

function U = omegau(L,u,nbhat)
// u      : sennal
// nbhat  : grado del polinomio
// L      : numero de datos usados para retardo
//
//          1          L+1-nbhat          L+1          2M+1+L
// u = [u(-M-L) ... u(-M-nbhat) ... u(-M) ... u(M)]
uv = u(L+1:$);
uh = u(L+1:-1:L+1-nbhat);
U = toeplitz(uv,uh);
endfunction

function X = omegax(L,x,nahat,i)
// x      : sennal
// nahat  : grado del polinomio
// L      : numero de datos usados para retardo
// i      : min([nbhat-nb nahat-na])
//
//          1          L+1-nahat          L-i          2M+L-i ... 2M+1+L
```

```
// x = [x(-M-L) ... x(-M-nahat) ... x(-M-i-1) ... x(M-i-1) ... x(M)]
xv = x(L-i:$-i-1);
xh = x(L-i:-1:L+1-nahat);
X = toeplitz(xv,xh);
endfunction
```

```
function o = omega(L,u,x,nbhat,nahat,i)
// nbhat : grado del numerador
// nahat : grado del denominador
// i      : min([nbhat-nb nahat-na])
U = omegau(L,u,nbhat);
X = omegax(L,x,nahat,i);
o = [U X];
endfunction
```

```
function u = gencos(p,N,K)
// p : componentes espectrales
// N : periodo fundamental
// K : largo de la secuencia
k = (1:K);
theta0 = 2*%pi/N;
i = modulo(p+1,2):floor(p/2);
// A = 4*rand(1,ceil(p/2),'uniform')-2;
A = ones(1,ceil(p/2));
u = A*cos(i'*theta0*k);
endfunction
```

```
function sv = hsv(ssG)
// ssG : representacion de estado de G
// sv   : valores singulares de hankel de G
P = ctr_gram(ssG);
Q = obs_gram(ssG);
sv = sort(sqrt(spec(P*Q)));
endfunction
```

```
function P = Pi(L,u,p)
// u : sennal
// p : persistencia a probar
// L : numero de datos usados para retardo
U = omegau(L,u,p-1);
s = size(U);
P = U'*U/s(1);
endfunction
```

F.1.2. dsp.sci

```
// dsp.sci
// version 2008.07.07
```

```
function [Y,theta] = espectro(y)
```

```

// ESPECTRO Obtiene el modulo de la transformada de Fourier de tiempo discreto.
//   [Y,theta] = ESPECTRO(y) calcula |Y(exp(j*theta))|, donde
//
//       Y(exp(j*theta)) = sum_{k=1}^{inf} y(k)*exp(-j*theta*k)
//
//   y entrega ademas el vector de eje de frecuencia angular normalizada
//   theta = omega*dt.

N = length(y);
Y = fftshift(abs(fft(y)));
dtheta = 2*pi/N;
theta = 1:N;
if modulo(N,2)==0 then theta = theta-(N+2)/2;
    else theta = theta-(N+1)/2;
end
theta = theta*dtheta;
endfunction

function C = espectroC(Y,dtheta)
// ESPECTROC Obtiene los coeficientes de la serie de Fourier de tiempo discreto.
//   C = ESPECTROC(Y,dtheta) calcula |C(theta)| a partir de |Y(exp(j*theta))|
//   y de dtheta, donde
//
//       C(theta) = Y(exp(j*theta))*dtheta/(2*pi)
//       y(k) = 1/(2*pi) int_{-pi}^{pi} Y(exp(j*theta))*exp(j*theta*k) dtheta

C = Y*dtheta/(2*pi);
endfunction

```

F.2. Ejemplos capítulo 1

F.2.1. Ejemplo 1.2: persis.sce

```

// persis.sce
// version 2008.08.05

clear; clc;
getf('funciones.sci');
getf('dsp.sci');
rand('seed',0);

p0 = 5;    // persistencia a probar
M = p0;
N = 2*M+1; // periodo fundamental
L = p0-1; // datos para retardo

printf(' p0 (2M+1) p det{Pu(p0)}\n');
printf('-----\n');
for p=1:2*p0,

```

```

u = gencos(p,N,N+L);
Pu = Pi(L,u,p0);
d = det(Pu);
printf(' %2d    %2d    %2d    %.2e\n',p0,N,p,d);
end
printf('-----\n\n');

[U,theta] = espectro(u(L+1:$));
cU = espectroC(U,theta(2)-theta(1));
scf(0); clf(0);
plot2d3(theta,cU);
xtitle('Componentes espectrales de u(k), con p=10','theta','Amplitud');

```

F.2.2. Ejemplo 1.4: sufinf-sr.sce

```

// sufinf-sr.sce
// version 2008.08.05

clc; clear;
z = poly(0,'z'); rand('seed',0);
getf('funciones.sci'); getf('dsp.sci');

// planta
B = 0.3; A = (z^2 - 2*(0.8)*cos(%pi/4)*z + (0.8)^2)*(z-0.7);
na = 3; nb = 3;
printf('G(1) = %f\n',horner(B/A,1));
disp(roots(A),'polos = ');
printf('\n');

