Lesson 16

State Estimation:
Prediction

In this lesson, algorithms for mean-squared predicted estimates are developed.

\[ \hat{x}_{Ms}(k|j), \quad k > j \]

Brief Aside (PG10 Text): (a parameter estimation problem)

Given the measurements \( z(1), z(2), \ldots, z(k) \), determine an estimator of \( \theta \), namely,

\[ \hat{\theta}_{Ms}(k) = \Phi[z(i), i = 1, 2, \ldots, k] \]

such that the conditional mean-squared error

\[ J, [\hat{\theta}_{Ms}(k)] = E \{ \tilde{\theta}_{Ms} - \tilde{\theta}_{Ms} / z(1), z(2), \ldots, z(k) \} \]

is minimized, where \( \tilde{\theta}_{Ms} = \theta - \hat{\theta}_{Ms} \).
The estimator that minimizes the mean-squared error is

$$\hat{\theta}_{MS}(k) = E\{\theta | Z(k)\}$$

where $Z(k) = CL(\varepsilon(1), \varepsilon(2), \ldots, \varepsilon(k))$

**Pf.**

$$J[\hat{\theta}_{MS}] = E \{ (\theta - \hat{\theta}_{MS})(\theta - \hat{\theta}_{MS}) | Z \}$$

$$= E \{ \theta' \theta - \theta' \hat{\theta}_{MS} - \hat{\theta}_{MS}' \theta + \hat{\theta}_{MS}' \hat{\theta}_{MS} | Z \}$$

$$= E\{\theta'\varepsilon | Z\} - E\{\hat{\theta}'_{MS} | Z\} - E\{\hat{\theta}' \theta | Z\} + E\{\hat{\theta}'_{MS} \hat{\theta}_{MS} | Z\}$$

$$= E\{\theta'\varepsilon | Z\} - E\{\theta' \hat{\theta}_{MS} | Z\} - E\{\hat{\theta}_{MS}' \theta | Z\} + E\{\hat{\theta}_{MS}' \hat{\theta}_{MS} | Z\}$$

$$= E\{\theta'\varepsilon | Z\} - E\{\theta' \hat{\theta}_{MS} | Z\} - E\{\hat{\theta}_{MS}' \theta | Z\} + E\{\hat{\theta}_{MS}' \hat{\theta}_{MS} | Z\}$$

$$= E\{\theta'\varepsilon | Z\} + [\hat{\theta}_{MS} - E\{\theta' \varepsilon | Z\}] [\hat{\theta}_{MS} - E\{\theta' \varepsilon | Z\} - E\{\theta' \hat{\theta}_{MS} | Z\} - E\{\hat{\theta}_{MS}' \theta | Z\}]$$

min when

$$\hat{\theta}_{MS} = E\{\theta | Z(k)\}$$

$$\hat{\theta}_{MS} = E\{\theta | Z(k)\}$$
Corollary. When $\Theta$ and $Z(k)$ are jointly Gaussian, the estimator that minimizes the mean-squared error is

$$\hat{\Theta}_{M_S} = \mu_\Theta + P_{\Theta Z}(k) P^{-1}_{Z}(k) [Z(k) - \mu_Z(k)].$$

**Pf.**

$$\hat{\Theta}_{M_S}(k) = E\{ \Theta \mid Z(k) \}$$

Aside \((\text{then } 12-1\text{ text})\):

(Recall)

Let $X$ and $Y$ be $n$- and $m$-dim vectors that are jointly Gaussian. Then

$$E(X \mid Y) = m_X + P_{XY} P_Y^{-1} (Y - m_Y).$$

Set $\Theta = X$

$Z(k) = Y$

then $\hat{\Theta}_{M_S}(k)$ follows.
Single-stage predictor:

\[ \hat{x}(k|k-1) = \text{single-stage predictor} \]

From previous this, we know

\[ \hat{x}(k|k-1) = E \left\{ x(k) \mid \mathcal{Z}(k-1) \right\} \]

where

\[ \mathcal{Z}(k-1) = \text{Col} \left( \mathcal{Z}(1), \mathcal{Z}(2), \ldots, \mathcal{Z}(k-1) \right) \]

