Pranaya Tomar & Martin Memmesheimer

Ex 3.2

a) We know $\mathcal{L}\left(\frac{C(\hat{b}-b)}{\sigma}\right) = \mathcal{N}(0, I_p)$

$$\Rightarrow \mathcal{L}\left(\left[\frac{C(\hat{b}-b)}{\sigma}\right]^2\right) = \chi^2_p \quad \text{(*explained below)}$$

Now, $\left[\frac{1 \cdot C(\hat{b}-b)}{\sigma}\right]^2 = \frac{(C(\hat{b}-b))^T (C(\hat{b}-b))}{\sigma^2}$

$$= \frac{(\hat{b}-b)^T C^T C (\hat{b}-b)}{\sigma^2}$$

and we know $C^T C = X^T X$

$$= (\hat{b}-b)^T \frac{X^T X}{\sigma^2} (\hat{b}-b)$$

$\Rightarrow \boxed{A = \frac{X^T X}{\sigma^2}}$, a symmetric $p \times p$ matrix as X is $N \times p$

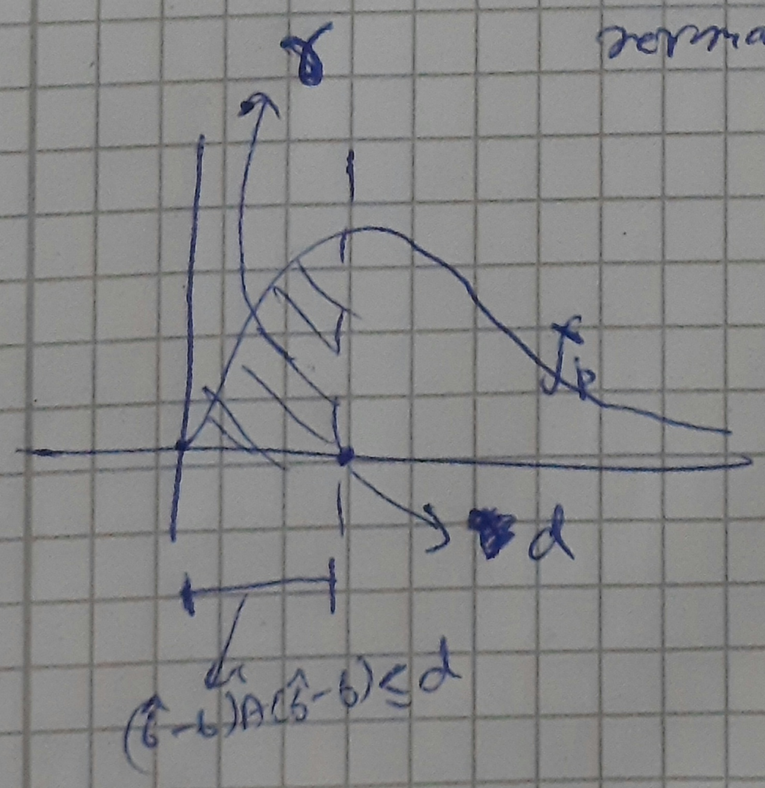
so $\mathcal{L}\left((\hat{b}-b)^T A (\hat{b}-b)\right) = \chi^2_p$

* p degrees of freedom since $Q = (\hat{b}-b)^T A (\hat{b}-b)$

and we know $\hat{b} = \underbrace{(X^T X)^{-1} X^T}_{\text{deterministic}} \underbrace{Y}_{\mathcal{L}(Y) = \mathcal{N}(Xb, \sigma^2 I_N)}$

so $\hat{b}$ is multivariate normally distributed $\Rightarrow$ each component of $\hat{b}$, $\hat{b}_1 \cdots \hat{b}_p$ is independently normally distributed

$\Rightarrow (\hat{b}-b)^T A (\hat{b}-b)$ is the sum of squares of p iid normally distributed RVs. $\Rightarrow \chi^2_p$ distribution



$$P\left[(\hat{b}-b)^T A (\hat{b}-b) \le d\right] = \gamma$$

Now, $(\hat{b}-b)^T A (\hat{b}-b)$ is an ellipsoid in p dimensions. Forming a confidence region of b centred at $\hat{b}$

but $A = \frac{X^T X}{\sigma^2}$ and σ^2 is an unknown quantity,

we know $\Sigma_b = \sigma^2 (X^T X)^{-1}$ (Thm 2.2.3 part 6))

$$\Rightarrow \Sigma_b^{-1} = \frac{X^T X}{\sigma^2}$$

$\Rightarrow$ Confidence region $(\hat{b}-b)^T A (\hat{b}-b)$

$$= \boxed{(\hat{b}-b)^T \Sigma_b^{-1} (\hat{b}-b)}$$

b) in markdown file

ex 3.3.

a) By 2.2.2 the LS-estimators coincides with the ML-estimators. By 2.2.3 we know that they are given by $\hat{b} = (X^T X)^{-1} X^T Y$ or equivalently as the solution of $X^T X b = X^T Y$.

$$X = \begin{pmatrix} 1 & \cos(x_1) & \sin(x_1) \\ 1 & \cos(x_2) & \sin(x_2) \\ \vdots & & \vdots \\ 1 & \cos(x_N) & \sin(x_N) \end{pmatrix} \quad \text{and}$$

$$X^T X = \begin{pmatrix} 1 & \sum_{i=1}^{N} \cos(x_i) & \sum_{i=1}^{N} \sin(x_i) \\ \sum_{i=1}^{N} \cos(x_i) & \sum_{i=1}^{N} \cos^2(x_i) & \sum_{i=1}^{N} \cos(x_i)\sin(x_i) \\ \sum_{i=1}^{N} \sin(x_i) & \sum_{i=1}^{N} \cos(x_i)\sin(x_i) & \sum_{i=1}^{N} \sin^2(x_i) \end{pmatrix}$$

