Leo Bremer
ESE 524
Homework 6

Problem 1

Problem statement

Show that the Fisher information matrix in the multivariate normal model, when $x = (x_1, x_2, \dots, x_M)$, $x_n \sim N(m, R)$ and R is known, is given by

$$J = M \frac{\partial}{\partial \theta} m^T R^{-1} \left(\frac{\partial}{\partial \theta} m^T\right)^T,$$

Evaluate J when $m = \theta$.

Solution

$= tr\left(R^{-1} \sum (x_i - \bar{x})(x_i - \bar{x})^T\right) + M(\bar{x} - m)^T R^{-1}($

Gaussian log-likelihood:

$$\Lambda = -\frac{Mn}{2} \log(2\pi) - \frac{M}{2} \log|R| - \frac{1}{2} \sum_{i=1}^{M} (x_i - m)^T R^{-1} (x_i - m)$$

$$J(\theta) = -E\left[\frac{\partial}{\partial \theta}\left(\frac{\partial}{\partial \theta} \Lambda\right)\right]^T$$

$$= -E\left[\frac{\partial}{\partial \theta}\begin{bmatrix} \frac{\partial}{\partial m} \Lambda \\ \frac{\partial}{\partial R^{-1}} \Lambda \end{bmatrix}\right]$$

$\hookrightarrow$ Hessian

$$\frac{\partial}{\partial m} \Lambda = -\frac{1}{2}(2) M (R^{-1})(\bar{x} - m) = -M R^{-1}(\bar{x} - m)$$

$$-E\left[\frac{\partial}{\partial \theta}\left(\frac{\partial}{\partial m} \Lambda\right)\right] = -E\left[\underbrace{(-MR^{-1})}_{\left(M \frac{\partial}{\partial \theta} m^T R^{-1}\right)} + \underbrace{(-M m^T)}_{M \frac{\partial}{\partial \theta} m^T}\right]$$

$$= MR^{-1} + M m^T$$

$$J(\theta) = J(m) = M \cdot R^{-1}$$

Problem 2

2. Problem Statement

Show that an efficient estimator whose Fisher information matrix is independent of θ is distributed as

$$\hat{\theta} \sim N[\theta, J^{-1}]$$

Solution

Efficient $\Rightarrow$ unbiased $\Rightarrow E[\hat{\theta}] = \theta$

$\hookrightarrow$ CRLB $\Rightarrow E[(\theta-\hat{\theta})(\theta-\hat{\theta})^T] = J^{-1}(\theta)$

use

However J is independent of θ, so

$E[(\theta-\hat{\theta})(\theta-\hat{\theta})^T] = J^{-1}$ where J^{-1} is a scalar matrix.

$\hookrightarrow$ covariance of $\hat{\theta}$

So $\hat{\theta} \sim N(\theta, J^{-1})$

Problem 3

3. <u>Problem statement</u>

Consider a random sample of scalar random variables

$X = (x_0, x_1, \ldots, x_{N-1})$

with $f(x_n|\theta) = \frac{1}{\sqrt{2\pi}\,\sigma} \cdot \exp\left(-(x_n-\theta)^2/2\sigma^2\right)$

The parameter of interest θ is normally distributed with $\theta \sim N[m, \sigma_\theta^2]$.

a) Find the conditional density of θ given x.

b) Find the conditional mean and variance of θ given x.

c) Compare $E[\theta|x]$ to $\hat{\theta}_{ML}$.

<u>Solution</u>

a) $f(\theta|x) = f(\theta) \cdot f(x|\theta)$

$$= \frac{1}{\sqrt{2\pi}\,\sigma_\theta} \cdot \exp\left(-(\theta-m)^2/2\sigma_\theta^2\right) \prod_{i=0}^{N-1} \frac{1}{\sqrt{2\pi}\,\sigma} \cdot \exp\left(-(x_i-\theta)^2/2\sigma^2\right)$$

$$= \left(\frac{(2\pi)^{-\left(\frac{N+1}{2}\right)}}{\sigma_\theta \cdot \sigma^N}\right) \cdot \exp\left(-(\theta-m)^2/2\sigma_\theta^2\right) \cdot \exp\left(-\sum_{i=0}^{\infty}(x_i-\theta)^2/2\sigma^2\right)$$

