

---

## Tercera Tarea IPD-462

06 de Mayo de 2009

---

### 1. Generalidades

- La tarea es de carácter individual.
- El formato es libre.
- Referencias sugeridas: [1, 2].
- **La notación utilizada más abajo corresponde a la notación definida en clases.**
- **Fecha de entrega (límite):** 25 de Mayo 2009, 10:00AM.

### 2. Cuestionario

#### Problema 1

Considere el sistema

$$x(k+1) = A(k)x(k) + v(k), \quad k \in \mathbb{N}_0, \quad x(0) = x_o, \quad (1)$$

$$y(k) = C(k)x(k) + e(k), \quad (2)$$

donde,  $\forall k \in \mathbb{N}_0$ ,  $A(k)$  y  $C(k)$  son matrices determinísticas finitas de dimensiones apropiadas, y  $v$  y  $e$  son procesos de ruido blanco de segundo orden con media y varianzas conocidas para todo  $k \in \mathbb{N}_0$ . Suponga, además, que

$$\mathcal{E} \left\{ \begin{bmatrix} v(k) - \mu_v(k) \\ e(k) - \mu_e(k) \end{bmatrix} \begin{bmatrix} v(i) - \mu_v(i) \\ e(i) - \mu_e(i) \end{bmatrix}^H \right\} = \begin{bmatrix} R_v(k)\delta(k-i) & 0 \\ 0 & R_e(k)\delta(k-i) \end{bmatrix}, \quad \forall k, i \in \mathbb{N}_0, \quad (3)$$

donde  $R_e(k) > 0 \forall k \in \mathbb{N}_0$ , y que  $x_o$  es una variable aleatoria de segundo orden no correlacionada con  $(v, e)$  tal que

$$\mathcal{E} \{x_o\} = \mu_o, \quad \mathcal{E} \{(x_o - \mu_o)(x_o - \mu_o)^T\} = P_o. \quad (4)$$

1. Halle una expresión, lo más compacta posible, para  $\hat{x}(k + \ell|k)$  cuando  $\ell \in \mathbb{N}_0$ .
2. Defina el proceso de innovaciones  $\tilde{y}$  a través de la relación

$$\tilde{y}(k) \triangleq y(k) - \hat{y}(k|k-1). \quad (5)$$

Demuestre que  $\tilde{y}$  es un proceso blanco de media cero, no correlacionado con  $y_0^{k-1}$ , y con matriz de varianza positiva definida dada por

$$R_{\tilde{y}}(k) = C(k)P(k|k-1)C(k)^T + R_e(k). \quad (6)$$

3. Suponga que (1) se reemplaza por la ecuación

$$x(k+1) = A(k)x(k) + B(k)u(k) + v(k), \quad (7)$$

donde  $B(k)$  es una matriz determinística finita de dimensiones apropiadas, y  $u$  es una señal medible y determinística o, en su defecto, tal que  $u(k) = f(y_0^{k-1})$ , donde  $f$  una función determinística. Determine una expresión recursiva para  $\hat{x}(k+1|k)$ .

**Solución:**

Vea el documento adjunto.

**Problema 2**

Considere el problema de estimar una constante real que obedece a una distribución uniforme en  $[-\Delta/2, \Delta/2]$ . Para ello, suponga que se cuenta con mediciones ruidosas de  $\theta$  dadas por

$$y(k) = \theta + e(k), \quad k \in \mathbb{N}_0, \quad (8)$$

donde  $e$  es un proceso de ruido blanco de media cero y varianza finita y constante en el tiempo. Suponga que  $\theta$  y  $e$  no están correlacionados.

