# 25 October 2018

**Review** $X$ a random vector with density $f_X(x_1, \dots, x_n)$

$(n \times 1)$ mean $= \mu =: E\{X\} := \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x f_X(x_1, \dots, x_n)\, dx_1 \cdots dx_n$

$(n \times n)$ Covariance $= \text{cov}\{X\} := E\{(X-\mu)\cdot(X-\mu)^T\} =: Q \geq 0$

$$Q_{ij} = [Q]_{ij} = E\{(X_i - \mu_i)(X_j - \mu_j)\}$$ covariance matrix

If $Q > 0$, $Q^{-1}$ called the information matrix

$(1 \times 1)$ Variance $= \text{var}\{X\} := E\{(X-\mu)^T(X-\mu)\}$

Recall: $A = n \times m$, $B = m \times n$ then

$$\text{tr}(A \cdot B) = \text{tr}(B \cdot A)$$

$$\therefore \underbrace{(X-\mu)^T(X-\mu)}_{1 \times 1} = \text{tr}\left((X-\mu)^T(X-\mu)\right) = \text{tr}\left((X-\mu)(X-\mu)^T\right)$$

Exercise: $\text{var}\{X\} = \text{tr}\, Q$

Idea: trace = a sum. Expectation = an integral



and we know "Integral of sum = sum of Integrals"

Matlab illustration of $Q$ diagonal vs $Q$ not diagonal $\left[ \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \mu, \; Q = \mathrm{cov}(X) \right]$

1) $Q = \begin{bmatrix} 10 & 0 \\ 0 & 1 \end{bmatrix}$

2) $Q = \begin{bmatrix} \frac{11}{2} & -9/2 \\ -9/2 & 11/2 \end{bmatrix} = O \begin{bmatrix} 10 & 0 \\ 0 & 1 \end{bmatrix} O^T$

$$O = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix}$$

3) $Y = OX$

---

Underdetermined: $y = Cx \quad x \in \mathbb{R}^n, \; y \in \mathbb{R}^m$

$S > 0, \; \|x\|^2 := x^T S x$, rows of $C$ linearly indep.

$\hat{x} := \arg\min\limits_{y = Cx} \|x\|^2 = S^{-1} C^T (C S^{-1} C^T)^{-1} y$

Overdetermined: $y = Cx + e, \; x \in \mathbb{R}^n, \; y \in \mathbb{R}^m$

$S > 0 \quad \|y\|^2 := y^T S y$, columns $C$ lin. indep.

$$\hat{x} := \underset{x\in\mathbb{R}^n}{\arg\min} \|\underbrace{y - Cx}_{e}\| = (C^T S C)^{-1} C^T S y$$

BLUE Best Linear Unbiased Estimator

(Goal: Find out how to choose the weight matrix in an overdetermined problem)

$$y = Cx + \varepsilon \qquad y \in \mathbb{R}^m, \; x \in \mathbb{R}^n, \; \varepsilon \in \mathbb{R}^m$$

Noise model $\mu = E\{\varepsilon\} = 0, \; \mathrm{cov}(\varepsilon) = Q > 0.$

$x$ is unknown constant. No noise model for $x$.

Vocabulary: one says $x$ is deterministic while $\varepsilon$ is stochastic

Seek $\hat{x} = Ky$, a linear estimator.

Assume: Columns of $C$ are linearly independent.

**Def.** $\hat{x}$ is an unbiased estimate if $E\{\hat{x}\} = x$, the value of the unknown constant.

$$
\begin{aligned}
x &= E\{\hat{x}\} \\
&= E\{Ky\} \\
&= E\{K(Cx+\varepsilon)\} \\
&= E\{KCx + K\varepsilon\} \\
&= E\{KCx\} + \underbrace{E\{K\varepsilon\}}_{\to 0} \\
&= KCx + 0 \\
&= KCx
\end{aligned}
$$

Unbiased $\iff$ $x = KCx$

Fact: $x = KCx \;\; \forall x \in \mathbb{R}^n \iff KC = I_{n\times n}$

**BEST**: Minimize variance of $\hat{x}$

Choose $K$ such that we minimize variance of $\hat{x}$ while constraining $K$ to give us an unbiased estimator.

$$\mathrm{var}(\hat{x}) = E\{(\hat{x}-x)^T(\hat{x}-x)\}$$

$$\begin{aligned}
\hat{x}-x &= Ky - x \\
&= K(Cx+\varepsilon) - x \\
&= \underbrace{KCx - x}_{0} + K\varepsilon \\
&= K\varepsilon
\end{aligned}$$

$$\begin{aligned}
\mathrm{var}(\hat{x}) &= \mathrm{var}(K\varepsilon) \\
&= E\{(K\varepsilon)^T(K\varepsilon)\} \\
&= E\{tr[(K\varepsilon)^T K\varepsilon]\} \\
&= E\{tr[(K\varepsilon)(K\varepsilon)^T]\} \\
&= E\{tr[K\varepsilon\varepsilon^T K^T]\} \\
&= tr(K\, E\{\varepsilon\varepsilon^T\}\, K^T) \\
&= tr\, KQK^T
\end{aligned}$$

!!!

