Minimum mean-square error (MMSE)

$$\underline{x} = U \cdot \underline{\theta} + \underline{w}$$

$\underline{\theta} \to \mathcal{N}(0, \sigma_\theta^2)$

$\underline{w} \to \mathcal{N}(0, \sigma^2)$

Bayesian PDF of $\underline{\theta}$

linear model

Solution:

MMSE

$$\hat{\underline{\theta}} = E[\underline{\theta} \mid \underline{x}]$$

difficult task

① • Gaussian case

② • linear estimator

$$\hat{\underline{\theta}} = A\underline{x} + \underline{b}$$

$$\hat{\underline{\theta}} = C_{\underline{\theta}\underline{x}} \, C_{\underline{x}\underline{x}}^{-1} (\underline{x} - \underline{\mu}_x) + \underline{\mu}_\theta$$

$E[(\underline{\theta} - \underline{\mu}_\theta)(\underline{x} - \underline{\mu}_x)^T]$

$E[(\underline{x} - \underline{\mu}_x)(\underline{x} - \underline{\mu}_x)^T] = C_{xx}$

how do we know this?

running estimate

$$\hat{C}_{xx} = \frac{1}{L}\sum_{l=1}^{L}(\underline{x}_l - \hat{\underline{\mu}}_x)(\underline{x}_l - \hat{\underline{\mu}}_x)^T$$

- assumption

Sample covariance matrix

Back to problem : mixing matrix

$$\underline{x} = U\underline{\theta} + \underline{w}$$

size P

$$\underline{x} = \begin{bmatrix} x[0] \\ \vdots \\ x[n-1] \end{bmatrix} = \begin{bmatrix} & U & \end{bmatrix} \underline{\theta} + \underline{w}$$

$x[n]$ $\qquad \underline{u}^T[n]$ $\qquad w[n]$

Orthogonality Principle :

$$E[\underbrace{(\hat{\theta}_i - \theta)}_{\text{error}}\underbrace{\underline{x}}_{\text{data}}] = 0 \quad \forall i$$

new estimate

$$\Theta_i[n] \leftarrow \Theta_i[n-1] \quad \forall i$$

$$\hat{\Theta}_i[n] = \hat{\Theta}_i[n-1] + C_i[n]$$

correction @ n
for estimate i

From the OP:

① $E[(\hat{\Theta}_i[n] - \Theta_i) \cdot \underline{X}] = 0$ n equations

and $E[(\hat{\Theta}_i[n] - \Theta_i) \cdot X[n]] = 0$ one equation

① $E[(\hat{\Theta}_i[n-1] + C_i[n]) \cdot \underline{X}] = 0$

a) $E[C_i[n] \cdot \underline{X}] = 0$

b) $E[(\hat{\Theta}_i[n-1] + C_i[n] - \Theta_i) X[n]] = 0$

i) does this work?

ii) compute $C_i[n]$

We will need $\hat{X}[n]$ $X \longrightarrow \hat{X}[n]$

$$\hat{x}[n] = E[x[n]] + C_{x[n]\cdot\underline{x}}^{T}\, C_{\underline{x}\underline{x}}^{-1} (\underline{x} - \underline{\mu}_x)$$

$x[n] = \underline{u}^T[n] \cdot \underline{\theta} + w[n]$

$U\underline{\theta} + n$

$$E[(x[n] - \overline{x[n]}) \cdot (\underline{x} - \underline{\mu}_x)]$$

$$= E[(\underline{u}^T[n]\underline{\theta} - \underline{u}^T[n] \cdot \underline{\mu}_\theta)(\underline{x} - \underline{\mu}_x)]$$

$$= \underline{u}^T[n]\, E[(\underline{\theta} - \underline{\mu}_\theta)(U\underline{\theta} - U\underline{\mu}_\theta)^T]$$

Bayesian term

$$= \underline{u}^T[n]\, \underbrace{E[(\underline{\theta} - \underline{\mu}_\theta)(\underline{\theta} - \underline{\mu}_\theta)^T]}\, U^T$$

$$\hat{x}[n] = \underline{u}^T[n] \cdot \underline{\mu}_\theta + \underline{u}^T[n]\, \underbrace{C_{\theta\theta}\, U^T C_{\underline{x}\underline{x}}^{-1} (\underline{x} - \underline{\mu}_x)}$$

$$\hat{x}[n] = \underline{u}^T[n]\, \underline{\mu}_\theta + \underline{u}^T[n]\, \hat{\underline{\theta}}[n-1]$$

