least-squares

goal: optimal estimation through linear observations

ref: Stengel Ch 4.1

previously: use observations $y = Cx$ to estimate $\hat{x}$

– solve linear equations – applying observer

now: use noisy observations $z = Cx + \eta$ to estimate $\hat{x}$

where η is an external disturbance / measurement noise

- we want to choose $\hat{x}$ to minimize $\underbrace{\|x - \hat{x}\|}$, i.e. estimation error

but: we don't know x, so can't evaluate

- instead, we'll minimize $\|z - C\hat{x}\|$, i.e. observation error

$\longrightarrow$ solve NLP $\min_{\hat{x}} \|z - C\hat{x}\|_2^2 = (z - C\hat{x})^T (z - C\hat{x})$

$\longrightarrow$ solve NLP $\min_{\hat{x}} \|z - C\hat{x}\|_2^2 = (z - C\hat{x})'(z - C\hat{x})$

— let $J(\hat{x}) = \|z - C\hat{x}\|_2^2$, then minima are <u>stationary</u>:

$$D_{\hat{x}} J = C^T C \hat{x} - C^T z$$

$$= 0 \iff \hat{x} = (C^T C)^{-1} C^T z \leftarrow \text{"least squares" estimate}$$

$\uparrow$ assume rank $C^T C = \dim \hat{x}$ — necessary that # rows $C \geq$ # cols C

— to confirm $\hat{x}$ is minimum: $D^2_{\hat{x}} J(\hat{x}) = C^T C > 0$

$\longrightarrow$ confirm that $C^T C > 0$, i.e. all eigenvalues are positive

$\longrightarrow$ compute $\hat{x}$ for $z = Cx + \eta$, $x \in \mathbb{R}$, $z \in \mathbb{R}^k$, $C = \mathbb{1} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$
(this should be familiar)

◦ now consider weighted least-squares objective:

$$J(\hat{x}) = (z - C\hat{x})' S^{-1} (z - C\hat{x}), \quad S = S^T > 0$$

$\longrightarrow$ verify that $S^{-1} = (S^{-1})^T > 0$

$\longrightarrow$ solve NLP $\min_{\hat{x}} J(\hat{x})$

— $\hat{x} = (C^T S^{-1} C)^{-1} C^T S^{-1} z \longleftarrow$ agrees with <u>previous answer</u> when $S = I$

— could have changed coordinates to ... $S^{(-1)}$? and applied when $S = \mathbb{1}$

– could have changed coordinates to $\varsigma = S^{(-1)?}z$ and applied ?

<u>fact:</u> if $S = \mathrm{Cov}[\eta]$, $E[\eta] = 0$, then

$\hat{x}$ has the minimum (co-)variance out of all <u>unbiased estimators</u>

– note that $E[\hat{x}] = (C^T S^{-1} C)^{-1} C^T S^{-1} E[z]$,

$$E[z] = E[Cx] + \cancel{E[\eta]} \to 0$$

so $E[\hat{x}] = (C^T S^{-1} C)^{-1} (C^T S^{-1} C) E[x] = E[x]$

◦ suppose we estimate $\hat{x}_1 \in \mathbb{R}^d$ using batch of measurements $z_1 = C_1 x + \eta_1 \in \mathbb{R}^{k_1}$

and subsequently obtain new measurements $z_2 = C_2 x + \eta_2$

◦ then we want to minimize $J(\hat{x}) = \| z - C\hat{x} \|_S^2$

$$= (z - C\hat{x})^T S^{-1} (z - C\hat{x})$$

where $z = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$, $C = \begin{bmatrix} C_1 & 0 \\ 0 & C_2 \end{bmatrix}$ $S = \begin{bmatrix} S_1 & 0 \\ 0 & S_2 \end{bmatrix}$, $\hat{x} = \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix}$

* we want to solve for $\hat{x}_2$ in terms of $\hat{x}_1$ that minimizes J

→ solve $D_{\hat{x}_2} J = 0$ for $\hat{x}_2$ in terms of $\hat{x}_1$

- $\hat{x}_2 = (C_1^T S_1^{-1} C_1 + C_2^T S_2^{-1} C_2)^{-1} (C_1^T S_1^{-1} z_1 + C_2^T S_2^{-1} z_2)$

→ apply matrix inversion Lemma with $P_1^{-1} = C_1^T S_1^{-1} C_1$

$$(A + UHV)^{-1} = A^{-1} - A^{-1} U (H^{-1} + V A^{-1} U)^{-1} V A^{-1}$$

- $(\underbrace{C_1^T S_1^{-1} C_1}_{P_1^{-1}} + C_2^T S_2^{-1} C_2)^{-1} = P_1 - P_1 C_2^T (C_2 P_1 C_2^T + S_2)^{-1} C_2 P_1$

→ substitute into formula for $\hat{x}_2$

- $\hat{x}_2 = \hat{x}_1 + K_2 (z_2 - C_2 \hat{x}_1)$, $K_2 = P_1 C_2^T (C_2 P_1 C_2^T + S_2)^{-1}$