## MINIMUM VARIANCE ESTIMATION

with an a priori state estimate

given

meas, given parameter

Meas: $\tilde{y} = Hx + v$ $\qquad v \sim N(0,R) \qquad R = E[vv^T]$

apriori est.: $\hat{x}_a = x + w$ $\qquad w \sim N(0,Q) \qquad Q = E[ww^T]$

$w = 0$

given

assuming a linear model

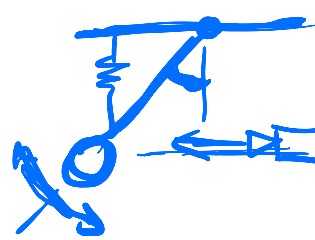
$$\hat{x}(\tilde{y},\hat{x}_a) = M\tilde{y} + N\hat{x}_a + n \quad \Leftarrow$$

Requirements:

- unbiased $E(\hat{x}) = E[x]$ (if noise is 0, $\hat{x} = x$)

$$E(\hat{x}) = MHx + M\cancel{v}^{0} + Nx + N\cancel{w}^{0} + n = E[x]$$

$$= \underbrace{(MH+N)}_{I}x + \underbrace{n}_{0} = E[x]$$



constraints:

$$MH + N = I \qquad n = 0$$

$$\Rightarrow \hat{x} = M\tilde{y} + N\hat{x}_a \quad \Leftarrow$$

Cost: $J = \frac{1}{2}Tr\left(E\left((\hat{x}-x)(\hat{x}-x)^T\right)\right) = \frac{1}{2}Tr\left(E(\hat{x}\hat{x}^T) - E(xx^T)\right)$

$$E(\hat{x}\hat{x}^T) = E\left(\left[\underbrace{(MH+N)}_{I}x + Mv + Nw\right]\left[\underbrace{(MH+N)}_{I}x + Mv + Nw\right]^T\right)$$

unbiased $MH + N = I$

$$E(\hat{x}\hat{x})^T = E(xx^T) + \underbrace{\cancel{E(xv^TM^T)}}_{0} + \underbrace{\cancel{E(xw^TN^T)}}_{0} +$$
$$\underbrace{\cancel{E(Mvx^T)}}_{0} + E(Mvv^TM^T) + \underbrace{\cancel{E(Mvw^TN^T)}}_{0} +$$
$$\underbrace{\cancel{E(Nwx^T)}}_{0} + \underbrace{\cancel{E(Nwv^TM^T)}}_{0} + E(Nww^TN^T)$$

for independent random variables $E(v) = 0$, $E(w) = 0$

$$\underline{E(vw^T)} = 0 \qquad E[vw^T] = \int vw^T \underbrace{p(v,w)}_{p(v)\,p(w)}\, dv\, dw$$

covariance between $v$ and $w$ $= 0$.

$$= \underbrace{\int v\, p(v)\, dv}_{0} \underbrace{\int w\, p(w)\, dw}_{0}$$

$$J = \frac{1}{2}\mathrm{Tr}\Big(\cancel{E(xx^T)} - \cancel{E(xx^T)} + M\underbrace{E(vv^T)}_{R}M^T + N\underbrace{E(ww^T)}_{Q}N^T\Big)$$

$$\min_{M,N} \; J = \frac{1}{2}\mathrm{Tr}(MRM^T + NQN^T) \qquad \text{Matrix dot product.}$$
$$\text{s.t.} \quad MH + N = I \qquad \langle \Lambda^T, I - MH - N \rangle$$

Lagrangian:

$$\mathcal{L}(M,N,\Lambda) = \frac{1}{2}\mathrm{Tr}\big((MRM^T) + NQN^T\big) + \mathrm{Tr}\big(\Lambda(I - MH - N)\big)$$

Optimality:

$$\frac{\partial \mathcal{L}}{\partial M} = \underline{MR - \Lambda^T H^T = 0} \qquad \frac{\partial \mathcal{L}}{\partial N} = NQ \underline{- \Lambda^T = 0} \qquad \frac{\partial \mathcal{L}}{\partial \Lambda} = I - MH - N = 0$$

Solving...