$$= \left(\frac{(2\pi)^{-\frac{N+1}{2}}}{\sigma_\theta \cdot \sigma^N}\right) \cdot \exp\left(-\frac{1}{2}\left(\frac{(\theta-m)^2}{\sigma_\theta^2} + \sum_{i=0}^{\infty}\frac{(x_i-\theta)^2}{\sigma^2}\right)\right)$$

$$= \underbrace{\left(\frac{(2\pi)^{-\frac{N+1}{2}}}{\sigma_\theta \cdot \sigma^N}\right)}_{C_1} \cdot \exp\left(-\frac{1}{2\sigma_\theta^2\sigma^2}\left(\sigma^2(\theta-m)^2 + \sigma_\theta^2\sum_{i=0}^{\infty}(x_i-\theta)^2\right)\right)$$

$$= C_1 \cdot \exp\left(-\frac{1}{2\sigma_\theta^2\sigma^2}\left(\sigma^2\theta^2 - 2\sigma^2\theta m + \sigma^2 m^2 + \sigma_\theta^2\sum_{i=0}^{N}(x_i^2 - 2x_i\theta + \theta^2)\right)\right)$$

$$= C_1 \cdot \exp\left(-\frac{1}{2\sigma_\theta^2\sigma^2}\left(\sigma^2\theta^2 - 2\sigma^2\theta m - 2\sigma_\theta^2 N \cdot \bar{x}\,\theta + N\sigma_\theta^2\theta^2\right)\right)$$

$$\underbrace{\cdot \exp\left(-\frac{1}{2\sigma_\theta^2\sigma^2}\left(\sigma_\theta^2\sum_{i=0}^{N-1} x_i^2 + \sigma^2 m^2\right)\right)}_{C_2}$$

$$= C_1 \cdot C_2 \cdot \exp\left(-\frac{1}{2\sigma_\theta^2\sigma^2}\left(\theta^2\left(N\sigma_\theta^2 + \sigma^2\right) - 2\theta\left(N\sigma_\theta^2\bar{x} + \sigma^2 m\right)\right)\right)$$

$$= C \cdot \exp\left(-\frac{1}{2\sigma_\theta^2\sigma^2}\left(\theta^2\left(\frac{N}{\sigma^2} + \frac{1}{\sigma_\theta^2}\right) - 2\theta\left(\frac{N\bar{x}}{\sigma^2} + \frac{m}{\sigma_\theta^2}\right)\right)\right)$$

$$= C \cdot \exp\left(-\frac{1}{2}\left(\frac{N}{\sigma^2} + \frac{1}{\sigma_\theta^2}\right)\left(\theta^2\left(\frac{\frac{N}{\sigma^2} + \frac{1}{\sigma_\theta^2}}{\frac{N}{\sigma^2} + \frac{1}{\sigma_\theta^2}}\right) - 2\theta\left(\frac{\frac{N\bar{x}}{\sigma^2} + \frac{m}{\sigma_\theta^2}}{\frac{N}{\sigma^2} + \frac{1}{\sigma_\theta^2}}\right)\right)\right)$$

$$= C \cdot \exp\left(-\frac{1}{2}\underbrace{\left(\frac{N}{\sigma^2} + \frac{1}{\sigma_\theta^2}\right)}_{\frac{1}{\sigma^2}}\left(\theta - \underbrace{\frac{\frac{N\bar{x}}{\sigma^2} + \frac{m}{\sigma_\theta^2}}{\frac{N}{\sigma^2} + \frac{1}{\sigma_\theta^2}}}_{\mu}\right)^2\right)$$

