

UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA

DEPARTAMENTO DE ELECTRÓNICA

**IDENTIFICACIÓN DE MODELOS DE ESTADO
USANDO ALGORITMO EM CON RESTRICCIONES**

Tesis de Grado presentada por

Felipe Eleazar Ávila Cárcamo

como requisito parcial para optar al título de

Ingeniero Civil Electrónico

y al grado de

Magíster en Ciencias de la Ingeniería Electrónica

Profesor Guía

Dr. Juan Yuz Eissmann

Profesor Co-guía

Dr. Juan Agüero Vásquez

Valparaíso, 2018.

TÍTULO DE LA TESIS:

IDENTIFICACIÓN DE MODELOS DE ESTADO USANDO ALGORITMO EM CON RESTRICCIONES

AUTOR:

Felipe Eleazar Ávila Cárcamo

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

Dr. Juan Yuz E.

Dr. Juan Agüero V.

Dr. Marcos Orchard O.

Valparaíso, Julio de 2018.

dedicado a mi familia

Stephie,

y amigos

AGRADECIMIENTOS

Para partir, quisiera agradecer en primera instancia a mi familia, en especial a mi padre y madre, pues siempre velaron por mí y mis hermanos, y nos motivan a apuntar a lo más alto. Nunca nos faltó algo, y sólo fue gracias al esfuerzo y sacrificio de ambos, el cual valoro día a día. A mis hermanos agradezco todas las alegrías y risas que me han entregado. Danilo, quien ha sido mi compañero de travesuras desde que soy niño, me ayudó a pasar muchas penas y siempre me hizo reír, tal como el Benjita y Vicente.

Me gustaría agradecer también a Stephanie, mi confidente en gran parte de mi paso por la Universidad. Desde que ella llegó a mi vida, me motivó a cambiar para bien muchos aspectos. Me acompañó en mis mejores momentos, como también en los malos, pero nunca dejó de creer en mí, y gracias a ella es en gran parte los resultados que he obtenido. Me entregó tantas alegrías, y entre esas alegrías nunca olvidaré la Patatina y Pepi.

También me gustaría agradecer a todos los amigos que pude sembrar en el paso de esta Universidad, los cuales siento que perdurarán para toda la vida. Hago mención especial a Julian, Rodrigo (Mister) y René, que fueron un gran apoyo y me dieron grandes momentos que jamás olvidaré.

Por último, agradezco al profesor Juan Yuz, quien siempre tuvo disponibilidad para ayudarme a sacar adelante esta investigación y siempre tuvo fe que lograríamos el objetivo, a pesar que mucho tiempo el escenario se vio bastante complicado. También agradezco al profesor Juan Carlos Agüero, quien también apoyó esta investigación y dio ideas geniales.

CONTENIDO

AGRADECIMIENTOS	I
RESUMEN	V
ABSTRACT	VII
1. INTRODUCCIÓN	1
1.1. Definición del problema	1
1.2. Estructura del Documento	2
2. IDENTIFICACIÓN DE MODELOS DE ESTADO USANDO ML	3
2.1. Máxima Verosimilitud	3
2.2. Algoritmo EM	4
2.2.1. Convergencia del algoritmo EM	7
2.3. Algoritmo EM en el dominio del tiempo	9
2.3.1. Paso E	10
2.3.2. Filtro y Suavizador de Kalman	12
2.3.3. Paso M	16
2.4. Sistemas en tiempo continuo y modelos discretos	18
2.4.1. Modelos en tiempo discreto	18
3. REGIONES DEFINIDAS POR DESIGUALDADES MATRICIALES	21
3.1. Desigualdades matriciales lineales	22
3.2. Regiones LMI	22
3.3. Regiones LMI de interés	24
4. ALGORITMO EM CON RESTRICCIONES	31
4.1. Algoritmo EM con restricciones	31
4.1.1. Criterio de Sylvester	32
4.1.2. Funciones de barrera logarítmica	35
5. SIMULACIONES	37
5.1. Ejemplo 1	37

5.2. Ejemplo 2	41
5.3. Ejemplo 3	45
6. CONCLUSIONES	47
6.1. Trabajo a futuro	48
REFERENCIAS	49

RESUMEN

El objetivo principal de la Identificación de Sistemas es construir modelos matemáticos que representen sistemas reales a partir de datos medidos de éste. La identificación de sistemas ha sido ampliamente estudiada y existen distintos algoritmos que sirven para cumplir el objetivo de identificación en función de algún criterio de optimización.

La presente Tesis tiene como objetivo estudiar la estimación de parámetros usando el algoritmo EM que permite obtener el estimador de Máxima Verosimilitud para modelos de lineales e invariantes en el tiempo cuando se imponen restricciones en la ubicación de los polos del sistema. El estimador de máxima verosimilitud obtenido mediante el algoritmo EM es un estimador consistente y asintóticamente no sesgado, lo que implica que cuando existe una cantidad limitada de datos (o baja relación señal a ruido) el modelo estimado puede no conservar características del sistema real, tales como estabilidad o comportamiento oscilatorio. Este problema es el que se intentará abordar mediante restricciones en la ubicación de los autovalores de la matriz del sistema.

Actualmente existen métodos de estimación con restricciones en las características del modelo obtenido cuando se realiza estimación mediante métodos de subespacios. Estos métodos fueron estudiados y de ellos se destaca la idea del uso de regiones basadas en desigualdades lineales matriciales. En esta Tesis se analiza el uso de estas desigualdades en el algoritmo EM.

La implementación del algoritmo EM con las restricciones mencionadas presenta problemas computacionales, por lo que en esta Tesis se presenta una sección donde se abordan estos problemas, mediante el uso del criterio de Sylvester para matrices reales y simétricas definidas positivas, y la aplicación de funciones de barrera logarítmicas para las restricciones, obteniendo así un algoritmo de rápida implementación.

El algoritmo desarrollado se pone a prueba analizando estimaciones obtenidas y comparándolas con el algoritmo EM sin modificar para las mismas mediciones. Se analiza el impacto de la cantidad de datos, relación señal a ruido a los mismos modelos con restricciones aplicadas.

ABSTRACT

The main objective in System Identification is to obtain a mathematical model that represents a real system from measured data. System Identification has been extensively studied. There are different algorithms that fulfill the objective based on different optimization criteria.

The objective of this Thesis is to study parametric estimation using the EM algorithm, that obtains the Maximum Likelihood estimator for linear time invariant models when location constraints are applied to the poles of the estimated model. The maximum likelihood estimator obtained by the EM algorithm is consistent and asymptotically unbiased. This implies that when a limited amount of data is available (or low signal-to-noise ratio) the estimated model may not preserve characteristics of the real system, such as stability or oscillatory behavior. This problem will be addressed by including constraints in the location of the eigenvalues of the system matrix.

There are estimation methods with constraints on the characteristics of the obtained model using subspace methods. Some of these methods use the idea of regions based on linear matrix inequalities. In this Thesis the application of constraints based on LMI region using EM algorithm is analyzed.

The implementation of the constrained EM algorithm presents computational problems, so in this Thesis we present this problem, and how to handle it, using the Sylvester Criterion for real and symmetric positive defined matrices, and the application of log-barrier functions for the constraints.

The proposed algorithm is tested by analyzing the obtained estimates and comparing them with the unconstrained EM algorithm. The impact of the amount of data, and signal-to-noise ratio is also studied.

INTRODUCCIÓN

1.1. Definición del problema

La estimación usando máxima verosimilitud (Maximum Likelihood, ML) ha sido ampliamente aplicada para estimar parámetros en el dominio del tiempo y de la frecuencia [1], [2], [3], [4], [5], [6]. Sin embargo, el problema de optimización asociado puede ser no convexo. El algoritmo de maximización de esperanza (Expectation Maximization EM) es un procedimiento iterativo para obtener el estimador ML [7].

El algoritmo EM introduce la noción del conjunto de dato completo, que incluye los datos observados y los ocultos o perdidos [7]. El algoritmo transforma el problema de maximizar la función de verosimilitud en maximizar una función auxiliar la cual es actualizada y maximizada en cada iteración. Para el caso de modelos en variables estados, una elección natural para los datos ocultos es la secuencia de estado.

El algoritmo EM converge a un punto estacionario de la función de verosimilitud, estos pueden ser puntos silla, o máximos globales o locales [8]. Condiciones para la convergencia del algoritmo fueron establecidos en [9]. Implementaciones numéricamente robustas del algoritmo EM para modelos de espacios de estados lineales de tiempo discreto han sido estudiadas en [10], y para modelos muestreados en [11] [12] [13].

El estimador de ML es consistente y asintóticamente no sesgado, sin embargo cuando se estiman parámetros siempre existe una cantidad limitada de datos a los que se pueden acceder. Como consecuencia, el modelo estimado puede no reflejar las características del sistema real, tales como, por ejemplo, estabilidad o comportamiento (no) oscilatorio. En particular, cuando se estiman modelos de tiempo continuo a partir de datos muestreados, conocimiento a priori de estas características del sistema físico tiene que ser reflejado en el modelo estimado. En esta tesis, consideramos este requerimiento y proponemos una estimación de máxima verosimilitud de los parámetros de modelos de espacios de estados sujeto a restricciones en los polos del sistema.

Identificación de sistemas en variables de estado incluyendo restricciones ha sido desarrollado usando principalmente estimación por métodos de subespacio. Por ejemplo, en [14] estimación por métodos de subespacios son aplicados y la matriz del modelo A es obtenida como el producto de una matriz de corrimiento con una pseudo-inversa. En este trabajo, se muestra que la matriz de corrimiento es formada incluyendo un bloque de ceros en una posición adecuada, entonces el modelo resultante es estable, sin embargo, con la desventaja

de distorsionar la matriz de observabilidad. Enfoques alternativos aumentan el conjunto de datos, en función de garantizar una estabilidad marginal [15]. En [16], un término de regularización es agregado al costo de mínimos cuadrados, para forzar restricciones en los autovalores de A . Una solución interesante se tiene en [17] la cual aplica algoritmos de estimación de subespacio, usando las desigualdades de Lyapunov, $P - APA^T > 0$ and $P > 0$, como restricciones. Estas desigualdades aseguran que el modelo estimado es asintóticamente estable. Este enfoque no distorsiona la matriz de observabilidad como en [14], ni modificada los estados estimados o la secuencia de entrada al sistema como en [15]. Más aún, en [18] una generalización de [17] es presentada usando restricciones en la forma de desigualdades cuadráticas matriciales. De hecho, el círculo unitario y otras regiones de interés pueden ser expresadas como desigualdades matriciales lineales (LMI) de las cuales se obtienen las restricciones que usa [18].

En esta tesis, se aplica restricciones basado en regiones LMI en la estimación de ML usando el algoritmo EM. En particular, usamos LMI para forzar restricciones en la localización de los polos del modelo. Problemas existentes en la implementación se solucionan proponiendo un algoritmo que ocupa el criterio de Sylvester para matrices simétricas definidas positivas, y las restricciones se aplican usando funciones de barrera logarítmica.

1.2. Estructura del Documento

Capítulo 1. Introducción al trabajo de Tesis. Se introducen de forma general los conceptos asociados al tema de estudio, el estado del arte, los problemas a los que responde esta investigación, contribuciones del trabajo y descripciones por capítulo.

Capítulo 2. Conceptos básicos de estimación paramétrica mediante Máximo Likelihood. Se presenta la idea detrás de la estimación ML y el algoritmo EM que permite estimar de manera iterativa el estimador ML para sistemas lineales invariantes en el tiempo en el dominio del tiempo.

Capítulo 3. Se presentan regiones LMI y restricciones basadas en estas regiones.

Capítulo 4. Algoritmo EM con restricciones. Se muestra como se restringirá el espacio de búsqueda en el algoritmo EM, y problemas computacionales y como se solucionan.

Capítulo 5. Simulaciones. Se muestran distintas simulaciones de los resultados obtenidos, comparando lo obtenido aplicando restricciones al espacio de búsqueda de EM con el algoritmo EM sin restricciones alguna.

Capítulo 6. Se presentan las conclusiones del trabajo realizado, y se discuten alcances de trabajos a futuro.

IDENTIFICACIÓN DE MODELOS DE ESTADO USANDO ML

2.1. Máxima Verosimilitud

La identificación de sistemas tiene por objetivo construir modelos matemáticos que representen sistemas reales a partir de datos disponibles, tales como mediciones del sistema o señales de entrada. Todo esto puede estar sujeto a perturbaciones, ruido, y por lo tanto, existe información no disponible. Existen entonces distintos métodos para estimar los parámetros de un sistema, tales como Mínimos Cuadrados o métodos de Predicción de error (PEM) (e.g. [2] [3]).

En este contexto, es que se desarrolla el método de identificación por máxima verosimilitud (ML por sus siglas en inglés Maximum Likelihood), el cual fue desarrollado en estadística por Fisher en [19], y aplicado en identificación de sistemas por [1]. Es así que este método de identificación ha sido aplicado en métodos de estimación de parámetros de sistemas en el dominio del tiempo y de la frecuencia [5] [20] [4] [6].

El método ML es a partir de un set de datos observados Y que representan una variable aleatoria de densidad condicional $p(Y|\theta)$, dependiendo de un parámetro desconocido $\theta \in \Theta$. Si se desea estimar el parámetro θ , éste se debe escoger de tal manera que se maximice la probabilidad condicional de los datos [21] [5] [22]. Esto se puede definir como

$$\hat{\theta}_{ML} \triangleq \underset{\theta \in \Theta}{\operatorname{argmax}} p(Y|\theta) \quad (2.1.1)$$

donde se puede definir $l(\theta) = p(Y|\theta)$ como la función de verosimilitud.

Dentro de las propiedades que presenta el estimador de máxima verosimilitud (MLE por sus siglas del inglés Maximum Likelihood Estimator) es que se destacan las siguientes

Teorema 2.1.1 (Principio de invarianza). *Si $\hat{\theta}_{ML}$ es MLE de $\theta \in \Theta \subset \mathbb{R}^\ell$, entonces $g(\hat{\theta}_{ML})$ es el MLE de $g(\theta)$ donde $g(\cdot)$ es una función tal $g : \Theta \rightarrow \mathbb{Z} \subset \mathbb{R}^n$ con $n \leq \ell$.*

Teorema 2.1.2. *Un estimador no sesgado que alcanza la cota de Cramer-Rao (es decir, que es un estimador eficiente y no sesgado), es también un estimador de Máxima Verosimilitud.*

Teorema 2.1.3 (Consistencia). *Sea $\hat{\theta}_{ML}$ el MLE de θ basado en N observaciones, variables aleatorias independientes e idénticamente distribuidas (i.i.d.), $y = (x_1, x_2, \dots, x_N)^T$,*

entonces, cuando $N \rightarrow \infty$, entonces $\hat{\theta}_{ML}$ converge a θ con probabilidad 1.

Teorema 2.1.4 (Normalidad Asintótica). Sea $\hat{\theta}_{ML}$ el MLE de θ basado en N variables aleatorias independientes e idénticamente distribuidas (i.i.d.), $y = (x_1, x_2, \dots, x_N)^T$, entonces $\hat{\theta}_{ML}^N$ converge en probabilidad a una variable aleatoria Gaussiana:

$$\sqrt{N} \left(\hat{\theta}_{ML}^N - \theta \right) \rightarrow \beta \quad (2.1.2)$$

donde

$$\beta \sim \mathcal{N}(\theta, \bar{M}_\theta^{-1}) \quad (2.1.3)$$

donde \bar{M}_θ es el promedio de la matriz de información de Fisher por muestra, es decir

$$\bar{M}_\theta = \frac{1}{N} M_\theta \quad (2.1.4)$$

donde M_θ es la matriz de información de Fisher dada por

$$M_\theta \triangleq \mathbf{E} \left\{ \left[\frac{\partial \log p(Y|\theta)}{\partial \theta} \right]^T \left[\frac{\partial \log p(Y|\theta)}{\partial \theta} \right] \right\} \quad (2.1.5)$$

Teorema 2.1.5 (Eficiencia). Dentro de los estimadores asintóticamente normales y uniformemente consistentes, $\hat{\theta}_{ML}$ es eficiente en el sentido de que alcanza en forma asintótica la cota inferior de Crámer-Rao.

2.2. Algoritmo EM

El algoritmo EM es un algoritmo iterativo que permite obtener el estimador de máxima verosimilitud. Este algoritmo introduce la idea de *conjunto de datos completo*, el cual indica que aparte de la existencia de los datos observados en el sistema, existen datos ocultos o perdidos que aportan información y los cuales no tenemos acceso. Esta idea resulta conveniente, incluso en el caso en que el problema de identificación no posee variables ocultas [7].

Para desarrollar el algoritmo EM, se define entonces $Y = \{y_0, y_1, \dots, y_N\}$ como el conjunto de datos observados, los cuales poseen una densidad de probabilidad $p(Y|\theta)$, con $\theta \in \Theta$ el vector de parámetros del sistema. La idea entonces del estimador de máxima verosimilitud es hallar un vector de parámetros θ que maximiza la función de densidad de probabilidad de los datos dadas las mediciones. De acuerdo a esto, consideremos que existe otro conjunto de datos a los cuales no tenemos acceso, los cuales llamaremos *variables ocultas*, y los denotaremos por X . Así, entonces existe un conjunto de *datos completos* Z , compuesto por las observaciones Y como las variables ocultas X , es decir, $Z = \{Y, X\}$.

De este modo, podemos trabajar con la función de densidad condicional

$$p(Z|Y, \theta) = \frac{p(Z|\theta)}{p(Y|\theta)} \Leftrightarrow p(Y|\theta) = \frac{p(Z|\theta)}{p(Z|Y, \theta)} \quad (2.2.1)$$

Aplicando logaritmo en ambas partes se obtiene

$$L(\theta) \triangleq \log p(Y|\theta) = \log p(Z|\theta) - \log p(Z|Y, \theta) \quad (2.2.2)$$

Note que en (2.2.2), $L(\theta)$ es el logaritmo de la función de verosimilitud de los datos observados, y por otra parte, $\log p(Z|\theta)$ es el logaritmo de la función de verosimilitud de los datos completos, es decir, el conjunto de datos observados más las variables ocultas.

Tomando la esperanza en ambos lados de la ecuación de (2.2.2) con respecto a la distribución condicional de los datos completos Z , dado los datos observados Y y para un $\hat{\theta}_k \in \Theta$ fijo, es que se tiene

$$\begin{aligned} \mathbf{E}_Z \left\{ \log p(Y|\theta)|Y, \hat{\theta}_k \right\} &= \mathbf{E}_Z \left\{ \log p(Z|\theta)|Y, \hat{\theta}_k \right\} - \mathbf{E}_Z \left\{ \log p(Z|Y, \theta)|Y, \hat{\theta}_k \right\} \\ L(\theta) &= \mathcal{Q}(\theta, \hat{\theta}_k) - \mathcal{H}(\theta, \hat{\theta}_k) \end{aligned} \quad (2.2.3)$$

donde

$$\mathcal{Q}(\theta, \hat{\theta}_k) = \mathbf{E}_Z \left\{ \log p(Z|\theta)|Y, \hat{\theta}_k \right\} \quad (2.2.4)$$

$$\mathcal{H}(\theta, \hat{\theta}_k) = \mathbf{E}_Z \left\{ \log p(Z|Y, \theta)|Y, \hat{\theta}_k \right\} \quad (2.2.5)$$

Note que en (2.2.4) se puede escribir como

$$\begin{aligned} \mathcal{H}(\theta, \hat{\theta}_k) &= \mathbf{E}_Z \left\{ \log p(Z|Y, \theta)|Y, \hat{\theta}_k \right\} \\ &= \mathbf{E}_Z \left\{ \log p(X, Y|Y, \theta)|Y, \hat{\theta}_k \right\} \end{aligned} \quad (2.2.6)$$

pero

$$p(X, Y|Y, \theta) = \frac{p(X, Y|\theta)}{p(Y|\theta)} = p(X|Y, \theta) \quad (2.2.7)$$

de este modo, combinando lo anterior en (2.2.5) se tiene

$$\mathcal{H}(\theta, \hat{\theta}_k) = \mathbf{E}_Z \left\{ \log p(X|Y, \theta)|Y, \hat{\theta}_k \right\} \quad (2.2.8)$$

Con estas definiciones es que se define el siguiente Lema

Lema 2.2.1 (Iteración del algoritmo EM). *Considere las expresiones (2.2.3) - (2.2.5). Suponga que se escoge una secuencia de parámetros $\{\hat{\theta}_k\}$ de tal suerte que*

$$\mathcal{Q}(\hat{\theta}_{k+1}, \hat{\theta}_k) \geq \mathcal{Q}(\theta, \hat{\theta}_k), \quad \forall \theta \in \Theta \quad (2.2.9)$$

entonces,

$$L(\hat{\theta}_{k+1}) \geq L(\hat{\theta}_k) \quad (2.2.10)$$

Demostración. Este lema es demostrado en distintos trabajos en la literatura, en particular se toma el desarrollo presentado en [23].