$$M = \Lambda^T H^T R^{-1}, \quad N = \Lambda^T Q^{-1},$$

$$\Rightarrow I - MH - N = I - \Lambda^T H^T R^{-1} H - \Lambda^T Q^{-1} = 0$$

$$\Rightarrow \Lambda^T = \left(H^T R^{-1} H + Q^{-1}\right)^{-1}$$

$$\Rightarrow M = \left(H^T R^{-1} H + Q^{-1}\right)^{-1} H^T R^{-1}$$

$$N = \left(H^T R^{-1} H + Q^{-1}\right)^{-1} Q^{-1}$$

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

before without apriori estimate

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

$Q^{-1}$ $\quad$ $Q^{-1}\hat{x}_a$

like weighted LS w weights $R^{-1}$

$R$ : measurement noise covariance "how much noise"

$Q$ : error the a prior estimate "how much was your initial estimate off"

$R \quad Q \nearrow \infty \Rightarrow$ original estimate was terrible

$$\Rightarrow Q^{-1} \rightarrow 0 \quad \Rightarrow \hat{x} = \left(H^T R H\right)^{-1} H^T R^{-1} \tilde{y}$$

ignoring $\hat{x}_a$

loose spring

$R \nearrow^{\infty}$ $Q$ $\Rightarrow$ tons of noise

$\Rightarrow R^{-1} \to 0$ $\Rightarrow \hat{x} = (Q^{-1})^{-1} Q^{-1} \hat{x}_a$ ignoring measurements

$Q Q^{-1}$ (because lots of noise) bad camera

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

Batch estimation:

$\hat{x}_a = \hat{x}_k$

$Q = P_k$

$\tilde{y}_{k+1}$

$H_{k+1}$

$R^{-1} = W_{k+1}$

treat batch estimation as an apriori estimate

LIMITS ON ESTIMATION ACCURACY:

CRAMER-RAO BOUND

$$P := E[(\hat{x}-x)(\hat{x}-x)^T] \geq ?$$

Covariance of estimator error.

$$P = E[(\hat{x}-x)(\hat{x}-x)^T] \geq F^{-1}$$

F: Fischer Information Matrix

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

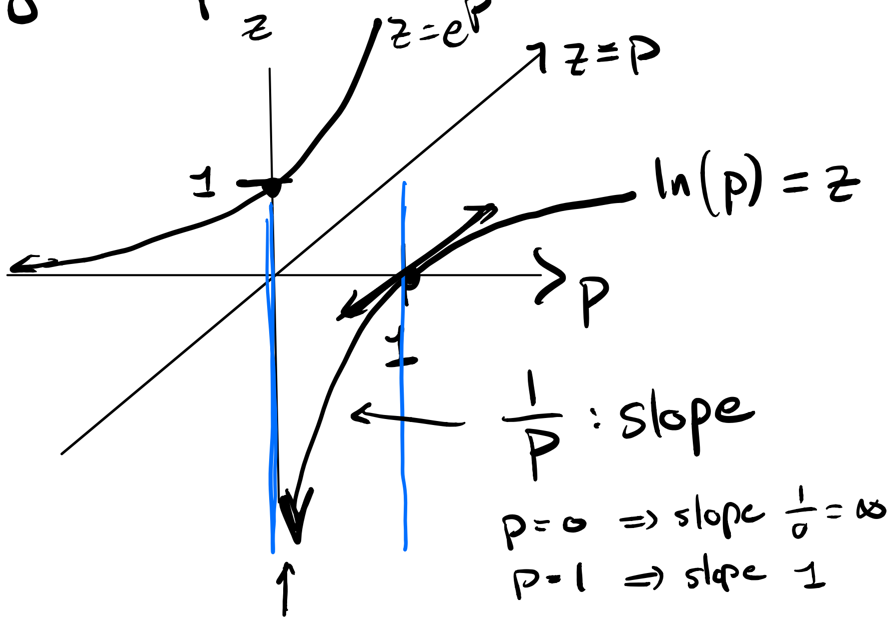
$\ln(p)$ $p \in [0,1]$

$p(\tilde{y}|x)$: density function of measurements given parameters $x$.