**Problem** $\hat{x} = \hat{K}y$ seek $\hat{K}$ satisfying

$$\hat{K} = \arg\min_{KC=I} tr(KQK^T)$$

$\therefore \hat{K}^\top = [\hat{k}_1^\top | \cdots | \hat{k}_n^\top] = Q^{-1}C(C^\top Q^{-1}C)^{-1}$.
Therefore,

$$\underline{\hat{K} = (C^\top Q^{-1}C)^{-1}C^\top Q^{-1}}$$

Know

**Theorem:** Let $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$, $y = Cx + \epsilon$, $E\{\epsilon\} = 0$, $E\{\epsilon\epsilon^\top\} =: Q > 0$, and $\text{rank}(C) = n$. The Best Linear Unbiased Estimator (BLUE) is $\hat{x} = \hat{K}y$ where

$$\hat{K} = \left(C^\top Q^{-1}C\right)^{-1} C^\top Q^{-1}. \checkmark$$

Moreover, the covariance of the error is

$$E\{(\hat{x} - x)(\hat{x} - x)^\top\} = \left(C^\top Q^{-1}C\right)^{-1}. \checkmark$$

Derive the variance.

**Remark:** Error covariance computation is an exercise. Solution (from previous calculations)

$$\begin{aligned}
E\{(\hat{x} - x)(\hat{x} - x)^\top\} &= KQK^\top \\
&= \left(C^\top Q^{-1}C\right)^{-1} C^\top Q^{-1}QQ^{-1}C\left(C^\top Q^{-1}C\right)^{-1} \\
&= \left(C^\top Q^{-1}C\right)^{-1}\left[C^\top Q^{-1}C\right]\left(C^\top Q^{-1}C\right)^{-1} \\
&= \left(C^\top Q^{-1}C\right)^{-1}
\end{aligned}$$

<u>Indeed</u>

$$\begin{aligned}
\hat{x} - x &= Ky - x \\
&= KCx + K\epsilon - x \\
&= K\epsilon \text{ (because } KC = I) \\
\therefore E\{(\hat{x} - x)(\hat{x} - x)^\top\} &= E\{(K\epsilon)(K\epsilon)^\top\} \\
&= E\{K\epsilon\epsilon^\top K^\top\} \\
&= KQK^\top
\end{aligned}$$

Key: BLUE = Weighted Least Squares with

weight matrix = Information matrix = Inverse of the covariance of the noise!

Deriving the answer and relating it to underdetermined equations

$$K = \begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{bmatrix} \qquad K^T = \left[ k_1^T \,\middle|\, k_2^T \,\middle|\, \cdots \,\middle|\, k_n^T \right]$$

$$tr(KQK^T) = tr\left( \begin{bmatrix} k_1 \\ \hline k_2 \\ \hline \vdots \\ \hline k_n \end{bmatrix} Q \left[ k_1^T \middle| k_2^T \middle| \cdots \middle| k_n^T \right] \right)$$

$$= tr\left( \begin{bmatrix} k_1 \\ \hline k_2 \\ \hline \vdots \\ \hline k_n \end{bmatrix} \left[ Qk_1^T \middle| Qk_2^T \middle| \cdots \middle| Qk_n^T \right] \right)$$

$$= \sum_{i=1}^{n} k_i Q k_i^T = \sum_{i=1}^{n} \| k_i \|_Q^2$$

$$KC = I_{n\times n} \iff C^T K^T = I_{n\times n}$$

$$\iff C^T \left[ k_1^T \,|\, k_2^T \,|\, \cdots \,|\, k_n^T \right] =$$

$$= \left[ e^1 \,|\, e^2 \,|\, \cdots \,|\, e^n \right]$$

$$e^i = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \leftarrow \text{ith row}$$

$$KC = I_{n\times n} \iff \boxed{C^T k_i^T = e^i}$$

$1 \leq i \leq n$

Problem: Decomposes into $n$ separate optimization problems

$$\hat{k}_i^T = \underset{C^T k_i^T = e^i}{\arg\min} \|k_i^T\|_Q^2$$

Underdetermined Least Squares

$$\hat{k}_i^T = Q^{-1} C \left( C^T Q^{-1} C \right)^{-1} e_i$$

$$\hat{K}^T = \left[\hat{k}_1^T \middle| \cdots \middle| \hat{k}_n^T\right]$$

$$= Q^{-1} C \left(C^T Q^{-1} C\right)^{-1}$$

$$I = \left[e^1 \middle| e^2 \middle| \cdots \middle| e^n\right]$$