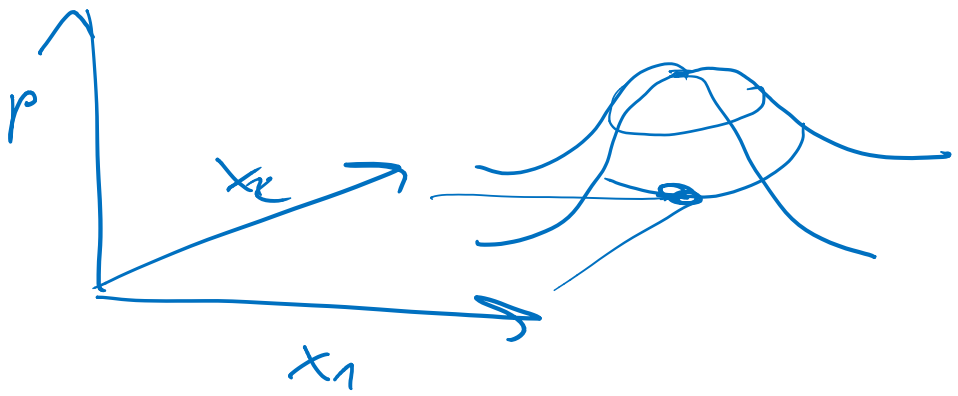
MULTIDIMENSIONAL NORMAL DISTR (GAUSSIAN)

SINGLE $\quad p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$

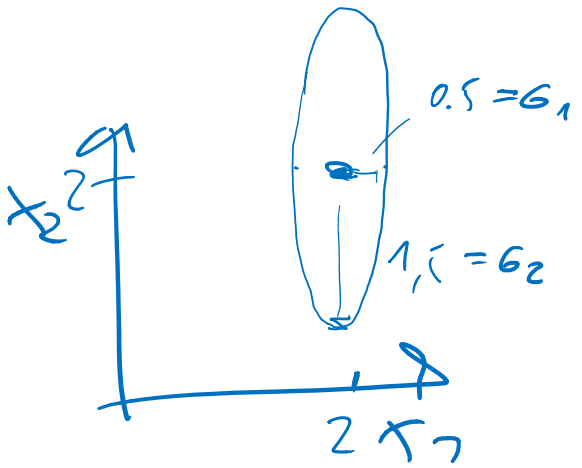
MULTI: $\quad \mu \in \mathbb{R}^n$, $\Sigma \in \mathbb{R}^{n \times n}$ $\quad \Sigma \geq 0$, SYMMETRIC

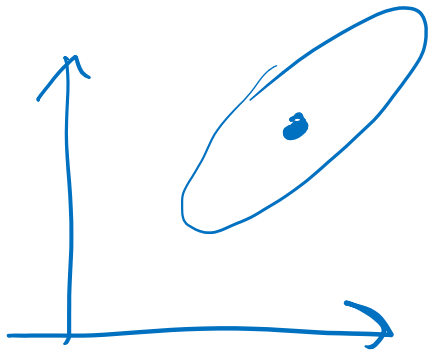
"COV. MATRIX"

$$p(x) = \frac{1}{\sqrt{\det(2\pi\Sigma)}} \cdot \exp\left(-\frac{(x-\mu)^T \Sigma^{-1} (x-\mu)}{2}\right)$$



$$\mu = \begin{pmatrix} 2 \\ 2 \end{pmatrix}, \Sigma = \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix} = \begin{pmatrix} 1/4 & 0 \\ 0 & 9/4 \end{pmatrix}$$



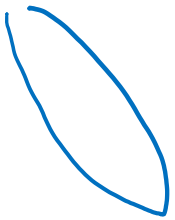


$$\Sigma = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}, \quad \mu = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$$

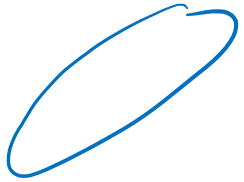
$$\Sigma = V D V^T$$

$$D = \begin{pmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_n \end{pmatrix} \geq 0$$

$$V^T V = I$$



a)



b)

$$\Sigma = \begin{pmatrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & \\ & & \ddots & \\ & & & \sigma_n^2 \end{pmatrix}$$

$\Sigma_{ij} \cdot \sigma_i \cdot \sigma_j$

Correlation $[-1, 1]$

RANDOM VARS.