// transferencias
sysG = syslin('d',B,A);
tran = 5*ceil(-inv(log(max(abs(roots(A))))));
printf('transiente = %d\n\n',tran);

// datos
M = 5; L = 7; p = 9;
N = 2*M+1+L;

// mediciones estado estacionario
u = gencos(p,2*M+1,tran+N);
x = flts(u,sysG);
u = u(tran+1:$);
x = x(tran+1:$);

for nbhat = 1:7,
    for nahat = 1:7,
        i = min([nbhat-nb nahat-na]);
        if (i<=0) then
            Oux = omega(L,u,x,nbhat,nahat,0);
        else

```

```

        Oux = omega(L,u,x,nbhat,naht,i);
    end
    clear Lxux;
    Lxux = Oux'*Oux/(2*M+1);
    printf('%.0e ',min( abs(spec(Lxux)) ) );
    end
    printf('\n');
end

[U,theta] = espectro(u(L+1:$));
cU = espectroC(U,theta(2)-theta(1));
scf(0); clf(0);
plot2d3(theta,cU);
xtitle('Componentes espectrales de u(k), con p=9','theta','Amplitud');

```

F.2.3. Ejemplo 1.5: sufinf-cr.sce

```

// sufinf-cr.sce
// version 2008.07.16

clear;
q = poly(0,'q'); rand('seed',0);
getf('funciones.sci'); getf('dsp.sci');

// planta
B = 0.3; A = (q^2 - 2*(0.8)*cos(%pi/4)*q + (0.8)^2)*(q-0.7);
na = 3; nb = 3;
printf('G(1) = %f\n',horner(B/A,1));
disp(roots(A),'polos = ');
printf('\n');

// ruido
D = 0.4*q; C = (q - 0.6);

// transferencias
sysG = syslin('d',B,A);
sysH = syslin('d',D,C);
tran = 5*ceil(-inv(log(max(abs([roots(A); roots(C)]))))) );
printf('transiente = %d\n\n',tran);

// datos
M = 500; L = 7; p = 7;
N = 2*M+1+L;

// mediciones estado estacionario
u = gencos(p,p,tran+N);
x = flts(u,sysG);
u = u(tran+1:$);
x = x(tran+1:$);

```

```

e = rand(1,tran+N,'normal');
v = flts(e,sysH);
v = v(tran+1:$);

y = x + v;

for nbhat = 1:7,
    for nahat = 1:7,
        i = min([nbhat-nb nahat-na]);
        Ouy = omega(L,u,y,nbhat,nahat,0);
        clear Lyuy;
        Lyuy = Ouy'*Ouy/(2*M+1);
        printf('%0e ',min(abs(spec(Lyuy))) );
    end
    printf('\n');
end

[U,theta] = espectro(u);
cU = espectroC(U,theta(2)-theta(1));
scf(0); clf(0);
plot2d3(theta,cU);
xtitle('Componentes espectrales de u(k), con p=7','theta','Amplitud');

```

F.3. Ejemplos capítulo 2

F.3.1. Ejemplo 2.1: Lxux.sce

```

// Lxux.sce
// version 2008.08.19

clc; clear;
q = poly(0,'q'); rand('seed',0);
getf('funciones.sci'); getf('dsp.sci');

// planta
n=3; B = 0.3; A = (q^2 - 2*(0.8)*cos(%pi/4)*q + 0.8^2)*(q-0.7);

// transferencias
sysG = syslin('d',B,A);
tran = 5*ceil(-inv(log(max(abs(roots(A))))));
printf('transiente = %d\n\n',tran);

// datos
nhatmax = 2*n;
p = 2*(nhatmax+1);
M = nhatmax+1; L = nhatmax+1;
N = 2*M+1+L;

```

```

// mediciones estado estacionario
u = gencos(p,2*M+1,tran+N);
x = flts(u,sysG);
u = u(tran+1:$);
x = x(tran+1:$);

[U,theta] = espectro(u(L+1:$));
cU = espectroC(U,theta(2)-theta(1));
scf(0); clf(0);
plot2d3(theta,cU);
xtitle('Componentes espectrales de u(k)', 'theta', 'Amplitud');

printf('\n vpmín(Lxux)    DR\n');
printf('-----\n');
for nhat = 1:nhatmax,
    Oux = omega(L,u,x,nhat,nhat,0);
    Lxux = Oux'*Oux/(2*M+1);
    vpmín = min(abs(spec(Lxux)));

    OOux = omega(L,u,x,nhat+1,nhat+1,0);
    LLxux = OOux'*OOux/(2*M+1);
    DR = abs(det(Lxux)/det(LLxux));

    printf('%2d  %.2e  %.2e\n',nhat,vpmín,DR);
end

```

F.3.2. Ejemplo 2.2: Lyuy.sce

```

// Lyuy.sce
// version 2008.08.17

clc; clear;
q = poly(0,'q'); rand('seed',0);
getf('funciones.sci'); getf('dsp.sci');
stacksize(50e6);

// planta
n=3; B = 0.3; A = (q^2 - 2*(0.8)*cos(%pi/4)*q + 0.8^2)*(q-0.7);