Usando (2.2.3), se puede definir la siguiente diferencia

$$L(\theta) - L(\hat{\theta}_k) = \left(\mathcal{Q}(\theta, \hat{\theta}_k) - \mathcal{Q}(\hat{\theta}_k, \hat{\theta}_k) \right) - \left(\mathcal{H}(\theta, \hat{\theta}_k) - \mathcal{H}(\hat{\theta}_k, \hat{\theta}_k) \right) \quad (2.2.11)$$

Ahora bien, de acuerdo a la definición dada en (2.2.5) y a las propiedades de logaritmo, es que la segunda diferencia del lado derecho de la igualdad en (2.2.11) se puede escribir como

$$\begin{aligned}
\mathcal{H}(\theta, \hat{\theta}_k) - \mathcal{H}(\hat{\theta}_k, \hat{\theta}_k) &= \mathbf{E}_Z \left\{ \log p(Z|Y, \theta) | Y, \hat{\theta}_k \right\} - \mathbf{E}_Z \left\{ \log p(Z|Y, \hat{\theta}_k) | Y, \hat{\theta}_k \right\} \\
&= \mathbf{E}_Z \left\{ \log p(Z|Y, \theta) - \log p(Z|Y, \hat{\theta}_k) | Y, \hat{\theta}_k \right\} \\
&= \mathbf{E}_Z \left\{ \log \left[\frac{p(Z|Y, \theta)}{p(Z|Y, \hat{\theta}_k)} \right] | Y, \hat{\theta}_k \right\}
\end{aligned} \tag{2.2.12}$$

usando la desigualdad de Jensen, se puede corroborar lo siguiente

$$\begin{aligned}
\mathbf{E}_Z \left\{ \log \left[\frac{p(Z|Y, \theta)}{p(Z|Y, \hat{\theta}_k)} \right] | Y, \hat{\theta}_k \right\} &\leq \log \left[\mathbf{E}_Z \left\{ \frac{p(Z|Y, \theta)}{p(Z|Y, \hat{\theta}_k)} | Y, \hat{\theta}_k \right\} \right] \\
&= \log \left[\int \frac{p(Z|Y, \theta)}{p(Z|Y, \hat{\theta}_k)} p(Z|Y, \hat{\theta}_k) dZ \right] \\
&= \log \left[\int p(Z|Y, \theta) dZ \right] \\
&= 0
\end{aligned} \tag{2.2.13}$$

Combinando entonces los resultados anteriores, es que se tiene que

$$\mathcal{H}(\theta, \hat{\theta}_k) - \mathcal{H}(\hat{\theta}_k, \hat{\theta}_k) \geq 0, \quad \forall \theta \in \Theta \tag{2.2.14}$$

Por lo tanto, con este resultado es que se puede afirmar que

$$\mathcal{Q}(\hat{\theta}_{k+1}, \hat{\theta}_k) \geq \mathcal{Q}(\hat{\theta}_k, \hat{\theta}_k) \Rightarrow L(\hat{\theta}_{k+1}) \geq L(\hat{\theta}_k) \tag{2.2.15}$$

Por lo que queda demostrado el resultado. ■

El lema anterior muestra que para encontrar un nuevo parámetro $\hat{\theta}_{k+1}$ que aumente el valor de la función de verosimilitud en cada iteración, basta con maximizar (o encontrar el valor que aumente numéricamente) $\mathcal{Q}(\theta, \hat{\theta}_k)$. A partir de esto es que definimos el corolario siguiente.

Corolario 2.2.2. *Sea la secuencia definida $\{\hat{\theta}_k\}$ generada por el algoritmo EM en cada iteración definido en el Lema 2.2.1. Entonces si la función $L(\theta)$ es acotada en Θ , esta convergerá monotónicamente a algún L^* .*

Demostración. La demostración resulta directa a partir de lo visto en el Lema 2.2.1, donde se vio que si se escoge una secuencia $\{\hat{\theta}_k\}$ de manera que maximice (o incremente) $\mathcal{Q}(\theta, \hat{\theta}_k)$, el logaritmo de la función de verosimilitud nunca decrecerá. ■

Como se vio anteriormente, el Lema 2.2.1 define los pasos que toma el algoritmo EM, traduciendo entonces el problema de maximizar la función de verosimilitud, lo cual implica un problema usualmente no-convexo y complicado de obtener, en maximizar una función auxiliar $\mathcal{Q}(\theta, \hat{\theta}_k)$, lo que puede simplificar el problema planteado. De hecho, es posible así obtener incluso una solución analítica para obtener $\hat{\theta}_{k+1} = \underset{\hat{\theta}}{\operatorname{argmax}} \mathcal{Q}(\theta, \hat{\theta}_k)$ para

algún problema planteado tanto en el dominio del tiempo o en frecuencia (ver por ejemplo [24], [10], [25], [11]).

Es así, que el algoritmo EM se resume en un algoritmo iterativo con los siguientes pasos

1. Comenzar con una estimación inicial de los parámetros $\hat{\theta}_0$.
2. **Paso E:** Obtener la función $\mathcal{Q}(\theta, \hat{\theta}_k)$ definido en (2.2.4).
3. **Paso M:** Calcular la estimación $\hat{\theta}_{k+1}$ que maximice $\mathcal{Q}(\theta, \hat{\theta}_k)$, o que aumente respecto $\mathcal{Q}(\hat{\theta}_k, \hat{\theta}_k)$.
4. Volver al paso 2, incrementando el índice $k \rightarrow k + 1$ en cada paso hasta cumplir algún criterio de convergencia.

2.2.1. Convergencia del algoritmo EM

En esta sección se analizará la convergencia del algoritmo EM al estimador de máxima verosimilitud ML. Del Corolario 2.2.2 se tiene que $L(\theta)$ converge monotónicamente a algún L^* , no obstante, no hay garantías que L^* sea el máximo global de $L(\theta) \in \Theta$. De hecho, si $L(\theta)$ tiene varios puntos estacionarios, la convergencia de una secuencia usando el algoritmo EM a algún punto estacionario, tal como un máximo local o global, como también un punto silla, depende de la elección de la condición inicial $\hat{\theta}_0$. Para ilustrar esto, supongamos que se quiere maximizar $L(\theta)$

$$\frac{\partial L(\theta)}{\partial \theta} = \frac{\partial \mathcal{Q}(\theta, \hat{\theta}_k)}{\partial \theta} - \frac{\partial \mathcal{H}(\theta, \hat{\theta}_k)}{\partial \theta} \quad (2.2.16)$$

pero se tiene de (2.2.14)

$$\left. \frac{\partial \mathcal{H}(\theta, \hat{\theta}_k)}{\partial \theta} \right|_{\theta=\hat{\theta}_k} = 0 \quad (2.2.17)$$

Ahora tómesese un punto arbitrario de la secuencia de EM $\hat{\theta}_k = \bar{\theta}$, entonces

$$\left. \frac{\partial L(\theta)}{\partial \theta} \right|_{\theta=\bar{\theta}} = \left. \frac{\partial \mathcal{Q}(\theta, \hat{\theta}_k)}{\partial \theta} \right|_{\theta=\bar{\theta}} - 0 \quad (2.2.18)$$

Si $\hat{\theta} = \theta^*$ es un punto estacionario de $L(\theta)$, por ejemplo un punto silla o máximo local, se tiene que

$$\left. \frac{\partial L(\theta)}{\partial \theta} \right|_{\theta=\theta^*} = \left. \frac{\partial \mathcal{Q}(\theta, \hat{\theta}_k)}{\partial \theta} \right|_{\theta=\theta^*} = 0 \quad (2.2.19)$$

lo que implica que el algoritmo de EM puede converger a puntos estacionarios θ^* (como puntos silla o máximos locales) si es que $\mathcal{Q}(\theta, \hat{\theta}^*)$ es maximizado en Θ en el punto θ^* ([8]).

Las condiciones de convergencia al máximo global fueron establecidas en [9], las cuales las veremos a continuación.

Como se mencionó anteriormente, en [9] se establecen las condiciones para que el algoritmo EM converja al máximo global. Esta sección recopila los resultados más importantes, tal como en [8].

Definición 2.2.3 (GEM). *Se define el algoritmo EM generalizado (GEM por las siglas en inglés Generalized EM algorithm) como el mapeo*

$$\hat{\theta}_{k+1} \triangleq M_{GEM}(\hat{\theta}_k) \quad (2.2.20)$$

tal que

$$\mathcal{Q}(M_{GEM}(\hat{\theta}_k), \hat{\theta}_k) \geq \mathcal{Q}(\hat{\theta}_k, \hat{\theta}_k) \quad (2.2.21)$$

Note que la condición para la iteración del algoritmo de EM normalmente es que $\hat{\theta}_{k+1} \triangleq M_{EM}(\hat{\theta}_k)$ maximice $\mathcal{Q}(\theta, \hat{\theta}_k)$ en cada iteración, siendo una condición más estricta, y, por tanto, un caso particular para el algoritmo GEM.

A continuación se presenta el teorema principal de convergencia establecido en [9] para algún algoritmo GEM

Teorema 2.2.4. *Sea $\{\hat{\theta}_k\}$ una secuencia generada por algoritmo GEM de acuerdo a la definición 2.2.3. Ahora bien, suponga que*

- *El mapeo $M_{GEM}(\hat{\theta}_k)$ es cerrado sobre el complemento de S , el conjunto de todos los puntos estacionarios de Θ .*
- *$L(\hat{\theta}_{k+1}) \geq L(\hat{\theta}_k)$ para todo $\hat{\theta}_k \notin S$.*

Entonces todos los puntos límites de $\{\hat{\theta}_k\}$ son puntos estacionarios (o máximos locales) de $L(\theta)$ y $L(\hat{\theta}_k)$ converge monótonicamente a $L^ = L(\theta^*)$, para algún punto estacionario θ^* .*

Demostración. La demostración se puede encontrar en [9]. ■

Para el algoritmo EM, se establece que la condición necesaria para la propiedad de clausura del mapeo del algoritmo EM, M_{EM} , es que

$$\mathcal{Q}(\theta, \xi) \text{ sea continuo en } \theta \text{ y } \xi \quad (2.2.22)$$

Estas condiciones son débiles y se cumple en la mayoría de las situaciones prácticas.

El siguiente Teorema que se encuentra en [9] establece la convergencia de una secuencia de algoritmo EM.

Teorema 2.2.5. *Suponga que $\mathcal{Q}(\theta, \hat{\theta}_k)$ satisface la condición de continuidad descrita en (2.2.22). Entonces, todos los puntos límites de una secuencia cualquiera $\{\hat{\theta}_k\}$ del algoritmo EM son puntos estacionarios de $L(\theta)$, y $L(\hat{\theta}_k)$ converge monótonicamente a $L^* = L(\theta^*)$, donde θ^* es algún punto estacionario.*

Demostración. La demostración de este teorema es consecuencia directa del Teorema 2.2.4, ya que la propiedad de clausura se cumple con la continuidad (2.2.22), mientras que la segunda condición es satisfecha automáticamente por EM en cada iteración. ■

Es así como las iteraciones de EM convergen a un punto estacionario, es decir, que no es necesariamente un máximo global, pudiendo ser máximo local o incluso un punto silla [9]. Es así como se establece la siguiente condición para asegurar convergencia a un máximo local

Teorema 2.2.6. *Suponga que $\mathcal{Q}(\theta, \hat{\theta}_k)$ satisface la condición de continuidad de 2.2.22 y a su vez*

$$\sup_{\theta \in \Theta} \mathcal{Q}(\theta, \hat{\theta}_k) > \mathcal{Q}(\hat{\theta}_k, \hat{\theta}_k) \quad (2.2.23)$$

Entonces todos los puntos límites de una secuencia $\{\hat{\theta}_k\}$ del algoritmo EM son máximos locales de $L(\theta)$ y $L(\hat{\theta}_k)$ converge monóticamente a $L^ = L(\theta^*)$ para algún máximo local θ^* .*

Demostración. Ver [9]. ■

Claramente la condición propuesta por el Teorema 2.2.6 es difícil de verificar, siendo entonces de utilidad limitada. A pesar de esto, en [9] se presentan clases de densidad de probabilidad exponenciales que satisfacen la condición propuesta.

Es así como la convergencia de una secuencia de valores de la función de verosimilitud $\{L(\hat{\theta}_k)\}$ a algún valor L^* no implica la convergencia de la secuencia $\{\theta_k\}$ a el punto θ^* .

2.3. Algoritmo EM en el dominio del tiempo

En esta sección se presentará el algoritmo EM en el dominio del tiempo, específicamente para modelos en variables de estados en tiempo discreto:

$$x_{k+1} = Ax_k + Bu_k + w_k \quad (2.3.1a)$$

$$y_k = Cx_k + Du_k + v_k \quad (2.3.1b)$$

donde las matrices del sistema son $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times n_u}$, $C \in \mathbb{R}^{n_y \times n}$, $D \in \mathbb{R}^{n_y \times n_u}$, y las señales $u_k \in \mathbb{R}^{n_u}$, $y_k \in \mathbb{R}^{n_y}$ y $x_k \in \mathbb{R}^n$ son la entrada, salida y estado del sistema respectivamente. El estado inicial se supone desconocido, pero independiente de las perturbaciones y distribuye de manera Gaussiana, definido como

$$x_0 \sim \mathcal{N}(\mu, P_0) \quad (2.3.2)$$

Además, se supone que w_k y v_k son el ruido del proceso y de medición, que distribuye de manera Gaussiana, los cuales poseen media cero y covarianza

$$\mathbf{E} \left\{ \begin{bmatrix} w_k \\ v_k \end{bmatrix} \begin{bmatrix} w_\ell^T \\ v_\ell^T \end{bmatrix} \right\} = \begin{bmatrix} Q & 0_{n \times n_y} \\ 0_{n_y \times n} & R \end{bmatrix} \delta_K(k - \ell) \quad (2.3.3)$$

en que δ_K es el delta de Kronecker definido en tiempo discreto. El problema principal es el de estimar a partir N mediciones que se tienen de la entrada y la salida los parámetros del sistema, siendo las matrices de estado A , B , C , D y las matrices de covarianza Q y R . Es así, que los parámetros a estimar son

$$\Gamma = \begin{bmatrix} A & B \\ C & D \end{bmatrix}, \quad \Pi = \begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix} \quad (2.3.4a)$$

$$\theta = \{\text{vec}(\Gamma), \text{vec}(\Pi)\} \quad (2.3.4b)$$

a partir de las señales de entrada y salida

$$U_N \triangleq \{u_0, u_1, \dots, u_{N-1}\} \quad (2.3.5a)$$

$$Y_N \triangleq \{y_0, y_1, \dots, y_{N-1}\} \quad (2.3.5b)$$

Para la definición del parámetro θ , la función de log-verosimilitud $L(\theta)$ para el sistema descrito en (2.3.1) es (sin tomar en cuenta los términos constantes) (ver [2] [10])

$$L(\theta) = \frac{-1}{2} \sum_{k=0}^{N-1} \log \det(CP_{k|k-1}C^T + R) - \frac{1}{2} \sum_{t=0}^{N-1} \epsilon_k^T(\theta)[CP_{k|k-1}C^T + R]^{-1}\epsilon_k(\theta) \quad (2.3.6)$$

donde (con $Y_0 \triangleq 0$)

$$\epsilon(\theta)_k = y_k - \hat{y}_{k|k-1}(\theta), \hat{y}_{k|k-1} = \mathbf{E}_\theta \{y_k | Y_{k-1}\} \quad (2.3.7)$$

siendo esta la predicción óptima en sentido cuadrático medio de la salida del sistema y a su vez

$$P_{k|k-1} \triangleq \mathbf{E}_\theta \{(x_k - \hat{x}_{k|k-1})(x_k - \hat{x}_{k|k-1})^T\} \quad (2.3.8)$$

es la matriz de covarianza asociada a la estimación del estado

$$\hat{x}_{k|k-1} \triangleq \mathbf{E}_\theta \{x_k | Y_{k-1}\} \quad (2.3.9)$$

Las dos expresiones anteriores pueden ser obtenidas mediante el predictor de Kalman, que se mencionará posteriormente. Una observación importante, es que si en el sistema descrito en (2.3.1) la secuencia de estado estuviera disponible, entonces podría obtenerse los parámetros θ usando técnicas simples de regresión lineal, por lo que la secuencia de estado es la que naturalmente se escoge como las variables ocultas para el algoritmo EM [10] [24]. Estas variables ocultas las definimos como

$$X = X_N = \{x_0, x_1, \dots, x_N\} \quad (2.3.10)$$

2.3.1. Paso E

El paso E consiste en determinar $\mathcal{Q}(\theta, \hat{\theta}_k)$ definido en (2.2.4) para el modelo descrito en (2.3.1). Esto se obtiene a partir del siguiente resultado

Lema 2.3.1. *Considere el modelo descrito en (2.3.1) - (2.3.3). Si las variables ocultas se escoge como la secuencia de estado $X \triangleq X_N \triangleq \{x_0, \dots, x_N\}$ y la data observada $Y \triangleq \{Y_N, U_N\}$, entonces para un $\Pi > 0$ la función $\mathcal{Q}(\theta, \hat{\theta}_k)$ es dada por*

$$\begin{aligned} -2\mathcal{Q}(\theta, \hat{\theta}_k) = & \log \det P_0 + \text{Tr} \left\{ P_0^{-1} \mathbf{E} \left\{ (x_0 - \mu)(x_0 - \mu)^T | Y_N, \hat{\theta}_k \right\} \right\} \\ & N \log \det \Pi + N \text{Tr} \left\{ \Pi^{-1} [\Phi - \Psi \Gamma^T - \Gamma \Psi^T + \Gamma \Sigma \Gamma^T] \right\} \end{aligned} \quad (2.3.11)$$

donde

$$z_t^T \triangleq [x_k^T, u_k^T], \xi_k^T \triangleq [x_{k+1}^T, y_k^T] \quad (2.3.12a)$$

$$\Phi \triangleq \frac{1}{N} \sum_{k=0}^{N-1} \mathbf{E} \left\{ \xi_k \xi_k^T | Y_N, \hat{\theta}_k \right\} \quad (2.3.12b)$$

$$\Psi \triangleq \frac{1}{N} \sum_{k=0}^{N-1} \mathbf{E} \left\{ \xi_k z_k^T | Y_N, \hat{\theta}_k \right\} \quad (2.3.12c)$$

$$\Sigma \triangleq \frac{1}{N} \sum_{k=0}^{N-1} \mathbf{E} \left\{ z_k z_k^T | Y_N, \hat{\theta}_k \right\} \quad (2.3.12d)$$

