

# 1 Modelado estocástico por estimación de mínimos cuadrados

## 1.1 Proceso de primer orden con tiempo muerto [1]

La función de transferencia en  $Z$  de un proceso de primer orden con periodo de muestreo  $T$  y un tiempo muerto de  $d$  periodos de muestreo se muestra en la ecuación 1.1

$$G(z) = \frac{C(z)}{M(z)} = \frac{k \cdot z^{-d}}{(z - e^{\frac{-T}{\tau}})} \quad (1.1)$$

La cual tiene una ecuación de diferencias que puede ser escrita como

$$c_n = e^{\frac{-T}{\tau}} \cdot c_{n-1} + k \cdot m_{n-1-d} \quad (1.2)$$

Para las muestras tomadas experimentalmente podemos asumir un error  $\varepsilon_n$ , sin correlación y con valor medio cero, producto de las varias fuentes de ruido en el proceso y reescribimos la ecuación de diferencias como

$$c_n = \alpha \cdot c_{n-1} + \beta \cdot m_{n-1-d} + \varepsilon_n \quad (1.3)$$

Donde

$$\alpha = e^{\frac{-T}{\tau}} \quad (1.4)$$

$$\beta = k \quad (1.5)$$

Si se asume que el tiempo muerto  $d$  es conocido, se pueden estimar los valores de mejor ajuste por mínimos cuadrados para  $\alpha$  y  $\beta$ , minimizando la suma de los errores al cuadrado.

Esta cantidad puede expresarse como

$$S(\alpha, \beta) = \sum_{n=1}^N \varepsilon^2 \quad (1.6)$$

Donde  $N$  es el número de muestras; o, a partir de la ecuación 1.3,

$$S(\alpha, \beta) = \sum_{n=1}^N (c_n - \alpha \cdot c_{n-1} - \beta \cdot m_{n-1-d})^2 \quad (1.7)$$

El valor mínimo de la función suma de cuadrados puede ser obtenido igualando a cero las dos derivadas parciales de  $S(\alpha, \beta)$ . Las derivadas parciales son

$$\frac{\partial S(\alpha, \beta)}{\partial \alpha} = -2 \sum_{n=1}^N (c_n - \alpha \cdot c_{n-1} - \beta \cdot m_{n-1-d}) \cdot c_{n-1} \quad (1.8)$$

$$\frac{\partial S(\alpha, \beta)}{\partial \beta} = -2 \sum_{n=1}^N (c_n - \alpha \cdot c_{n-1} - \beta \cdot m_{n-1-d}) \cdot m_{n-1-d} \quad (1.9)$$

Al igualar estas derivadas a cero, se encuentra un conjunto de ecuaciones llamadas **ecuaciones normales**. Después de simplificar, manteniendo los mismos límites para las sumatorias, el resultado es

$$\hat{\alpha} \sum c_{n-1}^2 + \hat{\beta} \sum m_{n-1-d} c_{n-1} = \sum c_n c_{n-1} \quad (1.10)$$

$$\hat{\alpha} \sum c_{n-1} m_{n-1-d} + \hat{\beta} \sum m_{n-1-d}^2 = \sum c_n m_{n-1-d} \quad (1.11)$$

Los “sombrecitos” en la notación  $\hat{\alpha}$  y  $\hat{\beta}$  denotan que estos valores son los estimados para  $\alpha$  y  $\beta$ .

Resolviendo las ecuaciones 1.10 y 1.11 para los valores  $\hat{\alpha}$  y  $\hat{\beta}$ , utilizando por ejemplo la regla de Cramer, producen el resultado:

$$\hat{\alpha} = \frac{\sum c_n c_{n-1} \sum m_{n-1-d}^2 - \sum c_{n-1} m_{n-1-d} \sum c_n m_{n-1-d}}{\Delta} \quad (1.12)$$

$$\hat{\beta} = \frac{\sum c_{n-1}^2 \sum c_n m_{n-1-d} - \sum c_n c_{n-1} \sum c_{n-1} m_{n-1-d}}{\Delta} \quad (1.13)$$

$$\Delta = \sum c_{n-1}^2 \sum m_{n-1-d}^2 - \left( \sum c_{n-1} m_{n-1-d} \right)^2 \quad (1.14)$$