// ruido
D = 0.4*q; C = (q - 0.6);

// transferencias
sysG = syslin('d',B,A);
sysH = syslin('d',D,C);
tran = 5*ceil(-inv(log(max(abs([roots(A); roots(C)]))))) );
printf('transiente = %d\n\n',tran);

// datos
nhatmax = 2*n;

```

```

M = [1e3 5e3 1e4 5e4]; Mmax = max(M);
L = nhatmax+1;
Nmax = 2.*Mmax+1+L;

// mediciones estado estacionario
u0 = rand(1,tran+Nmax,'normal');
x0 = flts(u0,sysG);

for Mi = M;
    N = 2*Mi+1+L;
    u = u0(tran+1:tran+N);
    x = x0(tran+1:tran+N);

    e = rand(1,tran+N,'normal');
    v = flts(e,sysH);
    v = v(tran+1:$);

    y = x + v;

    SNR = 20*log10(stdev(x)/stdev(v));

    printf('M = %.0e, SNR = %.2f\n',Mi,SNR);
    printf('~n vpmín(Lyuy)   DR\n');
    printf('-----\n');
    for nhat = 1:nhatmax,
        Ouy = omega(L,u,y,nhat,nhat,0);
        Lyuy = Ouy'*Ouy/(2*Mi+1);
        vpmín = min(abs(spec(Lyuy)));

        OOuy = omega(L,u,y,nhat+1,nhat+1,0);
        LLyuy = OOuy'*OOuy/(2*Mi+1);
        DR = abs(det(Lyuy)/det(LLyuy));

        printf('%2d  %.2e  %.2e\n',nhat,vpmín,DR);
    end
    printf('\n');
end

```

F.3.3. Ejemplo 2.3: Lzuy.sce

```

// Lzuy.sce
// version 2008.08.19

clc; clear;
q = poly(0,'q'); rand('seed',0);
getf('funciones.sci'); getf('dsp.sci');
stacksize(50e6);

// planta
n=3; B = 0.3; A = (q^2 - 2*(0.8)*cos(%pi/4)*q + 0.8^2)*(q-0.7);

```

```

// ejemplo 3, (2002) al-smadi, wilkes "robust and accurate..."
// n=3; B = q^3 - 0.45*q^2 - q^1; A = q^3 - 2.2*q^2 + 1.77*q - 0.52;

// ejemplo 4, (1993) liang, wilkes, cadzow "arma model order..."
// n=6; B = q^6 + 0.3452*q^5 + 0.53*q^4 + 0.3985*q^3 + 0.8138*q^2;
// A = q^6 + 0.7907*q^5 + 0.042*q^4 - 0.5556*q^3 - 0.0247*q^2 + 0.3846*q + 0.3026;

// ejemplo 1, (1996) al-smadi, wilkes "on estimating arma model orders"
// n=7; B = q^7 - 0.1292*q^6 - 0.3893*q^5 + 0.3163*q^4 + 0.2549*q^3 - 0.5823*q^2;
// A = q^7 - 0.2459*q^6 - 0.3786*q^5 - 0.0707*q^4 - 0.0944*q^3 - 0.2881*q^2 + 0.107*q^1 + 0.4695;

// ruido
D = 0.4*q; C = (q - 0.6);

// transferencias
sysG = syslin('d',B,A);
sysH = syslin('d',D,C);
tran = 5*ceil(-inv(log(max(abs([roots(A); roots(C)]))))));
printf('transiente = %d\n\n',tran);

// datos
nhatmax = 2*n;
M = [1e3 5e3 1e4 5e4];
Mmax = max(M); L = nhatmax+1;
Nmax = 2*Mmax+1+L;

// mediciones estado estacionario
u0 = rand(1,tran+Nmax,'normal');
x0 = flts(u0,sysG);

for Mi = M,
    N = 2*Mi+1+L;

    u = u0(tran+1:tran+N);
    x = x0(tran+1:tran+N);

    e = rand(1,tran+N,'normal');
    v = flts(e,sysH);
    v = v(tran+1:$);

    y = x + v;

    SNR = 20*log10(stdev(x)/stdev(v));

    printf('M = %.0e, SNR = %.2f\n',Mi,SNR);
    printf('~n vpmín(Lzux) vpmín(Lzuy)   IDR   | ||Lzuy|| \n');
    printf('-----+-----\n');
    for nhat = 1:2*n,
        // filtro instrumental
        m = nhatmax+1;
        F = q^(-nhat)+q^(-m);

```

```

// F = q^(-m)+q^(-2*m);
sysF = syslin('d',F);
z = flts(u0,sysF);
z = z(tran+1:tran+N);

// identificacion
Ouz = omega(L,u,z,nhat,nhat,0);

Oux = omega(L,u,x,nhat,nhat,0);
Lzux = Ouz'*Oux/(2*Mi+1);
clear Oux;
vpminLzux = min(abs(spec(Lzux)));

Ouy = omega(L,u,y,nhat,nhat,0);
Lzuy = Ouz'*Ouy/(2*Mi+1);
clear Ouz; clear Ouy;
vpminLzuy = min(abs(spec(Lzuy)));
normLzuy = norm(Lzuy);