Demostración. La demostración es similar a la presentada en [25], y que se muestra en [10]. Para comenzar, se debe aplicar la reglas de Bayes, y usando las propiedades de Markov entonces en (2.3.1) se tiene

$$\begin{aligned} p(Y_N, X_N) &= p(Y_N | X_{N-1}) p(X_{N-1}) \\ &= p(x_0) \prod_{k=0}^{N-1} p(x_{k+1}, y_k | x_k) \end{aligned} \quad (2.3.13)$$

dado que las perturbaciones tienen distribución Gaussiana implica que

$$\begin{aligned} p(x_0) &\sim \mathcal{N}(\mu, P_0) \\ p \left(\begin{bmatrix} x_{k+1} \\ y_k \end{bmatrix} | x_k \right) &\sim \mathcal{N} \left(\begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} x_k \\ u_k \end{bmatrix}, \begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix} \right) = \mathcal{N}(\Gamma z_k, \Pi) \end{aligned} \quad (2.3.14)$$

aplicando logaritmo a (2.3.13) se tiene entonces

$$\begin{aligned} \log p(Y_N, X_{N+1}) &= \log p(x_0) + \sum_{k=0}^{N-1} \log p(x_{k+1}, y_k | x_k) \\ &= -\frac{n2\pi}{2} - \frac{1}{2} \log \det P_0 - \frac{1}{2} (x_0 - \mu)^T P_0^{-1} (x_0 - \mu) \\ &\quad - \frac{(n + n_y)2\pi}{2} - \frac{1}{2} \log \det \Pi \\ &\quad - \sum_{k=0}^{N-1} \left[\left(\begin{bmatrix} x_{k+1} \\ y_k \end{bmatrix} - \Gamma \begin{bmatrix} x_k \\ u_k \end{bmatrix} \right)^T \Pi^{-1} \left(\begin{bmatrix} x_{k+1} \\ y_k \end{bmatrix} - \Gamma \begin{bmatrix} x_k \\ u_k \end{bmatrix} \right) \right] \end{aligned} \quad (2.3.15)$$

Usando propiedades de traza [26] es que se tiene

$$u^T A v = \text{Tr} \{ u^T A v \} = \text{Tr} \{ A v u^T \} \quad (2.3.16a)$$

$$\text{Tr} \{ A + B \} = \text{Tr} \{ A \} + \text{Tr} \{ B \} \quad (2.3.16b)$$

donde A, B son matrices y u, v vectores de dimensiones adecuadas. Es así entonces que la expresión (2.3.15) la amplificamos por -2 y quitamos los valores constantes (dado que no afectan en la optimización) quedando

$$\begin{aligned} -2 \log p(Y_N, X_{N+1}) &= \log \det P_0 + \text{Tr} \{ P_0^{-1} (x_0 - \mu)(x_0 - \mu)^T \} \\ &\quad + \log \det \Pi \\ &\quad + \text{Tr} \left\{ \sum_{k=0}^{N-1} \Pi^{-1} \left[\left(\begin{bmatrix} x_{k+1} \\ y_k \end{bmatrix} + \Gamma \begin{bmatrix} x_k \\ u_k \end{bmatrix} \right) \left(\begin{bmatrix} x_{k+1} \\ y_k \end{bmatrix} - \Gamma \begin{bmatrix} x_k \\ u_k \end{bmatrix} \right)^T \right] \right\} \end{aligned} \quad (2.3.17)$$

Aplicando ahora la esperanza condicional y usando la definición dada por (2.3.12a) se tiene que

$$\begin{aligned}
-2 \mathbf{E} \left\{ \log p(Y_N, X_{N+1}) | Y_N, \hat{\theta}_k \right\} &= -2\mathcal{Q}(\theta, \hat{\theta}_k) \\
&= \log \det P_0 + \frac{1}{2} \text{Tr} \left\{ \mathbf{E} \left\{ P_0^{-1} (x_0 - \mu)(x_0 - \mu) | Y_N, \hat{\theta}_k \right\} \right\} \\
&\quad + \log \det \Pi \\
&\quad + \text{Tr} \left\{ \Pi^{-1} \sum_{k=0}^{N-1} \mathbf{E} \left\{ [(\xi_k + \Gamma z_k)(\xi_k - \Gamma z_k)^T] | Y_N, \hat{\theta}_k \right\} \right\}
\end{aligned} \tag{2.3.18}$$

Usando por último la linealidad del operador esperanza y amplificando la traza que contiene la sumatoria por N/N se tiene

$$\begin{aligned}
-2\mathcal{Q}(\theta, \hat{\theta}_k) &= \log \det P_0 + \frac{1}{2} \text{Tr} \left\{ \mathbf{E} \left\{ P_0^{-1} (x_0 - \mu)(x_0 - \mu) | Y_N, \hat{\theta}_k \right\} \right\} \\
&\quad + \log \det \Pi \\
&\quad + N \text{Tr} \left\{ \Pi^{-1} [\Phi - \Psi \Gamma^T - \Gamma \Psi^T + \Gamma \Sigma \Gamma^T] \right\}
\end{aligned} \tag{2.3.19}$$

Donde las matrices Φ , Ψ , Σ son las definidas en (2.3.12), por lo que queda demostrado. ■

Ahora bien, las expresiones de (2.3.12) son obtenidas mediante el suavizador de Kalman, el cual se verá a continuación

2.3.2. Filtro y Suavizador de Kalman

En esta sección se presentará el filtro de Kalman y el suavizador de Kalman, dado que ambos son necesarios de implementar para el paso E de EM. Las ecuaciones del suavizador de Kalman requieren computar anteriormente el filtro de Kalman. Ahora bien, consideremos el sistema de tiempo discreto dado en (2.3.1) - (2.3.3). Se denota la estimación de un estado x_k hasta el instante n por

$$\hat{x}_{k|n} \triangleq \mathbf{E} \{ x_k | Y_n \} \tag{2.3.20}$$

Donde Y_n son todas las mediciones que se poseen hasta el instante n (tanto las mediciones en la salida como la señal de entrada). A su vez, la matriz de covarianza de error se define por

$$P_{k|n} \triangleq \mathbf{E} \{ (x_k - \hat{x}_{k|n})(x_k - \hat{x}_{k|n})^T \} \tag{2.3.21}$$

Y sea el error de estimación denotado por

$$\tilde{x}_{k|n} \triangleq x_k - \hat{x}_{k|n} \tag{2.3.22}$$

El filtro de Kalman es entonces el mejor estimador de lineal de la secuencia de estados en sentido cuadrático medio [27] [28] [29] y define este teorema

Teorema 2.3.2. *El filtro óptimo respecto a la covarianza de estimación para estimar la secuencia de estados del sistema (2.3.1) - (2.3.3) esta dado por*

$$\hat{x}_{k|k-1} = A \hat{x}_{k-1|k-1} + B u_{k-1} \tag{2.3.23a}$$

$$P_{k|k-1} = AP_{k-1|k-1}A^T + Q \quad (2.3.23b)$$

$$K_k = P_{k|k-1}C^T(CP_{k|k-1}C^T - R)^{-1} \quad (2.3.23c)$$

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k(y_k - Cx_{k|k-1} - Du_k) \quad (2.3.23d)$$

$$P_{k|k} = (I - K_kC)P_{k|k-1} \quad (2.3.23e)$$

Demostración. La demostración de las ecuaciones se puede encontrar en [29] [27]. ■

Con esto, resta revisar el suavizador de Kalman y ver el uso de este (y del filtro de Kalman) en el paso E del algoritmo EM. Entre las implementaciones del suavizador de Kalman la más conocida es la del algoritmo RTS que se encuentra desarrollada en [30] (las siglas hacen referencia a los autores Rauch-Tung-Striebel).

Teorema 2.3.3 (Kalman Smoother). *El suavizador de Kalman correspondiente al modelo de estados definido en (2.3.1) - (2.3.3), válido para $k = 1, \dots, N$ con condiciones finales $\hat{x}_{n|n}$ y $P_{N|N}$ obtenidos por el filtro de Kalman visto anteriormente esta dado por*

$$\hat{x}_{k-1|N} = \hat{x}_{k-1|k-1} + S_{k-1}(\hat{x}_{k|N} - \hat{x}_{k|k-1}) \quad (2.3.24a)$$

$$P_{k-1|N} = P_{k-1|k-1} + S_{k-1}(P_{k|N} - P_{k|k-1})S_{k-1}^T \quad (2.3.24b)$$

donde

$$S_{k-1} = P_{k-1|k-1}A^T P_{k|k-1}^{-1} \quad (2.3.25)$$

Demostración. La demostración se puede encontrar en [30], como también otra demostración más directa se encuentra en [31] quien toma los principales resultados de [32, p. 367] ■

Cabe señalar que, en las ecuaciones descritas en el paso E mostradas en (2.3.12), es necesario obtener una expresión para la covarianza entre dos errores de estimación sucesivos, es decir

$$M_{k+1|N} \triangleq \mathbf{E} \{ (x_{k+1} - \hat{x}_{k+1|N})(x_k - \hat{x}_{k|N})^T | Y_N \} \quad (2.3.26)$$

Para esto, se define el siguiente Lema [10]

Lema 2.3.4. *Sean las ecuaciones del Kalman Smoother descrito anteriormente. La covarianza cruzada entre dos errores de estimación sucesivos es definido por*

$$M_{k|N} = P_{k|k}S_{k-1}^T + S_k(M_{k+1|N} - AP_{k|k})S_{k-1}^T \quad (2.3.27)$$

para $k = N - 1, N - 2, \dots, 1$. La condición final es

$$M_{N|N} = (I - K_N C)AP_{N-1|N-1} \quad (2.3.28)$$

Demostración. Usando la definición dada en (2.3.26) y las definiciones dadas por (2.3.23d) y (2.3.24a) se obtiene

$$\begin{aligned} M_{k|k} &= \mathbf{E} \{ \tilde{x}_{k|k} \tilde{x}_{k-1|k} \} = \mathbf{E} \{ (x_k - \hat{x}_{k|k})(x_{k-1} - \hat{x}_{k-1|k})^T \} \\ &= \mathbf{E} \{ (\tilde{x}_{k|k-1} - K_k(C\tilde{x}_{k|k-1} + v_k))(\tilde{x}_{k-1|k-1} + S_{k-1}K_k(C\tilde{x}_{k|k-1} + v_k))^T \} \end{aligned} \quad (2.3.29)$$

desarrollando lo expresado se tiene

$$\begin{aligned}
M_{k|k} &= M_{k|k-1} + P_{k|k-1}(S_{k-1}K_kC)^T - K_kCM_{k|k-1} \\
&\quad - K_kCP_{k|k-1}(S_{k-1}K_kC)^T - K_kRK_k^TS_{k-1}^T \\
&= (I - K_kC)M_{k|k-1} + (I - K_kC)P_{k|k-1}C^TK_k^TS_{k|k-1}^T \\
&\quad - K_kRK_k^TS_{k-1}^T
\end{aligned} \tag{2.3.30}$$

Combinando esta expresión con (2.3.23c) se obtiene

$$\begin{aligned}
M_{k|k} &= (I - K_kC)M_{k|k-1} + P_{k|k-1}C^T(CP_{k|k-1}C^T + R)^{-1}CP_{k|k-1}S_{k-1}^T \\
&\quad - P_{k|k-1}C^T(CP_{k|k-1}C^T + R)^{-1}(CP_{k|k-1}C^T + R) \\
&\quad \times (CP_{k|k-1}C^T + R)^{-1}CP_{k|k-1}S_{k-1}^T \\
&= (I - K_kC)M_{k|k-1} + P_{k|k-1}C^T(CP_{k|k-1}C^T + R)^{-1}CP_{k|k-1}S_{k-1}^T \\
&\quad - P_{k|k-1}C^T(CP_{k|k-1}C^T + R)^{-1}CP_{k|k-1}S_{k-1}^T \\
&= (I - K_kC)M_{k|k-1}
\end{aligned} \tag{2.3.31}$$

a su vez

$$\begin{aligned}
M_{k|k-1} &= \mathbf{E} \left\{ \tilde{x}_{k|k-1} \tilde{x}_{k-1|k-1}^T \right\} = \mathbf{E} \left\{ (A\tilde{x}_{k-1|k-1} + \tilde{w}_{k-1})(\tilde{x}_{k-1|k-1}^T) \right\} \\
&= A_{k-1}P_{k-1|k-1}
\end{aligned} \tag{2.3.32}$$

así (2.3.31) queda como

$$M_{k|k} = (I - K_kC)AP_{k-1|k-1} \tag{2.3.33}$$

y para $k = N$ se tiene

$$M_{N|N} = (I - K_NC)AP_{N-1|N-1} \tag{2.3.34}$$

probando la condición final (2.3.28).

Usando (2.3.24a) se puede tener

$$\begin{aligned}
x_k - \hat{x}_{k|N} &= \tilde{x}_{k|N} = \tilde{x}_{k|k} - S_k(\hat{x}_{k+1|N} - \hat{x}_{k+1|k}) \\
\Leftrightarrow \tilde{x}_{k|N} + S_k\hat{x}_{k+1|N} &= \tilde{x}_{k|k} + S_k\hat{x}_{k+1|k}
\end{aligned} \tag{2.3.35}$$

Así mismo

$$\tilde{x}_{k-1|N} + S_{k-1}\hat{x}_{k|N} = \tilde{x}_{k-1|k-1} + S_{k-1}\hat{x}_{k|k-1} \tag{2.3.36}$$

Multiplicando (2.3.35) por el transpuesto de (2.3.36) por la derecha se obtiene

$$\begin{aligned}
&(\tilde{x}_{k|N} + S_k\hat{x}_{k+1|N})(\tilde{x}_{k-1|N}S_{k-1}\hat{x}_{k|N})^T \\
&= (\tilde{x}_{k|k} + S_k\hat{x}_{k+1|k})(\tilde{x}_{k-1|k-1}S_{k-1}\hat{x}_{k|k-1})^T \\
&= [(I - K_kC)\tilde{x}_{k|k-1} - K_k\tilde{v}_k + S_k(A\tilde{x}_{k|k} + Bu_k) \\
&\quad \times [\tilde{x}_{k-1|k-1} + S_{k-1}(A\hat{x}_{k-1|k-1} + Bu_{k-1})]^T
\end{aligned} \tag{2.3.37}$$

Aplicando esperanza y trabajando la expresi3n algebraica se tiene

$$\begin{aligned}
M_{k|N} + S_k \mathbf{E} \left\{ \hat{x}_{k+1|N} \hat{x}_{k|N}^T \right\} S_{k-1}^T \\
&= \mathbf{E} \{ [(I - K_k C + S_k A K_k C) \tilde{x}_{k|k-1} + (S_k A - I) K_k \tilde{v}_k + \\
&\quad + S_k (A \hat{x}_{k|k-1} + B u_k)] [\tilde{x}_{k-1|k-1} + S_{k-1} (A \hat{x}_{k-1|k-1} + B u_{k-1})] \} \\
&= (I - K_k C + S_k A K_k C) M_{k|k-1} + S_k [A_k \mathbf{E} \{ \hat{x}_{k|k-1} \hat{x}_{k-1|k-1}^T \} A^T \\
&\quad + A \mathbf{E} \{ \hat{x}_{k|k-1} u_{k-1}^T \} B^T + B \mathbf{E} \{ u_k \hat{x}_{k-1|k-1} \} A^T \\
&\quad + B \mathbf{E} \{ u_k u_{k-1}^T \} B^T] S_{k-1}
\end{aligned} \tag{2.3.38}$$

y a su vez

$$\begin{aligned}
\mathbf{E} \left\{ \hat{x}_{k+1|N} \hat{x}_{k|N}^T \right\} &= \mathbf{E} \{ (x_{k+1} - \tilde{x}_{k+1|N})(x_k - \tilde{x}_{k|N})^T \} \\
&= \mathbf{E} \{ x_{k+1} x_k^T \} - \mathbf{E} \{ (\tilde{x}_{k+1|N} + \hat{x}_{k|N}) \tilde{x}_{k|N}^T \} \\
&\quad - \mathbf{E} \{ \tilde{x}_{k+1|N} (\tilde{x}_{k|N} + \hat{x}_{k|N})^T \} + \mathbf{E} \{ \tilde{x}_{k+1|N} x_{k|N}^T \} \\
&= \mathbf{E} \{ x_{k+1} x_k^T \} - M_{k+1|N} - M_{k+1|N} + M_{k+1|N} \\
&= \mathbf{E} \{ x_{k+1} x_k^T \} - M_{k+1|N}
\end{aligned} \tag{2.3.39}$$

tambi3n se cumple que

$$\begin{aligned}
\mathbf{E} \{ x_{k+1} x_k^T \} &= \mathbf{E} \{ [A x_k + B u_k + \tilde{w}_k] [A x_{k-1} + B u_{k-1} + w_{k-1}]^T \} \\
&= A \mathbf{E} \{ x_k x_{k-1}^T \} A^T + A \mathbf{E} \{ x_k u_{k-1}^T \} B^T + A Q_{k-1} \\
&\quad + B \mathbf{E} \{ u_k x_{k-1}^T \} A^T + B \mathbf{E} \{ u_k u_{k-1}^T \} B^T
\end{aligned} \tag{2.3.40}$$

Procediendo de manera similar a antes se puede llegar a

$$\mathbf{E} \left\{ \hat{x}_{k|k-1} \hat{x}_{k-1|k-1}^T \right\} = \mathbf{E} \{ x_k x_{k-1}^T \} - M_{k|k-1} \tag{2.3.41}$$

y finalmente

$$\mathbf{E} \{ \hat{x}_{k|k-1} u_{k-1}^T \} = \mathbf{E} \{ (x_k \tilde{x}_{k|k-1}) u_{k-1}^T \} = \mathbf{E} \{ x_k u_{k-1}^T \} \tag{2.3.42}$$

Usando asi las expresiones que se obtuvieron en (2.3.39) - (2.3.42) en (2.3.38) es que se puede llegar simplificando t3rminos a

$$\begin{aligned}
M_{k|N} + S_k (A Q_{k-1} - M_{k+1|N}) S_{k-1}^T &= (I - K_k C + S_k A K_k C) M_{k|k-1} \\
&\quad - S_k (A M_{k|k-1} A^T) S_{k-1}
\end{aligned} \tag{2.3.43}$$

usando en esta expresi3n la obtenida en (2.3.32) se puede obtener

$$\begin{aligned}
M_{k|N} &= (I - K_k C) A P_{k-1|k-1} + S_k A K_k C A P_{k-1|k-1} \\
&\quad + S_k (M_{k+1|N} - A (A P_{k-1|k-1} A^T + Q)) S_{k-1}^T \\
&= P_{k|k} P_{k|k-1}^{-1} A P_{k-1|k-1} + S_k A K_k C A P_{k-1|k-1} \\
&\quad + S_k (M_{k+1|N} - A P_{k|k-1}) S_{k-1}^T
\end{aligned} \tag{2.3.44}$$

Usando la definición de S_k se puede despejar de la siguiente manera

$$S_k = P_{k|k} A^T P_{k+1|k}^{-1} \Leftrightarrow P_{k+1|k} S_k^T = A_k P_{k|k} \quad (2.3.45)$$

Esto se logra multiplicando ambos lados de la ecuación por la derecha por $P_{k+1|k}$ y luego transponiendo ambas partes. Es así entonces que (2.3.44) queda como

$$\begin{aligned} M_{k|N} &= P_{k|k} S_{k-1}^T + S_k A K_k C P_{k|k-1} S_{k-1}^T + S_k (M_{k+1|N} - A P_{k|k-1}) S_{k-1}^T \\ &= P_{k|k} S_{k-1}^T + S_k (M_{k+1|N} - A(I - K_k C) P_{k|k-1}) S_{k-1}^T \\ &= P_{k|k} S_{k-1}^T + S_k (M_{k+1|N} - A P_{k|k}) S_{k-1}^T \end{aligned} \quad (2.3.46)$$

Quedando así (2.3.27) demostrado. ■

2.3.3. Paso M

En esta sección se analiza el paso M, en el cual se obtiene un $\hat{\theta}_{k+1}$ con la función obtenida en el paso E y obteniendo la estimación de la secuencia de estados. En este caso, dado que (2.3.11) incluye explícitamente la distribución del estado inicial (su media μ y varianza P_0), es que estos también se pueden considerar parámetros a estimar. Antes de definir el Lema que muestra las expresiones obtenidas en el paso M, notemos que

Observación 2.3.5. Sea $\theta \in \Theta$ el vector de parámetros definido por

$$\theta \triangleq [\beta^T, \eta^T]^T \quad (2.3.47a)$$

$$\beta \triangleq [\text{vec}(\Gamma), \text{vec}(\mu)] \quad (2.3.47b)$$

$$\eta \triangleq [\text{vec}(\Pi), \text{vec}(P_0)] \quad (2.3.47c)$$

Si $\Theta = \Theta_1 \times \Theta_2$, con $\beta \in \Theta_1$ y $\eta \in \Theta_2$, donde Θ_1 es un espacio cerrado en \mathbb{R}^ℓ , con $\ell = n^2 + n \cdot n_u + n \cdot n_y + n_y^2 + n$, y Θ_2 un subespacio cerrado de R^v , con $v = n^2 + (n + n_y)^2$, se asume que todos los $\eta \in \Theta_2$ son tales que implican matrices simétricas y positivas definidas para Π y P_0 .