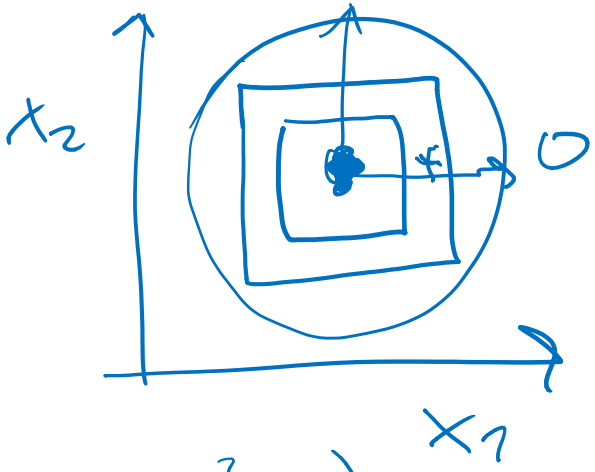
CORRELATION $\Rightarrow$ DEPENDENCE

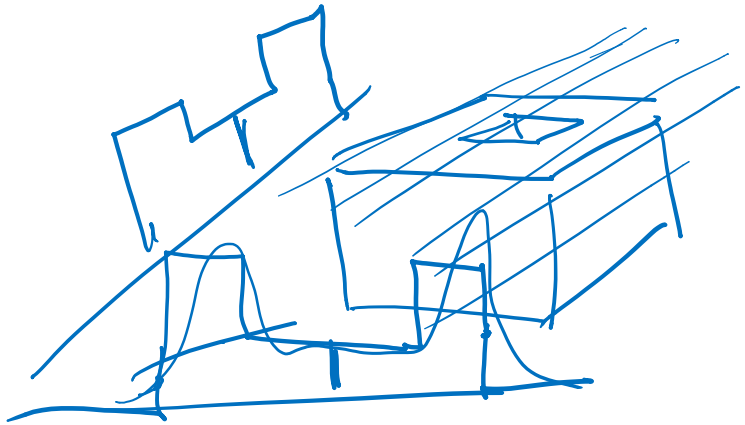
MEAN = EXPECTATION

VARIANCE

COVARIANCE MATRIX

LINEAR (AFFINE) MAP : MEAN ? COV. M. ?





UNCORRELATED ? YES / NO

INDEPENDENT ? YES / NO

$p(x_1, x_2) = p(x_1) \cdot p(x_2)$

$$\Sigma = \begin{pmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{pmatrix}$$

a)

b)

c)

d)

e)

BAYES THM

$$p(x_2|x_1) = \frac{p(x_1|x_2)\cdot p(x_2)}{p(x_1)}$$

PAR ← x_2, OBS. ← x_1

$p(x_2)$: A-PRIORI

$p(x_1|x_2)$: COND. PROB. OF x_1 GIVEN x_2

$p(x_2|x_1)$: A-POSTERIORI PDF

CONDITIONAL PDF: $p(x_1|x_2) = \frac{p(x_1, x_2)}{p(x_2)} \iff p(x_1, x_2) = p(x_1|x_2)\cdot p(x_2)$

ESTIMATOR: GIVEN DATA $Y_N \in \mathbb{R}^N$, EST. SOME $\theta \in \mathbb{R}^1$. Y_N IS RANDOM

$\hat{\theta}_N(Y_N)$ MAP FROM MEAS. TO POINT ESTIMATE

$\Rightarrow \hat{\theta}_N$ RANDOM

EX: "SAMPLE MEAN": $\hat{\theta} = \frac{\sum Y(i)}{N} =: M(Y_N)$

"SAMPLE VARIANCE": $\frac{\sum_i (Y(i) - M(Y_N))^2}{(N-1)}$ $\quad \frac{(N-1)}{(N)}$