OOuz = omega(L,u,z,nhat+1,nhat+1,0);
OOuy = omega(L,u,y,nhat+1,nhat+1,0);
LLzuy = OOuz'*OOuy/(2*Mi+1);
clear OOuz; clear OOuy;
IDR = abs(det(Lzuy)/det(LLzuy));

printf('%2d  %.2e  %.2e  %.2e | %.2e\n',nhat,vpminLzux,vpminLzuy,IDR,normLzuy);
end
printf('\n');
end

sz = stacksize();
printf('\nmemoria total: %d, memoria usada: %d\n',sz(1),sz(2));
stacksize(sz(2)*2);

```

F.3.4. Ejemplo 2.4: Lxux2.sce

```

// Lxux2.sce
// version 2008.08.23

clc; clear;
q = poly(0,'q'); rand('seed',0);
getf('funciones.sci'); getf('dsp.sci');
stacksize(50e6);

// planta
// n=3; B = 0.3; A = (q^2 - 2*(0.8)*cos(%pi/4)*q + 0.8^2)*(q-0.7);

// ejemplo 3, (2002) al-smadi, wilkes "robust and accurate..."
n=3; B = q^3 - 0.45*q^2 - q^1; A = q^3 - 2.2*q^2 + 1.77*q - 0.52;

```

```

// ejemplo 1, (1993) liang, wilkes, cadzow "arma model order..."
// n=6; B = q^6 - 0.2997*q^5 + 0.4147*q^4 - 0.2794*q^3 + 0.4973*q^2;
// A = q^6 - 1.2798*q^5 + 0.7805*q^4 - 0.1635*q^3 + 0.7566*q^2 - 1.0621*q + 0.7821;

// ejemplo 4, (1993) liang, wilkes, cadzow "arma model order..."
// n=6; B = q^6 + 0.3452*q^5 + 0.53*q^4 + 0.3985*q^3 + 0.8138*q^2;
// A = q^6 + 0.7907*q^5 + 0.042*q^4 - 0.5556*q^3 - 0.0247*q^2 + 0.3846*q + 0.3026;

// ejemplo 1, (1996) al-smadi, wilkes "on estimating arma model orders"
// n=7; B = q^7 - 0.1292*q^6 - 0.3893*q^5 + 0.3163*q^4 + 0.2549*q^3 - 0.5823*q^2;
// A = q^7 - 0.2459*q^6 - 0.3786*q^5 - 0.0707*q^4 - 0.0944*q^3 - 0.2881*q^2 + 0.107*q^1 + 0.4695;

// transferencias
sysG = syslin('d',B,A);
tran = 5*ceil(-inv(log(max(abs(roots(A))))));
printf('transiente = %d\n\n',tran);

// datos
nhatmax = 2*n;
M = 500; L = nhatmax;
N = 2*M+1+L;

// mediciones estado estacionario
u = rand(1,tran+N,'normal');
x = flts(u,sysG);
u = u(tran+1:tran+N);
x = x(tran+1:tran+N);

printf('\n\nna) M = %.0e\n',M);
for nbhat = 1:nhatmax,
    for nahat = 1:nhatmax,
        // identificacion
        Oux = omega(L,u,x,nbhat,nahat,0);
        Lxux = Oux'*Oux/(2*M+1);
        clear Oux;
        vpLxux = abs(spec(Lxux));
        vpmiLxux = min(vpLxux);

        printf('%.2e ',vpmiLxux);
    end
    printf('\n');
end
printf('\n');

sz = stacksize();
printf('\nmemoria total: %d, memoria usada: %d\n',sz(1),sz(2));

```

F.3.5. Ejemplos 2.5 y 2.6: Lzuy2.sce

```
// Lzuy2.sce
```

```

// version 2008.08.19

clc; clear;
q = poly(0,'q'); rand('seed',0);
getf('funciones.sci'); getf('dsp.sci');
stacksize(50e6);

// planta
// n=3; B = 0.3; A = (q^2 - 2*(0.8)*cos(%pi/4)*q + 0.8^2)*(q-0.7);

// ejemplo 3, (2002) al-smadi, wilkes "robust and accurate..."
// n=3; B = q^3 - 0.45*q^2 - q^1; A = q^3 - 2.2*q^2 + 1.77*q - 0.52;

// ejemplo 1, (1993) liang, wilkes, cadzow "arma model order..."
// n=6; B = q^6 - 0.2997*q^5 + 0.4147*q^4 - 0.2794*q^3 + 0.4973*q^2;
// A = q^6 - 1.2798*q^5 + 0.7805*q^4 - 0.1635*q^3 + 0.7566*q^2 - 1.0621*q + 0.7821;

// ejemplo 4, (1993) liang, wilkes, cadzow "arma model order..."
n=6; B = q^6 + 0.3452*q^5 + 0.53*q^4 + 0.3985*q^3 + 0.8138*q^2;
A = q^6 + 0.7907*q^5 + 0.042*q^4 - 0.5556*q^3 - 0.0247*q^2 + 0.3846*q + 0.3026;