La observación anterior es importante, dado que η define las estimaciones de las matrices Π y P_0 , las cuales al representar matrices de covariancia deben ser simétricas y positivas definidas. Ahora definimos el Lema para el paso M

Lema 2.3.6 (Paso M). Sea Σ, Ψ definidos en (2.3.12) que satisface $\Sigma > 0$, es así que como los términos que maximizan $\mathcal{Q}(\theta, \hat{\theta}_k)$ son

$$\Gamma \triangleq \begin{bmatrix} \hat{A}_{k+1} & \hat{B}_{k+1} \\ \hat{C}_{k+1} & \hat{D}_{k+1} \end{bmatrix} = \Psi \Sigma^{-1} \quad (2.3.48a)$$

$$\hat{\mu} \triangleq \hat{x}_{0|N} \quad (2.3.48b)$$

A su vez, para los parámetros restantes (usando Φ definido en (2.3.12))

$$\Pi \triangleq \begin{bmatrix} \hat{Q}_{k+1} & 0 \\ 0 & \hat{R}_{k+1} \end{bmatrix} = \Phi - \Psi \Sigma^{-1} \Psi^T \quad (2.3.49a)$$

$$\hat{P}_0 \triangleq P_{0|N} \quad (2.3.49b)$$

Entonces el vector η definido en (2.3.47) es un punto estacionario de $\mathcal{Q}([\hat{\beta}^T, \eta^T]^T, \hat{\theta}_k)$ con respecto a η , y los estimados de Π como P_0 definidos anteriormente son no negativos por construcción. A su vez, si los estimados de Π y P_0 son positivos definidos, si la secuencia de entrada $\{u_k\}$ es tal que

$$\sum_{k=0}^{N-1} u_k u_k^T > 0 \quad (2.3.50)$$

si $\hat{\theta}_k$ implica un sistema controlable y observable, y si $\eta \in \Theta_2$, entonces

$$\hat{\eta} = \operatorname{argmax}_{\eta \in \Theta_2} \mathcal{Q} \left(\begin{bmatrix} \hat{\beta} \\ \eta \end{bmatrix}, \hat{\theta}_k \right) \quad (2.3.51)$$

En combinación con las expresiones obtenidas en (2.3.48), esto otorga un único máximo global de $\mathcal{Q}(\theta, \hat{\theta}_k)$ para $\theta \in \Theta$.

Demostración. Considere en primera instancia maximizar $\mathcal{Q}(\theta, \hat{\theta}_k)$ con respecto a β . Es así que solo hay que tomar las expresiones dependientes de Γ y μ , las cuales pueden ser escritas como

$$\mathcal{Q}(\mu, \hat{\theta}_k) = -\operatorname{Tr} \left\{ P_0^{-1} \mathbf{E} \left\{ (x_0 - \mu)(x_0 - \mu)^T | Y_N, \hat{\theta}_k \right\} \right\} = -\operatorname{Tr} \left\{ P_0^{-1} [(\hat{x}_{0|N} - \mu)(\hat{x}_{0|N} - \mu)^T + P_{0|N}] \right\} \quad (2.3.52)$$

y

$$\begin{aligned} \mathcal{Q}(\Pi, \hat{\theta}_k) &= \operatorname{Tr} \left\{ \Pi^{-1} [\Phi - \Psi \Gamma^T - \Gamma \Psi^T + \Gamma \Sigma^{-1} \Gamma^T] \right\} = \\ &= \operatorname{Tr} \left\{ \Pi^{-1} [(\Gamma - \Psi \Sigma^{-1}) \Sigma (\Gamma - \Psi \Sigma^{-1})^T + \Phi - \Psi \Sigma^{-1} \Psi^T] \right\} \end{aligned} \quad (2.3.53)$$

Claramente ambos términos anteriores son maximizados de manera global con respecto a los elementos de β con las elecciones tomadas en (2.3.48).

Ahora bien, maximizando respecto a los términos de η , teniendo en cuenta el $\hat{\beta}$ escogido anteriormente, es que se puede derivar matricialmente e igualar a cero

$$\begin{aligned} \frac{\partial}{\partial \Pi} \mathcal{Q}(\Pi, \hat{\theta}_k) &= \frac{\partial}{\partial \Pi} \log \det \Pi + \frac{\partial}{\partial \Pi} \operatorname{Tr} \left\{ \Pi^{-1} (\Phi - \Psi \Sigma^{-1} \Psi^T) \right\} \\ &= \Pi^{-1} - \Pi^{-1} (\Phi - \Psi \Sigma^{-1} \Psi^T) \Pi^{-1} \end{aligned} \quad (2.3.54)$$

Esto es claramente cero para la elección para Π mostrada en (2.3.49), el cual es definido como un punto estacionario de $\mathcal{Q}(\theta, \hat{\theta}_k)$ con respecto a una parametrización simétrica y definida positiva de Π , con lo que se completa las demostraciones para (2.3.48) - (2.3.49). Las demostraciones que quedan en el Lema aparecen en [10]. ■

Con todo lo expuesto anteriormente, es que se tiene completo el algoritmo EM para un sistema lineal invariante en el tiempo, en tiempo discreto y en variables de estado definido en (2.3.1) - (2.3.3).

2.4. Sistemas en tiempo continuo y modelos discretos

Todo lo analizado anteriormente es para sistemas discretos. Cuando los datos de entrada y salida se obtienen a partir de un sistema en tiempo continuo sujeto a perturbaciones estocásticas es posible modelarlo como un sistema discreto considerado en la Sección 2.3. La relación entre ambos dominios (continuo y discreto) se encuentra en [11] mediante el siguiente modelo de ecuaciones diferenciales estocásticas (SDE-model, por sus siglas en inglés) [33]

$$dx(t) = A_c x(t)dt + B_c u(t)dt + dw(t) \quad (2.4.1)$$

$$dz(t) = C_c x(t)dt + D_c u(t)dt + dv(t) \quad (2.4.2)$$

donde $u(t) \in \mathbb{R}$, $y(t) \in \mathbb{R}$, y $x(t) \in \mathbb{R}^n$ son la entrada, salida y estados respectivamente; como también las matrices del sistema son $A_c \in \mathbb{R}^{n \times n}$, $B_c \in \mathbb{R}^{n \times 1}$, $C_c \in \mathbb{R}^{1 \times n}$ y $D_c \in \mathbb{R}$; y las perturbaciones incrementales de estado $dw(t)$ y la perturbación incremental en la salida $dv(t)$ son procesos estocásticos independientes, de media cero y distribución Gaussiana (esto es, w y v son procesos de Wiener) tal que

$$\mathbf{E} \left\{ \begin{bmatrix} dw(t) \\ dv(t) \end{bmatrix} \begin{bmatrix} dw(s) \\ dv(s) \end{bmatrix}^T \right\} = \begin{cases} \begin{bmatrix} Q_c & 0 \\ 0 & R_c \end{bmatrix} dt & t = s \\ 0 & t \neq s \end{cases} \quad (2.4.3)$$

El estado inicial es independiente de $dw(t)$ y $dv(t)$, de distribución Gaussiana con media \bar{x}_0 y covarianza P_0 . Se asume entonces que existe conocimiento a priori de algunas características del sistema, tal como podría ser la estabilidad del sistema. Se asume también que el periodo de muestreo es constante $T_s = \Delta$.

2.4.1. Modelos en tiempo discreto

El problema de muestrear señales de tiempo continuo sujeta a perturbaciones ha sido ampliamente analizado en la literatura [11] [34] [35] [36] [37]. Se asume entonces que la entrada en tiempo continuo $u(t)$ se genera por alguna secuencia de entrada u_k mediante un retentor de orden cero (ZOH, por sus siglas en inglés), esto quiere decir que

$$u(t) = u_k, \quad k\Delta < t < (k+1)\Delta \quad (2.4.4)$$

donde Δ es el período de muestreo constante. A su vez, se incluye un filtro de integración y reseteo (IRF, por sus siglas en inglés) a la señal de la salida del sistema antes del muestreo instantáneo, tal como se describe en [11] [36] [38]. Bajo estas condiciones, es que se puede definir el siguiente modelo en tiempo discreto en una forma incremental, la cual posee las mismas propiedades de segundo orden que el sistema continuo muestreado [37]

$$dx_k^+ = A_\delta x_k \Delta + B_\delta u_k \Delta + dw_k^+ \quad (2.4.5)$$

$$\bar{y}_{k+1} \Delta = dz_k^+ = C_\delta x_k \Delta + D_\delta u_k \Delta + dv_k^+ \quad (2.4.6)$$

donde los incrementos se definen por

$$df_k^+ = f_{k+1} - f_k \quad (2.4.7)$$

Las matrices son dadas por

$$A_\delta = \frac{e^{A_c \Delta} - I}{\Delta}, \quad B_\delta = \left[\frac{1}{\Delta} \int_0^\Delta e^{A_c \eta} d\eta \right] B_c \quad (2.4.8)$$

$$C_\delta = C_c \left[\frac{1}{\Delta} \int_0^\Delta e^{A_c \eta} d\eta \right] \quad (2.4.9)$$

$$D_\delta = D_c + C_c \left[\frac{1}{\Delta} \int_0^\Delta \int_0^\xi e^{A_c \eta} d\eta d\xi \right] B_c \quad (2.4.10)$$

y la estructura de la covarianza de los vectores de ruidos es dada por

$$\mathbf{E} \left\{ \begin{bmatrix} dw_l^+ \\ dv_l^+ \end{bmatrix}^T \begin{bmatrix} dw_k^+ \\ dv_k^+ \end{bmatrix} \right\} = \begin{bmatrix} Q_\delta & S_\delta \\ (S_\delta)^T & R_\delta \end{bmatrix} \Delta \delta_k(l-k) \quad (2.4.11)$$

donde

$$\begin{aligned} \begin{bmatrix} Q_\delta & S_\delta \\ (S_\delta)^T & R_\delta \end{bmatrix} &= \frac{1}{\Delta} \int_0^\Delta \begin{bmatrix} e^{A_c \eta} & 0 \\ C_c \int_0^\eta e^{A_c \xi} d\xi & I \end{bmatrix} \\ &\quad \times \begin{bmatrix} Q_c & 0 \\ 0 & R_c \end{bmatrix} \begin{bmatrix} e^{A_c \eta} & 0 \\ C_c \int_0^\eta e^{A_c \xi} d\xi & I \end{bmatrix}^T d\eta \end{aligned} \quad (2.4.12)$$

Demostración. La demostración se encuentra en [36], y una implementación estable en términos numéricos para obtener las integrales de las exponenciales matriciales se presenta en [39]. Se observa que si el periodo de muestreo Δ tiende a cero, este modelo incremental (2.4.5)-(2.4.12) converge a su representación SDE (2.4.1)-(2.4.2). ■

Alternativamente, este modelo en tiempo discreto en forma incremental (2.4.5)-(2.4.12) puede ser expresado usando el operador de adelanto q :

$$qx_k = x_{k+1} = A_q x_k + B_q u_k + \tilde{w}_k \quad (2.4.13)$$

$$\bar{y}_k = C_q x_k + D_q u_k + \tilde{v}_k \quad (2.4.14)$$

donde la estructura de la covarianza de los vectores de ruido es

$$\mathbf{E} \left\{ \begin{bmatrix} \tilde{w}_l \\ \tilde{v}_l \end{bmatrix}^T \begin{bmatrix} \tilde{w}_k \\ \tilde{v}_k \end{bmatrix} \right\} = \begin{bmatrix} Q_q & S_q \\ (S_q)^T & R_q \end{bmatrix} \delta_K(l-k) \quad (2.4.15)$$

Con esto es que se puede definir las siguientes transformaciones afines entre las matrices del modelo en tiempo discreto en su modo incremental con el descrito mediante el operador de adelanto q tal como se ve a continuación

$$A_q = I + \Delta A_\delta, \quad B_q = \Delta B_\delta, \quad C_q = C_\delta, \quad D_q = D_\delta \quad (2.4.16)$$

$$Q_q = \Delta Q_\delta, \quad S_q = S_\delta, \quad R_q = \frac{1}{\Delta} R_\delta \quad (2.4.17)$$

Con esto podemos definir la siguiente observación

Observación 2.4.1. *El modelo con muestreo de las señales en forma incremental es ideal para aplicaciones con altas tasas de muestreo, debido a su convergencia cuando $\Delta \rightarrow 0$ [35] [36] [37]. A su vez, este modelo puede ser preferido para estimar los parámetros de tiempo continuo de manera directa a partir de datos muestreados [37].*

Lema 2.4.2. *Las condiciones de estabilidad sobre las matrices A para los sistemas discretos definidos en (2.4.5)-(2.4.12) y (2.4.13)-(2.4.15) son*

- *Para el sistema discreto definido por el operador de adelanto q es que los autovalores de A_q se encuentren en $z \in \mathbb{C}$ tal que $|z| < 1$, o sea, el círculo unitario.*
- *Para el sistema discreto en forma incremental, es que los autovalores de A_δ es que se encuentren en $\gamma \in \mathbb{C}$ tal que $|\gamma + \frac{1}{\Delta}| < \frac{1}{\Delta}$.*

Demostración. Para el sistema discreto definido mediante el operador de adelanto se tiene que (2.4.13) se puede reescribir como

$$x_{k+1} = A_q^{k+1}x_0 + \sum_{k=0}^{n-1} A_q^k B_q u_{n-1-k} + \sum_{k=0}^{n-1} A_q^k \bar{w}_{k-1-k} \quad (2.4.18)$$

Se puede apreciar que las sumatorias que contienen A^k pueden tomarse como sumatorias geométricas matriciales, por lo tanto la condición de estabilidad es que los autovalores de A_q se encuentren dentro del círculo unitario. Más detalles se encuentran por ejemplo en [37].

Para el sistema discreto definido en forma incremental, es que se puede usar las ecuaciones de equivalencia entre los modelos (2.4.16). Aquí se ve que el mapeo entre autovalores de A_q a A_δ es $z = 1 + \Delta\gamma$, por tanto si la condición de estabilidad para A_q es que sus autovalores se encuentren en la región $|z| < 1$, entonces la condición para autovalores de A_δ es que sus autovalores se encuentren en la región $|1 + \Delta\gamma| < 1 \Leftrightarrow |\frac{1}{\Delta} + \gamma| < 1/\Delta$, quedando demostrado. ■

Observación 2.4.3. *Note que las restricciones cuando se estima A_q o A_δ nacen a partir de conocimiento que se tenga a partir de A_c .*

- *Por ejemplo, si A_c es estable, entonces A_q y A_δ son estables (esto quiere decir que sus autovalores se encuentran dentro de las regiones de estabilidad definidas para cada uno).*
- *A su vez, si A_c posee autovalores complejos conjugados (o sea, polos resonantes), entonces A_q y A_δ poseen polos resonantes.*

Detalles se pueden encontrar en [37].

REGIONES DEFINIDAS POR DESIGUALDADES MTRICIALES

Una amplia gama de problemas asociados a control e identificación de sistemas han podido simplificarse a problemas de optimización convexa usando desigualdades matriciales lineales (LMI, por sus siglas en inglés). El uso de LMI en el análisis de sistemas dinámicas se remonta a las desigualdades de Lyapunov, que establecen que un sistema continuo y lineal es estable sí y sólo sí existe una matriz P simétrica y definida positiva tal que la matriz A del sistema cumpla con $AP + PA^T < 0$.

En teoría de control existen varios ejemplos de usos de LMI. Entre los principales destacan la de fijar los polos de lazo cerrado cumpliendo ciertos criterios de desempeño. Por ejemplo, en [40] [41] [42] se muestra que la síntesis de controladores que cumplan cierto desempeño en H_∞ puede ser formulado como un problema de optimización convexo usando LMI. Estas LMI corresponden a desigualdades que son similares a la resolución de la ecuación de Riccati. Dado que las LMI reflejan principalmente una restricción en la ubicación de los polos, es que ofrece mayor flexibilidad en la combinación de esto con otras restricciones en el sistema de lazo cerrado [43]. En [44] se muestra como diseñar controladores de retroalimentación en H_∞ para cumplir distintas clases de restricciones en la ubicación de los polos para regiones convexas en el plano complejo usando LMI. Por otra parte, en [45] se analiza la ubicación de los polos de lazo cerrado de manera robusta en regiones definidas por LMI.

Ejemplos de uso de LMI en identificación de sistemas están asociados a estimación usando métodos de subespacios. Por ejemplo [17] se usan las desigualdades de Lyapunov (LMI) para restringir los polos del modelo identificado a que sean estables. A su vez, dentro de este marco se encuentra el trabajo de [46] [47], en el cual el uso de LMI similar al del [17], apunta a obtener un sistema con polos estrictamente reales. Por su parte [18] extiende estos resultados para regiones basadas en LMI, lo cual es una generalización de los resultados anteriores.

Las LMI pueden ser resueltas de manera eficiente usando algoritmos de punto interior, tal como se describe en [48] [49] [50], ofreciendo una simplificación del problema original en términos numéricos y computacionales.

3.1. Desigualdades matriciales lineales

Una desigualdad matricial lineal tiene la forma de

$$F(x) \triangleq F_0 + \sum_{i=1}^m x_i F_i > 0 \quad (3.1.1)$$

donde $x \in \mathbb{R}^m$ y las matrices $F_i \in \mathbb{R}^{n \times n}$, $i = 0, \dots, m$ son matrices constantes dadas. El símbolo de desigualdad en (3.1.1) hace referencia a que $F(x)$ es una matriz definida positiva, lo que significa que $u^T F(x) u > 0$ para todo vector no nulo $u \in \mathbb{R}^{n \times n}$. La desigualdad dada en (3.1.1) es equivalente a un conjunto de n desigualdades polinomiales en x , es decir, que los principales menores de $F(x)$ sean positivos (ver Criterio de Sylvester en Sección 4).

$F(x)$ es una restricción convexa en x , o sea que el conjunto $\{x \in \mathbb{R}^m | F(x) > 0\}$ es convexo. A su vez, problemas como desigualdades lineales, desigualdades cuadráticas lineales, desigualdades de normas matriciales, y restricciones que aparecen en teoría de control, tal como desigualdad de Lyapunov o desigualdades cuadráticas matriciales lineales pueden llevarse a una forma de una LMI.

Más características de LMI pueden encontrarse en [43].