Se puede presentar un pequeño problema si los valores de  $c$  y  $m$  son registrados iniciando en  $n = 1$  con  $d > 0$ , se requieren muestras adicionales anteriores a los valores tomados en los datos. Existen dos formas de resolver el problema. Una es asumir que los valores iniciales son cero. La otra es iniciar las sumatorias con  $n = d + 1$  en lugar de  $n = 1$ . Esto disminuye la cantidad de datos usados realmente para los cálculos; pero, generalmente no constituye un problema. Los estimados para  $\tau$  y  $k$  se obtienen de la ecuaciones 1.4 y 1.5

$$\hat{\tau} = \frac{-T}{\ln \hat{\alpha}} \quad (1.15)$$

$$\hat{k} = \hat{\beta} \quad (1.16)$$

### 1.1.1 Algoritmo

Se asumió que el retardo  $d$  era conocido para calcular los valores de  $\hat{\alpha}$  y  $\hat{\beta}$ . En general, este no es el caso, y  $d$  es un parámetro adicional que debe ser determinado. El siguiente algoritmo puede ser usado para determinar el modelo del proceso a partir de los datos muestreados.

1. Encuentre  $\hat{\alpha}$ ,  $\hat{\beta}$  y  $\sum \varepsilon_n^2$  para  $d = 0$ .
  2. Usando  $d + 1$  encuentre nuevos valores para  $\hat{\alpha}$ ,  $\hat{\beta}$  y  $\sum \varepsilon_n^2$ .
  3. Si el valor  $\sum \varepsilon_n^2$  para  $d + 1$  es menor que el valor para  $d$ , entonces  $d + 1$  es un mejor valor.
  4. Repita los pasos 2 y 3 incrementando  $d$  hasta que el valor de  $\sum \varepsilon_n^2$  para  $d + 1$  sea mayor que el que aquel para  $d$ .
  5. Los valores de  $\hat{\alpha}$ ,  $\hat{\beta}$  para  $d$  pueden ser asumidos como los de mejor ajuste.
- Luego puede determinar si lo desea  $\hat{\tau}$  y  $\hat{k}$  o escribir directamente la ecuación de diferencias o la función de transferencia en tiempo discreto.

### 1.1.2 Ejemplos

Un programa para Matlab® llamado *stochastic* que implementa este método se encuentra en [2], usaremos este programa para estimar el modelo para un conjunto de datos.

