MINIMUM VARIANCE ESTIMATION: $\bar{w}$ PRIOR

meas: $\left\{ \tilde{y} = Hx + v \right.$ $\quad v \sim N(0, R)$ } ind

$\Rightarrow \left\{ x = \underset{\sim}{\hat{x}_a} + w \right.$ $\quad w \sim \underline{N(0, Q)}$

$\hat{x}_a = x + w$

Estimator: $M, N, n \rightarrow$ we pick.

$$\hat{x} = M\tilde{y} + N\hat{x}_a + n$$ } linear estimator

UNBIASED $E[\hat{x}] = x$

$$E[\hat{x}] = E[MHx + \cancel{Mv}^{\,0} + Nx + \cancel{Nw}^{\,0} + \cancel{n}] = x$$

set to 0

$$= E[(MH+N)x]$$

- $n = 0$
- $MH + N = I$ } $\rightarrow$ modification of $MH = I$

Covariance:

$$E[(\hat{x}-x)(\hat{x}-x)^T] = E((Mv+Nw)(Mv+Nw)^T)$$

where

$$\hat{x} = \underbrace{(MH+N)}_{I}x + Mv + Nw$$

$$= ME[vv^T]M^T + ME[\cancel{vw^T}]N^{0} + NE[ww^T]N^T + NE[\cancel{wv^T}]M^{T\,0}$$

$$= MRM^T + NQN^T$$

$$J = \frac{1}{2}\mathrm{Tr}\left(MRM^T + NQN^T\right)$$

$$\min_{M,N} \quad J = \frac{1}{2}\mathrm{Tr}\left(MRM^T + NQN^T\right)$$

$$\text{s.t.} \quad MH + N = I$$

SOLN:

meas prior

$$\hat{x} = \left(\underline{H^T R^{-1} H} + \underline{Q^{-1}}\right)^{-1}\left[\underline{H^T R^{-1}} \tilde{y} + \underline{Q^{-1} \hat{x}_a}\right]$$

$Q \to \infty$ more uncertain about priori

Limiting cases

$Q^{-1} \to 0$

$$\Rightarrow \hat{x} = \left(H^T R^{-1} H\right)^{-1} H^T R^{-1} \tilde{y}$$

$R \to \infty$ more uncertain about the measurements

$R^{-1} \to 0$

$$\hat{x} = \left(Q^{-1}\right)^{-1} Q^{-1} \hat{x}_a = \hat{x}_a$$

CRAMER-RAO BOUND:

for an estimator $\hat{x}(\tilde{y})$ or $\hat{x}(\tilde{y}, \hat{x}_a)$

unbiased : $E[\hat{x}] = x$ $\Leftarrow$

efficient : the covariance

$$E\left[(\hat{x}-x)(\hat{x}-x)^T\right]$$

is "as small/tight as possible."

Intuition: $\hat{x}(\tilde{y})$

$\tilde{y} = Hx + v$ $\quad v \sim N(0, R)$

how "good" or accurate $\hat{x}(\tilde{y})$ is should depend on R

$\hat{x} \doteq M\tilde{y} + n$

- unbiased $\Rightarrow$ $\underset{\text{multiple}}{MH = I}$, $n = 0$

- $E\left[(\hat{x}-x)(\hat{x}-x)^T\right] = MRM^T$

"pick M s.t. $MH = I$ & MRM^T

$Tr(MRM^T)$

as tight as possible

CRAMER-RAO BOUND:

Given distribution $p(\tilde{y}|x)$

unbiased estimator $\hat{x}(\tilde{y})$

s.t. $E[\hat{x}(\tilde{y})] = x$

then
$$\boxed{E\left[(\hat{x}-x)(\hat{x}-x)^T\right] \succeq F^{-1}}$$

where F is

Fischer Information Matrix.

$$\boxed{F = E\left[\underbrace{\left(\frac{\partial}{\partial x}\ln(p(\tilde{y}|x))\right)^T}_{\text{col}} \underbrace{\frac{\partial}{\partial x}\ln(p(\tilde{y}|x))}_{\text{row}}\right]}$$

$$= E\left[\frac{\partial^2}{\partial x \partial x^T}\ln[p(\tilde{y}|x)]\right]$$

needs to be proven

F^{-1}: gives us a lower bound on Covariance.