Defina

$$\hat{\theta}(k) \triangleq \arg \min_{\substack{\theta = Ay_0^{k-1} + b \\ A \in \mathbb{R}^{1 \times k}, b \in \mathbb{R}}} \mathcal{E} \{ (\theta - \bar{\theta})(\theta - \bar{\theta})^T \} \quad (9)$$

1. Demuestre que,  $\forall k \in \mathbb{N}_0$ ,

$$\hat{\theta}(k+1) = \left( 1 - \frac{P(k)}{P(k) + R_e} \right) \hat{\theta}(k) + \frac{P(k)}{P(k) + R_e} y(k), \quad \hat{\theta}(0) = 0, \quad (10)$$

donde

$$P(k+1) = P(k) \left( 1 - \frac{P(k)}{P(k) + R_e} \right), \quad P(0) = \frac{\Delta^2}{12}. \quad (11)$$

2. Demuestre que<sup>1</sup>

$$\lim_{k \rightarrow \infty} \mathcal{E} \{ \hat{\theta}(k) - \theta \} = 0, \quad \lim_{k \rightarrow \infty} P(k) = 0. \quad (12)$$

Concluya que, como consecuencia de lo anterior y dejando de lado tecnicismos y casos patológicos,  $\hat{\theta}(k) \rightarrow \theta$  cuando  $k \rightarrow \infty$ .

---

<sup>1</sup>La formulación original de la tarea contenía error, pues se pedía probar  $\lim_{k \rightarrow \infty} \mathcal{E} \{ \hat{\theta}(k) \} = \theta$ . Esto es claramente falso. De hecho, (10) implica  $\mathcal{E} \{ \hat{\theta}(k) \} = 0, \forall k \in \mathbb{N}_0$ .

3. Considere el filtro con entrada  $y$  y salida  $\hat{\theta}$  descrito por (10) para cada valor de  $k \in \mathbb{N}_0$ . Determine si este filtro es pasa-bajos o pasa-altos, su ganancia a continua, polos, etc. En base a lo anterior explique, de modo intuitivo, la forma en que (10) construye  $\hat{\theta}$  y el efecto que la elección de  $P(0)$  y  $R_e$  tienen en la estimación de  $\theta$ .<sup>2</sup>
4. Use Matlab/Simulink para ilustrar los puntos anteriores. Comente. (Grafique  $y$ ,  $\hat{\theta}$ ,  $P$  y los polos del filtro mencionado en la parte anterior para al menos tres conjuntos de valores de  $P(0)$  y  $R_e$ .)
5. ¿Cuáles son las ventajas del esquema de estimación en (10) por sobre un filtro lineal e invariante en el tiempo? (Apóyese en sus simulaciones.)

### Solución

1. El problema en cuestión corresponde al problema de estimar, utilizando un estimador afín en las mediciones  $y(0), y(1), \dots, y(k-1)$ , el estado del sistema

$$x(k+1) = x(k), \quad x(0) = \theta, \quad (13)$$

$$y(k) = x(k) + e(k), \quad (14)$$

donde  $e$  corresponde a ruido de medición. Dadas la caracterización de  $e$  y de  $\theta$ , concluimos que el estimador buscado puede calcularse utilizando un Filtro de Kalman en que la varianza del ruido de proceso  $R_v(k) = 0$  y la varianza del ruido de medición es  $R_e(k)$  es constante para todo  $k$ , digamos  $R_e(k) = R_e$ . Para inicializar el Filtro de Kalman, y dado que  $x_o = \theta$  se halla uniformemente distribuido en  $[-\Delta/2, \Delta/2]$ , basta con elegir  $\hat{x}(0| - 1) = \mathcal{E}\{\theta\} = 0$  y  $P(0) = \mathcal{E}\{\theta^2\} = \Delta^2/12$ . Explotando el que todas las variables involucradas son escalares, es posible reescribir las ecuaciones usuales del Filtro de Kalman en la forma sugerida.