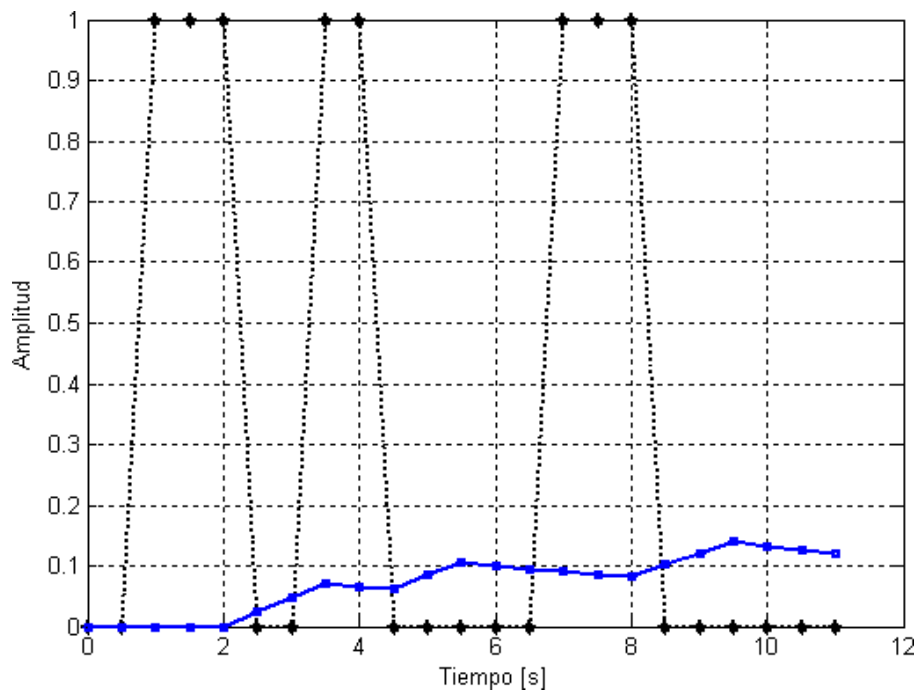


Figura 1.1: Estímulo y respuesta del sistema para el ejemplo 1.1

**Ejemplo 1.1:** Estimaremos el modelo con los siguientes datos, la entrada es llamada x, la salida es y.

Tabla 1.1: Datos de un proceso de primer orden con  $T = 0.5s$

n	0	1	2	3	4	5	6	7	8	9	10	11
x	0.0	0.0	1.0	1.0	1.0	0.0	0.0	1.0	1.0	0.0	0.0	0.0
y	0.0	0.0	0.0	0.0	0.0	0.0244	0.0476	0.0696	0.0662	0.0630	0.0843	0.1046

Invocamos la función *stochastic* en Matlab® de la siguiente forma

```
>> x = [0 0 1 1 1 0 0 1 1 0 0 0];
>> y = [0 0 0 0 0 0.0244 0.0476 0.0696 0.0662 0.0630 0.0843 0.1046];
>> [alfa, beta, d, funcion] = stochastic(y, x, T)

alfa = 0.9512
beta = 0.0244
d = 2
función =  $\frac{0.024385}{z^2(z-0.9512)}$ 

Tiempo de muestreo = 0.5s
```

**Ejemplo 1.2:** Aplicaremos un ruido aleatorio a las mediciones de la salida y observaremos como esto afecta la estimación de parámetros.

Tabla 1.2: Datos de un proceso de primer orden con  $T = 0.5s$  con ruido

n	0	1	2	3	4	5	6	7	8	9	10	11
x	0.0	0.0	1.0	1.0	1.0	0.0	0.0	1.0	1.0	0.0	0.0	0.0
y	0.0007	-0.0007	-0.0006	-0.0008	-0.0019	0.0249	0.0501	0.0685	0.0652	0.0645	0.0844	0.1077

Al procesar estos datos de igual forma que en el ejemplo 1.1 obtenemos una función ligeramente diferente, la cual se muestra como  $\hat{G}(z)$ . En la figura 1.2 se muestra la gráfica de la salida y la salida con ruido y en la figura 1.3 se muestra la comparación de respuestas ante escalón de las dos funciones estimadas.