↳ ony depends on $p(\tilde{y}|x)$]

$\tilde{y} = Hx + v \Rightarrow p(\tilde{y}|x) = \left(\frac{1}{*}\right) e^{-\frac{1}{2}(\tilde{y}-Hx)^T R^{-1} (\tilde{y}-Hx)}$

$v \sim N(0, R)$

↳ $v = \tilde{y} - Hx \qquad p(v) = \left(\frac{1}{*}\right) e^{-\frac{1}{2} v^T R^{-1} v}$

$$\boxed{F = E\left[\left(\frac{\partial}{\partial x}\ln(p(\tilde{y}|x))\right)^T \frac{\partial}{\partial x}\ln(p(\tilde{y}|x))\right]}$$

$$\frac{\partial}{\partial x}\ln p = \boxed{\frac{1}{p}\frac{\partial p}{\partial x}} \Leftarrow \quad *$$

$F : \frac{1}{p}\frac{\partial p}{\partial x}$

F	BIGGER	SMALLER
	$\frac{\partial p}{\partial x}$ ↑ increasing	$\frac{\partial p}{\partial x}$ ↓ decreasing
	p ↓ decreasing	p ↑ increasing

F^{-1}

LOWER BOUND UP

$\frac{\partial p}{\partial x} \downarrow$ decreases

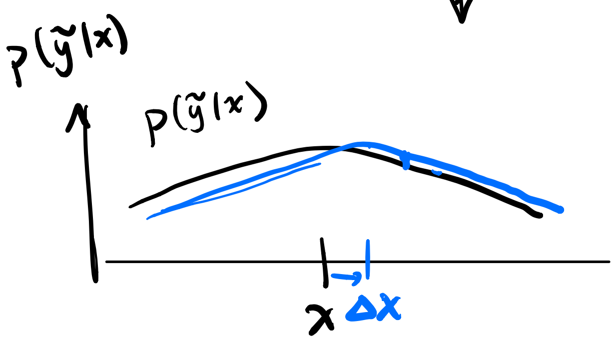
$\rightarrow p \uparrow$ increases

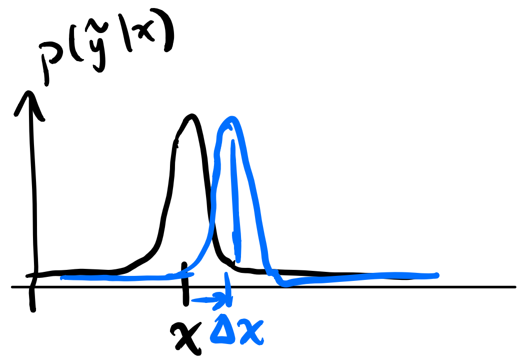
LOWER BOUND DOWN

$\frac{\partial p}{\partial x} \uparrow$ increases

$\rightarrow p \downarrow$ decreases

Ex. $\tilde{y} = x + v \quad | \quad x \in \mathbb{R} \quad v \sim \mathcal{N}(0, \sigma^2)$





PROOF:

$p(\tilde{y}|x)$

$\hat{x}(\tilde{y})$: unbiased $\quad E[\hat{x}] = x \qquad$ 1 unbiased

$$\rightarrow E[\hat{x} - x] = \int_{-\infty}^{\infty} (\hat{x} - x)\, p(\tilde{y}|x)\, d\tilde{y} = 0$$

take $\frac{d}{dx}$:
$$\frac{d}{dx} \int_{-\infty}^{\infty} (\hat{x} - x)\, p(\tilde{y}|x)\, d\tilde{y} = 0$$

$$\int_{-\infty}^{\infty} \frac{d}{dx} \Big(\underline{(\hat{x} - x)}\, \underline{p(\tilde{y}|x)} \Big)\, d\tilde{y} = 0$$

$$\int_{-\infty}^{\infty}(\hat{x}-x)\frac{d}{dx}p(\tilde{y}|x)\,d\tilde{y} - \int_{-\infty}^{\infty} I\,p(\tilde{y}|x)\,d\tilde{y} = 0$$

$$[\hat{x}-x]\ [\ \_\_\_\_\ ] \qquad I$$

② identity

$$\frac{dp}{dx} = \frac{\partial}{\partial x}\ln(p)\,p \qquad \frac{\partial}{\partial x}\ln p = \frac{1}{p}\frac{\partial p}{\partial x}$$

$$\int_{-\infty}^{\infty}(\hat{x}-x)\frac{\partial}{\partial x}\ln(p)\,p(\tilde{y}|x)\,d\tilde{y} = I$$

$$\Rightarrow \int_{-\infty}^{\infty} a b^T\,d\tilde{y} = I$$

- $a = (\hat{x}-x)\,p(\tilde{y}|x)^{1/2}$
- $b = \frac{\partial}{\partial x}\ln(p)\,p(\tilde{y}|x)^{1/2}$

$p(\tilde{y}|x)^{1/2}$: coordinate change on function space

$$P := E[(\hat{x}-x)(\hat{x}-x)^T] = \int_{-\infty}^{\infty} a a^T\,d\tilde{y}$$

$$F := E\left[\left(\frac{\partial}{\partial x}\ln(p)\right)^T\frac{\partial}{\partial x}\ln(p)\right] = \int_{-\infty}^{\infty} b b^T\,d\tilde{y}$$

