parameter

Θ → $\underline{X}$ → MMSE

$E((\hat{\Theta}-\theta)^2)$

$f_{\Theta}(\theta)$

$$\boxed{\hat{\Theta} = E(\Theta|\underline{X})}$$

② $\int_{-\infty}^{\infty} f_{\Theta|\underline{X}}(\theta|\underline{x})\cdot\theta$ ①

Classical case

best $\hat{\Theta}$ = ?

CRLB: $var(\hat{\Theta}) \geq \frac{1}{I(\theta)}$

$E((\hat{\theta}-\theta)^2)$

$E((\hat{\theta}-E(\hat{\theta}))^2)$

$f_{X|\Theta}(x|\theta)$

Example:

$X_i = a + n_i$ ~ $\sim N(0,\sigma^2)$

⇓

$\underline{X} = (X_1 \cdots X_n)$

Classical:

$var(\hat{\theta}) = \frac{\sigma^2}{n}$

$\hat{\Theta} = \frac{1}{n}\sum^{n} X_i$

Bayesian: $f_{\underline{X}|a}(\underline{x}|a) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_i-a)^2}{2\sigma^2}}$

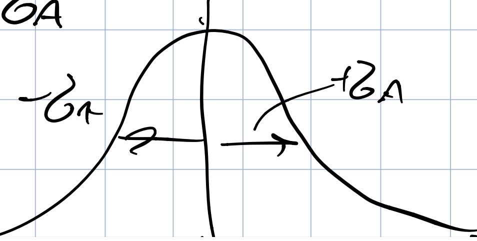
$$f_{\underline{X},A}(\underline{x},a) = f_{\underline{X}|A}(\underline{x}|a)\cdot f_A(a)$$

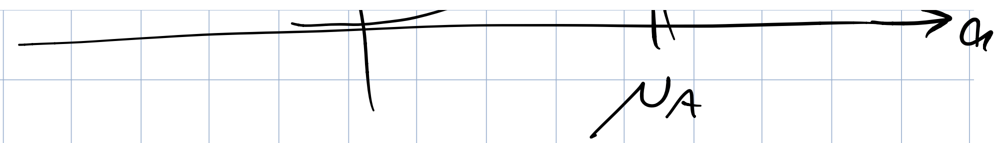
must have
- good assumption
- performance depends on prior

$$E(A|\underline{x}) = \int_{-\infty}^{\infty} A\cdot f_{A|\underline{X}}(a|\underline{x}) = \int_{-\infty}^{\infty} A \frac{f_{\underline{X}|A}(\underline{x}|a)\, f_A(a)}{f_{\underline{X}}(\underline{x})}$$

• compute also

Assume: $f_A(a) = \frac{1}{\sqrt{2\pi\sigma_A^2}} e^{-\frac{(a-\mu_A)^2}{\sigma_A^2}}$





Multivariate Gaussian Distribution

$$\underline{x} = (x_1, x_2, \cdots\cdots x_n)$$

individually $\;\;\hookrightarrow\;\; x_i = N(\mu_i, \sigma_i^2)$

$$\frac{1}{\sqrt{2\pi}\,\sigma_i} e^{-\frac{|x_i-\mu_i|^2}{2\sigma_i^2}}$$

jointly $\;\; f_{x_i | (x_j, x_{j+1} \cdots)} = N(\mu, \sigma^2)$

det

$$f_{\underline{x}}(\underline{x}) = \frac{|C_{xx}|^{-1/2}}{(2\pi)^{n/2}} \exp\left(-\frac{(\underline{x}-\mu_A)^T C_{xx}^{-1} (\underline{x}-\mu_A)}{2}\right)$$

quadratic form

$$\underbrace{(\underline{x}-\underline{\mu}_x)^T}_{\text{row}} \left[C_{xx}^{-1} \right] \left[\underbrace{(\underline{x}-\underline{\mu}_x)}_{\text{column}} \right]$$

mean

correlation matrix

mean

$$[\;\;\;\;\;\;\;\;] \left[\;\right] = \text{scalar}$$

$$C_{xx} = \begin{array}{cc} & \begin{array}{cccc} \phantom{C_{11}} & j & & \end{array} \\ i & \begin{bmatrix} C_{11} & C_{12} & \cdots & \\ C_{21} & C_{22} & \cdots & \\ C_{31} & C_{ij} & \cdots & \\ & & & C_{nn} \end{bmatrix} \end{array}$$

$$C_{ij} = E\left[(X_i - \overset{\text{mean}}{\mu_i})(X_j - \overset{\text{mean}}{\mu_j})\right]$$