$$\hat{G}(z) = \frac{0.024356}{z^2(z-0.9485)}$$

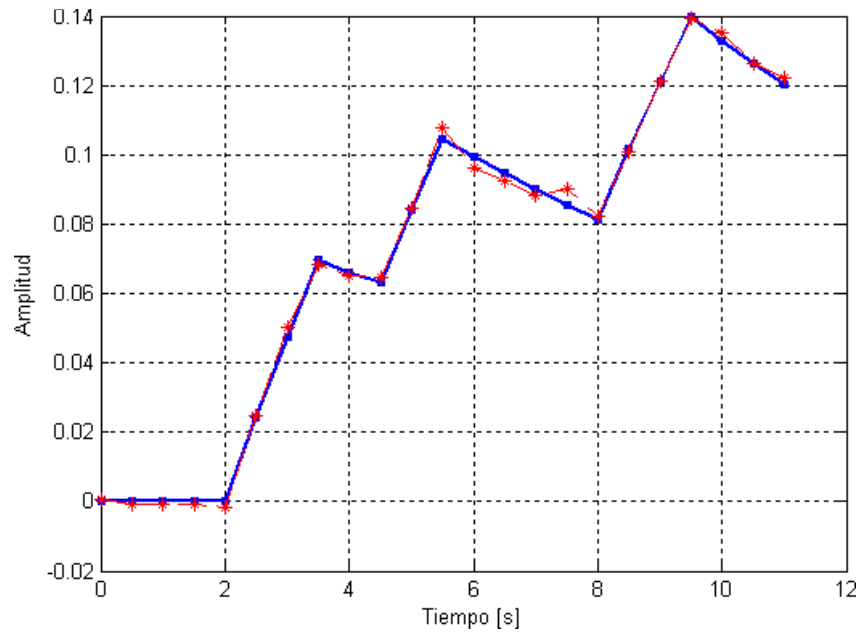


Figura 1.2: Respuesta normal y con ruido

En la figura 1.2, la curva con \* en las muestras, es la salida con ruido aditivo agregado.

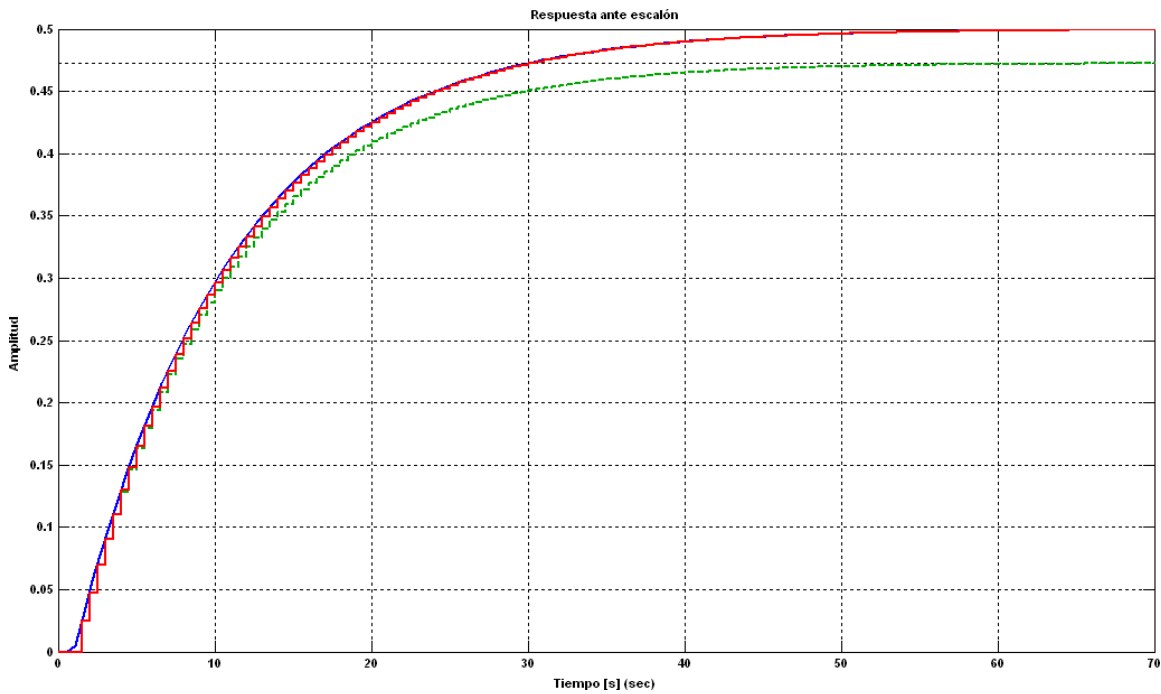


Figura 1.3: Comparación de los modelos obtenidos con datos con y sin ruido

Se puede observar en simulación mostrada en la figura 1.3, que el modelo estimado inicialmente, de los datos sin ruido, corresponde fielmente al modelo discretizado con ZOH del sistema continuo, sistema del cual se tomaron los datos del ejemplo. También se observa que la curva para el modelo obtenido de los datos con ruido aditivo (curva

punteada), no alcanza el valor final de 0.5, como si lo hace la curva del modelo estimado de los datos sin ruido; sino que presenta un error de aproximadamente un 10%.