3.2. Regiones LMI

Las regiones LMI son regiones del plano complejo que puede ser representado por una desigualdad de matrices lineales. Sea \mathcal{D} una región LMI subconjunto del plano complejo definido por

$$\mathcal{D} = \{z \in \mathbb{C} : L + zM + \bar{z}M^T < 0\} \quad (3.2.1)$$

donde $L^T = L \in \mathbb{R}^{m \times m}$ y $M \in \mathbb{R}^{m \times m}$. La función matricial

$$f_{\mathcal{D}}(z) = L + zM + \bar{z}M^T \quad (3.2.2)$$

es llamada la función característica de \mathcal{D} . Note que la función característica $f_{\mathcal{D}}$ es una función que toma valores en las matrices Hermitianas en el espacio de $m \times m$. A su vez, una matriz $P < 0$ indica que $-P$ es definida positiva.

En otras palabras, una región LMI es un subconjunto del plano complejo que es representado por una LMI en z y \bar{z} , o equivalente, una LMI en $x = \text{Re}(z)$ como $y = \text{Im}(z)$, teniendo así la forma dada por (3.1.1). Como resultado, es que la región LMI es convexa y más aún, estas regiones son simétricas con respecto al eje real dado que para cualquier $z \in \mathcal{D}$ se cumple que

$$f_{\mathcal{D}}(z) = \overline{f_{\mathcal{D}}(z)} < 0 \quad (3.2.3)$$

Con esto, se tiene la siguiente definición

Definición 3.2.1 (\mathcal{D} -estabilidad). *Una matriz $A \in \mathbb{R}^{n \times n}$ es \mathcal{D} -estable si sus autovalores se encuentran dentro de la región del plano complejo \mathcal{D} .*

Teorema 3.2.2. *Sea una matriz $A \in \mathbb{R}^{n \times n}$. La matriz A es \mathcal{D} -estable si y sólo si existe una matriz simétrica $P > 0$ tal que*

$$M_{\mathcal{D}}(A, P) < 0 \quad (3.2.4a)$$

$$\begin{aligned} M_{\mathcal{D}}(A, P) &\triangleq L \otimes P + M \otimes (AP) + M^T \otimes (PA^T) \\ &= [L_{k,l}P + M_{k,l}(AP) + M_{l,k}(PA^T)]_{1 \leq k, l \leq m} \end{aligned} \quad (3.2.4b)$$

Demostración. Si $X \in \mathbb{C}^{n \times n}$ es una matriz Hermitiana definida positiva entonces su parte real $\operatorname{Re}(X)$ es una matriz real simétrica y definida positiva. La demostración es que si $X = \operatorname{Re}(X) + \operatorname{Im}(X)j$ y $X = X^H$, es claro que $\operatorname{Re}(X)$ tiene que ser simétrica y que $\operatorname{Im}(X)$ es antisimétrica. Como resultado, si $v^T X v = v^T \operatorname{Re}(X) v$ para cualquier $v \in \mathbb{R}^n$, por lo tanto $\operatorname{Re}(X) = \operatorname{Re}(X)^T > 0$.

Con esto en mente, se empieza la demostración de suficiencia. Sea λ algún autovalor de A , y sea $v \in \mathbb{C}^n$ ser un vector no negativo tal que $v^H A = \lambda v^H$. Usando entonces la identidad

$$(I_m \otimes v^H) M_{\mathcal{D}}(A, P) (I_m \otimes v) = (v^H P v) f_{\mathcal{D}}(\lambda) \quad (3.2.5)$$

es inmediato que si se cumple que $M_{\mathcal{D}}(A, P) < 0$ y $P > 0$ implica que $f_{\mathcal{D}}(\lambda) < 0$, o equivalentemente, que $\lambda \in \mathcal{D}$. Por tanto A es \mathcal{D} -estable. Ahora resta demostrar la necesidad. supongamos que A es \mathcal{D} -estable. Se debe demostrar la existencia de una matriz $P > 0$ tal que $M_{\mathcal{D}} < 0$. Para esto, es útil definir matrices complejas $A \in \mathbb{C}^{n \times n}$ y $P = P^H \in \mathbb{C}^{n \times n}$ como sigue

$$M_{\mathcal{D}}(A, P) \triangleq L \otimes P + M \otimes (AP) + M^T \otimes (AP)^H \quad (3.2.6)$$

Sea el caso donde A es diagonal $\Delta = \operatorname{Diag}(\lambda_i)$, con $\lambda_i \in \mathcal{D}$. Es fácil verificar en tal caso que

$$M_{\mathcal{D}}(\Delta, I) = U^T \operatorname{Diag}(f_{\mathcal{D}}(\lambda_i)) U \quad (3.2.7)$$

donde U es alguna matriz de permutación. Por lo tanto, $M_{\mathcal{D}}(\Delta, P) < 0$ se cumple para $P = I$.

Ahora bien, el caso general sea Δ la matriz diagonal de los autovalores de A (contando la multiplicidad del polinomio característico). De lo visto anteriormente es que tenemos que $M_{\mathcal{D}}(\Delta, I) < 0$ se cumple. Trabajando entonces con la forma canónica de Jordan para A es que se puede construir una secuencia de matrices invertibles T_k tal que

$$\lim_{k \rightarrow \infty} T_k^{-1} A T_k = \Delta \quad (3.2.8)$$

Por ejemplo, se puede tomar

$$T_k = \begin{bmatrix} k & 0 \\ 0 & 1 \end{bmatrix} \quad (3.2.9)$$

si se tiene

$$A = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix} \quad (3.2.10)$$

Dado entonces que $M_{\mathcal{D}}(Y, I)$ es una función continua en Y , se cumple que

$$\lim_{k \rightarrow \infty} M_{\mathcal{D}}(T_k^{-1} A T_k, I) = M_{\mathcal{D}}(\Delta, I) < 0 \quad (3.2.11)$$

Por lo tanto, $M_{\mathcal{D}}(T_k^{-1} A T_k, I) < 0$ para un k lo suficientemente grande. Entonces, dado un k definido y sea $T \triangleq T_k$. Usando la identidad $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$ y la definición dada por (3.2.4b) se puede verificar que

$$(I_m \otimes T) M_{\mathcal{D}}(T^{-1} A T, I) (I_m \otimes T^H) = M_{\mathcal{D}}(A, T T^H) \quad (3.2.12)$$

Juntando esto con que $M_{\mathcal{D}}(T^{-1}AT, I) < 0$, es que se debe cumplir que $M_{\mathcal{D}}(A, P) < 0$ para algún $P = TT^H > 0$.

Para finalizar la demostración, se observa que a pesar de que $P = TT^H$ no es necesariamente real, su parte real satisface $\text{Re}(P) > 0$ y que $M_{\mathcal{D}}(A, \text{Re}(P)) = \text{Re}(M_{\mathcal{D}}(A, P)) < 0$ como consecuencia de lo descrito al principio de la demostración y tomando un A real. ■

Adicionalmente, propiedades importantes de las regiones LMI que se pueden mencionar son [44] [45] [18]

- Intersecciones de regiones LMI son regiones LMI.
- Cualquier región convexa y simétrica con respecto al eje real puede ser descrita por una región LMI.

3.3. Regiones LMI de interés

En esta sección se describirán regiones LMI de interés en lo que respecta a identificación de sistemas.

Lema 3.3.1 (Circunferencia). *Se desea que los autovalores de A se ubiquen dentro de una circunferencia centrada en $(q, 0)$ y de radio r . Para este caso, se tiene que las matrices para (3.2.1) son*

$$L = \begin{bmatrix} -r & -q \\ -q & -r \end{bmatrix}, \quad M = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad (3.3.1)$$

Demostración. Para demostrar esto, es que se toma (3.2.1) como

$$f_{\mathcal{D}} = \begin{bmatrix} -r & -q + z \\ -q + \bar{z} & -r \end{bmatrix} < 0 \Leftrightarrow \begin{bmatrix} r & q - z \\ q - \bar{z} & r \end{bmatrix} > 0 \quad (3.3.2)$$

Dado que la matriz anterior es Hermitiana es que se puede usar complemento de Schur para demostrar que es una matriz definida positiva si se cumple que

$$r > 0 \quad (3.3.3a)$$

$$r - (q - \bar{z})r^{-1}(q - z) > 0 \quad (3.3.3b)$$

Desarrollando la segunda desigualdad de (3.3.3) se tiene que

$$\begin{aligned} r - (q - \bar{z})r^{-1}(q - z) &> 0 \\ r^2 - (q - \bar{z})(q - z) &> 0 \\ r^2 - (q^2 - (z + \bar{z})q + \|z\|^2) &> 0 \\ r^2 - q^2 + 2xq - x^2 - y^2 &> 0 \\ (x - q)^2 + y^2 &< r^2 \end{aligned} \quad (3.3.4)$$

donde $x = \text{Re}(z)$ e $y = \text{Im}(z)$ y siendo esta última la ecuación de circunferencia en el plano complejo, de radio r y centrado en $(q, 0)$. ■

Por tanto, para que los autovalores de una matriz $A \in \mathbb{R}^n$ estén dentro de esta circunferencia se requiere que

$$\begin{bmatrix} -r & -q \\ -q & -r \end{bmatrix} \otimes P + \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \otimes (AP) + \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \otimes (PA^T) < 0_{2n} \quad (3.3.5)$$

Lo cual queda como

$$\begin{bmatrix} -rP & -qP + AP \\ -qP + (AP)^T & -rP \end{bmatrix} < 0_{2n} \quad (3.3.6)$$

Como se desean soluciones cerradas, es que se desea un $P \geq 0_n$ y que la inecuación anterior sea cerrada. Para esto basta escribirla como

$$\begin{bmatrix} r(P - \delta I_n) & -(qP + Q) \\ -(qP + Q^T) & rP \end{bmatrix} \geq 0_{2n} \quad (3.3.7)$$

Lema 3.3.2 (Semiplano). *Se desea que los autovalores de A se ubiquen a la izquierda de la recta $\text{Re}(z) = \alpha$. En tal caso se tiene que las matrices para (3.2.1) son*

$$L = -2\alpha \quad , \quad M = 1 \quad (3.3.8)$$

Demostración. La región en el plano complejo que se encuentra a la izquierda de la recta $\text{Re}(z) = \alpha$ se puede escribir como

$$\text{Re}(z) < \alpha \Leftrightarrow (z + \bar{z}) < 2\alpha \Leftrightarrow -2\alpha + z + \bar{z} \quad (3.3.9)$$

Es así como se encuentra rápidamente las matrices (en este caso, constantes) para la ecuación (3.2.1)

$$L = -2\alpha \quad , \quad M = 1 \quad (3.3.10)$$

quedando así demostrado. ■

Lema 3.3.3 (Banda Imaginaria). *En este caso se quiere que los autovalores de A se ubiquen en la zona del plano complejo $|\text{Im}(z)| < \alpha$. En tal caso las matrices para definir (3.2.1) son*

$$L = \begin{bmatrix} -2\alpha & 0 \\ 0 & -2\alpha \end{bmatrix} \quad , \quad M = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \quad (3.3.11)$$

Demostración. Para la demostración la región en $z \in \mathbb{C}$ tal que

$$|\text{Im}(z)| < \alpha \quad , \quad \alpha > 0 \quad (3.3.12)$$

la podemos definir como

$$w = \frac{z - \bar{z}}{2j} = \text{Im}(z) \quad (3.3.13)$$

entonces se puede trabajar como

$$\begin{aligned} \|w\| &< \alpha \\ \|w\|^2 &< \alpha^2 \\ w \cdot w^* &< \alpha^2 \\ \left(\frac{z - \bar{z}}{2j} \right) \left(\frac{-z + \bar{z}}{-2j} \right) &< \alpha^2 \\ (z - \bar{z})(-z + \bar{z}) &< 4\alpha^2 \\ 4\alpha^2 - (z - \bar{z})(-z + \bar{z}) &> 0 \end{aligned} \quad (3.3.14)$$

Ahora bien, el equivalente a lo anterior es lo siguiente

$$\begin{aligned}
4\alpha^2 - (z - \bar{z})(-z + \bar{z}) &> 0 \quad , \quad \alpha > 0 \\
\Leftrightarrow \begin{bmatrix} 2\alpha & (z - \bar{z}) \\ (-z + \bar{z}) & 2\alpha \end{bmatrix} &> 0 \\
\Leftrightarrow \begin{bmatrix} -2\alpha & -(z - \bar{z}) \\ -(-z + \bar{z}) & -2\alpha \end{bmatrix} &< 0
\end{aligned} \tag{3.3.15}$$

Esto es equivalente dado que se puede usar el criterio de Sylvester para matrices definidas positivas (o negativas), encontrando las ecuaciones equivalentes mostradas. Por lo tanto, se tiene entonces que

$$L = \begin{bmatrix} -2\alpha & 0 \\ 0 & -2\alpha \end{bmatrix} \quad , \quad M = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \tag{3.3.16}$$

quedando entonces demostrado. ■

En el siguiente Lema se obtiene la combinación de regiones LMI

Lema 3.3.4 (Combinación de regiones LMI). *Sean dos regiones LMI generalizadas escritas como*

$$\begin{aligned}
f_{\mathcal{D}_1}(z) &= L_1 + zM_1 + \bar{z}M_1^T \\
f_{\mathcal{D}_2}(z) &= L_2 + zM_2 + \bar{z}M_2^T
\end{aligned} \tag{3.3.17}$$

Es así que la intersección entre ambos se representará con una función característica

$$f_{\mathcal{D}_3}(z) = L_3 + zM_3 + \bar{z}M_3^T \tag{3.3.18}$$

tal que

$$M_3 = \begin{bmatrix} M_1 & 0 \\ 0 & M_2 \end{bmatrix} \quad , \quad L_3 = \begin{bmatrix} L_1 & 0 \\ 0 & L_2 \end{bmatrix} \tag{3.3.19}$$

Demostración. La demostración que se debe hacer es que

$$f_{\mathcal{D}_3}(z) < 0 \Leftrightarrow f_{\mathcal{D}_1}(z) < 0 \wedge f_{\mathcal{D}_2}(z) < 0 \tag{3.3.20}$$

Se parte primero demostrando la suficiencia. Sea entonces

$$L_3 + zM_3 + \bar{z}M_3^T = \begin{bmatrix} L_1 + zM_1 + \bar{z}M_1^T & 0 \\ 0 & L_2 + zM_2 + \bar{z}M_2^T \end{bmatrix} \tag{3.3.21}$$

por lo tanto se tiene que

$$L_3 + zM_3 + \bar{z}M_3^T < 0 \tag{3.3.22}$$

Esto quiere decir que si la función característica es definida negativa, entonces para cualquier vector v de dimensiones apropiadas se cumplirá

$$v^T(L_3 + zM_3 + \bar{z}M_3^T)v = v^T \begin{bmatrix} L_1 + zM_1 + \bar{z}M_1^T & 0 \\ 0 & L_2 + zM_2 + \bar{z}M_2^T \end{bmatrix} v < 0 \tag{3.3.23}$$

Es así que si $v^T = [u^T \ 0]$, con u vector cualquiera de dimensiones apropiadas y 0 representa un vector de ceros de dimensiones apropiadas, entonces

$$\begin{aligned} [u^T \ 0] \begin{bmatrix} L_1 + zM_1 + \bar{z}M_1^T & 0 \\ 0 & L_2 + zM_2 + \bar{z}M_2^T \end{bmatrix} \begin{bmatrix} u \\ 0 \end{bmatrix} < 0 \\ \Rightarrow u^T(L_1 + zM_1 + \bar{z}M_1^T)u < 0 \end{aligned} \quad (3.3.24)$$

por lo que $L_1 + zM_1 + \bar{z}M_1^T$ es definido negativo, representando la primera región LMI. Análogamente, si $v^T = [0 \ w^T]$, con w vector cualquiera de dimensiones apropiadas y 0 representa un vector de ceros de dimensiones apropiadas, entonces

$$\begin{aligned} [0 \ w^T] \begin{bmatrix} L_1 + zM_1 + \bar{z}M_1^T & 0 \\ 0 & L_2 + zM_2 + \bar{z}M_2^T \end{bmatrix} \begin{bmatrix} 0 \\ w \end{bmatrix} < 0 \\ \Rightarrow w^T(L_2 + zM_2 + \bar{z}M_2^T)w < 0 \end{aligned} \quad (3.3.25)$$

siendo así $L_2 + zM_2 + \bar{z}M_2^T$ definido negativo, representando la segunda región LMI y demostrando la suficiencia. Resta demostrar la necesidad. Entonces si se cumple

$$L_1 + zM_1 + \bar{z}M_1^T < 0 \wedge L_2 + zM_2 + \bar{z}M_2^T < 0 \quad (3.3.26)$$

usando la definición de (3.3.21) y tomando un vector u y w cualquiera real de dimensiones apropiadas es que

$$\begin{aligned} [u^T \ w^T] \begin{bmatrix} L_1 + zM_1 + \bar{z}M_1^T & 0 \\ 0 & L_2 + zM_2 + \bar{z}M_2^T \end{bmatrix} \begin{bmatrix} u \\ w \end{bmatrix} \\ = u^T(L_1 + zM_1 + \bar{z}M_1^T)u + w^T(L_2 + zM_2 + \bar{z}M_2^T)w \end{aligned} \quad (3.3.27)$$

Es claro así que, como las matrices que representan $f_{\mathcal{D}_1}$ y $f_{\mathcal{D}_2}$ son definidas negativas, es que

$$u^T(L_1 + zM_1 + \bar{z}M_1^T)u + w^T(L_2 + zM_2 + \bar{z}M_2^T)w < 0 \quad (3.3.28)$$

demostrando que $f_{\mathcal{D}_1} < 0 \wedge f_{\mathcal{D}_2} < 0 \Rightarrow f_{\mathcal{D}_3} < 0$, quedando demostrado (3.3.20), por lo que $f_{\mathcal{D}_3}$ con las matrices de (3.3.19) representa la ecuación característica de la intersección de las dos regiones LMI y siendo así una región LMI. ■

Con esto, sea la intersección de un disco centrado en $(-q, 0)$ con radio r y un semiplano $Re(z) > \alpha \Leftrightarrow -z + -\bar{z} + 2\alpha < 0$. Se cumple para el disco que

$$L_1 = \begin{bmatrix} -r & q \\ q & -r \end{bmatrix}, \quad M_1 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad (3.3.29)$$

y para el semiplano $Re(z) > \alpha$

$$L_2 = 2\alpha, \quad M_2 = -1 \quad (3.3.30)$$

Por lo que la intersección tendrá matrices

$$L_3 = \begin{bmatrix} -r & q & 0 \\ q & -r & 0 \\ 0 & 0 & 2\alpha \end{bmatrix}, \quad M_3 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad (3.3.31)$$

En las siguientes figuras se simulan regiones en el plano imaginario son resultado para distintos semiplanos y discos en Wolfram Mathematica

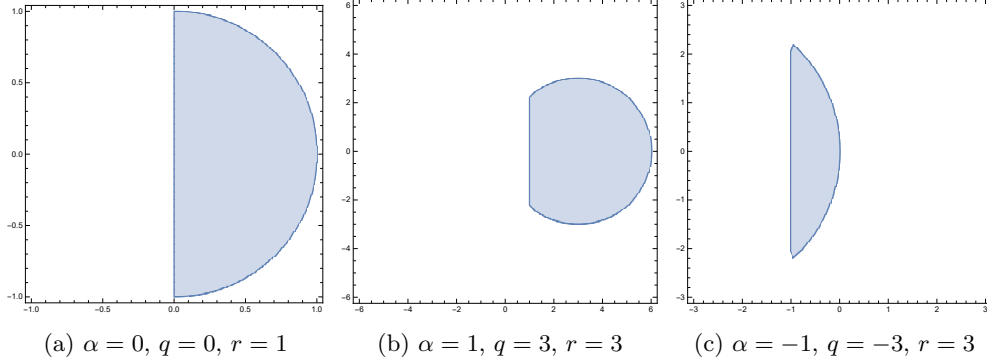


Figura 3.1: Intersección de un semiplano $Re(z) > \alpha$ y un disco centrado en $(q, 0)$ y radio r

por lo que las matrices mostradas en (3.3.31) son efectivamente las que representan la intersección de un disco con un semiplano.