covariance

correlation $E(X_i X_j)$

X_j, X_i

uncorrelated: $\quad C_{ij} = 0 \quad j \neq i$

identically distributed: $\quad C_{ii} = \sigma^2$

$$C_{xx} = \sigma^2 I \longrightarrow \text{diagonal}$$

$$\underline{x}^T C_{xx}^{-1} \underline{x} = \frac{1}{\sigma^2} \cdot \underline{x}^T \cdot x = \frac{1}{\sigma^2} \cdot \sum_{i=1}^{n} x_i^2$$

$$f_{\underline{x}}(\underline{x}) = \frac{\sigma^{-\frac{n}{2}}}{\sqrt{2\pi}^{\,n}} \exp\left(-\frac{1}{\sigma^2} \sum_{i=1}^{n} x_i^2\right)$$

$$= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x_i^2}{\sigma^2}} = \prod_{i=1}^{n} N(0,\sigma^2)$$

then joint = product of independent RVs

$$\underline{X} = [\underline{y}, \underline{z}]$$ — arbitrary partitioning

$\underline{y}$: hidden, $\underline{z}$: known

$$f_{\underline{y}|\underline{z}}(\underline{y}|\underline{z}) = \text{work with}$$

$$\Downarrow$$

$$f_{\underline{y}|\underline{z}} \cdot f_{\underline{z}} = f_{\underline{X}} \qquad f_{\underline{y}|\underline{z}} = \frac{f_{\underline{X}}}{f_{\underline{z}}}$$

$$E(\underline{Y}|\underline{z}) = \int_{\underline{Y}} \underline{Y} \cdot f_{\underline{Y}|\underline{z}}(\underline{y}|\underline{z})$$

$$\iint_{\underline{y}} f_X \iint f_{y,z}$$

Central formula:

$$E(\underline{Y}|\underline{z}) = E[\underline{Y}] + C_{yz}\, C_{zz}^{-1} (\underline{z} - E(\underline{z}))$$

inverse computationally

$\cdot\ C_{\underline{z}\underline{z}} \;:\; E[(z_i - \mu_{z_i})(z_j - \mu_{z_j})]$

$\cdot\ C_{y|\underline{z}} \;:\; E((\underline{y} - \underline{\mu}_{y|z})(\underline{y} - \underline{\mu}_{y|z}) \mid \underline{z})$

$\underline{X}$ = jointly Gaussian $(\underline{y}, \underline{z})$

$$\underline{z} \longrightarrow \boxed{\text{Estimator}} \longrightarrow \hat{\underline{y}} = E[\underline{y} \mid \underline{z}]$$

$$= M \cdot (\underline{z} - \underline{\mu}_z)$$

$$\underline{z} \longrightarrow \underset{\substack{\uparrow \\ -\underline{\mu}_z}}{\oplus} \longrightarrow \boxed{\text{Estimator}} \qquad \hat{y} = M \cdot \underline{z}$$

Example: $\quad x_i = a + n_i$

$$\text{set } [\underline{y}, \underline{z}] = [\overset{y}{a}, \overset{z}{\underline{x}}]$$

$E[a \mid \underline{x}] \Rightarrow$ need $C_{\underline{z}\underline{z}} = C_{\underline{x}\underline{x}}$

$$E((X_i - \mu_i)(X_j - \mu_j)] \qquad a \sim N(\mu_A, \sigma_A^2)$$

$$= \underset{(a+n_i)(a+n_j)}{E[X_i X_j]} - \underset{(a+n_i)\mu_A}{E[X_i \mu_j]} - E[\mu_i X_j] + \mu_i \mu_j$$

$= E[a^2] + E[n_i n_j] - \cancel{\mu_A^2} - \cancel{\mu_A^2} + \cancel{\mu_A^2}$

$= \underbrace{E[a^2] - \cancel{\mu_A^2}}_{E[a^2] - (E[a])^2} + E[n_i n_j]$

$\delta_{ij} = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$

$C_{ij} = \sigma_A^2 + \delta_{ij} \sigma_2$

$$C_{\underline{xx}} = \begin{bmatrix} \sigma_A^2 + \sigma^2 & \sigma_A^2 & \cdots & \\ \sigma_?^2 & \sigma_A^2 + \sigma^2 & & \vdots \\ \sigma_?^2 & & \ddots & \\ \vdots & \cdots & & \end{bmatrix}$$

$[1\,1\,1\,1]$

$\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & \cdots \\ 1 & 1 & \cdots \\ \cdots \end{bmatrix}$