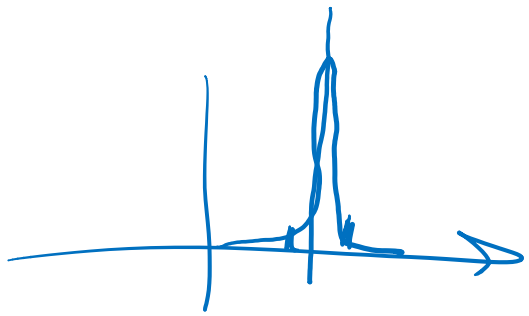
UNBIASED ESTIMATORS FOR (TRUE) MEAN AND VARIANCE OF $Y(i)$ (i.i.d.)

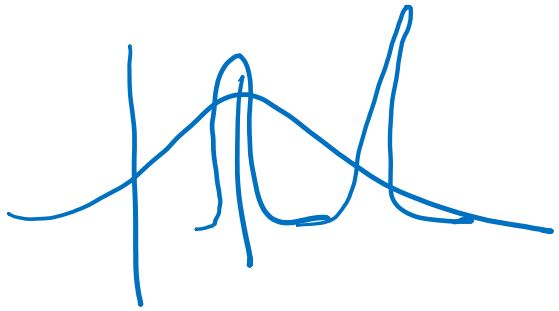
UNBIASED ESTIMATOR: EXP. OF $\hat{\theta}$ = TRUE VAL.

ASYMPTOTIC UNBIASED EST.: EXP OF $\lim_{N\to\infty} \hat{\theta}_N$ = ...

CONSISTENCY: $\forall \varepsilon > 0 : \lim_{N\to\infty} P(\hat{\theta}_N \in [\theta_0 - \varepsilon, \theta_0 + \varepsilon]) = 1$

↑ AS. UNBIASEDNESS & $\mathrm{cov}(\hat{\theta}) \to 0$ ($N \to \infty$)





LS-ESTIMATORS (LINEAR LS)

GIVEN MODEL $\boxed{y(i) = \varphi(i)^T \cdot \theta_0 + \varepsilon(i)}$ (MODEL PREDICTION)

i.i.d, ZERO MEAN

$\varphi(i)$: REGRESSION VECTORS

$$\varphi(i) = \begin{bmatrix} * \\ \vdots \\ * \end{bmatrix} \in \mathbb{R}^d$$

$$\hat{\theta}_{LS} = \arg\min_{\theta} \| y_N - \Phi_N \cdot \theta \|_2^2$$

$$y_N = \begin{bmatrix} y(1) \\ \vdots \\ * \end{bmatrix}, \quad \Phi_N = \begin{bmatrix} \varphi(1)^T \\ \vdots \\ * \end{bmatrix}$$

$$\hat{\theta}_{LS} = \phi_N^+ \cdot y_N$$

ϕ_N^+ : PSEUDO-INVERSE =

$$\begin{cases} (\phi_N^T \phi_N)^{-1} \phi_N^T & \text{IF } \phi_N^T \phi_N \text{ INVERTIBLE } (\Leftrightarrow \text{RANK } \phi_N = d) \\ V \cdot S^+ U^T & \text{ELSE IF } \phi_N = U S V^T \text{ (SVD)} \end{cases}$$

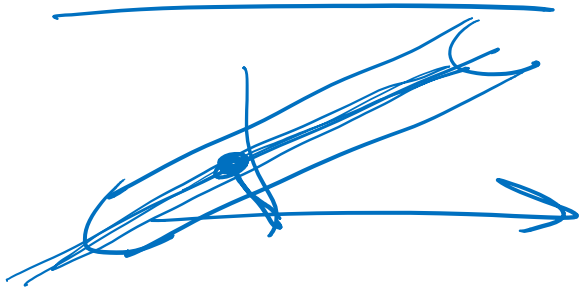
$$S = \begin{pmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & \\ & & & 0 \\ & & & & \ddots \\ & & & & & 0 \end{pmatrix} \qquad S^+ = \begin{pmatrix} \sigma_1^{-1} & & & \\ & \ddots & & \\ & & \sigma_r^{-1} & \\ & & & 0 \\ & & & & \ddots \\ & & & & & 0 \end{pmatrix}$$