Lema 3.3.5 (Sección Circular). *Si se desea que los autovalores de una matriz A se ubiquen dentro de una sección circular, la cual se define dentro de una circunferencia de radio r y centro $(q, 0)$ y a la derecha de las rectas $x = \frac{|y|}{m} + q$ con $m > 0$, $x = Re(z)$ e $y = Im(z)$, es que se usa la combinación de ambas regiones. Las matrices que definen la región compleja a la derecha de las rectas $x = \frac{|y|}{m} + q$, con $x = Re(z)$ e $y = Im(z)$ son*

$$L_1 = \begin{bmatrix} -2q & 0 \\ 0 & -2q \end{bmatrix}, \quad M_1 = \begin{bmatrix} 1 & \frac{-1}{m} \\ \frac{1}{m} & 1 \end{bmatrix} \quad (3.3.32)$$

y combinándola con las matrices que definen la circunferencia

$$L_2 = \begin{bmatrix} -r & -q \\ -q & -r \end{bmatrix}, \quad M_3 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad (3.3.33)$$

se puede definir entonces las matrices para (3.2.1) que representan la sección circular

$$L = \begin{bmatrix} L_1 & 0 \\ 0 & L_2 \end{bmatrix}, \quad M = \begin{bmatrix} M_1 & 0 \\ 0 & M_2 \end{bmatrix} \quad (3.3.34)$$

Esto por el Lema 3.3.4.

Demostración. Esta sección circular debe ser simétrica respecto al eje real y convexa. Para esto es que se considera $z \in \mathbb{C}$ con $z = x + iy$. Una sección circular se compone de dos rectas de pendientes m y $-m$, y un radio máximo r . Si nos centramos únicamente en las rectas (que pasan por el centro de la circunferencia) es que las rectas son

$$y = \pm m(x - q) \quad (3.3.35)$$

Esto se puede expresar como

$$x = \frac{|y|}{m} + q \quad (3.3.36)$$

Ahora bien, como se quiere una sección definida a la derecha de estas rectas para ser convexa es que se cumple

$$x - \left(\frac{|y|}{m} + q \right) \geq 0 \quad (3.3.37)$$

Resumiendo, las restricciones son

$$x \geq q \quad (3.3.38)$$

$$x - \left(\frac{|y|}{m} + q \right) \geq 0 \quad (3.3.39)$$

Esto se puede reescribir como

$$(x - q)^2 \geq 0 \quad (3.3.40)$$

$$(x - q)^2 - \left(\frac{y}{m} \right)^2 \geq 0 \quad (3.3.41)$$

Como x es la parte real e y la imaginaria, esto puede llevarse a

$$\left(\frac{z + \bar{z}}{2} - q \right)^2 + \left(\frac{z - \bar{z}}{2m} \right)^2 \geq 0 \quad (3.3.42)$$

Esto llevándolo a matrices es equivalente a lo siguiente

$$\begin{bmatrix} \frac{z + \bar{z}}{2} - q & -\frac{z - \bar{z}}{2m} \\ \frac{z - \bar{z}}{2m} & \frac{z + \bar{z}}{2} - q \end{bmatrix} \geq 0 \quad (3.3.43)$$

Esto se puede demostrar usando el complemento de Schur, y llegaríamos a las ecuaciones de las rectas mostradas. Es así entonces que (multiplicando la ecuación matricial por 2) las matrices de LMI que representan la zona descrita son

$$L_1 = \begin{bmatrix} -2q & 0 \\ 0 & -2q \end{bmatrix}, \quad M_1 = \begin{bmatrix} 1 & \frac{-1}{m} \\ \frac{1}{m} & 1 \end{bmatrix} \quad (3.3.44)$$

Quedando así demostrado. Las matrices para sección circular ya se analizaron en el Lema 3.3.1 y la combinación de regiones LMI se analizó en el Lema 3.3.4. ■

Usando Mathematica se hace pruebas para graficar la zona LMI con las matrices descritas L y M , y obviamente agregandolas a las de circunferencia centrada en $(q, 0)$ y radio r . Tomando entonces un $q = 1$, radio $r = 1$ y pendiente $m = 1$ es que se tiene entonces

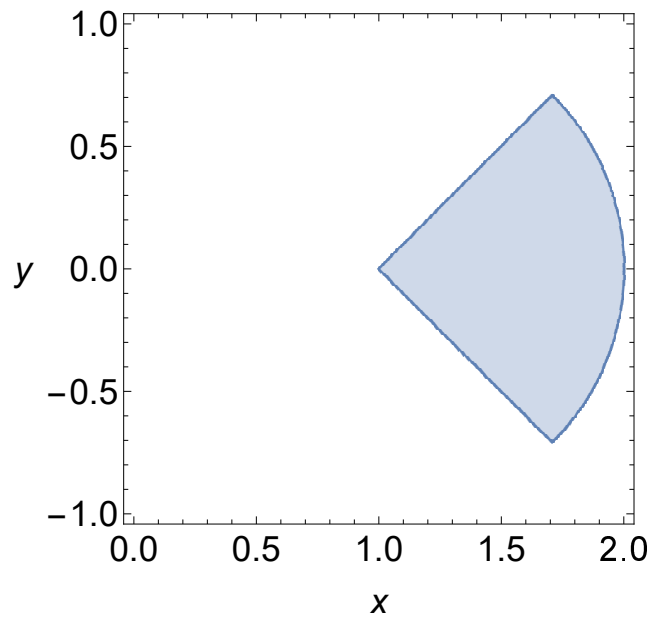


Figura 3.2: Región circular centrada en $(1, 0)$, de radio $r = 1$ y pendiente $m = 1$.

En la Figura 3.2 se aprecia que lo mostrado corresponde a la región circular buscada.

ALGORITMO EM CON RESTRICCIONES

En este Capítulo se abordará el algoritmo EM propuesto con restricciones definidas por LMI. Este algoritmo cumple el objetivo de preservar características que se saben a priori del sistema que se está modelando. Se mostrará la manera de simplificar los cálculos computacionales mediante el uso del criterio de Sylvester, y los problemas que se presentan al usar algoritmos de optimización no lineal con restricciones presentes en la plataforma Matlab.

4.1. Algoritmo EM con restricciones

En esta sección se presenta el algoritmo EM con restricciones en la ubicación de los polos del modelo obtenido. En lo presentado en la Sección 2.3.1 se puede ver en las ecuaciones (2.3.11)-(2.3.12) la función a optimizar $Q(\theta, \hat{\theta}_k)$ en el algoritmo EM. Estas ecuaciones muestran la función auxiliar a optimizar para un modelo en tiempo discreto descrito con el operador de adelanto (2.3.1) - (2.3.3). Tal como se menciona en la Sección 2.4, se puede también describir un modelo en tiempo discreto en forma incremental, a partir del modelo discreto descrito con operador adelanto usando las ecuaciones (2.4.16) - (2.4.17). Esto sirve enormemente, ya que se puede modificar así la función auxiliar $Q(\theta, \hat{\theta}_k)$ para obtener los parámetros descritos en el modelo (2.4.5) - (2.4.12), el cual responde mejor altas tasas de muestreo.

De esta manera, lo que se tiene para el algoritmo EM sin restricción es equivalente a

$$\hat{\theta}_{k+1} = \underset{\theta}{\operatorname{argmin}} \quad -Q(\theta, \hat{\theta}_k) \quad (4.1.1)$$

donde θ posee los parámetros para el modelo en el cual se esté trabajando (en forma incremental o con operador de adelanto).

Tomando en cuenta lo visto en el capítulo 3, es que mediante la siguiente función

$$f_{\mathcal{D}}(z) = \{z \in \mathbb{Z} : L + Mz + M^T \bar{z} < 0\} \quad (4.1.2)$$

donde $L, M \in \mathbb{R}^{m \times m}$ y L es simétrica, $f_{\mathcal{D}}(z)$ define una región del plano complejo. Es así que para una matriz cualquiera $A \in \mathbb{R}^{n \times n}$, si existe $P = P^T \in \mathbb{R}^{n \times n}$ definido positivo y

se cumple lo siguiente

$$L \otimes P + M \otimes AP + M^T \otimes (AP)^T < 0 \quad (4.1.3)$$

Entonces A tiene sus autovalores dentro de la región del plano complejo definido por (4.1.2) usando las matrices L y M . Por lo tanto, podemos modificar el criterio de optimización (4.1.1) de la siguiente manera

$$\hat{\theta}_{k+1} = \underset{\theta}{\operatorname{argmin}} \quad -\mathcal{Q}(\theta, \hat{\theta}_k) \quad (4.1.4a)$$

Sujeto a

$$L \otimes P + M \otimes AP + M^T \otimes (AP)^T < 0 \quad (4.1.4b)$$

$$P = P^T > 0 \quad (4.1.4c)$$

Donde L y M definen la región donde se desea que la matriz A tenga sus autovalores (A puede ser la matriz para el modelo discreto en forma incremental (2.4.5) o definido con el operador de adelanto (2.3.1)).

Se debe recordar que las desigualdades de (4.1.4b) y (4.1.4c) indican que las matrices que se encuentran a mano izquierda sean definidas negativa y positiva respectivamente. El enfoque usando regiones LMI para restringir los autovalores de A es ventajoso, dado que en vez de calcular directamente los autovalores de A y restringir su parte real e imaginario, es que se usan la matrices a la izquierda de la desigualdades (4.1.4b)-(4.1.4c) las cuales son simétricas y reales, por lo que se puede usar el criterio de Sylvester para definirla negativa (o positiva).

4.1.1. Criterio de Sylvester

El criterio de Sylvester indica que una matriz simétrica (en general, Hermititiana) es definida positiva sí y sólo sí sus menores principales son todos positivos. Cuando se aborda la demostración de suficiencia de menores principales positivos en textos de álgebra lineal, es usual basarlo en eliminación Gaussiana [51, pp. 331-332] o en teoría de reducción de formas cuadráticas [52, pp. 328-329]. Aquí se tomará una demostración [53] que hace uso de ideas de teoría de espacios vectoriales. Esta demostración es corta y requiere sólo conceptos simples de álgebra lineal. Demostraciones menos directas son dadas en [54].

Teorema 4.1.1. *Sea W_1 y W_2 subespacios de dimensión finita de un espacio vectorial V . Entonces se cumple*

$$\dim(W_1 + W_2) + \dim(W_1 \cap W_2) = \dim W_1 + \dim W_2 \quad (4.1.5)$$

Demostración. Extendiendo una base para $W_1 \cap W_2$ a una base de W_1 y una base de W_2 . Estas bases se combinan para formar una base de $W_1 + W_2$. Más detalles pueden encontrarse en [52, p. 36] o [51, pp. 199-200]. ■

Dado que los determinantes de matrices similares son iguales, entonces se tiene la siguiente observación

Observación 4.1.2. *El determinante de una matriz diagonalizable es el producto de sus autovalores.*

Teorema 4.1.3 (Teorema Espectral). *Sea A una matriz simétrica real $\mathbb{R}^{n \times n}$. Entonces todos los autovalores de A son reales y existe una base ortonormal de $\mathbb{R}^{n \times n}$ siendo los autovectores de A .*

Demostración. Revisar [52, pp. 312-314] o [51, pp. 295-296]. ■

Teniendo todo lo necesario definido para la demostración, procedemos el análisis para matrices definidas positivas.

Definición 4.1.4. *Una matriz real simétrica $A \in \mathbb{R}^{n \times n}$ es definida positiva si para todo vector no nulo $v \in \mathbb{R}^n$ cumple*

$$v^T A v > 0 \quad (4.1.6)$$

Definimos el siguiente teorema

Teorema 4.1.5. *Una matriz real y simétrica A es definida positiva si y sólo si todos sus autovalores son definidos positivos*

Demostración. La demostración se basa en que dado que es definida positiva, implica que es diagonalizable. Como es simétrica, sus autovectores pueden elegirse ortonormales entre sí, por lo tanto $A = V D V^T$, con V la matriz con los autovectores y D la matriz de autovalores. Entonces, para un vector $u \in \mathbb{R}^n$ se tiene que:

$$u^T A u = u^T V D V^T u = (V^T u)^T D (V^T u) \quad (4.1.7)$$

Se observa que $V^T u \in \mathbb{R}^n$ es un vector, por lo tanto si definimos $V^T u = s$, se tiene

$$(V^T u)^T D (V^T u) = s^T D s = \text{vec}(s^T D s) = s^T \otimes s \text{vec}(D) \quad (4.1.8)$$

Como D es la matriz diagonal de autovalores, la expresión anterior es positiva sí y sólo si todos los elementos de D son mayor a cero. ■

Ahora, definamos los menores principales

Definición 4.1.6 (Menores Principales). *Sea $A \in \mathbb{R}^{n \times n}$. Para $1 \leq k \leq n$, la k -ésima submatriz principal de A es la matriz $k \times k$ formada por las primeras k filas y k columnas. Su determinante se define como el k -ésimo menor principal.*

Lema 4.1.7. *Sea v_1, \dots, v_n una base de un espacio vectorial V . Suponga que W es un subespacio de V de dimensión k . Si $m < k$, entonces existe un vector no nulo $w \in W$ el cual es combinación lineal de v_{m+1}, \dots, v_n .*

Demostración. Dado que

$$\dim W + (n - m) = k + n - m > n \quad (4.1.9)$$

Entonces, W tiene intersección no trivial con el espacio generado de v_{m+1}, \dots, v_n por el Teorema 4.1.1. ■

El siguiente Lema es un caso especial del Teorema "min-max" de Courant-Fischer [54, p. 179].

Lema 4.1.8. *Sea $A \in \mathbb{R}^{n \times n}$ una matriz real y simétrica. Si $w^T A w > 0$ para todos los vectores no nulos w en un subespacio W de dimensión k , tal que W es subespacio de \mathbb{R}^n , entonces A posee al menos k autovalores positivos (tomando en cuenta la multiplicidad).*

Demostración. Sea v_1, \dots, v_n una base ortonormal de \mathbb{R}^n construida con los autovectores de A . Sea $\lambda_1, \dots, \lambda_n$ los autovalores correspondientes de A . Suponga que los primeros m autovalores son positivos y el resto no lo es. Si $m < k$, entonces el Lema 4.1.7 implica que existe un vector no nulo $w \in W$ el cual puede ser escrito como

$$w = c_{m+1}v_{m+1} + \dots + c_nv_n \quad (4.1.10)$$

Por lo que se tiene

$$w^T A w = \sum_{j,k=m+1}^n c_j c_k \lambda_k v_j^T \cdot v_k = c_{m+1}^2 \lambda_{m+1} + \dots + c_n^2 \lambda_n \quad (4.1.11)$$

dado que v_1, \dots, v_n es un conjunto ortonormal. Pero entonces se tiene $w^T A w \leq 0$, lo cual es una contradicción, por lo que $m \geq k$, tal como se busca. ■

Con todas estas definiciones y lemas, es que se presenta el Criterio de Sylvester.

Teorema 4.1.9 (Criterio de Sylvester). *Una matriz real simétrica es definida positiva sí y solo sí todos sus principales menores son positivos.*

Demostración. Es directa la demostración que los principales menores positivos es condición necesaria. Esto se logra mostrando que cada k -ésima submatriz principal es positiva definida, por lo tanto tiene determinante positivo por el Teorema 4.1.5 y la Observación 4.1.2. Nuestra demostración que los principales menores positivos implican una matriz positiva definida procede mediante dos Lemas vistos 4.1.7 y 4.1.8. Usando estos Lemas, es que la demostración de suficiencia se completa mediante inducción: Para $n = 1$, el resultado es trivial. Asuma entonces la suficiencia de los menores principales positivos para matrices reales y simétricas $(n - 1) \times (n - 1)$. Si $A \in \mathbb{R}^{n \times n}$ es matriz real y simétrica con menores principales positivos, entonces, por hipótesis inductiva, sus $(n - 1)$ submatrices principales son positivas definidas. Sea W el subespacio $(n - 1)$ -dimensional de \mathbb{R}^n consistente con los vectores los cuales sus últimos valores son cero. Entonces para cualquier vector no negativo $w \in W$, $w^T A w > 0$. El Lema 4.1.8 implica entonces que A posee al menos $n - 1$ autovalores positivos (contando multiplicidad). Se aplica entonces la Observación 4.1.2 y el hecho que $\det(A) > 0$ para concluir que A posee n autovalores positivos. ■

Con esto, definamos entonces el problema de optimización considerando lo siguiente.

Definición 4.1.10 (Constrained EM). *Sea una región del plano complejo \mathcal{D} , la cual es convexa y simétrica con respecto al eje real. Esta región la definimos mediante el uso de LMI usando la función característica (4.1.2), obteniendo las matrices $L, M \in \mathbb{R}^{m \times m}$. A su vez, sea un sistema a identificar usando modelos de tiempo discreto dados por (2.3.1)-(2.3.3), o equivalentemente (2.4.5)-(2.4.12). Si se desea que la matriz $A \in \mathbb{R}^{n \times n}$ de estos*

modelos estimados usando el algoritmo EM posea autovalores dentro de la región del plano complejo \mathcal{D} , es que el paso M se define como

$$\hat{\theta}_{k+1} = \underset{\theta}{\operatorname{argmin}} \quad -\mathcal{Q}(\theta, \hat{\theta}_k) \quad (4.1.12a)$$

Sujeto a

$$c_i > 0 \quad , \quad i = 1, \dots, n \cdot m \quad (4.1.12b)$$

$$p_j > 0 \quad , \quad j = 1, \dots, n \quad (4.1.12c)$$

Donde se tiene que (4.1.12b)-(4.1.12c) son los menores principales de las matrices en (4.1.3), es decir:

$$c_i = \det(-(L \otimes P + M \otimes AP + M^T \otimes (AP)^T)_{1:i,1:i}) \quad , \quad i = 1, \dots, n \cdot \quad (4.1.13a)$$

$$p_j = \det(P_{1:j,1:j}) \quad , \quad j = 1, \dots, n \quad (4.1.13b)$$

Siendo el operador $P_{1:j,1:j}$ la submatriz principal de P definida por los primera j filas y columnas.

4.1.2. Funciones de barrera logarítmica

El problema de optimización no lineal definido en el paso M para el algoritmo de estimación usando regiones LMI puede ser resuelto, por ejemplo, usando $fmincon(\cdot)$ en MatLab. Sin embargo, esta herramienta de optimización no lineal no garantiza que las restricciones son estrictamente satisfechas en cada iteración. De acuerdo a la información presente del desarrollador en [55] solo restricciones de acotación (*bound constraints* en inglés) se cumplen estrictamente en cada iteración. Estas restricciones son del tipo

$$lb \leq x \leq ub \quad (4.1.14)$$

Donde x es la variable a encontrar en el problema de optimización, lb y ub son vectores constantes que corresponden a las cotas inferior y superior respectivamente de x . Claramente esta restricción no es la que se muestra en el problema de optimización mostrado en (4.1.12). Como consecuencia, el no satisfacer estas restricciones indica que en algún paso de EM no se obtenga los autovalores de la matriz A dentro de la región de interés del plano complejo. Este es un problema, dado que a pesar que la violación de las restricciones puede ser bastante pequeña (en términos numéricos del orden de 10^{-6}), esta cantidad no tiene relación alguna con que tan cerca o lejos de la región de interés se encuentran los autovalores de la matriz A estimada, lo cual puede implicar que en alguna situación se tenga autovalores inestables, presentando problemas numéricos al usar el filtro y suavizador de Kalman en el paso E.

Las funciones barreras son ampliamente usadas en problemas de optimización con restricción, en especial en el algoritmo de punto interior [43].