$C_{\underline{xx}} = \sigma^2 \cdot I + \sigma_A^2 \cdot 1 \cdot 1^T$

$C_{y|z} = C_{yy} - C_{yz} C_{zz}^{-1} C_{zy}$

$[a, x] \qquad = C_{aa} - C_{a\underline{x}} C_{xx}^{-1} C_{\underline{x}a}$

$\Big\} \; \cancel{S_{ij}}$

$E[(a-\mu_A)(a-\mu_A+n_i)]$

$C_{yz} = C_{a\cdot\underline{x}} = \big[E[a \cdot x_i] \cdot$

$= \big[E[a(a+n_i)] \big]$

$$x_i = a + n_i$$

$$E[a^2] + \cancel{E[a\, n_i]}$$

$$= [\sigma_A^2 \cdots\cdots \sigma_A^2]$$

$$C_{a\underline{x}} = \sigma_A^2 \cdot \underline{1}^T$$

$$E[a|\underline{x}] = \mu_A + C_{a\underline{x}} \cdot C_{\underline{x}\underline{x}}^{-1} (\underline{x} - \underline{\mu}_A)$$

$$= \mu_A + \sigma_A^2 \underline{1}^T \cdot [\sigma_2 I + \sigma_A \underline{1}\underline{1}^T]^{-1} \times (\underline{x} - \mu_A \cdot \underline{1})$$

prior information

measurement

$$E[a|\underline{x}] = \mu_A + \frac{\sigma_A^2}{\sigma^2} \underline{1}^T \left(I - \frac{\frac{\sigma_A^2}{\sigma^2} \underline{1}\underline{1}^T}{1 + n \frac{\sigma_A^2}{\sigma^2}} \right) (\underline{x} - \mu_A \underline{1})$$

$$\mu_A + \frac{\sigma_A^2}{\sigma^2} \underline{1}^T \underbrace{\left[\quad \right]}_{M} (\underline{x} - \mu_A \underline{1})$$

$$\mu_A - \frac{\sigma_A^2}{\sigma^2} M \cdot \underline{1} \cdot \mu_A + \frac{\sigma_A^2}{\sigma^2} \underline{1}^T M \cdot \underline{x}$$

$$\overbrace{\left(1 - \frac{\sigma_A^2}{\sigma^2} M \cdot \underline{1}\right) \cdot \mu_A} + \overbrace{\frac{\sigma_A^2}{\sigma^2} 1^T M \cdot \underline{X}} \longleftarrow$$

$$\hat{a} = \underbrace{(1-\alpha) \cdot \mu_A}_{\text{a priori part}} + \underbrace{\alpha \cdot \frac{1}{n}\sum_{i=1}^{n} X_i}_{\text{measurement}} \quad \longleftarrow$$

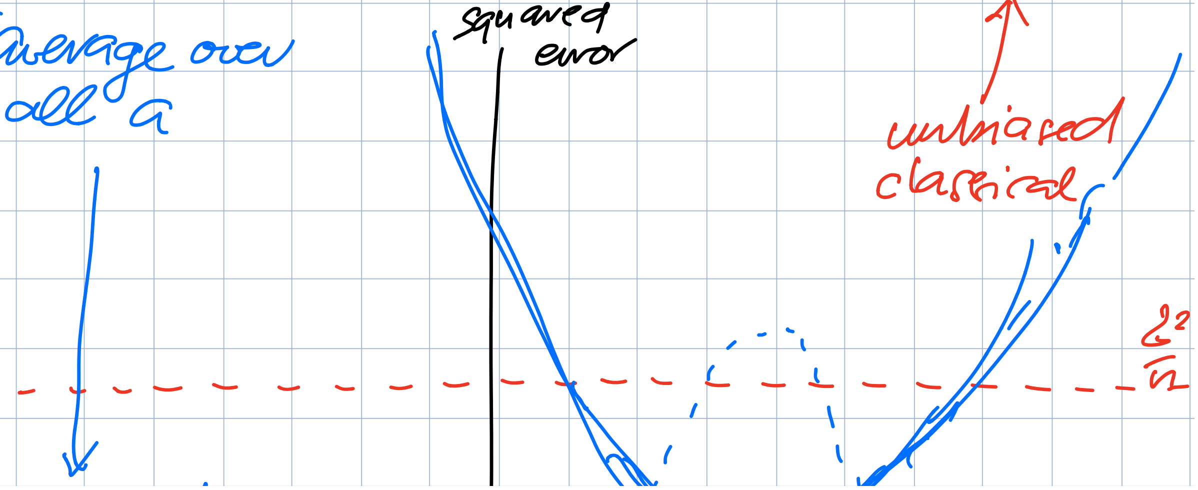
$$\alpha = \frac{\sigma_A^2}{\sigma_A^2 + \sigma^2} < 1 \qquad \text{Classical: } \hat{a} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

$\sigma_A^2 \to \infty \quad \alpha \to 1 \quad \longrightarrow$ Classical case

$\sigma_A^2 \to 0 \quad \alpha \to 0$

$$\overline{E((\hat{a} - a)^2)} = (\alpha - 1)^2 (a - \mu_A)^2 + \alpha^2 \frac{\sigma^2}{n}$$

average over all a



Sensitive to prior assumptions