measurement prediction

MMSE estimate of new measurement from $\underline{x}$

$e[n] \cdot \begin{bmatrix} x[0] \\ \vdots \\ x[n-1] \end{bmatrix}$

$$OP: \quad E[(\underbrace{\hat{x}[n] - x[n]}_{\text{error}}) \cdot \underline{x}] = 0$$

if I use

$$c_i[n] = K_i[n]\underbrace{(\hat{x}[n] - x[n])}_{\text{measurement error}}$$

$\Rightarrow$ condition a) is satisfied

Find $K_i[n]$

compute at every n for every i

Kalman gain

$$\underline{K}[n] = \begin{bmatrix} K_1[n] \\ \vdots \\ K_p[n] \end{bmatrix}$$

2nd condition; b)

$$E[(\theta_i - \hat{\theta}_i[n-1] - c_i[n])x[n]] = 0$$

scalar

$$E[(\theta_i - \hat{\theta}_i[n-1] - k_i[n](x[n] - \hat{x}[n])x[n]] + 0$$

$$\underbrace{E[(\theta_i - \hat{\theta}_i[n] - k_i[n](x[n] - \hat{x}[n]) \cdot \hat{x}[n]]}_{=0}$$

$$E\left[\left(\theta_i - \hat{\theta}_i[n-1]\right)\left(x[n] - \hat{x}[n]\right)\right]$$

$$= k_i[n]\, E\left[\left(x[n] - \hat{x}[n]\right)^2\right]$$

$$E\left[\left(\theta_i - \hat{\theta}_i[n-1]\right)\left(\underline{v}^T[n]\,\underline{\theta} - \underline{v}^T[n]\,\hat{\underline{\theta}}[n-1]\right)\right]$$

$$= k_i[n]\, E\Big[\left(\underline{v}^T[n]\,\underline{\theta} - \underline{v}^T[n]\,\hat{\underline{\theta}}[n-1]\right)$$
$$\left(\underline{v}^T[n]\,\underline{\theta} - \underline{v}^T[n]\,\hat{\underline{\theta}}[n-1]\right)^T\Big]$$

$$\underline{v}^T[n]\, E\left[\left(\theta_i - \hat{\theta}_i[n-1]\right)\left(\underline{\theta} - \hat{\underline{\theta}}[n-1]\right)\right]$$

$$= k_i[n]\, \underbrace{\underline{v}^T[n]\, E\left[\left(\underline{\theta} - \hat{\underline{\theta}}[n-1]\right)\left(\underline{\theta} - \hat{\underline{\theta}}[n-1]\right)^T\right] \underline{v}[n]}_{= \underline{v}^T[n]\, M[n-1]\, \underline{v}[n]}$$

- parameter estimation error matrix at time $n-1$

$$k_i[n] = \frac{\left(\underline{v}^T[n]\, E\left[\left(\theta_i - \hat{\theta}_i[n-1]\right)\left(\underline{\theta} - \hat{\underline{\theta}}[n-1]\right)\right]\right)}{}$$

$$\underline{u}^T[n]\, M[n-1]\, \underline{u}[n]$$

scalar

$$\underline{K}[n] = \begin{bmatrix} K_1[n] \\ \vdots \\ K_p[n] \end{bmatrix} = \frac{\text{object} \quad \underline{u}[n]}{\underline{u}^T[n]\, M[n-1]\, \underline{u}[n]}$$

$$\left(\underline{\begin{bmatrix} | & | & | & | & | \end{bmatrix}} \right)^T = \begin{bmatrix} | & | & & | \\ E[(\theta_1-\hat{\theta}_1[n])(\underline{\theta}-\hat{\underline{\theta}})], & E[(\theta_2-\hat{\theta}_2[n])(\underline{\theta}-\hat{\underline{\theta}})] & \cdots & \\ | & | & & | \end{bmatrix}^T$$

$$\begin{bmatrix} — E[(\theta_1-\hat{\theta}_1[n])(\underline{\theta}-\hat{\underline{\theta}})]^T — \\ — E[(\theta_2-\hat{\theta}_2[n])(\underline{\theta}-\hat{\underline{\theta}})^T] — \\ \\ \text{———} \\ \text{——} \end{bmatrix}$$

$$E\left[(\underline{\theta}-\underline{\hat{\theta}}[n-1])(\underline{\theta}-\underline{\hat{\theta}}[n-1])^T\right]$$

$$\underline{K}[n] = \frac{M[n-1]\, \underline{u}[n]}{\underline{u}^T[n]\, M[n-1]\, \underline{u}[n]}$$

vector

scalar

Kalman gain

$$\underline{X}' = \begin{bmatrix} \underline{X} \\ x[n] \end{bmatrix} = \begin{bmatrix} U \\ \underline{u}^T[n] \end{bmatrix} \underline{\theta} + \begin{bmatrix} \underline{W} \\ w[n] \end{bmatrix}$$