ϕ_N^+ : PINV(ϕ)

ASS. RANK(ϕ_N) $< d$

$$\hat{\theta}_{PINV} = \phi_N^+ \cdot y_N = \arg\min_{\theta} \|\theta\|_2^2 \text{ s.t. } \phi_N^T \phi_N \theta = \phi_N^T y_N$$

MINIMUM NORM SOLUTION AMONG ALL OPTIMIZERS OF LS-PROBLEM



WLS : $W \succeq 0$

$$\hat{\theta}_{WLS} = \arg\min_{\theta} \|Y_N - \Phi_N \cdot \theta\|_W^2$$

$$= \arg\min_{\theta} \|W^{\frac{1}{2}}(Y_N - \Phi_N \cdot \theta)\|_2^2$$

$$= \arg\min_{\theta} \|\underbrace{W^{\frac{1}{2}} Y_N}_{\tilde{Y}_N} - \underbrace{(W^{\frac{1}{2}}\Phi_N)}_{\tilde{\Phi}_N}\theta\|_2^2$$

$$= (W^{\frac{1}{2}}\Phi_N)^+ \cdot W^{\frac{1}{2}} Y_N$$

(W)LS GIVES UNBIASED ESTIMATE IF ε IS ZERO MEAN

OPTIMAL WEIGHTING W : $W = \Sigma_\varepsilon^{-1}$ WITH $\Sigma_\varepsilon = \mathrm{Cov}(\varepsilon)$

↓ LEADS TO SMALLEST $\mathrm{COV}(\hat{\theta}_{WLS})$ (= ML EST.)

FOR i.i.d. NOISE : LS IS OPT. (NO WEIGHTING NEC.)

CLOSED BOOK EXAM

(TWO SHEETS WITH 4 PAGES OF HANDWRITTEN OR NOT-TOO SMALL FONT (8 pt))

PEN

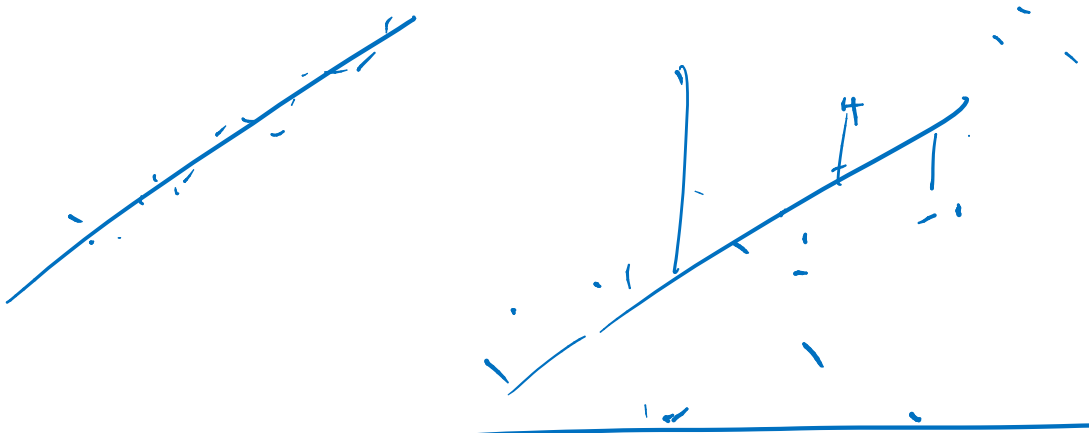
DATE: MARCH 16, 14:00 – 17:00

MEASURE OF FIT — R-SQUARED

$$R^2 = 1 - \frac{\|\varepsilon\|^2}{\|\tilde{y}\|^2} = \frac{\|\hat{y}\|^2}{\|y\|^2} \qquad \in [0,1]$$

RES. AFTER FIT / RES.