Note that

\[ x(k) = \Phi(k, k-1) x(k-1) + \Gamma(k, k-1) w(k-1) + \Psi(k, k-1) u(k-1) \]

take the conditional mean of both sides of the state equation:

\[ E \{ x(k) \mid \mathcal{Z}(k-1) \} = \Phi(k, k-1) E \{ x(k-1) \mid \mathcal{Z}(k-1) \} + \Gamma(k, k-1) E \{ w(k-1) \mid \mathcal{Z}(k-1) \} + \Psi(k, k-1) E \{ u(k-1) \mid \mathcal{Z}(k-1) \} \]

\[ \hat{x}(k|k-1) = \Phi(k, k-1) \hat{x}(k-1|k-1) + \Gamma(k, k-1) E \{ w(k-1) \mid k-1 \} + \Psi(k, k-1) E \{ u(k-1) \mid k-1 \} \]

Notice that this equation should be initialized by

\[ \hat{x}(0|0) = E \{ x(0) \mid \text{no measurements} \} = m_x(0) \]
let $p(k|k-1) = \text{error - covariance matrix}$

$$p(k|k-1) = \text{Cov}[\tilde{x}(k|k-1)]$$

**:**

$$\because P(k|k-1) = E\{[\tilde{x}(k|k-1) - m_x(k|k-1)][\tilde{x}(k|k-1) - m_x(k|k-1)]^T]\}$$

when $\tilde{x}(k|k-1) = x(k) - \hat{x}(k|k-1)$.

Additionally, (filtering error - covariance matrix):

$$P(k-1|k-1) = E\{[\tilde{x}(k-1|k-1) - m_x(k-1|k-1)][\tilde{x}(k-1|k-1) - m_x(k-1|k-1)]^T]\}$$

For unbiasedness $m_x(k|k-1) = m_x(k-1|k-1) = 0$. (prop. 7.11.2).

**:**

$$P(k|k-1) = E\{\tilde{x}(k|k-1)\tilde{x}'(k|k-1)\} \quad (1)$$

$$P(k-1|k-1) = E\{\tilde{x}(k-1|k-1)\tilde{x}'(k-1|k-1)\} \quad (2)$$

Also

$$\tilde{x}(k|k-1) = x(k) - \hat{x}(k|k-1) = \Phi(k, k-1)x(k-1) + \Gamma(k, k-1)w(k-1)$$

$$+ \Psi(k, k-1)u(k) - \Phi(k, k-1)\hat{x}(k-1|k-1) - \Psi(k, k-1)u(k)$$

$$= \Phi(k, k-0)[x(k-0) - \hat{x}(k-1|k-0)] + \Gamma(k, k-1)w(k-1).$$

**:**

$$\tilde{x}(k|k-1) = \Phi(k, k-1)\tilde{x}(k-1|k-1) + \Gamma(k, k-1)w(k-1) \quad (3)$$
\[ P(k|k-1) = E \left\{ \bar{\Phi} \bar{X}(k-1|k-1) + \Gamma W(k-1) \right\} \left[ \bar{\Phi} \bar{X}(k-1|k-1) + \Gamma W(k-1) \right]' \]

\[ = \bar{\Phi} E \left\{ \bar{X}(k-1|k-1) \bar{X}(k-1|k-1)' \right\} \bar{\Phi}' + \Gamma E \left\{ W(k-1) W(k-1)' \right\} \Gamma' \]

\[ \therefore P(k|k-1) = \bar{\Phi} (k; k-1) P(k-1|k-1) \bar{\Phi}' (k; k-1) + \Gamma (k; k-1) \Omega (k-1) \Gamma' (k; k-1) \]

it is observed that this step should be

initiated by

\[ P(0|0) = E \left\{ \tilde{X}(0) \tilde{X}'(0) \right\} = E \left\{ \left[ \bar{X}(0) - \bar{X}(0) \right] \left[ \bar{X}(0) - \bar{X}(0) \right]' \right\} \]

\[ = E \left\{ \left[ X(0) - X(0) \right] \left[ X(0) - X(0) \right]' \right\} \]

\[ = P_x(0) = \text{covariance matrix of } X(0) \]

Remark: \( P(k|k-1) \) and \( P_x(k) \) have the same form!
General state Predictor:

the objective is to determine \( \hat{x}(klj) \), \( k > j \), under the assumption that filtered state estimate \( \hat{x}(jlj) \) and its error covariance matrix \( P(jlj) = E[\tilde{x}(jlj)\tilde{x}^T(jlj)] \) are known for some \( j = 0, 1, \ldots \).