Thus the conditional mean is $\boxed{\left(\frac{\frac{N\bar{x}}{\sigma^2} + \frac{m}{\sigma_\theta^2}}{\frac{N}{\sigma^2} + \frac{1}{\sigma_\theta^2}}\right)}$

and the conditional variance is $\boxed{\frac{1}{\frac{N}{\sigma^2} + \frac{1}{\sigma_\theta^2}}}$

and $f(\theta|x) \sim N\left(\frac{\frac{N\bar{x}}{\sigma^2} + \frac{m}{\sigma_\theta^2}}{N/\sigma^2 + \frac{1}{\sigma_\theta^2}}, \frac{1}{\frac{N}{\sigma^2} + \frac{1}{\sigma_\theta^2}}\right)$

---

c) The ML estimator $\hat{\theta}_{ML} = \frac{1}{N}\sum_{i=0}^{N} x_i = \bar{x}$ ← sample mean

$E[\theta|x] = \frac{\frac{N\bar{x}}{\sigma^2} + \frac{m}{\sigma_\theta^2}}{\frac{N}{\sigma^2} + \frac{1}{\sigma_\theta^2}}$. Consider $\lim_{N\to\infty} E[\theta|x]$:

$$= \lim_{N\to\infty} \frac{\frac{\partial}{\partial N}\left(\frac{N\bar{x}}{\sigma^2} + \frac{m}{\sigma_\theta^2}\right)}{\frac{\partial}{\partial N}\left(\frac{N}{\sigma^2} + \frac{1}{\sigma_\theta^2}\right)} = \frac{\frac{\bar{x}}{\sigma^2}}{\frac{1}{\sigma^2}} = \boxed{\bar{x} = \hat{\theta}_{ML}}$$

Problem 4

4. Problem Statement

Define loss function $L[\theta, \hat{\theta}(x)] = \pi^T a[\theta, \hat{\theta}(x)]$

with $\pi = [\pi_1, \pi_2, \ldots, \pi_p]^T : \pi_i > 0 \ \forall i$,

$a[\theta_i, \hat{\theta}_i(x)] = |\theta_i - \hat{\theta}_i(x)|$

show that the conditional risk may be written as

$$\int d\theta\, f(\theta|x)\, \pi^T a[\theta, \hat{\theta}(x)] = \sum_{i=1}^{p} \pi_i \beta$$

where $\beta = -\int_{-\infty}^{\hat{\theta}_i(x)} d\theta_i\, f(\theta_i|x)[\theta_i - \hat{\theta}_i(x)] + \int_{\hat{\theta}_i(x)}^{\infty} d\theta_i\, f(\theta_i|x)[\theta_i - \hat{\theta}_i(x)]$

a) Conditional risk $= E[L(\theta, \hat{\theta}(x))]$

$$= \int_{-\infty}^{\infty} f(\theta|x)\, L(\theta, \hat{\theta}(x))\, d\theta$$

$$= \int_{-\infty}^{\infty} d\theta\, f(\theta|x)\, \pi^T \cdot a[\theta, \hat{\theta}(x)] = \int_{-\infty}^{\infty} d\theta\, f(\theta|x)\, \pi^T \cdot |\theta_i - \hat{\theta}_i(x)|$$

For $\hat{\theta}_i > \theta_i$:

$$-\int_{-\infty}^{\hat{\theta}_i} d\theta_i\, f(\theta_i|x)\, \pi^T(\theta_i - \hat{\theta}_i(x))$$

For $\hat{\theta}_i < \theta_i$:

$$\int_{\hat{\theta}_i(x)}^{\infty} d\theta_i\, f(\theta_i|x)\, \pi^T(\theta_i - \hat{\theta}_i(x))$$

We combine these integrals by adding them:

$$-\int_{-\infty}^{\hat{\theta}_i} d\theta_i \cdot f(\theta_i|x) \underbrace{\pi^T(\theta_i - \hat{\theta}_i(x))}_{\text{inner product}} + \int_{\hat{\theta}_i}^{\infty} d\theta_i \cdot f(\theta_i|x)\, \pi^T(\theta_i - \hat{\theta}_i(x)) =$$

$$= -\int_{-\infty}^{\hat{\theta}} d\theta_i\, f(\theta_i|x) \sum_{i=1}^{p} \pi_i (\theta_i - \hat{\theta}_i(x)) + \int_{\hat{\theta}_i}^{\infty} d\theta_i \cdot f(\theta_i|x) \sum_{i=1}^{p} \pi_i (\theta_i - \hat{\theta}_i(x))$$

