Module – 6.5

FROM GAUSS TO KALMAN: SEQUENTIAL, LINEAR MINIMUM VARIANCE ESTIMATION

S. Lakshmivarahan

School of Computer Science

University of Oklahoma

Norman, Ok – 73069, USA

varahan@ou.edu

LINEAR MINIMUM VARIANCE ESTIMATE

- Similar to Gauss-Markov theorem in Chapter 14 (LLD (2006))
- $z = Hx + v$
- Assumptions:
  - $E(v) = 0$, $COV(v) = \Sigma_v$ – SPD
  - $E(x) = m_x$, $COV(x) = \Sigma_x$ – SPD
  - x, v are not correlated
- Seek $\hat{x} = \Phi(z) = Az + b$
  - Linear, unbiased, min. variance estimate

MINIMUM VARIANCE - CONTINUED

- Let $\tilde{x} = x - \hat{x}$
- We seek to minimize mean squared error:

$$
\begin{aligned}
E[\tilde{\mathbf{x}}^T \tilde{\mathbf{x}}] &= E[(\mathbf{x} - \hat{\mathbf{x}})^T (\mathbf{x} - \hat{\mathbf{x}})] \\
&= E[tr[(\mathbf{x} - \hat{\mathbf{x}})^T (\mathbf{x} - \hat{\mathbf{x}})] \\
&= E[tr[(\mathbf{x} - \hat{\mathbf{x}})(\mathbf{x} - \hat{\mathbf{x}})^T] \\
&= tr[E(\mathbf{x} - \hat{\mathbf{x}})(\mathbf{x} - \hat{\mathbf{x}})^T] \\
&= tr(\mathbf{P})
\end{aligned}
$$

where COV($\tilde{x}$) = P

CONDITION FOR UNBIASEDNESS

- $m = E(\hat{x}) = E(b + Az) = b + A \cdot E(Hx + v) = b + AHm$
- $b = (I - AH)m$
- $\hat{x} = b + Az = (I - AH)m + Az = m + A(z - Hm)$
    - Look at this structure – we saw it in Bayesian framwork!
- $P = COV(\tilde{x}) = E[(x - \hat{x})(x - \hat{x})^T]$
- $x - \hat{x} = (x - m) - A(z - Hm)$
- $(x - \hat{x})^T(x - \hat{x}) = [(x - m) - A(z - Hm)][(x - m) - A(z - Hm)]^T$

$$= (x - m)(x - m)^T - (x - m)(z - Hm)^T A^T$$

$$- A(z - Hm)(x - m)^T + A(z - Hm)(z - Hm)^T A^T$$

- But $z = Hx + v$, $z - Hm = H(x - m) + v$

EXPRESSION FOR THE VARIANCE

- $\therefore$ $P = E(x - \hat{x})(x - \hat{x})^T$

  $= E[(x - m)(x - m^T]$ $\rightarrow$ Σ_x

  $- E[(x - m)[(x - m)^T H^T + v^T]]A^T$ $\rightarrow$ $-\Sigma_x H^T A^T$

  $- E[AH(x - m) + v][x - m]^T$ $\rightarrow$ $-AH\Sigma_x$

  $+ A^T E[(H(x - m) + v)(H(x - m) + v)^T]A$ $\rightarrow$ ADA^T

- $P = \Sigma_x + ADA^T - AH\Sigma_x - \Sigma_x H^T A^T$, $D = (H\Sigma_x H^T + \Sigma_v)$
- Thus P is a quadratic function of $A|_{nxm}$

MINIMIZING THE TOTAL VARIANCE

- Minimize trace of P (total variance) w.r.t. A
- tr(P) $= \sum_{i=1}^{n} \mathrm{P}_{\mathrm{ii}}$
- $\mathrm{P}_{\mathrm{ii}} = (\Sigma_x)_{ii} + A_{i*}DA_{i*}^T - A_{i*}b_{i*}^T - b_{i*}A_{i*}^T$
  - b_{i*} = ith row of n x m matrix $\Sigma_x H^T$
- $\mathrm{P}_{\mathrm{ii}} = A_{i*}DA_{i*}^T - 2b_{i*}^T A_{i*} + (\Sigma_x)_{ii}$

  $= y^T D y - 2by + c$
  - $y = A_{i*}$ i^{th} row of A, $y^T = A_{i*}^T$
  - $b = b_{i*}^T$, $b^T = b_{i*}$