## 1.2 Sistemas de orden superior con tiempo muerto

Colocando el modelo discreto del proceso como la ecuación de diferencias de la forma

$$c_n = \alpha_1 c_{n-1} + \alpha_2 c_{n-2} + \cdots + \alpha_r c_{n-r} + \beta_1 m_{n-1-d} + \beta_2 m_{n-2-d} + \cdots + \beta_s m_{n-s-d} + \varepsilon_n \quad (1.17)$$

Podremos obtener los estimados de mejor ajuste por el método de mínimos cuadrados para los parámetros  $\alpha_i$ , con  $i = 1, 2, 3, \dots, r$  y  $\beta_j$ ,  $j = 1, 2, 3, \dots, s$ .

Cada medición  $c_n$  para  $n = 1, 2, 3, \dots, N$  da lugar a una ecuación tal como la ecuación 1.17. El conjunto de  $N$  ecuaciones puede ser escrito de forma matricial como

$$\mathbf{c} = \mathbf{X}\mathbf{b} + \boldsymbol{\varepsilon} \quad (1.18)$$

Donde

$$\mathbf{c} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \\ c_N \end{bmatrix}, \quad \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \vdots \\ \varepsilon_N \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \vdots \\ \alpha_r \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \vdots \\ \beta_s \end{bmatrix}$$

Y

$$\mathbf{X} = \begin{bmatrix} c_0 & c_{-1} & \cdots & c_{1-r} & m_{-d} & m_{-1-d} & \cdots & m_{1-s-d} \\ c_1 & c_0 & \cdots & c_{2-r} & m_{1-d} & m_{-d} & \cdots & m_{2-s-d} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ c_{N-1} & c_{N-2} & \cdots & c_{N-r} & m_{N-1-d} & m_{N-2-d} & \cdots & m_{N-s-d} \end{bmatrix}$$

Los datos  $c_{1-r}, c_{2-r}, \dots, c_{N-1}, c_N$  y  $m_{1-s-d}, m_{2-s-d}, \dots, m_{N-2-d}, m_{N-1-d}$  deben ser obtenidos para llenar los elementos de las matrices y vectores en la ecuación 1.18; para lo cual podemos proceder como en el punto 1.1 asumiendo ceros o cambiando el índice de inicio.

Los errores residuales son

$$\boldsymbol{\varepsilon} = \mathbf{c} - \mathbf{X}\mathbf{b} \quad (1.19)$$

Y la suma de los errores al cuadrado es

$$S(\mathbf{b}) = \boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon} = (\mathbf{c} - \mathbf{X}\mathbf{b})^T (\mathbf{c} - \mathbf{X}\mathbf{b}) \quad (1.20)$$

El conjunto de ecuaciones normales se encuentra igualando las derivadas parciales a cero

$$\frac{\partial S(\hat{\mathbf{b}})}{\partial \hat{b}_k} = \mathbf{0} \quad (1.21)$$

para cada uno de los elementos  $\hat{b}_k$ ,  $k = 1, 2, \dots, r + s$  en  $\hat{\mathbf{b}}$ , el vector de los parámetros estimados para el modelo.

Encontrando las derivadas indicadas en la ecuación 1.21 tenemos

$$\frac{\partial S(\hat{\mathbf{b}})}{\partial \hat{b}_k} = \frac{\partial (\boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon})}{\partial \hat{b}_k} = 2\boldsymbol{\varepsilon}^T \frac{\partial (\boldsymbol{\varepsilon})}{\partial \hat{b}_k} \quad (1.22)$$

Combinando todas las ecuaciones obtenidas usando la ecuación 1.22 tenemos

$$\mathbf{X}^T [\mathbf{c} - \mathbf{X}\hat{\mathbf{b}}] = \mathbf{0} \quad (1.23)$$

Resolviendo la ecuación 1.23 para  $\hat{\mathbf{b}}$  obtenemos

$$\hat{\mathbf{b}} = [\mathbf{X}^T \mathbf{X}]^{-1} \mathbf{X}^T \mathbf{c} \quad (1.24)$$

Se requiere invertir la matriz  $\mathbf{X}^T \mathbf{X}$ . Esta matriz es cuadrada de dimensiones  $r + s$ . Una vez obtenido el producto  $\mathbf{X}^T \mathbf{X}$ , el trabajo necesario para calcular su inversa está relacionado únicamente al número de parámetros en el modelo de la ecuación 1.17 y no al número de mediciones.