$\ln(\cdot)$:

always negative

$-\infty \to 0$




$$\frac{\partial}{\partial x}\ln\big(p(\tilde{y}|x)\big) = \frac{d\ln}{\partial p}\frac{\partial p}{\partial x} = \frac{1}{p}\frac{\partial p(\tilde{y}|x)}{\partial x} \quad \star$$

$$F := \int \frac{1}{p(\tilde{y}|x)}\frac{\partial p}{\partial x}\frac{\partial p}{\partial x}^T \frac{1}{\cancel{p(\tilde{y}|x)}}\cancel{p(\tilde{y}|x)}\, d\tilde{y}$$

$$= \int \underline{\frac{\partial p}{\partial x}}\frac{\partial p}{\partial x}^T \frac{1}{\underline{p(\tilde{y}|x)}}\, d\tilde{y}$$

F:

**big**

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

$p(\tilde{y}|x) \searrow 0$ small density

**small**

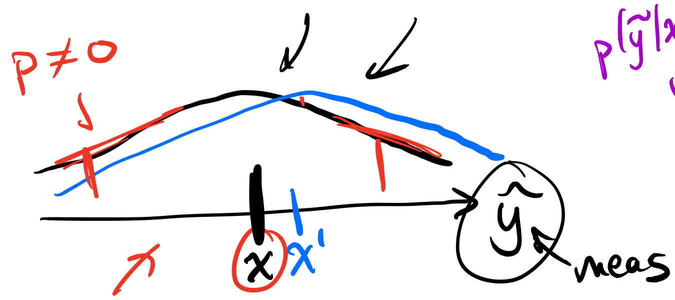
$\frac{\partial p}{\partial x} \searrow$ flat

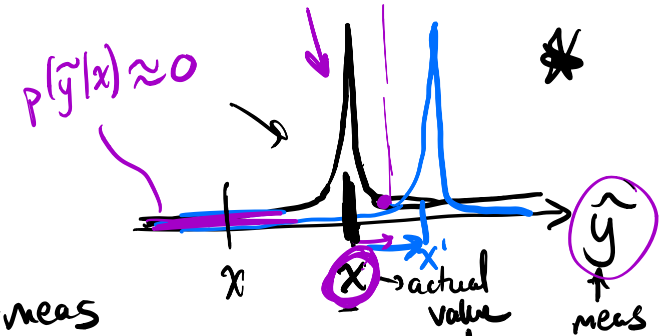
$p(\tilde{y}|x) \nearrow 1$ more density

$F^{-1}$:

| | big | small |
|---|---|---|
| | $\frac{\partial p}{\partial x} \searrow$ flat | $\frac{\partial p}{\partial x} \nearrow$ steep |
| | $p(\tilde{y}\|x) \nearrow$ more mass | $p(\tilde{y}\|x) \searrow 0$ less mass |

1D example:




$$P = E[(\hat{x}-x)(\hat{x}-x)^T] \geq F^{-1}$$

$\hat{x}(\tilde{y})$ if $p(\tilde{y}|x)$ represents more accurate measurements than $F^{-1} \to 0$

Notes:

- only good for unbiased estimators
- an efficient estimator $\hat{x}(\tilde{y})$ · estimator you pick

s.t. $$P = E((\hat{x}-x)(\hat{x}-x)^T) = F^{-1}$$

DERIVATION:

$$\int_{-\infty}^{\infty} p(\tilde{y}|x)\, d\tilde{y} = 1.$$

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

Note: can pull derivative into integral because the boundaries are const.
- Leibnitz integral rule

unbiased estimator cond:

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

$\boxed{\frac{\partial}{\partial x}:}$ $$\int_{-\infty}^{\infty} (\hat{x}-x)\, \underbrace{\frac{dp(\tilde{y}|x)}{\partial x}}_{\text{row vector}}\, d\tilde{y} - I = 0$$

ln properties:

$$\left[\ \frac{dp}{dx} = \frac{\partial}{\partial x} \underbrace{\ln(p(\tilde{y}|x))}_{\frac{1}{p}\frac{dp}{dx}}\ \cancel{p}(\tilde{y}|x) \right.$$