MINIMIZATION

- Minimize P_{ii} w.r.t y – a standard quadratic form
- $\nabla P_{ii} = 2Dy - 2b = 0$
- $\Rightarrow y = D^{-1}b$
- $\Rightarrow A_{i*}^T = D^{-1}b_{i*}^T$
- $[A_{1*}^T \; A_{2*}^T \; \ldots\ldots \; A_{m*}^T] = D^{-1}[b_{1*}^T \; b_{2*}^T \ldots. \; b_{m*}^T]$

OPTIMAL P

- $A^T = D^{-1}H\Sigma_x$
- $A = \Sigma_x H^T D^{-1}$

  $= \Sigma_x H^T[H\Sigma_x H^T + \Sigma_v]^{-1}$
- $\therefore\ \hat{x} = m + \Sigma_x H^T[H\Sigma_x H^T + \Sigma_v]^{-1}[z - Hm]$
- Substituting A in P
- $\Rightarrow P = \Sigma_x - \boxed{\Sigma_x H^T[H\Sigma_x H^T + \Sigma_v]^{-1}H\Sigma_x}$

Subtracted

RELATION BETWEEN BAYES L.S. SOLUTION AND LINEAR MIN. VARIANCE SOLUTION - DUALITY

- Bayesian – state space
  - $\hat{x}_{MS} = \Sigma_e[H^T\Sigma_v^{-1}z + \Sigma_e^{-1}m_x] \rightarrow$ (16.2.26)
  - $\Sigma_e = [H^T\Sigma_e^{-1}H + \Sigma_e]^{-1} = COV(\hat{x}_{MS}) \rightarrow$ (16.2.25)
  - State-space, used for n < m
- L.M.V. – observation space
  - $\hat{x} = m + \Sigma_x H^T[H\Sigma_x H^T + \Sigma_v]^{-1}[z - Hm] \rightarrow$ (17.1.15)
  - $P = \Sigma_x – \Sigma_x H^T[H\Sigma_x H^T + \Sigma_v]^{-1}H\Sigma_x \rightarrow$(17.1.11)
  - Observation space, used for m < n
- They are the same!

BRIDGE: SHERMAN-MORRISON-WOODBURY LEMMA IN MATRIX THEORY (APPENDIX B)

- LMV
  - Recall: $D = (H\Sigma_x H^T + \Sigma_v)$
  - $D^{-1} = (H\Sigma_x H^T + \Sigma_v)^{-1}$

    $= \Sigma_v^{-1} - \Sigma_v^{-1}H[H^T\Sigma_v^{-1}H + \Sigma_x^{-1}]^{-1}H^T\Sigma_v^{-1}$
  - Multiply both side by $\Sigma_x H^T$

    $\Sigma_x H^T[H\Sigma_v H^T + \Sigma_v]^{-1}$

    $= \Sigma_x H^T\Sigma_v^{-1} – \Sigma_x H^T\Sigma_v^{-1}H[H^T\Sigma_v^{-1}H + \Sigma_x^{-1}]^{-1}H^T\Sigma_v^{-1}$

    $=\{\Sigma_x – \Sigma_x H^T\Sigma_v^{-1}H[H^T\Sigma_v^{-1}H + \Sigma_v^{-1}]^{-1}\}H^T\Sigma_v^{-1}$

    $=\{\Sigma_x[H^T\Sigma_v^{-1}H+\Sigma_x^{-1}]-\Sigma_x H^T\Sigma_v^{-1}H\}(H^T\Sigma_v^{-1}H+\Sigma_x^{-1})^{-1}H^T\Sigma_v^{-1}$

    $= I\cdot(H^T\Sigma_v^{-1}H+\Sigma_x^{-1})^{-1}H^T\Sigma_v^{-1}$

SHERMAN-MORRISON-WOODBURY CONTINUED

- Now, substitute in (17.1.15)

$$\hat{x} = m + \Sigma_x H^T[H\Sigma_x H^T + \Sigma_v]^{-1}(z - Hm)$$
$$= m + (H^T\Sigma_v^{-1}H + \Sigma_x^{-1})^{-1}H^T\Sigma_v^{-1}(z - Hm)$$
$$= (H^T\Sigma_v^{-1}H + \Sigma_x^{-1})^{-1}H^T\Sigma_v^{-1}z$$
$$+ \{I - (H^T\Sigma_v^{-1}H + \Sigma_x^{-1})^{-1}H^T\Sigma_v^{-1}H\}m$$