The system can then be solved for example by Cholesky-Zerlegung

b) is in R Marshdam

c) we cannot apply lm() since c_2 enters in a nonlinear fashion.

$$Y_j = c_0 + c_1 \sin(x_j - c_2) + Z_j$$
$$= c_0 - c_1 \cos(x_j) \cdot \sin(c_2) + c_1 \cdot \sin(x_j) \cdot \cos(c_2) + Z_j$$

if we set $a_1 = -c_1 \sin(c_2)$ and

$a_2 = c_1 \cdot \cos(c_2)$ we see some sort of equivalence between the two

d) from c) we get $\hat{c}_0 = \hat{a}_0$

$$c_1 = \frac{a_2}{\cos(c_2)} \Rightarrow a_1 = -\frac{a_2}{\cos(c_2)} \cdot \sin(c_2)$$

$$\Rightarrow \tan(c_2) = -\frac{a_1}{a_2} \quad (a_2 \neq 0) \Rightarrow \hat{c}_2 = \tan^{-1}\left(-\frac{\hat{a}_1}{\hat{a}_2}\right)$$

and $c_2 = \cos^{-1}\left(\frac{a_2}{c_1}\right) \Rightarrow a_1 = -c_1 \cdot \sin\left(\cos^{-1}\left(\frac{a_2}{c_1}\right)\right)$

$$\Leftrightarrow a_1 = -c_1 \cdot \sqrt{1 - \underbrace{\cos^2\left(\cos^{-1}\left(\frac{a_2}{c_1}\right)\right)}_{= \frac{a_2^2}{c_1^2}}} \Leftrightarrow \frac{a_1^2}{c_1^2} = 1 - \frac{a_2^2}{c_1^2}$$

$$\Leftrightarrow c_1 = \pm\sqrt{a_1^2 + a_2^2} \Rightarrow \hat{c}_1 = \pm\sqrt{\hat{a}_1^2 + \hat{a}_2^2} \quad \square$$

ex 4

a) Claim: $\bar{h} := \frac{1}{N} \cdot \sum_{i=1}^{N} H_{ii} = \frac{P}{N}$

Proof: The trace is invariant under permutations

$\Rightarrow \quad \text{trace}(H) = \text{trace}(X(X^TX)^{-1}X^T)$

$= \text{trace}(X^TX(X^TX)^{-1}) = \text{trace}(I_p) = p$

$\Rightarrow \quad \bar{h} = \frac{1}{N} \cdot \text{trace}(H) = \frac{P}{n}$

q.e.d.

b) $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$

$H = X \underbrace{(X^TX)^{-1}}_{= \frac{1}{\|X\|_2^2}} X^T = \frac{1}{\|X\|_2^2} (XX^T)$

since X^TX is inv $x \neq 0$ and therefore $\frac{1}{\|X\|_2^2}$ is well defined

$\Rightarrow \quad H_{ii} = \frac{1}{\|X\|_2^2} \cdot x_i^2$ and $\bar{h} = \frac{1}{N}$ (by a))

i) Claim: data point u is an outlier

Proof: $H_{uu} = \frac{1}{\|X\|_2^2} \cdot \left(\frac{N}{N}\right)^2 = \frac{1}{\|X\|_2^2}$

calculate $\|X\|_2^2 = \left(\frac{1}{N}\right)^2 + \dots + \left(\frac{N}{N}\right)^2 = \frac{1}{N^2} \sum_{i=1}^{N} i^2$

$= \frac{1}{N^2} \cdot \frac{N(N+1)(2N+1)}{6} = \frac{N}{3} + \frac{1}{2} + \frac{1}{6N}$

for $N = 6$: $\frac{1}{\|X\|_2^2} = \frac{1}{\frac{6}{3} + \frac{1}{2} + \frac{1}{36}} \approx 0{,}4 > \frac{1}{3} = \frac{2}{6} = \frac{2}{N}$

ii) Claim: ~~for $j \in [\frac{N}{2}+1, N]$ j is an outlier~~ There is no outlier.

Proof: * $H_{jj} = \frac{1}{\|X\|_2^2} = \frac{1}{\sum_{i=1}^{\frac{N}{2}+1} i} = \frac{1}{\sum_{i=1}^{\frac{N}{2}} i} = \frac{2}{(\frac{N}{2})(\frac{N}{2}+1)}$

* for $j \in [\frac{N}{2}+1, N]$

$= \frac{2}{\frac{N^2}{4} + \frac{N}{2}} = \frac{8}{N^2+2N}$

for $N = 6$ $H_{jj} = \frac{8}{36+12} = \frac{1}{6} \leq \frac{2}{6} = \frac{2}{N} = 2 \cdot \bar{h}$

and for N larger H_{jj} gets smaller faster because of the N^2 than $\frac{2}{N}$

and for $j \in [0, \frac{N}{2}]$ $H_{jj} = 0 \leq \frac{2}{N} = 2 \cdot \bar{h}$

$\Rightarrow$ there are no outliers.

iii) Claim: Data points $N-1$ and N are outliers.

Proof: $\|X\|_2^2 = \sqrt{2}$

$j \in \{N-1, N\}$ $\quad H_{jj} = \frac{1}{\sqrt{2}} > \frac{2}{N}$ for $N \geq 6$

since $\frac{1}{\sqrt{2}} > \frac{2}{N} \Leftrightarrow N > 2\sqrt{2} \approx 2{,}82$

q. e. d.