$$\frac{d}{dx}\ln(p) = \frac{1}{p}\frac{dp}{dx}$$

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

$$a = p(\tilde{y}|x)^{1/2}(\hat{x} - x)$$

$$b = p(\tilde{y}|x)^{1/2} \frac{d}{dx} \ln(p(\tilde{y}|x))$$

---

$$P = \int_{-\infty}^{\infty} (a a^T) \, d\tilde{y}$$

$$F = \int_{-\infty}^{\infty} b b^T \, d\tilde{y}$$

vector row vector $\forall \alpha, \ \beta = F^{-1}\alpha$

$$\alpha^T \left[ I = \int_{-\infty}^{\infty} a b^T \, d\tilde{y} \right] \beta \Leftarrow$$

$$\alpha^T \beta = \int_{-\infty}^{\infty} \underbrace{\alpha^T a}_{\text{scalars}} \; \underbrace{b^T \beta}_{\text{scalars}} \; d\tilde{y} \Rightarrow \left( \int_{-\infty}^{\infty} \alpha^T (a b^T) \beta \, d\tilde{y} \right)^2 \le \int_{-\infty}^{\infty} \underbrace{(\alpha^T a)^2}_{\alpha^T a a^T \alpha} d\tilde{y} \int_{-\infty}^{\infty} \underbrace{(\beta^T b)^2}_{\beta^T b b^T \beta} d\tilde{y}$$

Cauchy Schwartz inequality ← infinite dim version

Fin dim $\quad x \cdot y = x^T y \quad (x^T y)^2 \le |x|^2 |y|^2$ Cauchy Schwarz fin dim

Inf dim: $\langle f(\cdot), g(\cdot) \rangle$

fin dim vector

$[x_1 \; \cdots \; x_n]$

index $i \Rightarrow x_i$

inf dim "vector"


"index": $t \Rightarrow f(t)$

$$x^T y = \sum_i x_i y_i \longrightarrow \langle f(t), g(t)\rangle = \int f(t) g(t)\, dt \qquad \circledast$$

cauchy schwarz for functions...

for $f(\cdot)$ $g(\cdot)$

$$\left(\int f(t) g(t)\, dt\right)^2 \le \int f(t)^2\, dt \int g(t)^2\, dt$$

forget

$f: t \mapsto \mathbb{R}^n \qquad g: t \mapsto \mathbb{R}^n$

$\langle f(t), g(t)\rangle \triangleq \int \sum_i f_i(t) g_i(t)\, dt$

$f: t \mapsto \mathbb{R}^{n\times m} \qquad g: t \mapsto \mathbb{R}^{n\times m}$

$\langle f(t), g(t)\rangle = \int Tr\left(f^T(t) g(t)\right) dt$

$$\alpha^T \beta = \int_{-\infty}^{\infty} \underbrace{\alpha^T a}_{\text{scalars}} \; \underbrace{b^T \beta}_{\text{scalars}} \; d\tilde{y} \Rightarrow$$

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

treating

$f = \alpha^T a \qquad g = b^T \beta$

$$\Big(\alpha^T \underbrace{\int_{-\infty}^{\infty} a b^T d\tilde{y}}_{I}\, \beta\Big)^2 \le \alpha^T \underbrace{\int_{-\infty}^{\infty} a a^T d\tilde{y}}_{P}\, \alpha\, \beta^T \underbrace{\int_{-\infty}^{\infty} b b^T d\tilde{y}}_{F}\, \beta$$

$$(\alpha^T \beta)^2 \le \alpha^T P \alpha\, \beta^T F \beta \qquad \underline{\beta = F^{-1}\alpha}$$

$$(\alpha^T F^{-1} \alpha)^2 \le \alpha^T P \alpha\, \alpha^T F^{-1} \alpha$$

$$\frac{1}{\underbrace{\alpha^T F^{-1} \alpha}} \quad (\underbrace{\alpha^T F^{-1} \alpha})(\alpha^T P \alpha - \alpha^T F^{-1} \alpha) \ge 0 \qquad \frac{1}{\alpha^T F^{-1} \alpha}$$

$F^{-1}$ is $>0$ pos def

$\alpha^T F^{-1} \alpha > 0$

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

$P - F^{-1}$ is pos semi def.

$$\boxed{P \ge F^{-1}}$$

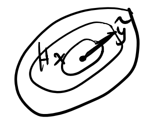
Ex. what is $F$ for a Gaussian?