// ejemplo 1, (1996) al-smadi, wilkes "on estimating arma model orders"
// n=7; B = q^7 - 0.1292*q^6 - 0.3893*q^5 + 0.3163*q^4 + 0.2549*q^3 - 0.5823*q^2;
// A = q^7 - 0.2459*q^6 - 0.3786*q^5 - 0.0707*q^4 - 0.0944*q^3 - 0.2881*q^2 + 0.107*q^1 + 0.4695;

// ruido
D = 0.4*q; C = (q - 0.6);

// transferencias
sysG = syslin('d',B,A);
sysH = syslin('d',D,C);
tran = 5*ceil(-inv(log(max(abs([roots(A); roots(C)]))))));
printf('transiente = %d\n\n',tran);

// datos
nhatmax = 2*n;
M = [1e4 5e4];
Mmax = max(M);
L = nhatmax;
Nmax = 2*Mmax+1+L;

// mediciones estado estacionario
u0 = rand(1,tran+Nmax,'normal');
x0 = flts(u0,sysG);

for Mi = M,
    N = 2*Mi+1+L;

    u = u0(tran+1:tran+N);
    x = x0(tran+1:tran+N);

    // e = rand(1,tran+N,'normal');

```

```

e = (rand(1,tran+N,'uniform')-0.5)*2*sqrt(3);
v = flts(e,sysH);
v = v(tran+1:$);

y = x + v;

SNR = 20*log10(stdev(x)/stdev(v));

printf('(^nb\\^na) M = %.0e, SNR = %.2f\n',Mi,SNR);
for nbhat = 1:nhatmax,
    for nahat = 1:nhatmax,
        // filtro instrumental
        m = nhatmax; d = nbhat; F = q^(-d)+q^(-m);
        sysF = syslin('d',F);
        z = flts(u0,sysF);
        z = z(tran+1:tran+N);

        // identificacion
        Ouz = omega(L,u,z,nbhat,nahat,0);
        Ouy = omega(L,u,y,nbhat,nahat,0);
        Lzuy = Ouz'*Ouy/(2*Mi+1);
        clear Ouz; clear Ouy;
        vpLzuy = abs(spec(Lzuy));
        vpmiLzuy = min(vpLzuy);

        printf('%.2e ',vpmiLzuy);
    end
    printf('\n');
end
printf('\n');
end

sz = stacksize();
printf('\nmemoria total: %d, memoria usada: %d\n',sz(1),sz(2));

```

F.4. Ejemplos capítulo 3

F.4.1. Ejemplos 3.1 y 3.2: Lxux.sce

```

// Lxux.sce
// version 2008.09.02

clc; clear;
q = poly(0,'q'); rand('seed',0);
getf('funciones.sci'); getf('dsp.sci');
stacksize(50e6);

// planta
// nb0=5; na0=3; B = 0.3; A = q^2*(q^2 - 2*(0.8)*cos(%pi/4)*q + 0.8^2)*(q-0.7);

```

```

// ejemplo 3, (2002) al-smadi, wilkes "robust and accurate..."
// nb0=2; na0=3; B = q^3 - 0.45*q^2 - q^1; A = q^3 - 2.2*q^2 + 1.77*q - 0.52;

// ejemplo 4, (1993) liang, wilkes, cadzow "arma model order..."
nb0=4; B = q^6 + 0.3452*q^5 + 0.53*q^4 + 0.3985*q^3 + 0.8138*q^2;
na0=6; A = q^6 + 0.7907*q^5 + 0.042*q^4 - 0.5556*q^3 - 0.0247*q^2 + 0.3846*q + 0.3026;

// ejemplo 1, (1996) al-smadi, wilkes "on estimating arma model orders"
// nb0=5; B = q^7 - 0.1292*q^6 - 0.3893*q^5 + 0.3163*q^4 + 0.2549*q^3 - 0.5823*q^2;
// na0=7; A = q^7 - 0.2459*q^6 - 0.3786*q^5 - 0.0707*q^4 - 0.0944*q^3 - 0.2881*q^2 + 0.107*q^1 + 0.4695;

// datos
nhatmax = 15; nbhat = 6; nahat = 12;
M = 250; L = nhatmax; N = 2*M+1+L;

// transferencias
sysG = syslin('d',B,A);
tran = 5*ceil(-inv(log(max(abs([roots(A)]))))));
printf('transiente = %d\n\n',tran);

// mediciones estado estacionario
u = rand(1,tran+N,'normal');
x = flts(u,sysG);
u = u(tran+1:tran+N);
x = x(tran+1:tran+N);