1step

K-step update

$$\underline{\hat{\theta}}[n] = \underline{\hat{\theta}}[n-1] + \underline{k}[n] \cdot \left(x[n] - \hat{x}[n] \right)$$

reuse old estimate

Kalman gain

Elements:

- Gain update:

order p^2

$$\underline{k}[n] = \frac{M[n-1]\,\underline{u}[n]}{\underline{u}^T[n]\, M[n-1]\, \underline{u}[n] + \sigma^2[n]}$$

$O(p^2)$

- **Error Variance Matrix:**

$$\underbrace{M[n]}_{} = \left(I - \underline{k}[n] \cdot \underline{u}^T[n]\right) M[n-1]$$

$O(p^2)$

- **Estimator Update:**

$O(p^2)$

$$\hat{\underline{\theta}}[n] = \hat{\underline{\theta}}[n-1] + \underline{k}[n] \underbrace{\left(x[n] - \underline{u}^T[n]\,\hat{\theta}[n-1]\right)}_{x[n]-\hat{x}[n]}$$

prediction error

The update has complexity $O(p^2)$
versus $O(n^3) \leftarrow C_{\underline{xx}}^{-1}$

- Sequential Linear Minimum Mean-square error estimator

Update $M[n] \longleftarrow M[n-1]$

$$M[n] = E\left[(\underline{\theta} - \hat{\underline{\theta}}[n])(\underline{\theta} - \hat{\underline{\theta}}[n])^T\right]$$

$$= E\left[\left(\underline{\theta} - \hat{\underline{\theta}}[n-1] - \underline{k}[n]\left(x[n] - \hat{x}[n]\right)\right)\left(\underline{\theta} - \hat{\underline{\theta}}[n-1] - \underline{k}[n]\left(x[n] - \hat{x}[n]\right)\right)^T\right]$$

$$= M[n-1] - E\left[(\underline{\theta} - \hat{\underline{\theta}}[n-1])(x[n] - \hat{x}[n])\right] \underline{k}^T[n]$$

$$- \underline{k}[n]\, E\left[(x[n] - \hat{x}[n])(\underline{\theta} - \hat{\underline{\theta}}[n-1])^T\right]$$

$$+ \underline{k}[n]\, E\left[(x[n] - \hat{x}[n])^2\right] \cdot \underline{k}^T[n]$$

drives estimate

Replace $$x[n] - \hat{x}[n] = \underbrace{\underline{v}^T[n]\left[\underline{\theta} - \hat{\underline{\theta}}[n-1]\right]}_{\text{scalar}} + w[n]$$

$$(x[n] - \hat{x}[n])^2 = \underline{v}^T[n](\underline{\theta} - \hat{\underline{\theta}}[n-1])(\underline{\theta} - \hat{\underline{\theta}}[n-1])^T \underline{v}[n]$$

$$+ \, w^2[n]$$

$$M[n] = M[n-1] - M[n-1]\,\underline{u}[n]\cdot\underline{k}^T[n]$$

$$- \underline{k}[n]\,\underline{u}^T[n]\,M[n-1]$$

$$\underbrace{+ \underline{k}[n]\,\underline{u}^T[n]\,M[n-1]\,\underline{u}[n]\,\underline{k}^T[n] + \underline{k}[n]\,\sigma^2\,\underline{k}^T[n]}$$

$$= \underline{k}[n]\underbrace{\left[\underline{u}^T[n]\,M[n-1]\,\underline{u}[n] + \sigma^2\right]}\underline{k}^T[n]$$

from gain update:

$$\underline{k}[n]\left[\underline{u}^T[n]\,M[n-1]\,\underline{u}[n] + \sigma^2\right] = M[n-1]\,\underline{u}[n]$$

$$M[n] = M[n-1]\left(I - \cancel{\underline{u}[n]}\,\underline{k}^T[n] + \cancel{\underline{u}[n]}\,\underline{k}^T[n]\right)$$

$$- \underline{k}[n]\,\underline{u}^T[n]\,M[n-1]$$

$$M[n] = \left(\underline{I} - \underline{k}[n]\,\underline{u}^T[n] \right) M[n-1]$$

Christian.Schlegel@Dal.ca