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

$\Rightarrow v = \tilde{y} - \underline{Hx}$

$$p(\tilde{y}|x) \sim e^{-\frac{1}{2} v^T R^{-1} v}$$

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




F:

$$\ln p(\tilde{y}|x) = -\frac{1}{2}(\tilde{y}-Hx)^T R^{-1}(\tilde{y}-Hx) - \frac{m}{2}\ln(2\pi) - \frac{1}{2}\det(R)$$

$$\frac{d}{dx}\ln p(\tilde{y}|x) = \frac{\partial}{\partial v}\left(-\frac{1}{2}v^T R^{-1} v\right)\Bigg|_{v=\tilde{y}-Hx} \frac{\partial v}{\partial x}$$

$$= 2\left(-\frac{1}{2}(\tilde{y}-Hx)^T R^{-1}\right)(-H)$$

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

$$F = E\left(\frac{d}{dx}\ln p \, \frac{d}{dx}\ln p^T\right)$$

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

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

$$= E\left(\left(\cancel{H^T R^{-1} H x} - \cancel{H^T R^{-1} H x} - H^T R^{-1} v\right)\left(\quad\right)^T\right)$$

$$= E\left(+ H^T R^{-1} v v^T R^{-1} H\right) = H^T R^{-1} \underbrace{E(vv^T)}_{R} R^{-1} H$$

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

$$E\left((\hat{x}-x)(\hat{x}-x)^T\right) = P \geq F^{-1} = (H^T R^{-1} H)^{-1}$$

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

$$P = E\left((\hat{x}-x)(\hat{x}-x)^T\right) = (H^T R^{-1} H)^{-1}$$

$$\boxed{\hat{x} = (H^T \underline{R}^{-1} H)^{-1} H^T \underline{R}^{-1} \tilde{y}} \Leftarrow$$ an efficient/unbiased estimator for linear meas model

**Review:**

meas model: ex linear

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

$v = \tilde{y} - Hx$

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

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

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

find an estimator: $\hat{x}(\tilde{y})$

WANT:

- $\boxed{E_{p(\tilde{y}|x)}\left[\hat{x}(\tilde{y}) - x\right] = 0}$

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

for example $\hat{x} = (H^T R^{-1} H)^{-1} H^T R^{-1} \tilde{y}$

$$E(\hat{x}) = E\left[(H^T R^{-1} H)^{-1} H^T R^{-1} (H \underbrace{x + v}_{\tilde{y}})\right] = E(x) + E(- \overset{0}{\cancel{v}})$$

$$E(\hat{x}) = E(x)$$

- $E\left((\hat{x}-x)(\hat{x}-x)^T\right) = P = F^{-1}$

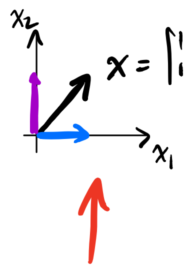
$\hat{x}(\tilde{y})$ efficient

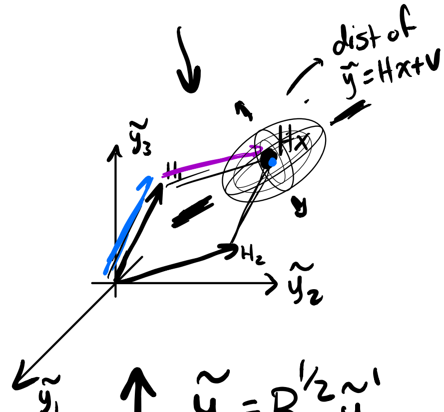
$\tilde{y} = Hx + v$ $\quad H \in \mathbb{R}^{3\times 2}$, $\quad H = [H_1 H_2]$