Oux = omega(L,u,x,nbhat,nahat,0)/sqrt(2*M+1);
Lxux = Oux'*Oux;
svOux = svd(Oux);
svLxux = svd(Lxux);

scf(0); clf(0);
titulo = strcat(['Perfil de valores singulares (^nb,^na) = (' string(nbhat) ', ' string(nahat) ')']);
xtitle(titulo,'valor singular','tamanno del valor singular');
plot2d(1:length(svLxux),svLxux,logflag='nl',style=-9);
plot2d(1:length(svOux),svOux,logflag='nl',style=-2);

epsilonL = norm(size(Lxux),'inf')*svLxux(1)*%eps;
epsilon0 = norm(size(Oux),'inf')*svOux(1)*%eps;
printf('epsilonL = %.2e, epsilon0= %.2e\n',epsilonL,epsilon0);
plot2d([1 length(svLxux)],epsilonL*[1 1],logflag='nl',style=5);
plot2d([1 length(svLxux)],epsilon0*[1 1],logflag='nl',style=2);

```

F.4.2. Ejemplos 3.3, 3.4 y 3.5: Lzuy.sce

```

// Lzuy.sce
// version 2008.08.13

clc; clear;

```

```

q = poly(0,'q'); rand('seed',0);
getf('funciones.sci'); getf('dsp.sci');
stacksize(50e6);

// planta
// n0=3; B = 0.3; A = q^2*(q^2 - 2*(0.8)*cos(%pi/4)*q + 0.8^2)*(q-0.7);

// ejemplo 3, (2002) al-smadi, wilkes "robust and accurate..."
// nb0=2; na0=3; B = q^3 - 0.45*q^2 - q^1; A = q^3 - 2.2*q^2 + 1.77*q - 0.52;

// ejemplo 4, (1993) liang, wilkes, cadzow "arma model order..."
// nb0=4; B = q^6 + 0.3452*q^5 + 0.53*q^4 + 0.3985*q^3 + 0.8138*q^2;
// na0=6; A = q^6 + 0.7907*q^5 + 0.042*q^4 - 0.5556*q^3 - 0.0247*q^2 + 0.3846*q + 0.3026;

// ejemplo 1, (1996) al-smadi, wilkes "on estimating arma model orders"
nb0=5; B = q^7 - 0.1292*q^6 - 0.3893*q^5 + 0.3163*q^4 + 0.2549*q^3 - 0.5823*q^2;
na0=7; A = q^7 - 0.2459*q^6 - 0.3786*q^5 - 0.0707*q^4 - 0.0944*q^3 - 0.2881*q^2 + 0.107*q^1 + 0.4695;

// ruido
// D = 0.4*q; C = (q - 0.6);
D = 0.0976*q^2 + 0.1953*q + 0.0976; C = q^2 - 0.9428*q + 0.3333; //thetac=pi/4

// datos
veces = 50;
nhatmax = 15; nbhat = 7; nahat = 14;
M = 3e3; L = nhatmax; N = 2*M+1+L;

// instrumentos
F = q^(-nbhat);

// transferencias
sysG = syslin('d',B,A);
sysH = syslin('d',D,C);
sysF = syslin('d',F);
tran = 5*ceil(-inv(log(max(abs([roots(A); roots(C)]))))));
printf('transiente = %d\n\n',tran);

// mediciones estado estacionario
u = rand(1,tran+N,'normal');
x = flts(u,sysG);
z = flts(u,sysF);
u = u(tran+1:tran+N);
x = x(tran+1:tran+N);
z = z(tran+1:$);

scf(0); clf(0);
titulo = strcat(['Perfil de valores singulares (^nb,^na) = (' string(nbhat) ', ' string(nahat) ')']);
xtitle(titulo,'valor singular','tamanno del valor singular');

Ouz = omega(L,u,z,nbhat,nahat,0);
Oux = omega(L,u,x,nbhat,nahat,0);
Lzux = Ouz'*Oux/(2*M+1);

```

```

svLzux = svd(Lzux); svLzux = svLzux(svLzux>1e-10);

for i=1:veces,
    e = rand(1,tran+N,'normal');
    //e = (rand(1,tran+N,'uniform')-0.5)*2*sqrt(3);
    v = flts(e,sysH);
    v = v(tran+1:$);

    y = x + v;

    Ouy = omega(L,u,y,nbhat,nahat,0);
    Lzuy = Ouz'*Ouy/(2*M+1);
    svLzuy = svd(Lzuy);

    plot2d(1:length(svLzuy),svLzuy,logflag='nl',style=-9);
end

av = stdev(v);
ezv = sqrt((1*av)**2/(2*M+1));
epsilon = sqrt((nbhat+1+nahat)*(nahat))*ezv;
SNR = 20*log10(stdev(x)/stdev(v));

printf('M = %.0e, av = %.2e, SNR = %.2f\n',M,av,SNR);
printf('epsilon = %.2e\n',epsilon);
plot2d([1 length(svLzuy)],epsilon*[1 1],logflag='nl',style=5);
plot2d(1:length(svLzux),svLzux,logflag='nl',style=-2);

```

F.5. Ejemplos capítulo 4

F.5.1. Ejemplos 4.1, 4.2, 4.3 y 4.4: Gzx.sce

```

// Gzx.sce
// version 2008.09.29

clc; clear;
q = poly(0,'q'); rand('seed',0);
getf('funciones.sci'); getf('dsp.sci');
stacksize(50e6);

// planta
// n0=3; B = 0.3; A = q^2*(q^2 - 2*(0.8)*cos(%pi/4)*q + 0.8^2)*(q-0.7);