La varianza se encuentra a partir de la ecuación 1.20 como

$$\hat{\sigma}^2 = \frac{S(\mathbf{b})}{N - r - s} \quad (1.25)$$

Para un valor particular del retardo del proceso  $d$ , la suma de los cuadrados de  $S(\hat{\mathbf{b}}_d)$  puede ser encontrada sustituyendo  $\hat{\mathbf{b}}_d$  en lugar de  $\mathbf{b}$  en la ecuación 1.20. La varianza mínima para todos los valores de  $d$  se encuentra de la forma

$$\hat{\sigma}_{min}^2 = \min[\hat{\sigma}_{min}^2, d = 0, 1, 2, \dots] \quad (1.26)$$

El valor de  $d$  que produce  $\hat{\sigma}_{min}^2$  es entonces la mejor aproximación para  $d$ . En la práctica, un conocimiento de los valores probables para el tiempo muerto puede ser usado para limitar el número de valores posibles de  $d$  a ser evaluados en la ecuación 1.26.

### Ejemplo 1.3 Solución matricial

Tabla 2: Datos de entrada y salida de un sistema con un  $T = 0.5s$

<b>n</b>	<b>0</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>
<b>m</b>	1.0	1.0	1.0	1.0	0.0	0.0	0.0	0.0	0.0
<b>c</b>	0.0	0.57	0.97	1.26	1.47	1.06	0.76	0.54	0.39

Supondremos un modelo de primer orden sin tiempo muerto,  $d = 0$ , **en el caso general hay que iterar hasta encontrar el valor para  $d$  que minimiza el error.** Con  $s$ , el número de parámetros del numerador, igual a uno y con  $r$ , el número de parámetros del denominador, el orden del sistema, también igual a 1. Disponemos los datos para solución matricial escribiendo la matriz  $\mathbf{X}$  y luego trasponiéndola. La matriz  $\mathbf{X}$  solamente tendrá dos, ( $r + s = 1 + 1 = 2$ ), columnas de 8 elementos. La primera columna inicia con  $c_0$ ; puesto que  $r = 1$ ; la segunda columna inicia con  $m_0$ ; puesto que  $d = 0$  y  $s = 1$ .

$$\mathbf{X}^T = \begin{bmatrix} 0.0 & 0.57 & 0.97 & 1.26 & 1.47 & 1.06 & 0.76 & 0.54 \\ 1.0 & 1.0 & 1.0 & 1.0 & 0.0 & 0.0 & 0.0 & 0.0 \end{bmatrix}$$

Y

$$\mathbf{c}^T = [0.57 \quad 0.97 \quad 1.26 \quad 1.47 \quad 1.06 \quad 0.76 \quad 0.54 \quad 0.39]$$

A partir de la ecuación 1.24

$$\hat{\mathbf{b}} = \begin{bmatrix} \hat{\alpha}_1 \\ \hat{\beta}_1 \end{bmatrix} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{c}$$

$$\hat{\mathbf{b}} = \begin{bmatrix} 0.718 \\ 0.565 \end{bmatrix}$$

El modelo discreto para el sistema es por lo tanto

$$\hat{G}(z) = \frac{0.565}{(z - 0.718)}$$

La solución del problema usando Matlab® tiene la siguiente forma

```
>> xt = [0 0.57 0.97 1.26 1.47 1.06 0.76 0.54; 1 1 1 1 0 0 0 0];
>> c = [0.57; 0.97; 1.26; 1.47; 1.06; 0.76; 0.54; 0.39];
>> x = transpose(xt);
>> b = inv(xt*x)*xt*c
b = 0.7179
    0.5650
```



**Referencias:**

- [1] Bollinger, John G., Duffie, Neil A.. “Computer Control of Machines and Processes”, Addison-Wesley, USA, 1988.
- [2] [www.ie.itcr.ac.cr/einteriano/control/TrabajosMatlab](http://www.ie.itcr.ac.cr/einteriano/control/TrabajosMatlab)