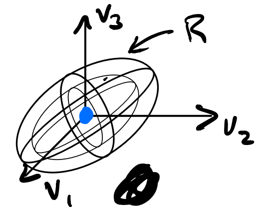
$v \sim N(0, R)$

$v \in \mathbb{R}^3$

$x \in \mathbb{R}^2$




$\tilde{y} = R^{1/2}\tilde{y}'$

$\tilde{y}, H, v \xrightarrow{R^{-1/2}} \tilde{y}', H', v'$

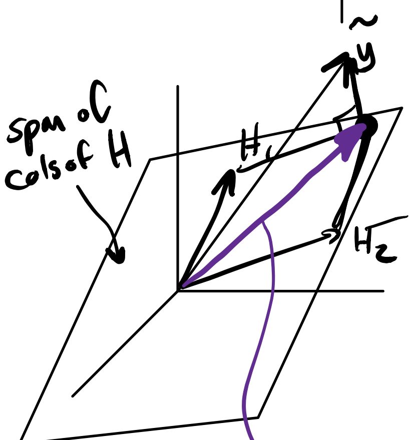
$\tilde{y}' = R^{-1/2}\tilde{y}$

$v' = R^{-1/2}v \quad E(v') = 0$

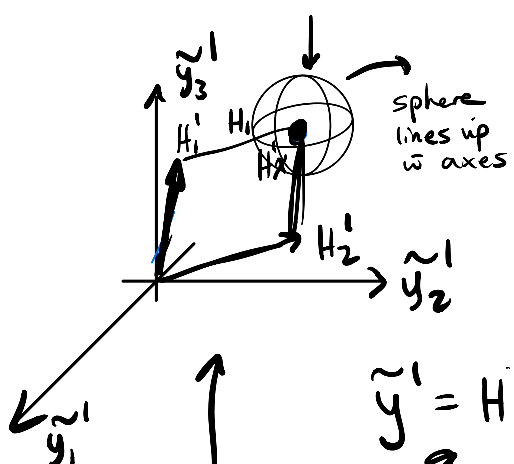
$E(v'v'^T) = R^{-1/2}\underbrace{E vv^T}_{R} R^{-1/2}$