$\alpha, \beta \in \mathbb{R}^n$

$$\alpha^T\int_{-\infty}^{\infty} a b^T\,d\tilde{y}\,\beta = \alpha^T\beta$$

③ Cauchy Schwartz

for functions

$$(\alpha^T\beta)^2 = \left(\int_{-\infty}^{\infty} \alpha^T a b^T \beta \, d\tilde{y}\right)^2 \leq \int_{-\infty}^{\infty} \alpha^T a a^T \alpha \, d\tilde{y} \int_{-\infty}^{\infty} \beta^T b b^T \beta \, d\tilde{y}$$

$$(\langle f, g \rangle)^2 \leq |f|^2 |g|^2 = \langle f, f \rangle \langle g, g \rangle$$

$$|f|^2 |g|^2 \cos\theta^2 \leq |f|^2 |g|^2$$

$$f = \alpha^T a \qquad g = \beta^T b \qquad \langle f, g \rangle = \int_{-\infty}^{\infty} f(\tilde{y}) g(\tilde{y}) \, d\tilde{y}$$

$$(\alpha^T\beta)^2 \leq \alpha^T \int_{-\infty}^{\infty} a a^T d\tilde{y} \, \alpha \;\; \beta^T \int_{-\infty}^{\infty} b b^T d\tilde{y} \, \beta$$

$$\leq \alpha^T P \alpha \;\; \beta^T F \beta$$

4 take $\beta = F^{-1}\alpha$

$$(\alpha^T F^{-1} \alpha)^2 \leq \alpha^T P \alpha \;\; \alpha^T F^{-1} \alpha$$

$$(\alpha^T F^{-1} \alpha)(\alpha^T F^{-1} \alpha - \alpha^T P \alpha) \leq 0$$

$$+$$

$$F \succ 0 \qquad \alpha^T (F^{-1} - P) \alpha \leq 0$$

$$\Rightarrow \boxed{P \succeq F^{-1}}$$

Ex. $\tilde{y} = \underline{Hx + v} \quad v \sim N(0, R)$

$$p(\tilde{y}|x) = \left(\frac{\ }{\ }\right) e^{-\frac{1}{2}(\tilde{y}-Hx)^T R^{-1}(\tilde{y}-Hx)}$$

$$\underline{F} = E\left[\frac{\partial}{\partial x}\ln(p)^T \frac{\partial}{\partial x}\ln(p)\right]$$

$$\ln(p) = -\frac{1}{2}(\tilde{y}-Hx)^T R^{-1}(\tilde{y}-Hx) + \underline{\text{const}}$$

$$\frac{\partial}{\partial x}\ln(p) = +\ (\tilde{y}-Hx)^T R^{-1} H$$

$$= \tilde{y}^T R^{-1} H - x^T H^T R^{-1} H$$

$$\tilde{y} = Hx + v$$

$$= v^T R^{-1} H + \cancel{x^T H^T R^{-1} H} - \cancel{x^T H^T R^{-1} H}$$

$$E\left[\frac{\partial}{\partial x}\ln p^T \frac{\partial}{\partial x}\ln p\right] = E\left[H^T R^{-1} v v^T R^{-1} H\right]$$

$$= H^T R^{-1} \underbrace{E[vv^T]}_{R} R^{-1} H$$

$$F = \underline{H^T R^{-1} H}$$

$$\underline{F^{-1}} = \underline{(H^T R^{-1} H)^{-1}} \leftarrow$$ lower bound on covariance

is the estimator $\hat{x} = (H^T R^{-1} H)^{-1} H^T R^{-1} \tilde{y}$

$$\hat{x} = (H^T R^{-1} H)^{-1} H^T R^{-1} (Hx + v)$$

$$= x + (H^T R^{-1} H)^{-1} H^T R^{-1} v$$

$$E[(\hat{x} - x)(\hat{x} - x)^T] = E[(H^T R^{-1} H)^{-1} H^T R^{-1} v v^T R^{-1} H (H^T R^{-1} H)^{-1}]$$

$$= (H^T R^{-1} H)^{-1} H^T \cancel{R^{-1} H} (\cancel{H^T R^{-1}} H)^{-1}$$

$$= (H^T R^{-1} H)^{-1}$$

Summary:

for model $\tilde{y} = Hx + v$, $v \sim N(0, R)$

the estimator $\hat{x} = (H^T R^{-1} H)^{-1} H^T R \tilde{y}$
is efficient.

its covariance is as small as possible.

Picture

$\tilde{y} = Hx + v$