2. Las propiedades de estimadores afines óptimos en sentido cuadrático medio permiten concluir que  $\mathcal{E}\{\hat{\theta}(k) - \theta\} = 0$  para cada  $k \in \mathbb{N}_0$ . Por lo tanto,  $\lim_{k \rightarrow \infty} \mathcal{E}\{\hat{\theta}(k) - \theta\} = 0$ . Por otro lado, tenemos que (11) implica que

$$P(k+1) - P(k) = -\frac{P(k)^2}{P(k) + R_e} \leq 0. \quad (15)$$

Como, además,  $P(k) \geq 0$  tenemos que  $P(k)$  converge a medida que  $k \rightarrow \infty$ . Para probar que la convergencia es a cero, note que  $P(0) = \Delta^2/12 > 0$  y  $R_e > 0$ <sup>3</sup> implican que  $P(k+1) - P(k) < 0 \forall k$ . Por lo tanto, la convergencia debe ser a cero. Como  $P(k)$  corresponde a la covarianza del error de estimación en función de  $k$ , lo anterior implica que, cuando  $k \rightarrow \infty$ , el error de estimación se convierte en una variable aleatoria de media cero y varianza cero. La conclusión es inmediata.

3. Para cada  $k \in \mathbb{N}_0$ , (10) define un filtro pasabajos, de ganancia a continua unitaria y con un polo real en

$$z = 1 - \frac{P(k)}{P(k) + R_e}, \quad 0 \leq 1 - \frac{P(k)}{P(k) + R_e} \leq 1. \quad (16)$$

<sup>2</sup>Obviamente, elegir  $P(0) \neq \frac{\Delta^2}{12}$  y  $R_e$  distinto a la varianza del ruido  $e$  implica pérdida de optimalidad. Sin embargo, como usuario del estimador en (10), Ud. puede optar por otras elecciones para  $P(0)$  y  $R_e$ .

<sup>3</sup>De otro modo el problema de estimar  $\theta$  sería trivial.

Como  $P(k)$  es decreciente, el polo del filtro en (10) varía monotónicamente entre un valor inicial y el valor  $z = 1$  a medida que  $k$  crece. Por lo tanto, el ancho de banda del filtro en (10) se hace progresivamente angosto a medida que  $k$  crece.

En esencia, para estimar  $\theta$  basta calcular el promedio de  $y$ . El filtro definido por (10), al tener ganancia a continua unitaria y ser un pasabajos cuyo ancho de banda se hace progresivamente angosto, va refinando progresivamente la estimación de este promedio.

Si  $P(0) \rightarrow \infty$  o  $R_e \rightarrow 0$ , el polo inicial corresponde a  $z = 0$ , i.e., ancho de banda  $\pi$  (máximo posible). En este caso es de esperarse que la estimación, aún cuando será muy ruidosa para  $k$  pequeños, converja rápidamente al valor medio de  $y$ , i.e.,  $\theta$ . Note que, en este caso,  $\hat{\theta}$  es igual a la medición  $y$  para  $k = 0$ , i.e., las estimaciones iniciales no se usan (como  $P(0) \rightarrow \infty$  estas estimaciones pasadas son “muy malas”). A medida que  $k$  aumenta, se comienzan progresivamente a usar las estimaciones pasadas y cada vez menos las mediciones. En el límite, las mediciones ya no se usan, y sólo la estimación pasada.

Si  $P(0) \rightarrow 0$  o  $R_e \rightarrow \infty$ , entonces el polo inicial será cercano a  $z = 1$ . La estimación será limpia (sin ruido), pero la convergencia muy lenta.

Con elecciones apropiadas para  $P(0)$  y  $R_e$  se puede balancear la tasa de convergencia con el nivel de ruido en la estimación.

4. Se consideró el caso  $\Delta = 1$  y una varianza del ruido  $e$  igual a 0,1. Para una realización de  $\theta$  y  $e$ , se corrieron tres Filtros de Kalman: El Caso 1 corresponde al filtro de Kalman óptimo (i.e.,  $P(0) = \Delta^2/12 = 1/12$ ,  $x(0| - 1) = 0$  y  $R_e = 0,1$ ), el Caso 2 corresponde a un caso en que se sobreestima la varianza del ruido de medición y se usa  $R_e = 100$ , y el Caso 3 corresponde a un caso en que se subestima la varianza del ruido de medición y se usa  $R_e = 10^{-9}$ . Los resultados se observan en la Figura 1. Se aprecia un comportamiento consistente con lo descrito en la parte anterior.
5. La principal ventaja de usar un Filtro de Kalman por sobre uno invariante en el tiempo radica en que el Filtro de Kalman ajusta, en forma dinámica, la velocidad de convergencia de la estimación y la influencia del nivel de ruido en la misma.