$= I$

$v' \sim N(0, I)$




$\underbrace{H(H^TH)^{-1}H^T\tilde{y}}_{x}$




sphere lines up w/ axes

coord trans on $\tilde{y}$ based on noise

do LS in this space

$\tilde{y}' = H'x + v'$

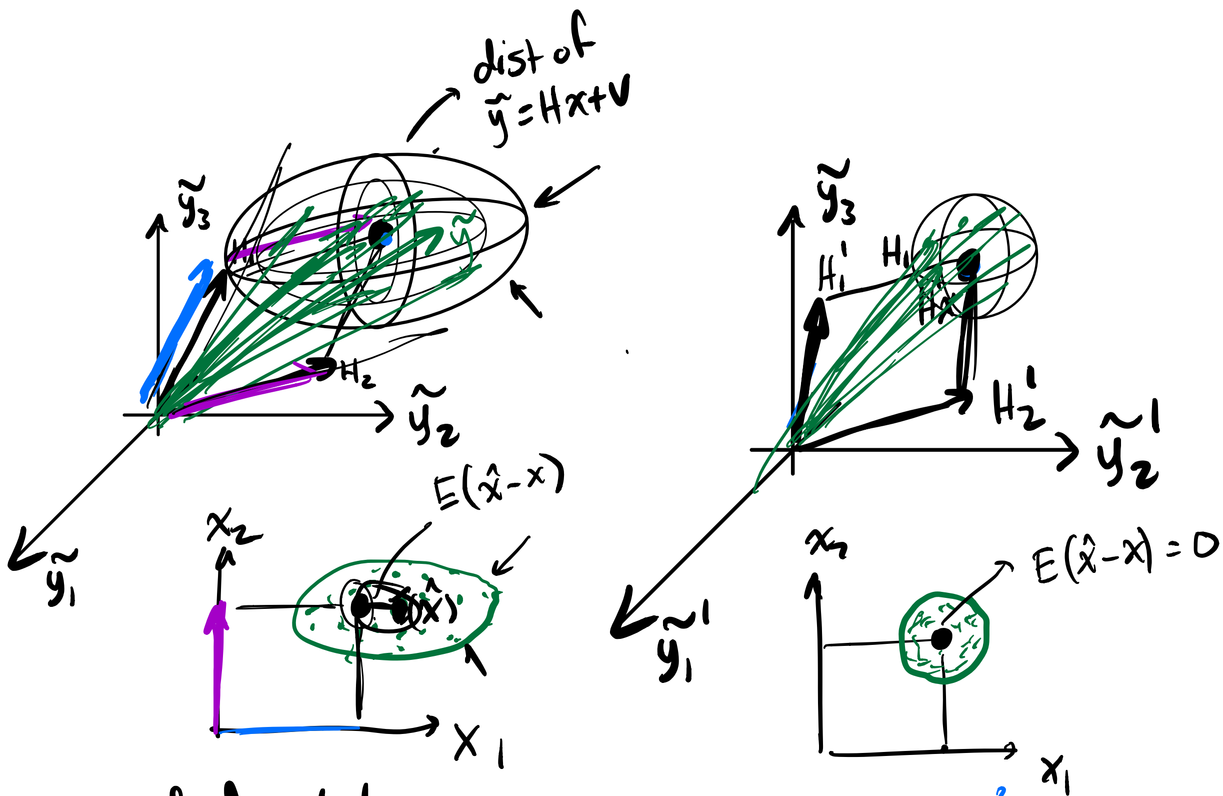
$v' \sim N(0, I)$

$\tilde{y}' = H'x + v'$

$x = (H'^T H')^{-1} H'^T \tilde{y}'$

$H' = R^{-1/2}H$

$x = (H^T R^{-1/2} R^{-1/2} H)^{-1} H^T R^{-1/2} R^{-1/2}\tilde{y} = (H^T R^{-1} H)^{-1} H^T R^{-1}\tilde{y}$




## Expected Values

Constants come out

Expected value is linear

$E(Mx) = M\,E(x)$

$E(c^T x) = c^T E(x)$

$E(xc^T) = E(x)c^T$

$E(x+y) = E(x) + E(y)$

$E((x+b)(x+b)^T)$

$E(xx^T + bx^T + xb^T + bb^T)$

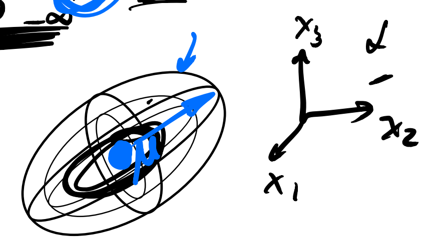
$E(xx^T) + E(bx^T) + E(xb^T) + E(bb^T)$

$p(x)$

$x^2,\ (x-\mu)^2$

$$\mu = E(x) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x\, p(x)\, dx_1 \cdots dx_n$$

integrate over all positions




$x \in \{1, 2, 3, 4, 5, 6\}$

$p(x) = 1/6 \quad \forall x$

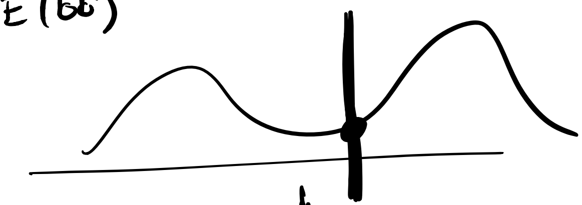
$$E(x) = \sum_{x=1}^{6} x \frac{1}{6}$$

sum over all dice possibilities

$= 1\frac{1}{6} + \cdots \quad 6\frac{1}{6}$

$= \frac{21}{6} = 3.5$

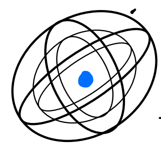
$E(xx^T) + bE(x)^T + E(x)b^T + E(bb^T)$

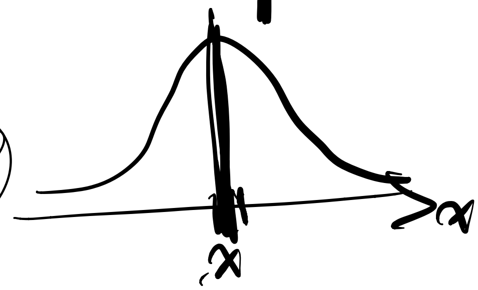


CoVariance

$\mu = E(x)$

$\Sigma = E((x-\mu)(x-\mu)^T)$






$E((x-\mu)^T(x-\mu))$

↓

average distance from the mean²

vector → $\begin{bmatrix}\mu_1\\ \vdots \\ \mu_n\end{bmatrix} = E(x) = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} \begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix} p(x_1 \ldots x_n)\, dx_1 \ldots dx_n$

$Tr(\Sigma) = E(Tr(x-\mu)(x-\mu)^T)$

$= E((x-\mu)^T(x-\mu))$

$\begin{bmatrix}x_1\\ x_n\end{bmatrix} p(x_1 \ldots x_n)$