The mean squared predictor estimator of \( x(k) \), \( \hat{x}(klj) \), is given by

\[
\hat{x}(klj) = \Phi(kj) \hat{x}(jlj) + \sum_{i=j+1}^{k} \Phi(k,i) \Psi(i,i-1) u(i-1); \quad k > j
\]

(b) the vector random process \( \{ \tilde{x}(klj), k = j+1, j+2, \ldots \} \) is

i. zero mean,

ii. Gaussian, and

iii. first order Markov, and

iv. its covariance matrix is governed by

\[
P(klj) = \Phi(k,k-1) P(k-1lj) \Phi'(k,k-1) + \Sigma(k,k-1) \Psi(k-1) \Sigma'(k,k-1)
\]
\[- (1)
\]
\[
X(k) = \Phi(k,0)X(0) + \sum_{i=1}^{k} \Phi(k,i) \left[ C(i,i-1)W(i-1) + \Psi(i,i-1)U(i-1) \right] \]
\[
\Phi(k,j) \Phi(j,0)
\]
\[
\sum_{i=1}^{j} + \sum_{i=j+1}^{k}
\]
\[
\therefore X(k) = \Phi(k,j) \Phi(j,0)X(0) + \sum_{i=1}^{j} \Phi(k,j) \Phi(j,i) \left[ \cdots \right] \]
\[
\sum_{i=1}^{j} + \sum_{i=j+1}^{k} \Phi(k,i) \left[ \cdots \right].
\]
\[
\therefore X(k) = \Phi(k,j) \left\{ \Phi(j,0)X(0) + \sum_{i=1}^{j} \Phi(j,i) \left[ \cdots \right] \right\} \]
\[
\sum_{i=1}^{j} + \sum_{i=j+1}^{k} \Phi(k,i) \left[ \cdots \right].
\]
\[
\therefore X(k) = \Phi(k,j)X(j) + \sum_{i=j+1}^{k} \Phi(k,i) \left[ C(i,i-1)W(i-1) + \Psi(i,i-1)U(i-1) \right].
\]

We apply the fundamental theorem of estimation theory to this eqn by taking the conditional expectation with respect to \( \mathcal{F}(j) \) on both sides of it.
\[ \hat{x}_{kl} = \hat{x}_{kj} \hat{x}_{jl} + \sum_{i=j+1}^{K} \Phi(k,i) \Phi(i, i-1) E \{ w(i-1) | z(j) \} \]

\[ + \Psi(i, i-1) E \left\{ u(i-1) \right\} \]

\[ E \{ w(i-1) | z(j) \} = E \{ w(i-1) | z(0), z(1), z(2), \ldots, z(j) \} \]

Note that \( z(j) \) depends at most on \( x(j) \) which, in turn, depends at most on \( w(j-1) \).

\[ E \{ w(i-1) | z(j) \} = E \{ w(i-1) | w(0), w(1), \ldots, w(j-1) \} \]

where \( j+1 \leq i \leq K \) (\( i = j+1, j+2, \ldots, K \)).

Because of this range of values on \( i \), \( w(i-1) \) is never included in the conditioning set of values \( w(0), w(1), \ldots, w(j-1) \); hence,

\[ E \{ w(i-1) | z(j) \} = E \{ w(i-1) \} = 0, \quad \forall i = j+1, j+2, \ldots, K. \]

\[ \hat{x}_{kl} = \Phi(k,i) \hat{x}_{kl} \Phi(i, i-1) w(i-1) \quad ; \quad K > j. \]
ii. \[ \hat{x}(k|l|j) = \Phi(k,j) \hat{x}(j|l|j) + \sum_{i=j+1}^{k} \Phi(k,i) \psi(i|i-1) w(i-1) \]

\[ \tilde{x}(k|l|j) = x(k) - \hat{x}(k|l|j) \]

\[ \tilde{x}(k|l|j) \text{ is Gaussian.} \]
iii. Know

\[ \tilde{x}(k|j) = \Phi(k,i) \tilde{x}(j|i) + \sum_{i=j+1}^{k} \Phi(k,i) \Gamma(i,i-1) w(i-1) \]

This is similar to the pole of state eqn when \( w(k) = 0 \).

\[ \tilde{x}(k|j) = \Phi(k,k-1) \tilde{x}(k-1|j) + \Gamma(k,k-1) w(k-1) \]

\( \tilde{x}(k|j) \) is first-order Markov.

iv. Because \( \tilde{x}(k|j) \) is a pole of the state-eqn, its covariance \( P(k|j) \) is similar to \( P_x(k) \). (eqn \( 15-29 \))

\[ P(k|j) = \Phi(k,k-1) P(k-1|j) \Phi'(k,k-1) + \Gamma(k,k-1) \Phi(k) \Gamma'(k,k-1) \]

See exp \( \frac{16.1}{144} \) for \( j = 0 \).
Remark

the last thr is quite limited because presently the only values of $\hat{x}(ji)$ and $P(ji)$ that we know are those at $j = 0$ (no measurements).