En particular, para nuestro problema definido en (4.1.12) se define una función de barrera logarítmica (*log-barrier function* en inglés) dada por

$$\phi(\theta) \triangleq -\sum_{i=1}^{n \cdot m} \log(-c_i(\theta)) - \sum_{j=1}^n \log(-p_j(\theta)) \quad (4.1.15)$$

Es así que el problema de optimización queda definido como

$$\theta_{i+1} \triangleq \underset{\theta}{\operatorname{argmin}} -t \cdot \mathcal{Q}(\theta, \hat{\theta}_i) + \phi(\theta) \quad (4.1.16a)$$

subject to

$$P = P^T \quad (4.1.16b)$$

donde escogemos un t numéricamente grande, así privilegiamos la optimización del costo original, y el logaritmo de las restricciones en el costo no influye mucho, sólo restringiendo el espacio de búsqueda de las variables y siendo una muy buena aproximación del problema original [56] [57, pp. 277-279].

SIMULACIONES

En el presente capítulo se presentan simulaciones mediante generación de datos de un sistema en tiempo continuo, suponiendo un retentor de orden cero en la entrada, y muestreando la salida usando un filtro de reseteo e integración (IRF). De esta forma se obtienen datos de entrada y salida, con ruido de medición y de proceso, y con ellos se identifica el sistema. Esta identificación se hace usando el modelo incremental, dada la alta tasa de muestreo, pero los resultados se muestran tanto para un modelo discreto incremental como para el modelo con operador de adelanto.

5.1. Ejemplo 1

Para proceder con las simulaciones, consideremos entonces el problema de estimación paramétrica usando el algoritmo EM para un sistema continuo dado por

$$G(s) = \frac{1.15(s + 5.609)}{(s + 6)(s + 3)} \quad (5.1.1)$$

En primera instancia se toma un periodo de muestreo de $\Delta = T_s = 0.008[s]$. Las matrices del modelo en espacio de estado con datos muestreados descrito con el operador de adelanto visto en (2.3.1)-(2.3.3) son

$$A_q = \begin{bmatrix} 0.9763 & 0 \\ 0 & 0.9531 \end{bmatrix}, B_q = \begin{bmatrix} 0.0079 \\ 0.0008 \end{bmatrix} \quad (5.1.2)$$

$$C_q = [0.9881 \quad 1.4646], D_q = 0.0046 \quad (5.1.3)$$

$$Q_q = \begin{bmatrix} 0.0039 & 0 \\ 0 & 0.0038 \end{bmatrix}, R_q = 1.5242 \quad (5.1.4)$$

Usando las ecuaciones de equivalencia (2.4.16) - (2.4.17) podemos tener el modelo de datos muestreados en su forma incremental definido en (2.4.5) - (2.4.12). Consideremos en principio un set de datos de largo $N = 300$. Los requerimientos son que el modelo estimado en tiempo discreto tiene que ser estable y no debe poseer un comportamiento oscilatorio, es decir, los autovalores deben ser reales. Por tanto, los autovalores de \hat{A}_q se deben ubicar en el segmento real $(0, 1)$ o equivalentemente, en el segmento real $(\frac{-1}{\Delta}, 0)$ para \hat{A}_δ (vea la ecuación de transformación (2.4.16)).

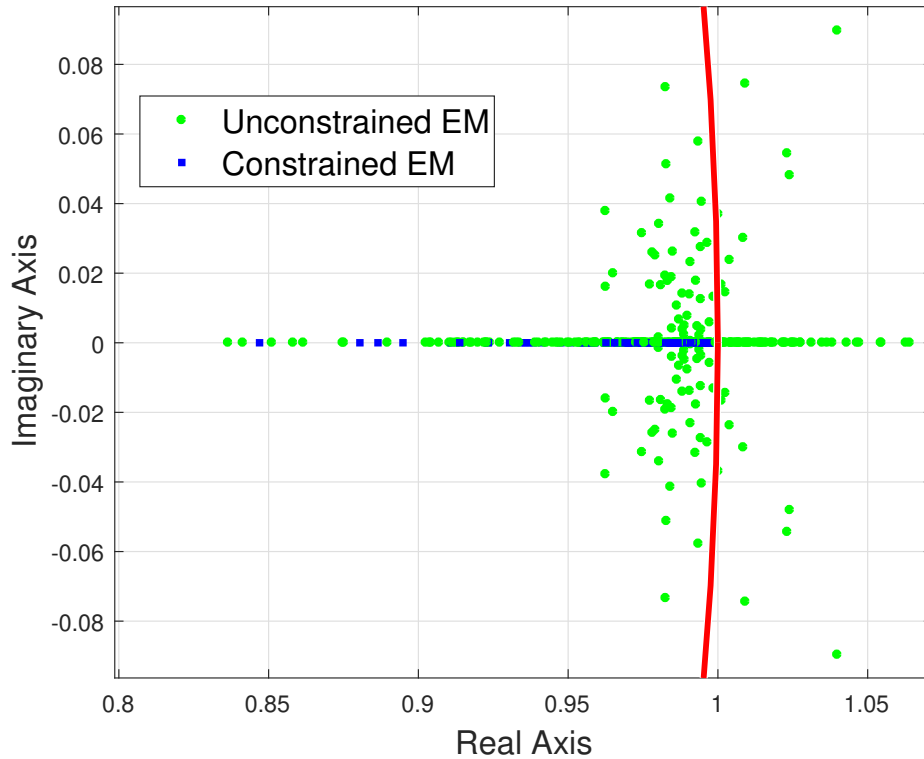


Figura 5.1: Autovalores del estimado de A_q para el modelo discreto con operador de corrimiento y para 200 realizaciones diferentes del ruido. La línea roja corresponde al círculo unitario (límite de estabilidad).

La Figura 5.1 muestra la ubicación de los autovalores de la matriz estimada \hat{A}_q usando tanto el algoritmo EM sin restricción (verde), y el algoritmo presentado en la sección anterior (azul) para un estudio de Monte-Carlo con 200 realizaciones de las secuencias de ruido. Como se puede observar, los autovalores de la matriz estimado considerando EM sin restricción son en algunos casos inestables o complejos conjugados. Sin embargo, incluyendo restricciones para el algoritmo EM como se mostró en el Capítulo 4, el modelo estimado es estable y no oscilatorio (o sea, con polos reales). La Figura 5.2 muestra la ubicación de los autovalores para la matriz \hat{A}_δ en el mismo estudio de Monte-Carlo.

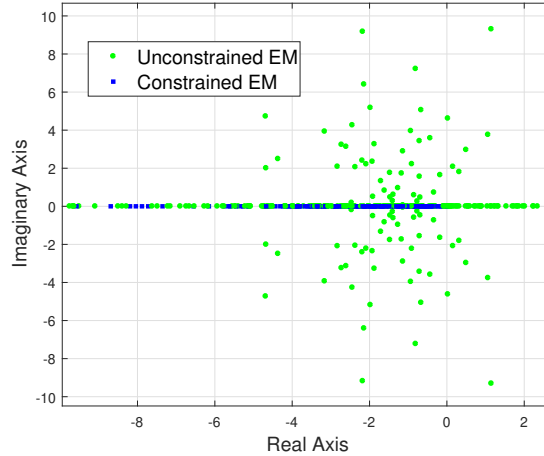


Figura 5.2: Autovalores de la matriz estimada A_δ para el modelo incremental y 200 realizaciones.

La Figura 5.3 muestra la evolución de la función de log-verosimilitud para una realización en particular usando tanto el algoritmo EM sin restricción y comparando con el algoritmo EM con restricciones basadas en regiones LMI. Para el cálculo de la función de (log)-verosimilitud, se usa (2.3.6) con un factor de escala de $\frac{1}{N}$. Como uno puede esperar, la introducción de restricciones en el espacio de parámetros admisible implica que la función de verosimilitud es menor que el caso de EM sin restricción. No obstante, para esta realización en particular, el estimado de EM sin restricción conlleva a un modelo oscilatorio, siendo el sistema real en tiempo continuo un sistema no oscilatorio y estable.

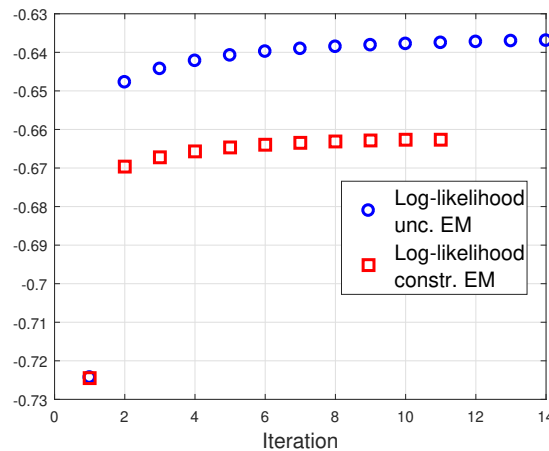


Figura 5.3: Evolución de la función de log-verosimilitud para cada iteración de los algoritmos de estimación.

La Figura 5.4 muestra la evolución de la función de log-verosimilitud para cuando el conjunto de datos que se tiene es grande ($N = 30.000$ datos). Se puede apreciar que casi no existe diferencia entre tanto la log-verosimilitud para el caso de EM sin restricción con el algoritmo propuesto de EM con restricción.

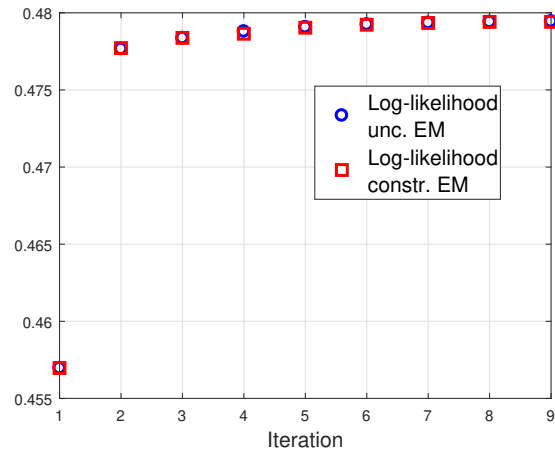


Figura 5.4: Evolución de la función de log-verosimilitud para cada iteración de los algoritmos de estimación.

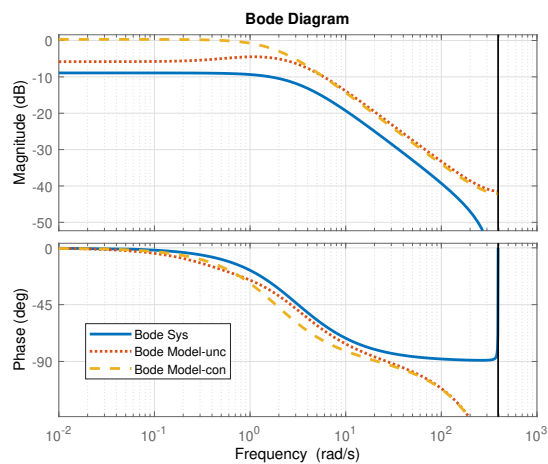


Figura 5.5: Diagramas de Bode del sistema y los modelos obtenidos.

Finalmente, la Figura 5.5 muestra los diagramas de Bode para el sistema real y los modelos estimados mediante EM sin restricción y EM con restricción en base a regiones LMI para una realización en específico. Uno podría pensar que para este particular caso, EM sin restricción entrega un modelo más preciso (tanto en magnitud como en fase), sin embargo,

este corresponde a un modelo en tiempo discreto inestable. Por otro lado, el modelo estimado obtenido usando el algoritmo de EM restringido con regiones LMI propuestos permite obtener un modelo estable y no oscilatorio, conservando las características del sistema real.

5.2. Ejemplo 2

Este ejemplo toma el mismo sistema simulado anteriormente, en el cual la cantidad de datos es poca, pero la diferencia radica en la zona del plano complejo donde se desea que se ubiquen los autovalores de los modelos obtenidos. En los casos anteriores, se restringió en el modelo incremental para estar dentro de la región de estabilidad (circunferencia centrada en $-\frac{1}{\Delta}$, 0 y radio $\frac{1}{\Delta}$ con $\Delta=0.008[s]$) y sobre el eje real. En este caso se restringe al semicírculo de la derecha de la circunferencia centrada en $(-8, 0)$ y radio 6.

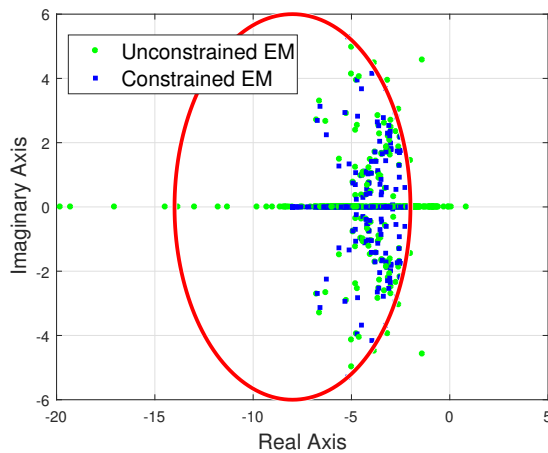


Figura 5.6: Autovalores de la matriz estimada A_δ para el modelo incremental y 200 realizaciones diferentes del ruido. La región de restricción es la del semicírculo derecho de la circunferencia en rojo.

Como se aprecia en la Figura 5.6, es que los autovalores efectivamente se encuentran en la región de restricción, validando el algoritmo propuesto. Dentro de esta restricción existen casos en que los modelos estimados son oscilatorios, siendo resultados similares a los obtenidos por el algoritmo de EM usual (sin restricción), y esto coincide con el hecho que si EM sin restricción no sale de los límites de restricción, es que el algoritmo propuesto EM con restricción arroja resultados similares. Se aprecia que los autovalores de los modelos obtenidos mediante el algoritmo de EM con restricción que se ubican sobre el eje real, se reparten entre los valores $(-8, -2)$, por lo que si se tiene una buena noción a priori de la ubicación de los autovalores del sistema real, se puede forzar los modelos obtenidos para EM y que sus características respondan de manera similar al sistema real. La Figura 5.7 muestra los autovalores de la matriz estimada A_q para este mismo caso.

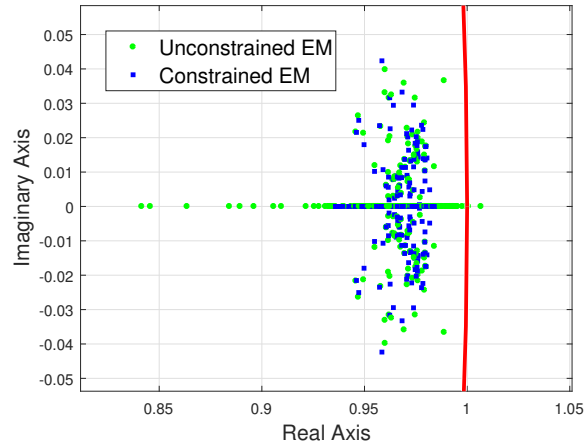


Figura 5.7: Autovalores del estimado de A_q para el modelo discreto con operador de corrimiento y para 200 realizaciones diferentes del ruido. La línea roja corresponde al círculo unitario (límite de estabilidad).

A continuación se presentan resultados al aumentar la cantidad de datos (o equivalentemente, aumentar la relación señal a ruido). Se aprecia en ambas figuras que los autovalores de los modelos obtenidos son equivalentes tanto para el algoritmo EM con restricción basado en LMI, como el algoritmo EM usual, con excepción de algunos casos particulares.

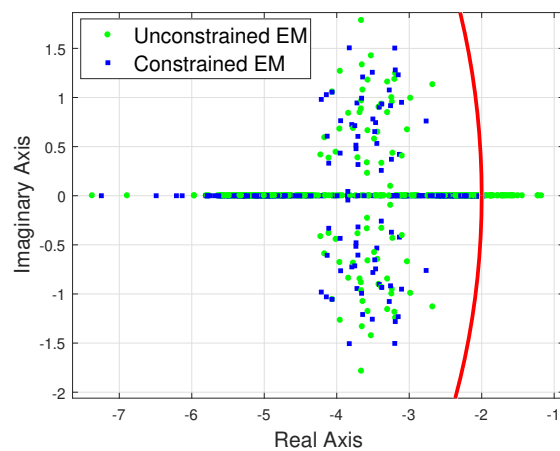


Figura 5.8: Autovalores de la matriz estimada A_δ para el modelo incremental y 200 realizaciones diferentes del ruido. La región de restricción es la del semicírculo derecho de la circunferencia en rojo.

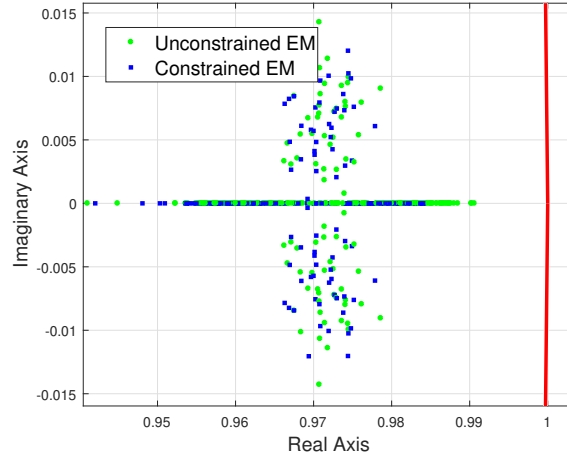


Figura 5.9: Autovalores del estimado de A_q para el modelo discreto con operador de corrimiento y para 200 realizaciones diferentes del ruido. La línea roja corresponde al círculo unitario (límite de estabilidad).

Por último, se pone el caso en que la región de restricción no incluye los autovalores reales. En este caso se aprecia claramente como el algoritmo EM usual (sin restricción) para una cantidad alta de datos $N = 2000$.

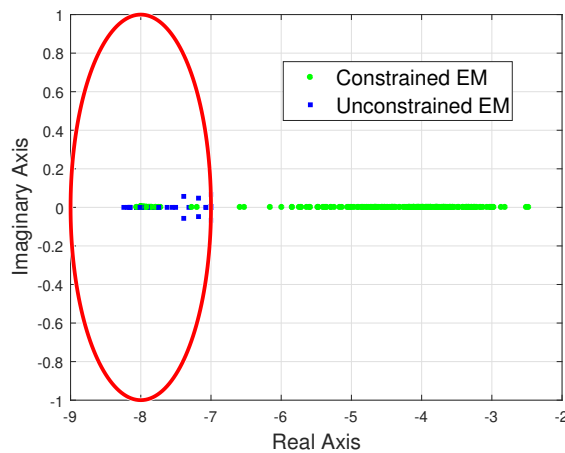


Figura 5.10: Autovalores de la matriz estimada A_δ para el modelo incremental y 200 realizaciones diferentes del ruido. La región de restricción es la que corresponde a la circunferencia en rojo.

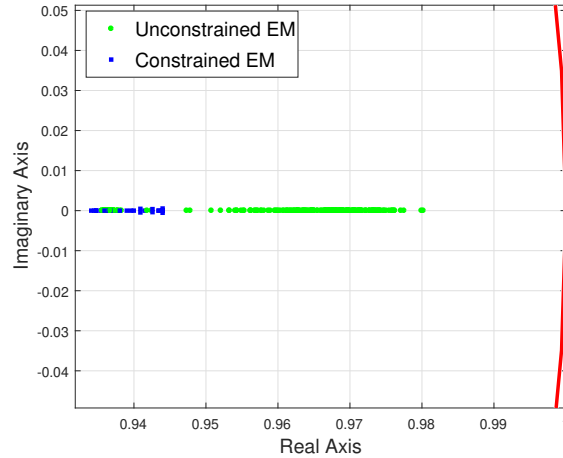


Figura 5.11: Autovalores del estimado de A_q para el modelo discreto con operador de corrimiento y para 200 realizaciones diferentes del ruido. La línea roja corresponde al círculo unitario (límite de estabilidad).