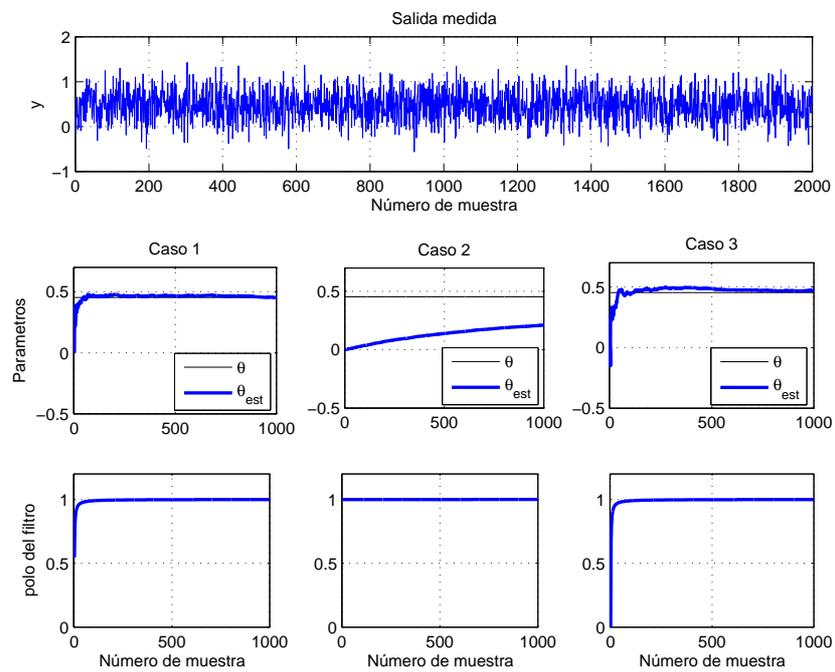


Figura 1: Estimación de  $\theta$  utilizando tres Filtros de Kalman (ver detalles en texto).

### Problema 3

Considere el sistema

$$x_G(k+1) = A_G x_G(k) + B_G u(k), \quad x_G(0) = x_{G,o}, \quad k \in \mathbb{N}_0, \quad (17a)$$

$$y_G(k) = C_G x_G(k), \quad (17b)$$

donde  $x_G$  ( $x_G(k) \in \mathbb{R}^n$ ) es el estado,  $u$  ( $u(k) \in \mathbb{R}^{n_u}$ ) la entrada e  $y_G$  ( $y_G(k) \in \mathbb{R}^{n_y}$ ) la salida. Suponga que  $A_G, B_G$  y  $C_G$  son matrices reales de dimensiones apropiadas, que  $x_o$  es una variable aleatoria de segundo orden, y que  $(A_G, B_G)$  es estabilizable y  $(C_G, A_G)$  es detectable.

A diferencia del caso analizado en la Pregunta 3 de la Tarea 2, suponga que *sólo se tiene acceso a la entrada de la planta y a una medición ruidosa de  $y_G$* . En particular, suponga que  $u(k)$  e

$$y(k) = y_G(k) + w(k) \quad (18)$$

son medibles para todo  $k$ , donde  $w$  es ruido blanco de media cero, varianza finita y constante para todo  $k$  e independiente de  $x_o$ .