$\Phi(K+1, K)$

144

$\mathbf{X}(K+1) = \frac{1}{\sqrt{2}} \mathbf{X}(K) + W(K)$

$Q = 25$

$P(0) = \text{variable}$

know ($j = 0$)

$\mathbf{X}(1|0) = \Phi(K, 0) \hat{\mathbf{X}}(0|0) = 0$

$\mathbf{X}(2|0) = \Phi(2, 0) \hat{\mathbf{X}}(0|0)$

Also

$P(K|0) = \frac{1}{2} P(K-1|0) + 25 = 0$

Let $P(0|0) = P(0) = 6 = 0$

$P(1|0) = \frac{1}{2} (6) + 25 = 28$

$P = 50.$

or

$P(0|0) = P(0) = 100 = 0$

$P(1|0) = \frac{1}{2} (100) + 25 = 75$

$P = 50.$

$\hat{\mathbf{X}}(K|0)$ (for large $K$) is zero with uncertainty of 50.
Figure 16-1  Prediction error variance $p(k|0)$. 
The Innovations Process:

The single-stage predicted estimate of \( \hat{x}(k+1) \) is

\[
\hat{x}(k+1 \mid k) = E \left\{ \hat{x}(k+1) \mid x(k) \right\}.
\]

Know

\( \tilde{x}(k+1) = H(k+1)x(k+1) + v(k+1) \)

\[E \{ \tilde{x}(k+1) \mid x(k) \} = H(k+1)E \{ x(k+1) \mid x(k) \} + E \{ v(k+1) \mid x(k) \} = \hat{x}(k+1 \mid k) \]

\[\hat{x}(k+1 \mid k) = H(k+1) \hat{x}(k+1 \mid k) \]

The innovations process (or prediction error process, or measurement residual process) is defined as follows:

\[\hat{z}(k+1 \mid k) = z(k+1) - \hat{x}(k+1 \mid k).\]
Thm.

a. \( \tilde{z}(k+1|k) \) can be represented as

i. \( \tilde{z}(k+1|k) = \tilde{z}(k+1) - \hat{x}(k+1|k) \)

ii. \( \tilde{z}(k+1|k) = z(k+1) - H(k+1)\hat{x}(k+1|k) \)

iii. \( \tilde{z}(k+1|k) = H(k+1)x(k+1|k) + v(k+1) \)

b. The innovations is a zero-mean, Gaussian, white noise sequence, with

\[
E\{ \tilde{z}(k+1|k) \tilde{z}'(k+1|k) \} = P_z(k+1|k) = H(k+1)p(k+1|k)H'(k+1) + R(k+1).
\]

Ps.

a. i. and ii are obvious.

iii. \( \tilde{z}(k+1|k) = \tilde{z}(k+1) - H(k+1)\hat{x}(k+1|k) \)

\[
= H(k+1)x(k+1) + v(k+1) - H(k+1)\hat{x}(k+1|k)
\]

\[
= H(k+1) [x(k+1) - \hat{x}(k+1|k)] + v(k+1) - \hat{x}(k+1|k).
\]
\[ \hat{z}(k+1|k) = H(k+1) \hat{x}(k+1|k) + v(k+1) \]

\[ E \{ \hat{z}(k+1|k) \} = H(k+1) E \{ \hat{x}(k+1|k) \} + E \{ v(k+1) \} = 0. \]

\[ \therefore \hat{z}(k+1|k) \text{ is z.m.} \]

Also,

\[ \hat{z}(k+1|k) = \hat{z}(k+1) - H(k+1) \hat{x}(k+1|k) \]

\[ \therefore \hat{z}(k+1|k) \text{ is Gaussian.} \]

To prove that \( \hat{z}(k+1|k) \) is white noise, we must show that

\[ E \{ \hat{z}(i+1|i) \hat{z}(j+1|j) \} = P_{\hat{z}\hat{z}} \delta_{ij} \]

let \( i > j \): (\( j > i \) is left as an exercise)

\[ \frac{16}{145} \]

\[ E \{ \hat{z}(i+1|i) \hat{z}(j+1|j) \} = E \left\{ \left[ H(i+1) \hat{x}(i+1|i) + v(i+1) \right] H(j+1) \hat{x}(j+1|j) + v(j+1) \right\} \]
\[
\begin{align*}
    &= H(i+1) E \left\{ \tilde{x}(i+1|i) \tilde{x}'(j+1|j) H'(j+1) \right\} + E \left\{ \nu(i+1) \nu'(j+1) \right\} \\
      &\quad + H(i+1) E \left\{ \tilde{x}(i+1|i) \nu'(j+1) \right\} + E \left\{ \nu(i+1) \tilde{x}(j+1|j) H'(j+1) \right\}.
\end{align*}
\]