Se observa de las Figuras 5.10 - 5.11 que para estas realizaciones particulares del ruido, es que la mayoría de las estimaciones de ambos algoritmos que se comparan entregan modelos con autovalores estables y reales. Pero se aprecia que EM usual converge en varios casos a la cercanía de los autovalores del sistema real, no así el algoritmo EM con restricción, lo cual se debe al tomar una zona del plano complejo que no posee estos autovalores.

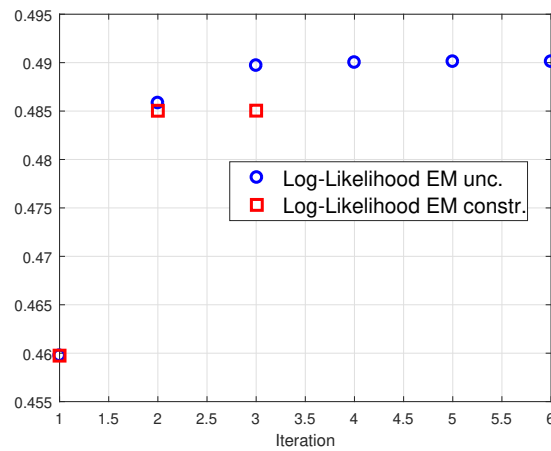


Figura 5.12: Evolución de la función de log-verosimilitud para cada iteración de los algoritmos de estimación.

En la figura 5.12 se muestra un caso particular de realizaciones del ruido para la evolución de la función de log-verosimilitud. Vemos que tanto para EM con restricción (rojo) como EM

sin restricción (azul) que ambos parten en el mismo punto, pero se diferencian enormemente desde la segunda iteración, lo cual se explica que en tal iteración EM sin restricción entrega parámetros para un modelo con autovalores fuera de la región de restricción.

5.3. Ejemplo 3

Este último ejemplo busca comparar nuestro algoritmo con identificación mediante subespacios usando el algoritmo N4SID de Matlab. Para esto, usamos el mismo sistema continuo de los ejemplos anteriores, con el mismo tiempo de muestreo $T_s = \Delta = 0.008[s]$, con $N = 500$ datos y alta relación señal a ruido. La región de restricción es que los autovalores sean estables. Los resultados se muestran a continuación:

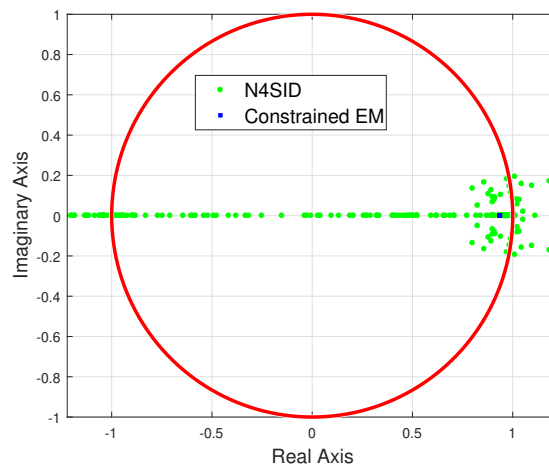


Figura 5.13: Autovalores de la matriz estimada A_q para el modelo discreto con operador de corrimiento y 200 realizaciones diferentes del ruido. La circunferencia en rojo es el límite de estabilidad.

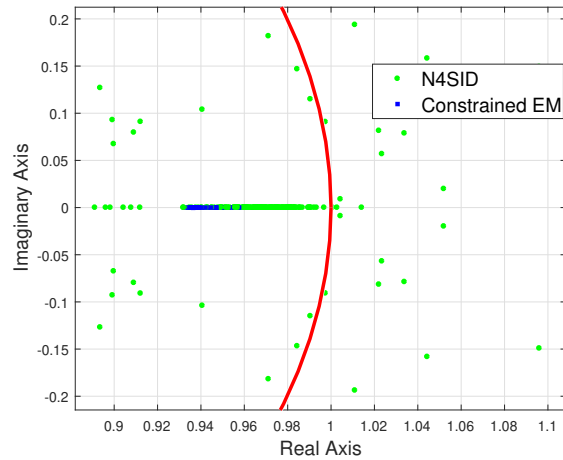


Figura 5.14: Zoom de autovalores de la matriz estimada A_q para el modelo discreto con operador de corrimiento y 200 realizaciones diferentes del ruido. La circunferencia en rojo es el límite de estabilidad.

En las Figuras 5.13 - 5.14 es que se puede apreciar como N4SID entrega modelos bastantes imprecisos. Esto se puede explicar que como existe una alta tasa de muestreo, los autovalores del sistema discreto real con operador de corrimiento se encuentran muy cercanos al límite de estabilidad, representando un problema numérico para N4SID, ya que se empieza a obtener una matriz A_q muy cercana a la identidad. Mientras tanto, el algoritmo de EM con restricción basado en LMI toma en cuenta esto, y al hacer estimación para modelos descritos en forma incremental, el período de muestreo no representa un problema mayor, obteniendo modelos precisos y que conservan las características del sistema real.

CONCLUSIONES

En esta Tesis se presentó una modificación al algoritmo EM para obtener el estimador de Máxima Verosimilitud de los parámetros de un sistema discreto en variables de estado, lineal e invariante en el tiempo, sujeto a perturbaciones. Este algoritmo EM restringe los modelos obtenidos mediante el uso de regiones basadas en LMI de manera que conserve características del sistema real.

Primero se estudió la estimación de Máxima Verosimilitud y el algoritmo EM. Se mostraron las expresiones para la estimación de los parámetros de los modelos del algoritmo EM y se analizó cada uno de los pasos que presenta, enfocándose en sistemas en variables de estado, lineales, invariantes en el tiempo y discretos.

Posteriormente se mostró un estudio de regiones basadas en desigualdades matriciales lineales. Se estudiaron las expresiones para definir estas regiones, las características que cumplen, y cómo es posible usar estas regiones para restringir la ubicación en el plano complejo de los autovalores de la matriz del sistema. En particular, esto se logra con las restricciones que se añadieron al paso M del algoritmo EM.

Posteriormente, se analizaron los problemas de implementación asociados. Se mostró que las restricciones añadidas al paso M son del tipo matrices definidas positivas, pero al ser matrices simétricas, se mostró el uso del Criterio de Sylvester, el cual simplifica la implementación del algoritmo. A su vez se expuso el problema de implementar este algoritmo en la conocida plataforma de simulación MatLab, donde su comando definido para optimizar funciones con restricciones no lineales no satisface de manera estricta estas restricciones. Como alternativa se propuso abordar este problema usando funciones de barrera logarítmica.

Asimismo, se presentan simulaciones que demuestran la utilidad el comportamiento del algoritmo presentado. Para estas simulaciones se tomó un sistema continuo pero con datos son muestreados. Se tomó un sistema con polos estables reales y se mostraron los resultados restringiendo a la región de estabilidad y luego a una región aproximada a la recta real y estable. Se comparó con el algoritmo EM sin restricciones para los mismos datos. Se mostró un buen desempeño del algoritmo, cumpliendo el objetivo deseado y mostrando una evolución de la función de verosimilitud para los casos presentados. De la comparación con el algoritmo EM original se rescata el hecho que las diferencias que presentan a veces las funciones de log-verosimilitud de ambos algoritmos se traducen claramente en las características del modelo obtenido, por ejemplo, en el Ejemplo 1 se aprecia que el algoritmo EM sin restricción cumple mejor el criterio de optimización comparado al algoritmo propuesto, pero el modelo obtenido mediante el algoritmo EM sin restricción es inestable, mientras

que con el algoritmo propuesto es un modelo estable y de polos reales, tal como el sistema original. También se rescata que si existe una buena idea de la ubicación de los polos, se puede ser más estricto, y no pedir solo estabilidad, siendo más preciso en la zona donde se desea que se encuentren los autovalores del modelo estimado, pero esto puede conllevar errores si es que se falla en definir la zona de restricción.

Cuando existe una alta cantidad de datos, o una alta relación señal a ruido, el algoritmo EM usual (sin restricción) cumple adecuadamente con el objetivo de estimar un modelo que aproxime bien al sistema original y que preserve las características como estabilidad o comportamiento (no) oscilatorio. El algoritmo EM con restricción propuesto se muestra particularmente útil justamente cuando se dispone de pocos datos, o bien, la relación señal a ruido es baja y, en este escenario, se desea preservar en el modelo estimado características del sistema original.

Por último se compara con métodos de estimación mediante subespacios usando el algoritmo N4SID. Como el caso analizado corresponde a una alta tasa de muestreo, se apreció como la matriz del sistema discreto y definido con el operador de adelanto tiende a una identidad. Esto está muy cercano al límite de estabilidad, por lo que N4SID no entrega resultados satisfactorios. Por su parte, el algoritmo EM con restricciones basadas en regiones LMI, usando modelos incrementales entrega buenos resultados y no existe problema a alta tasa de muestreo, obteniendo modelos adecuados pudiendo restringir la ubicación de sus polos dentro de regiones complejas definidas por LMI.

6.1. Trabajo a futuro

Una primera extensión al trabajo presentado en esta Tesis es cómo incluir restricciones más simples para la ubicación de los autovalores del modelo estimado. En [18] [17] aprovechan el uso de las matrices de peso que existen en los costos definidos para estimación mediante subespacios para definir las restricciones con LMI y transformándolas en restricciones lineales. En nuestro caso las restricciones son no lineales, pero se podría estudiar una manera de escribir de manera más simple las restricciones buscadas.

Otro trabajo a futuro es aplicar estas restricciones al algoritmo EM en el dominio de la frecuencia. Para esto se debería estudiar [24] [4] [58] que muestra el desarrollo de EM para sistemas discretos, y se afirma que los problemas del dominio del tiempo y frecuencia son equivalentes.

Finalmente una extensión interesante sería aplicar restricciones a las matrices de observabilidad y controlabilidad del modelo obtenido, permitiendo obtener directamente un modelo de realización mínima, o a su vez restricciones en las estructuras de las matrices.

REFERENCIAS

- [1] K. J. Åström and T. Bohlin, “Numerical identification of linear dynamic systems from normal operating records,” *IFAC Proceedings Volumes*, vol. 2, no. 2, pp. 96–111, 1965.
- [2] L. Ljung, *System identification*. Wiley Online Library, 1999.
- [3] T. Söderström and P. Stoica, *System Identification*. Prentice-Hall International, 1989.
- [4] J. C. Agüero, J. I. Yuz, G. C. Goodwin, and R. A. Delgado, “On the equivalence of time and frequency domain maximum likelihood estimation,” *Automatica*, vol. 46, no. 2, pp. 260–270, 2010.
- [5] G. C. Goodwin and R. L. Payne, *Dynamic system identification: experiment design and data analysis*. Academic press, 1977.
- [6] R. Pintelon and J. Schoukens, “Box–Jenkins identification revisited – Part I: Theory,” *Automatica*, vol. 42, no. 1, pp. 63–75, 2006.
- [7] A. P. Dempster, N. M. Laird, and D. B. Rubin, “Maximum likelihood from incomplete data via the EM algorithm,” *Journal of the royal statistical society. Series B (methodological)*, pp. 1–38, 1977.
- [8] G. J. McLachlan and T. Krishnan, *The EM algorithm and extensions*. Wiley, 2008.
- [9] C. J. Wu, “On the convergence properties of the EM algorithm,” *The Annals of statistics*, pp. 95–103, 1983.
- [10] S. Gibson and B. Ninness, “Robust maximum-likelihood estimation of multivariable dynamic systems,” *Automatica*, vol. 41, no. 10, pp. 1667–1682, 2005.
- [11] J. Yuz, J. Alfaro, J. Agüero, and G. Goodwin, “Identification of continuous-time state-space models from non-uniform fast-sampled data,” *IET Control Theory & Applications*, vol. 5, no. 7, pp. 842–855, 2011.
- [12] R. P. Aguilera, B. I. Godoy, J. C. Agüero, G. C. Goodwin, and J. I. Yuz, “An EM-based identification algorithm for a class of hybrid systems with application to power electronics,” *International Journal of Control*, vol. 87, no. 7, pp. 1339–1351, 2014.

- [13] F. Chen, J. C. Agüero, M. Gilson, H. Garnier, and T. Liu, "EM-based identification of continuous-time ARMA models from irregularly sampled data," *Automatica*, vol. 77, pp. 293–301, 2017.
- [14] J. M. Maciejowski, "Guaranteed stability with subspace methods," *Systems & Control Letters*, vol. 26, no. 2, pp. 153–156, 1995.
- [15] N. L. C. Chui and J. M. Maciejowski, "Realization of stable models with subspace methods," *Automatica*, vol. 32, no. 11, pp. 1587–1595, 1996.
- [16] T. Van Gestel, J. A. Suykens, P. Van Dooren, and B. De Moor, "Identification of stable models in subspace identification by using regularization," *IEEE Transactions on Automatic control*, vol. 46, no. 9, pp. 1416–1420, 2001.
- [17] S. L. Lacy and D. S. Bernstein, "Subspace identification with guaranteed stability using constrained optimization," *IEEE Transactions on automatic control*, vol. 48, no. 7, pp. 1259–1263, 2003.
- [18] D. N. Miller and R. A. De Callafon, "Subspace identification with eigenvalue constraints," *Automatica*, vol. 49, no. 8, pp. 2468–2473, 2013.
- [19] R. A. Fisher, "On the mathematical foundations of theoretical statistics," *Phil. Trans. R. Soc. Lond. A*, vol. 222, no. 594-604, pp. 309–368, 1922.
- [20] R. A. Delgado, J. I. Yuz, J. C. Agüero, and G. C. Goodwin, "A numerical study of time and frequency domain maximum likelihood estimation," *IFAC Proceedings Volumes*, vol. 42, no. 10, pp. 1133–1138, 2009.
- [21] K. Åström, "Maximum likelihood and prediction error methods," *IFAC Proceedings Volumes*, vol. 12, no. 8, pp. 551–574, 1979.
- [22] J. P. Norton, *An introduction to identification*. Dover Publications Inc., 2009.
- [23] G. McLachlan and T. Krishnan, *The EM algorithm and extensions*, vol. 382. John Wiley & Sons, 2007.
- [24] J. C. Agüero, J. I. Yuz, and G. C. Goodwin, "Frequency domain identification of MIMO state space models using the em algorithm," in *Control Conference (ECC), 2007*.
- [25] R. H. Shumway and D. S. Stoffer, "An approach to time series smoothing and forecasting using the em algorithm," *Journal of time series analysis*, vol. 3, no. 4, pp. 253–264, 1982.
- [26] D. S. Bernstein, *Matrix mathematics: Theory, facts, and formulas with application to linear systems theory*. Princeton University, 2005.
- [27] B. D. Anderson and J. B. Moore, "Optimal filtering," *Englewood Cliffs*, 1979.
- [28] R. H. Shumway and D. S. Stoffer, "Time series analysis and its applications," *Studies In Informatics And Control*, vol. 9, no. 4, pp. 375–376, 2000.

- [29] T. Kailath, “Lectures on Wiener and Kalman filtering,” Springer, 1981.
- [30] H. E. Rauch, C. Striebel, and F. Tung, “Maximum likelihood estimates of linear dynamic systems,” *AIAA journal*, vol. 3, no. 8, pp. 1445–1450, 1965.
- [31] C. F. Ansley and R. Kohn, “A geometrical derivation of the fixed interval smoothing algorithm,” *Biometrika*, vol. 69, no. 2, pp. 486–487, 1982.
- [32] A. P. Sage, J. L. Melsa, and W. Steinway, “Estimation theory with applications to communication and control,” *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 4, no. 1, 1971.
- [33] B. Øksendal, “Stochastic differential equations,” Springer, 2003.
- [34] G. C. Goodwin, J. C. Aguero, M. E. Cea-Garrido, M. E. Salgado, and J. I. Yuz, “Sampling and sampled-data models: The interface between the continuous world and digital algorithms,” *IEEE Control Systems*, vol. 33, no. 5, pp. 34–53, 2013.
- [35] R. H. Middleton and G. C. Goodwin, *Digital Control and Estimation: A Unified Approach*. Prentice Hall Englewood Cliffs, NJ, 1990.
- [36] A. Feuer and G. Goodwin, *Sampling in digital signal processing and control*. Springer, 2012.
- [37] J. I. Yuz and G. C. Goodwin, *Sampled-data models for linear and nonlinear systems*. Springer, 2016.
- [38] K. J. Åström, *Introduction to stochastic control theory*. Courier Corporation, 2012.
- [39] L. Ljung and A. Wills, “Issues in sampling and estimating continuous-time models with stochastic disturbances,” *Automatica*, vol. 46, no. 5, pp. 925–931, 2010.
- [40] P. Gahinet and P. Apkarian, “A linear matrix inequality approach to H_∞ control,” *International journal of robust and nonlinear control*, vol. 4, no. 4, pp. 421–448, 1994.
- [41] T. Iwasaki and R. E. Skelton, “All controllers for the general H_∞ control problem: LMI existence conditions and state space formulas,” *Automatica*, vol. 30, no. 8, pp. 1307–1317, 1994.
- [42] A. Packard, “Gain scheduling via linear fractional transformations,” *Systems & control letters*, vol. 22, no. 2, pp. 79–92, 1994.
- [43] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear matrix inequalities in system and control theory*. Siam, 1994.
- [44] M. Chilali and P. Gahinet, “ H_∞ design with pole placement constraints: an LMI approach,” *IEEE Transactions on automatic control*, vol. 41, no. 3, pp. 358–367, 1996.
- [45] M. Chilali, P. Gahinet, and P. Apkarian, “Robust pole placement in LMI regions,” *IEEE Transactions on Automatic Control*, vol. 44, no. 12, pp. 2257–2270, 1999.

- [46] J. B. Hoagg, S. L. Lacy, R. S. Erwin, and D. S. Bernstein, "Subspace identification with lower bounded modal frequencies," in *American Control Conference, 2004*.
- [47] T. McKelvey and R. Moheimani, "Estimation of phase constrained mimo transfer functions with application to flexible structures with mixed collocated and non-collocated actuators and sensors," in *Proc. 16th IFAC World Congress, 2005*.
- [48] Y. Nesterov and A. Nemirovskii, *Interior-point polynomial algorithms in convex programming*. Siam, 1994.
- [49] Y. E. Nesterov and A. Nemirovskii, "An interior-point method for generalized linear-fractional programming," *Mathematical Programming*, vol. 69, no. 1-3, pp. 177–204, 1995.
- [50] S. Boyd and L. El Ghaoui, "Method of centers for minimizing generalized eigenvalues," *Linear algebra and its applications*, vol. 188, pp. 63–111, 1993.
- [51] G. Strang, *Linear Algebra and Its Applications*. Harcourt Brace Jovanovich, Inc., San Diego, 1988.
- [52] K. Hoffman, R. Kunze, and H. E. Finsterbusch, *Álgebra lineal*. Prentice-Hall Hispanoamericana, 1973.
- [53] G. T. Gilbert, "Positive definite matrices and Sylvester's criterion," *The American Mathematical Monthly*, vol. 98, no. 1, pp. 44–46, 1991.
- [54] R. A. Horn, R. A. Horn, and C. R. Johnson, *Matrix analysis*. Cambridge university press, 1990.
- [55] MathWorks contributors, "Iterations can violate constraints." Available at <https://www.mathworks.com/help/optim/ug/iterations-can-violate-constraints.html>; accessed August 1st, 2018.
- [56] S. Wright and J. Nocedal, "Numerical optimization," *Springer Science*, vol. 35, no. 67-68, p. 7, 1999.
- [57] V. Robert *et al.*, "Linear programming: Foundations and extensions," *Department of Operations Research and Financial Engineering*, 2001.
- [58] A. Wills, B. Ninness, and S. Gibson, "Maximum likelihood estimation of state space models from frequency domain data," *IEEE Transactions on Automatic Control*, vol. 54, no. 1, pp. 19–33, 2009.