- Consider the second term:

$$\{I - (H^T\Sigma_v^{-1}H + \Sigma_x^{-1})^{-1}H^T\Sigma_v^{-1}H\}m$$
$$= (H^T\Sigma_v^{-1}H + \Sigma_x^{-1})^{-1}[(H^T\Sigma_v^{-1} H + \Sigma_x^{-1}) - H^T\Sigma_v^{-1}H]m$$
$$= (H^T\Sigma_v^{-1}H + \Sigma_x^{-1})[\Sigma_x^{-1}m]$$

- Combining

$$\hat{x} = (H^T\Sigma_v^{-1}H + \Sigma_x^{-1})^{-1}[\Sigma_x^{-1}m + H^T\Sigma_v^{-1}z]$$

KALMAN FILTERS - STATIC CASE

- $x \in R^n$ – unknown, constant
- x^- is an unbiased estimate of x if no observation.
- $E(x^-) = x$
- (x^-, Σ_-) – prior information
  - $z = Hx + v$, $E(v) = 0$, $COV(v) = \Sigma_v$
  - Usual conditions on v
- Linear Min. Variance Approach
  - $x^+ = Lx^- + Kz$ (x^+: posterior)
  - Find L, K such that x^+ us unbiased and has minimum variance

STATIC CASE - CONTINUED

- (a) Unbiasedness:
  - $x^+ = Lx^- + Kz$
  - $x = E(x^+)$

    $= E[Lx^- + Kz]$

    $= E[Lx^- + K(Hx + v)]$

    $= LE(x^-) + KHx + \cancel{KE(v)}$

    $= Lx + KHx = (L + KH)x$
  - $\therefore L + KH = I$ or $L = I - KH$
  - $\therefore x^+ = Lx^- + Kz$

    $= (I - KH)x^- + Kz \longleftarrow$ structure of the unbiased estimate

    $= x^- + K[z - Hx^-]$

STATIC CASE - CONTINUED

- We now need to compute the total variance of x^+
- $\text{var}(x^+) = E[(x^+ - x)^T(x^+ - x)]$

$$= E[tr[(x^+ - x)^T(x^+ - x)]]$$

$$= E[tr[(x^+ - x)(x^+ - x)^T]]$$

$$= tr(\Sigma^+)$$

- Recall:
  - $x^+ = (I - KH)x^- + Kz$

    $= (I - KH)x^- + KHx + Kv$
  - $\therefore x^+ - x = (I - KH)x^- + KHx - x + Kv$

    $= (I - KH)(x^- - x) + Kv$

STATIC CASE - CONTINUED

- $\therefore \Sigma^{+} = E[(I - KH)(x^{-} - x) + Kv]\,[(I - KH)(x^{-} - x) + Kv]^{T}$

  $= (I - KH)E[(x^{-} - x)(x^{-} - x)^{T}](I - KH)^{T} + KE(vv^{T})K^{T}$

  $= (I - KH)\Sigma_{-}(I - KH)^{T} + K\Sigma_{v}K^{T}$

  $= \Sigma_{-} + KDK^{T} - KH\Sigma_{-} - \Sigma_{-}H^{T}K^{T}$

  $D = (H\Sigma_{-}H^{T} + \Sigma_{v})$

- Choose K to minimize $tr(\Sigma^{+})$
- Similar to the problem we just solved.
  - $\Rightarrow K = \Sigma_{-}H^{T}D^{-1}$

    $= \Sigma_{-}H^{T}[H\Sigma_{-}H^{T} + \Sigma_{v}]^{-1}$

Kalman gain

STATIC CASE - CONTINUED

- $\therefore\ x^+ = x^- + \Sigma_- H^T[H\Sigma_- H^T + \Sigma_x]^{-1}[z - Hx^-]$

  $\Sigma^+ = \Sigma_- - \Sigma_- H^T[H\Sigma_- H^T + \Sigma_v]^{-1}H\Sigma_-$

EXERCISES

1. Σ_+ does not depend on observations and hence can be precomputed – Verify this claim
2. Reformulate as 3-D Var

   (x^-, Σ_-) and (z, Σ_v) $z = Hx + v$

   $f(x) = ½(z - Hx)^T \Sigma_v^{-1}(z - Hx) + ½(x^- - x)^T \Sigma_-^{-1} (x^- - x)$

   Min. f(x) w.r. to x and find the solution

REFERENCE

- A. P. Sage and J. L. Melsa (1971) Estimation Theory and its application to communications and Control, *McGraw Hill*
- Also refer to chapter 17 in LLD (2006)