La salida  $y_G$  del sistema en (17) debe seguir una referencia estocástica  $r$ . Suponga que  $r$  puede modelarse como un proceso estacionario en sentido amplio de segundo orden y media cero, que posee una densidad espectral de potencia racional y sin ceros sobre el círculo unitario. Suponga que  $r$  es independiente de  $(x_o, w)$ . Para medir la calidad del seguimiento de la referencia  $r$ , se propone el funcional

$$J \triangleq \sigma_e^2 + \rho \sigma_u^2, \quad \rho \in \mathbb{R}^+, \quad (19)$$

donde  $e \triangleq r - y_G$  es el error de seguimiento y

$$\sigma_i^2 \triangleq \lim_{k \rightarrow \infty} \mathcal{E} \left\{ [i(k) - \mu_i(k)]^T [i(k) - \mu_i(k)] \right\}, \quad i \in \{e, u\}. \quad (20)$$

1. Determine la ley de control lineal óptima.
2. Use simulaciones para comparar el desempeño de la estrategia por Ud. propuesta en la parte anterior y aquel de las estrategias por Ud. diseñadas en la Tarea 2. Comente.

### Solución

1. La situación bajo consideración se ilustra en la Figura 2, donde  $G \triangleq C_G(zI - A_G)^{-1}B_G$  es la función de transferencia de la planta,  $\Omega_r \in \mathcal{RH}_2$  es el sistema generador de la referencia, y  $w_r$  es una secuencia de ruido blanco de media cero y varianza unitaria. Suponga que

$$x_r(k+1) = A_r x_r(k) + B_r w_r(k), \quad x_r(0) = x_{r,o}, \quad k \in \mathbb{N}_0 \quad (21)$$

$$r(k) = C_r x_r(k) \quad (22)$$

corresponde a una realización mínima de  $\Omega_r$ .

Defina<sup>4</sup>

$$x(k) \triangleq \begin{bmatrix} x_G(k) \\ x_r(k) \end{bmatrix}, \quad Q \triangleq C_Q^T C_Q, \quad C_Q \triangleq \begin{bmatrix} -C_G & C_r \end{bmatrix}, \quad (23)$$

$$A \triangleq \begin{bmatrix} A_G & 0 \\ 0 & A_r \end{bmatrix}, \quad B \triangleq \begin{bmatrix} B_G \\ 0 \end{bmatrix}, \quad C \triangleq \begin{bmatrix} C_G & 0 \end{bmatrix}, \quad v(k) \triangleq \begin{bmatrix} 0 \\ B_r \end{bmatrix} w_r(k). \quad (24)$$

<sup>4</sup>Usamos la notación en el enunciado de la pregunta.

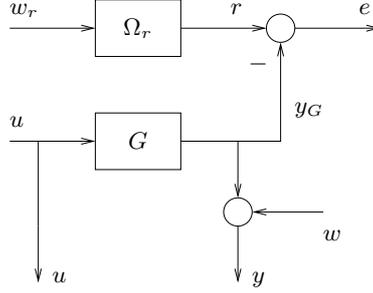


Figura 2: Situación considerada en la Pregunta 3.

Con lo anterior, el problema bajo consideración puede escribirse como el problema de hallar la ley de control lineal que minimiza

$$J \triangleq \lim_{N \rightarrow \infty} \frac{1}{N} \mathcal{E} \left\{ \sum_{k=0}^{N-1} \{x(k)^T Q x(k) + \rho u(k)^T u(k)\} \right\}, \quad (25)$$

sujeto a

$$x(k+1) = Ax(k) + Bu(k) + v(k), \quad x(0) = x_o \triangleq \begin{bmatrix} x_{G,o} \\ x_{r,o} \end{bmatrix}, \quad (26)$$

y a la restricción de que sólo son medible las señales  $u$  e  $y$ , donde

$$y(k) = Cx(k) + w(k). \quad (27)$$

Las condiciones del problema implican que  $x_o$  independiente de  $(w, w_r)$  y  $w_r$  es independiente de  $w$ . Además,  $\rho > 0$  y  $R_w > 0$ . Estas condiciones, junto a la detectabilidad y estabilizabilidad de  $(A_G, B_G, C_G, 0)$  y al hecho de que  $A_r$  es estable, permiten utilizar el principio de separación y concluir que la ley de control lineal estacionaria que minimiza  $J$  existe y obedece a<sup>5</sup>