// ejemplo 3, (2002) al-smadi, wilkes "robust and accurate..."
// nb0=2; na0=3; B = q^3 - 0.45*q^2 - q^1; A = q^3 - 2.2*q^2 + 1.77*q - 0.52;

// ejemplo 4, (1993) liang, wilkes, cadzow "arma model order..."
nb0=4; B = q^6 + 0.3452*q^5 + 0.53*q^4 + 0.3985*q^3 + 0.8138*q^2;
na0=6; A = q^6 + 0.7907*q^5 + 0.042*q^4 - 0.5556*q^3 - 0.0247*q^2 + 0.3846*q + 0.3026;

```

```

// ejemplo 1, (1996) al-smadi, wilkes "on estimating arma model orders"
// nb0=5; B = q^7 - 0.1292*q^6 - 0.3893*q^5 + 0.3163*q^4 + 0.2549*q^3 - 0.5823*q^2;
// na0=7; A = q^7 - 0.2459*q^6 - 0.3786*q^5 - 0.0707*q^4 - 0.0944*q^3 - 0.2881*q^2 + 0.107*q^1 + 0.46

// datos
veces = 50;
nhatmax = 15; df = 6; nahat = 12;
M = 1e5; L = nhatmax; N = 2*M+1+L;

// instrumentos
F = q^(-df);

// transferencias
sysG = syslin('d',B,A);
sysF = syslin('d',F);
tran = 5*ceil(-inv(log(max(abs(roots(A))))));
printf('transiente = %d, M=%d\n',tran,M);

scf(0); clf(0);
titulo = strcat(['Perfil de valores singulares (df,^na)= ( ' string(df) ', ' string(nahat) ')']);
xtitle(titulo,'valor singular','tamanno del valor singular');

for i=1:veces,
    // mediciones estado estacionario
    // u = rand(1,tran+N,'normal');
    u = (rand(1,tran+N,'uniform')-0.5)*2*sqrt(3);
    x = flts(u,sysG);
    z = flts(u,sysF); // caso u(k) medible
    // z = flts(x,sysF); // caso ARMA
    x = x(tran+1:tran+N);
    z = z(tran+1:tran+N);

    Oz = omegax(L,z,nahat,0);
    Ox = omegax(L,x,nahat,0);
    Gzx = Oz'*Ox/(2*M+1);
    svGzx = svd(Gzx);

    plot2d(1:length(svGzx),svGzx,logflag='nl',style=-9);
end

```

F.5.2. Ejemplos 4.5, 4.6 y 4.7: hsv.sce

```

// hsv.sce
// version 2008.08.20

clc; clear;
q = poly(0,'q'); rand('seed',0);
getf('funciones.sci');
stacksize(50e6);

```

```

// planta
// n=3; B = 0.3; A = (q^2 - 2*(0.8)*cos(%pi/4)*q + 0.8^2)*(q-0.7);

// ejemplo 3, (2002) al-smadi, wilkes "robust and accurate..."
// n=3; B = q^3 - 0.45*q^2 - q^1; A = q^3 - 2.2*q^2 + 1.77*q - 0.52;

// ejemplo 4, (1993) liang, wilkes, cadzow "arma model order..."
// n=6; B = q^6 + 0.3452*q^5 + 0.53*q^4 + 0.3985*q^3 + 0.8138*q^2;
// A = q^6 + 0.7907*q^5 + 0.042*q^4 - 0.5556*q^3 - 0.0247*q^2 + 0.3846*q + 0.3026;

// ejemplo 1, (1996) al-smadi, wilkes "on estimating arma model orders"
n=7; B = q^7 - 0.1292*q^6 - 0.3893*q^5 + 0.3163*q^4 + 0.2549*q^3 - 0.5823*q^2;
A = q^7 - 0.2459*q^6 - 0.3786*q^5 - 0.0707*q^4 - 0.0944*q^3 - 0.2881*q^2 + 0.107*q^1 + 0.4695;

// transferencias
sysG = syslin('d',B,A);
tran = 5*ceil(-inv(log(max(abs(roots(A))))));

// datos
conjM = [1e2 1e3 1e4 1e5];
nhat = 10; L = nhat;

// instrumentos
F = q^(-nhat);
sysF = syslin('d',F);

// respuesta a impulso
d(1,1) = 1; d(1,2*nhat) = 0;
g = flts(d,sysG);
// matriz de Hankel
H = hank(nhat,nhat,g(2:$));
svH = svd(H);

printf('%0.4f ',svH);
printf('\n-----\n');
// matrices
for M = conjM,
    N = 2*M+1+L;
    u = rand(1,tran+N,'normal');
    x = flts(u,sysG);
    z = flts(u,sysF);
    u = u(tran+1:$);
    x = x(tran+1:$);
    z = z(tran+1:$);