1 : PERFECT FIT
0 : BAD FIT



COVARIANCE FROM SINGLE EXP. (LS)

$$\mathrm{cov}(\hat{\theta}) \approx \sigma_\varepsilon^2 \cdot (\Phi_N^T \Phi_N)^{-1} \qquad \sigma_\varepsilon^2 = \frac{\|\varepsilon\|_2^2}{N-d}$$

FOR NLS

$$\Phi_N \approx J = \frac{\partial r}{\partial \theta}$$

$$\theta_i = \hat{\theta}_i \pm \sqrt{\sigma_i^2} \qquad \sigma_i^2 : \text{DIAGONAL ENTRIES OF } \mathrm{cov}(\hat{\theta})$$

ALWAYS : GIVE CONFID. INTERVALS FOR ESTIMATES

MAX. LIKELIHOOD (ML) EST.

GIVEN: PREDICTION MODEL

$$M_N(\theta) + \varepsilon_N = y_N$$

- NOISE ASSUMPTIONS ON ε_N (ZERO MEAN,)

PDF : $p(y_N|\theta)$

"LIKELIHOOD FCN"

"NEG. LOG. LIKELIHOOD" $-\log(p(y_N|\theta))$

ε_N GAUSSIAN $\varepsilon_N \sim N(0, \Sigma)$

$$p(y_N|\theta) = \frac{1}{\sqrt{\dots}} \exp\left(\frac{(y_N - M(\theta))^T \Sigma^{-1} (\dots)}{2}\right)$$

$$-\log p = \dots + \frac{1}{2}\|y_N - M(\theta)\|^2_{\Sigma^{-1}}$$

BAYESIAN EST. : ADD "A-PRIORI KNOWLEDGE" $p(\theta)$

MAP : $\hat{\theta}_{MAP} = \arg\max_\theta \; p(y_N|\theta) \cdot p(\theta)$

ML : $\hat{\theta}_{ML} = \arg\max_\theta \; p(y_N|\theta) = \arg\min_\theta \; -\log p(y_N|\theta)$

NONLIN. LEAST SQ.

(lsq nonlin)

(GAUSS-NEWTON)

CRAMER-RAO INEQUALITY: (LOWER BOUND)

$$L(\theta, y) = -\log(y|\theta)$$

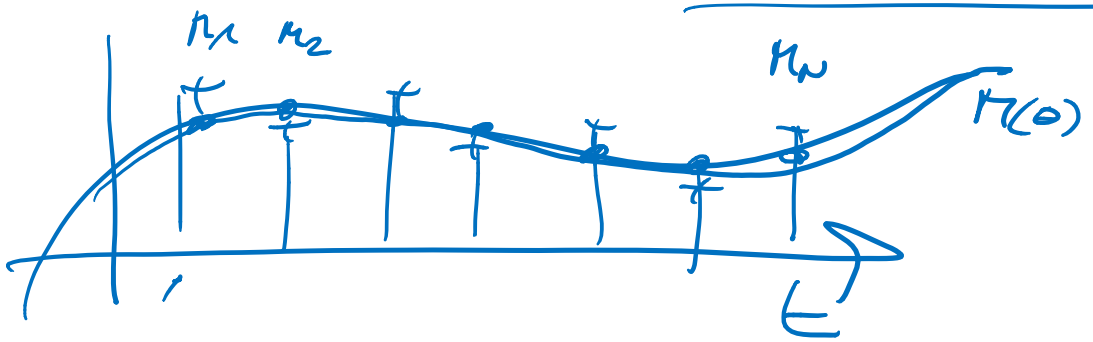
$$\mathbb{E}\{\nabla_\theta^2 L(\theta, y)\} =: M = \text{FISHER INF. MATRIX}$$

FOR ANY UNBIASED ESTIM. $\hat{\theta}$:

$$\boxed{\text{COV}(\hat{\theta}) \geq M^{-1}}$$

FOR $N \to \infty$, ML COV $\longrightarrow$ CR- (ML ASYMPTOT. OPT.)

$M(\theta)$ CAN BE "FORWARD SIMULATION MAP"



NOISE: OUTPUT ERRORS
INPUT ERRORS
→ EQU. ERRORS (LS)

MODELS: LIN, ARX

ODE SIMULATION

CTL "