$$u(k) = -K\hat{x}(k|k-1), \quad (28)$$

donde

$$K \triangleq (\rho I + B^T L B)^{-1} B^T L A, \quad (29)$$

$L$  es la única solución de

$$L = A^T L A - A^T L B (\rho I + B^T L B)^{-1} B^T L A + Q, \quad (30)$$

y

$$\hat{x}(k+1|k) = A\hat{x}(k|k-1) + Bu(k) + J(y(k) - C\hat{x}(k|k-1)), \quad (31)$$

<sup>5</sup>Esto supone que  $u(k)$  puede construirse sólo en términos de  $y^{k-1}$ .

donde

$$J \triangleq APC^T (CPC^T + R_w)^{-1} \quad (32)$$

y  $P$  es la única solución de

$$P = APA^T - APC^T (CPC^T + R_w)^{-1} CPA^T + \begin{bmatrix} 0 & 0 \\ 0 & B_r B_r^T \end{bmatrix}. \quad (33)$$

2. Considérese el ejemplo de la Pregunta 3 de la Tarea 2. Suponga, además, que el ruido de medición  $w$  posee varianza  $R_w = 0,1$ . Con estos parámetros, la ley de control óptima puede calcularse como sigue:

```
G=tf([1 -0.8],poly([2 0.5]),1);
G=minreal(ss(G));
Omega_r = tf([1-0.9],[1 -0.9],1);
Omega_r = minreal(ss(Omega_r));
A = blkdiag(G.a, Omega_r.a);
B = [G.b; zeros(size(Omega_r.a,1), size(G.b,2))];
C=[G.c 0];
Rv = blkdiag(0, 0, Omega_r.b*Omega_r.b' );
Rw=0.1;
rho=0.1;
CQ = [-G.c Omega_r.c];
Q = CQ'*CQ;
R = rho*eye(size(G.b,2));
Kopt = dlqr(A, B, Q, R);
Jopt = dlqr(A', C', Rv, Rw)';
```

La Figura 3 muestra estimaciones del costo óptimo para 100 realizaciones de las condiciones iniciales y de las señales exógenas, y las compara con los costos medidos en los casos de la Pregunta 3 de la Tarea 2. Se aprecia que la estrategia de control que mide los estados de  $G$  y  $\Omega_r$  obtiene mejor desempeño que la estrategia óptima diseñada en la Parte 1, así como también mejor desempeño que el controlador PI propuesto en la Pregunta 3 de la Tarea 2. Esto es consistente con la intuición, pues la situación considerada en este caso considera una fuente adicional de ruido (el ruido de medición  $w$ ) y, además, se deben estimar los estados de  $G$  y  $\Omega_r$ . Nuevamente, el controlador PI está lejos de alcanzar el desempeño de la propuesta de la Parte 1.

## Referencias

- [1] H. Kwakernaak and R. Sivan. *Linear Optimal Control Systems*. Wiley-Interscience, 1972.
- [2] T. Söderström. *Discrete-time stochastic systems*. Prentice Hall, 1994.

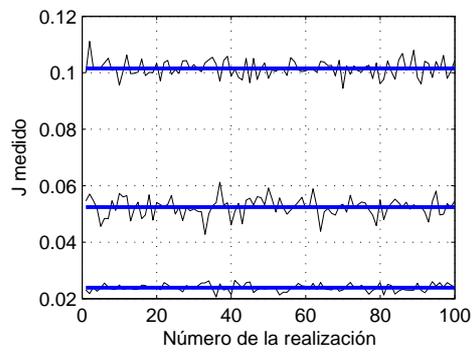


Figura 3: Valor medido del funcional  $J$  cuando se usa la ley de control óptima de la Parte 1 y los dos controladores de la Tarea 2.