\[
\begin{align*}
    &= H(i+1) E \left\{ \tilde{x}(i+1|i) \left[ H'(j+1) \tilde{x}'(j+1|j) + \nu'(j+1) \right] \right\} \\
      &\quad \underbrace{\tilde{X}(j+1|j)}_{\tilde{z}(j+1|j)} \\
      &\quad z(j+1) - H(j+1) \tilde{x}(j+1|j)
\end{align*}
\]

**Brief Aside** (Orthogonality principle) pg 131

Suppose \( f[z(k)] \) is any function of the data \( z(k) \). Then the error in the mean-squared estimator is orthogonal to \( f[z(k)] \) in the sense that

\[
E \left\{ [\Theta - \hat{\Theta}_m(k)] f'[z(k)] \right\} = 0.
\]

Using the fact that (verify):

\[
E \left\{ x g(\beta) \right\} = E \left\{ E \left[ x | \beta \right] g(\beta) \right\}
\]
we have:

\[ E \{ [\theta - \hat{\theta}_{MS}(k)] f'(x) \}^2 = E \{ E \{ \theta - \hat{\theta}_{MS} | x \}^2 f'(x) \} \]

\[ = 0. \]

**Remark:**

Note that if \( f[z(k)] = \hat{\theta}_{MS}(k) \)

then

\[ E \{ \hat{\theta}_{MS}(k) \hat{\theta}'_{MS}(k) \} = 0. \]

\[ \therefore \]

then Cont'd:

\[ \therefore \]

\[ E \{ \tilde{x}(i+1|1) \tilde{x}'(j+1|1) \} = H(i+1) E \{ \tilde{x}(i+1|1) \tilde{x}'(j+1|1) \} \]

\[ = H(i+1) E \{ \tilde{x}(i+1|1) \tilde{x}'(j+1|1) \} H'(j+1|1). \]

or the orthogonality principle

When \( i = j \)

\[ \text{P}_{xx}(i+1|i) = E \{ [H(i+1) \tilde{x}(i+1|i) + v(i+1)] [H(i+1) \tilde{x}(i+1|i) + v(i+1)] \} \]

\[ = H(i+1) E \{ \tilde{x}(i+1|i) \tilde{x}'(i+1|i) \} H'(i+1) + E \{ v(i+1) v'(i+1) \} \]

\[ = H(i+1) \text{P}(i+1|i) H'(i+1) + R(i+1). \]
Lesson 17

State Estimation:
Filtering (the Kalman Filter)

\[ \hat{x}(k+1|k+1) = E\{x(k+1)|z(k+1)\} \]

\[ = E\{x(k+1)|z(1), z(2), \ldots, z(k), z(k+1)\} \]

\[ = E\{x(k+1)|z(k), z(k+1)\} \]

Note that \( z(k) \) and \( z(k+1) \) are dependent, because

\[ z(k) = \{z(1), z(2), \ldots, z(k)\} \rightarrow \{x(1), x(2), \ldots, x(k)\} \]

\[ z(k+1) = H(k+1)x(k+1) + v(k+1) \]

\[ \hat{x}(k) = \Phi x(k) + \Gamma v(k) + \Phi u(k) \]

\( z(k) \) and \( z(k+1) \) both depend on \( x(k) \) and are, therefore, dependent.
Recall that $x(k+1)$, $z(k)$ and $\tilde{z}(k+1)$ are jointly Gaussian random vectors; hence,

$$\hat{x}(k+1 | k+1) = \mathbb{E} \{ x(k+1) | z(k), \tilde{z}(k+1 | k) \}$$

where

$$\tilde{z}(k+1 | k) = z(k+1) - \mathbb{E} \{ \tilde{z}(k+1 | k) \}$$

(innovations process)

- Theorem 12-4

Recall

If $x$, $y$, and $z$ are jointly Gaussian and if $y$ and $z$ are dependent,

$$\mathbb{E} \{ x | y, z \} = \mathbb{E} \{ x | y, \tilde{z} \}$$

where

$$\tilde{z} = z - \mathbb{E} \{ z | y \}$$

So that

$$\mathbb{E} \{ x | y, \tilde{z} \} = \mathbb{E} \{ x | y \} + \mathbb{E} \{ x | \tilde{z} \} - \mathbb{E} \{ x \}.$$