    Ox = omegax(L,x,nhat,0);
    Oz = omegax(L,z,nhat,0);
    Gzx = Oz'*Ox/(2*M+1);
    svGzx = svd(Gzx);
    printf('%0.4f ',svGzx); printf('\n');
end

```

Apéndice G

Publicaciones basadas en esta tesis

LOS resultados obtenidos durante el desarrollo de esta tesis han permitido generar un artículo, que ha sido aceptado en una conferencia internacional. Este artículo es detallado a continuación

C. Guajardo and R. Rojas. *Rank of finite-sample covariance matrices for model order estimation*. 11th IASTED International Conference on Intelligent Systems and Control, ISC-2008, Orlando, Florida, USA, 16-18 November 2008.

Referencias

- [1] A. Al-Smadi y D. Wilkes (1996). *On estimating ARMA model orders*. IEEE Int. Symp. on Circuits and Systems, Vol. 2, pp. 505-508.
- [2] A. Al-Smadi y D. Wilkes (2002). *Robust and accurate ARX and ARMA model order estimation of non-gaussian processes*. IEEE Transactions on Signal Processing, Vol. 50, Issue 3, pp. 759-763.
- [3] B. Anderson y J. Moore (1990). *Optimal control: linear quadratic methods*. Prentice-Hall, Inc. Upper Saddle River, NJ, USA.
- [4] J. Cadzow (1982). *Spectral estimation: An overdetermined rational model equation approach*. Proc IEEE, Vol. 70, pp. 907-939.
- [5] J. Cadzow y O. Solomon (1986). *Algebraic approach to system identification*. IEEE Trans. on Acoustics, Speech and Signal Processing, Vol. 34, Issue 3, pp. 462-469.
- [6] Y. Chan y J. Wood (1984). *A new order determination technique for ARMA processes*. IEEE Transactions on Acoustics, Speech and Signal Processing, Vol. 32, Issue 3, pp. 517-521.
- [7] C. Davila y H. Chiang (1994). *An algorithm for pole-zero system model order estimation*. <http://citeseer.ist.psu.edu/354492.html>.
- [8] H. Duong y I. Landau (1994). *On statistical properties of a test for model structure selection using the extended instrumental variable approach*. IEEE Transactions on Automatic Control, Vol. 39, Issue 1, pp. 211-215.
- [9] J. Fuchs (1987). *ARMA order estimation via matrix perturbation theory*. IEEE Transactions on Automatic Control, Vol. AC-32, pp. 358-361.
- [10] K. Glover (1984). *All optimal Hankel-norm approximations of linear multivariable systems and their L^∞ -error bounds*. Int. J. Control, vol. 39, no. 6, pp. 1115-1193.
- [11] G. Golub y C. van Loan (1996). *Matrix computations* (3.^a edición). The John Hopkins University Press, Baltimore, Maryland.
- [12] T. Kailath (1980). *Linear systems*. Prentice Hall.
- [13] G. Liang, D. Wilkes y J. Cadzow (1993). *ARMA model order estimation based on the eigenvalues of the covariance matrix*. IEEE Transactions on Signal Processing, Vol. 41, Issue 10, pp. 3003-3009.
- [14] L. Ljung (1999). *System identification: theory for the user*. Prentice-Hall, Inc. Upper Saddle River, NJ, USA.
- [15] Mathematics Resource Center (2003-2006). *From geometry to algebra, an introduction to linear algebra*. Department of Mathematics, Indian Institute of Technology Bombay, India. <http://www.mathresource.iitb.ac.in/linear%20algebra/>

- [16] A. Papoulis (1991). *Probability, random variables and stochastic processes* (3.^a edición). McGraw-Hill.
- [17] I. Rezek y S. Roberts (1997). *Parametric Model Order Estimation: A Brief Review*. Proceedings of The 1997 IEE Colloquium on the Use of Model Based Digital Signal Processing Techniques in the Analysis of Biomedical Signals.
- [18] R. Rojas (1985). *Instrumental determinant ratio test for multivariable structure identification*. 2nd IASTED international congress on telecommunication and control, Telecon85, Rio de Janeiro, Brasil, vol. 1, pp. 230-233.
- [19] T. Söderström y P. Stoica (1983). *Instrumental variable methods for system identification*. Springer-Verlag, Berlin, Germany.
- [20] T. Söderström y P. Stoica (1989). *System identification*. Prentice-Hall International, Hemel Hempstead, UK.
- [21] P. Stoica (1981). *On a procedure for testing the order of time series*. IEEE Transactions on Automatic Control, Vol. 26, Issue 2, pp. 572-573.
- [22] P. Stoica, M. Cedervall and A. Eriksson (1995). *Combined instrumental variable and subspace fitting approach to parameter estimation of noisy input-output systems*. IEEE Transactions on Signal Processing, Vol. 43, Issue 10, pp. 2386-2397.
- [23] P. Wellstead (1978). *An instrumental product moment test for model order estimation*. Automatica vol. 14, pp. 89-91.
- [24] P. Wellstead y R. Rojas (1982). *Instrumental product moment model-order testing: extensions and applications*. International Journal of Control, vol. 35, no. 6, pp. 1013-1027.
- [25] X. Zhang y Y. Zhang (1993). *Determination of the MA order of an ARMA process using sample correlations*. IEEE Trans. Signal Processing, vol. 41, pp. 2277-2280.