$$= \sum_{i=1}^{p} \pi_i \left(-\int_{-\infty}^{\hat{\theta}} d\theta_i\, f(\theta_i|x)(\theta_i - \hat{\theta}_i(x)) + \int_{\hat{\theta}_i}^{\infty} d\theta_i \cdot f(\theta_i|x)(\theta_i - \hat{\theta}_i(x)) \right)$$

$$= \sum_{i=1}^{p} \pi_i \beta$$

b) Minimizing estimator gives $E[L(\theta, \hat{\theta}(x))] = 0$

$$\Rightarrow \sum_{i=1}^{p} \pi_i B = 0$$

$$\Rightarrow \sum_{i=1}^{p} \pi_i \left(- \int_{-\infty}^{\hat{\theta}_i(x)} d\theta_i \cdot f(\theta_i | x) (\theta_i - \hat{\theta}_i(x)) + \int_{\hat{\theta}_i(x)}^{\infty} d\theta_i \cdot f(\theta_i | x) (\theta_i - \hat{\theta}_i(x)) \right) = 0$$

$\pi_i = 0$ is trivial. Only other way for the sum to be 0 if the two integrals are equal:

$$\int_{-\infty}^{\hat{\theta}_i(x)} d\theta_i \cdot f(\theta_i | x) (\theta_i - \hat{\theta}_i(x)) = \int_{\hat{\theta}_i(x)}^{\infty} d\theta_i \cdot f(\theta_i | x) (\theta_i - \hat{\theta}_i(x))$$

Thus $\hat{\theta}_i$ must be the median so that both integrals are equal.

Problem 5

5. Problem statement

Simplify the Kalman Filter for

a) $A = A$, $u_t = 0 \quad \forall t \geq 0$

and

b) $A = I$, $R_0 = r_0 I$, $r_0 \to \infty$

Kalman Equations:

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + k_t (y_t - c^T \hat{x}_{t|t-1})$$

$$\hat{x}_{t|t-1} = A \hat{x}_{t-1|t-1}$$

$$k_t = P_{t|t-1} c \gamma_t^{-1}$$

$$\gamma_t = c^T P_{t|t-1} c + \sigma_n^2$$

$$P_{t|t} = P_{t|t-1} - \gamma_t k_t k_t^T$$

$$P_{t+1|t} = A P_{t|t} A^T + \sigma_u^2 b b^T.$$

Solution

a) $A = A$, $u_t = 0 \quad \forall t > 0$

$$x_{t+1} = A x_t + b \cancel{u_t}^{0}$$

$$P_{t+1|t} = A P_{t|t} A^T + 0$$

b) $A = I$, $R_0 = r_0 I$, $r_0 \to \infty$

$$R_{t+1} = r_0 I + \sigma_u^2 b b^T$$

If $r_0 \to \infty$, $R_{t+1} \to \infty$

Problem 6

6. Problem Statement

Minimize the quadratic form $(y-x)^T(y-x)$ subject to the constraint $x = H\Theta$. Should find $\hat{x} = P_{H}y$, where

$$P_{H} = H(H^TH)^{-1}H^T.$$

Then apply this result to the quadratic form $(y-x)^T W (y-x)$ for $W \sim$ full rank, symmetric. Interpret the result.

Solution

$$\min (y-H\Theta)^T(y-H\Theta) = y^Ty - \underbrace{y^TH\Theta}_{\to\, \Theta^TH^Ty} - \Theta^TH^Ty + \Theta^TH^THH\Theta = f(\Theta)$$

$$\frac{\partial f}{\partial \Theta} = -2y^TH + \Theta^T(H^TH + H^TH) = 2\Theta^TH^TH - 2y^TH = 0$$

$$\Theta^TH^TH = y^TH$$

$$H^TH\cdot\Theta = H^Ty$$

$$\Theta = (H^TH)^{-1}H^Ty$$

$$x = H\Theta$$

$$\Rightarrow \hat{x} = H(H^TH)^{-1}H^Ty$$

Now, use the same approach on the following minimization:

$$\min \underbrace{(y-H\Theta)^TW(y-H\Theta)}_{f(\Theta)}$$

$$f(\Theta) = y^TWy - y^TWH\Theta - \Theta^TH^TWy + \Theta^TH^TWH\Theta$$

$$\frac{\partial f}{\partial \Theta} = -y^TWH - y^TW^TH + \Theta^T(H^TW^TH + H^TWH) = 0$$

since W is symmetric, $W^T = W$

$$\text{So } \frac{\partial f}{\partial \Theta} = -2y^TWH + 2\Theta^TH^TWH = 0$$

$$y^TWH = \Theta^TH^TWH$$

$$H^TWy = H^TWH\Theta$$

$$\Theta = (H^TWH)^{-1}H^TWy$$

$$\hat{x} = H(H^TWH)^{-1}H^TWy$$

Problem 7

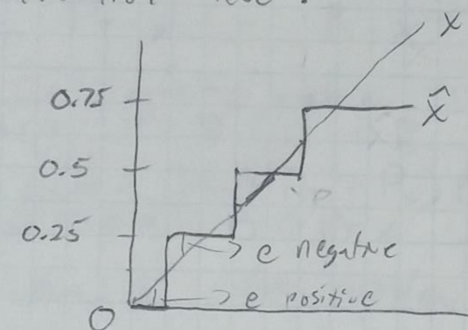
7. <u>Problem Statement</u>

Suppose that a uniform rounding quantizer is used to quantize an arbitrary random variable x to L quantization levels, with B bits. Find the pmf for the quantizer error $e = x - \hat{x}$ for large L. What is the noise power of e in terms of L and B?

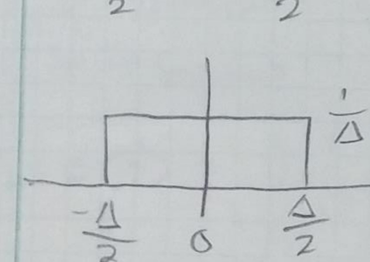
<u>Solution</u>

Assume that $0 \leq x \leq a$, let $a = 1$ without loss of generality. Then the number of quantization levels is $2^B = L$ and the difference between quantization levels $\Delta = \frac{1}{L}$. Assume that when a value is sample it is quantized to the nearest quantization level:

$B = 2$
$L = 4$



Then the error $e = x - \hat{x}$ is distributed uniformly from $-\frac{\Delta}{2}$ to $\frac{\Delta}{2}$:



As L is increased, Δ decreases and this distribution tightens toward 0.

The noise power $E[e^2] = \int_{-\infty}^{\infty} f_e(e) \cdot e^2 \cdot de = \int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}} \frac{1}{\Delta} e^2 de$

$$= \frac{1}{\Delta}\left[\frac{1}{3}e^3 \Big|_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}}\right] = \frac{1}{\Delta}\left[\frac{\Delta^3}{24} - \frac{\Delta^3}{24}\right] = \boxed{\frac{\Delta^2}{12}}$$

Problem 8

8. Problem statement

Suppose X is an n-dimensional vector with $X \sim N(\mu m, \sigma^2 I)$. Find a test to determine whether or not $\mu \leq 0$ or $\mu > 0$. Note that X is corrupted with noise.

$$Y = \gamma Q_{m\perp} X$$

Where $\gamma > 0$ and $Q_{m\perp}$ is an unknown rotation orthogonal to m.

Solution → Get X in terms of its orthogonal components

$S = \text{span}(m)$ and let $T = S^\perp$, the orthogonal complement of S in $\mathbb{R}^n$. Let P_T be a projection matrix of $X \to T$ and P_S be a projection $X \to S$.

Then $X = P_T \cdot x + P_S \cdot x$.

Let Q be a rotation matrix s.t. $QQ^T = I$.
Then $Q_{m\perp} = P_T Q P_T + P_S$.

Note: $Q \cdot P_S x = (P_T Q P_T + P_S) P_S x = P_S x$
and $Q \cdot P_T x = P_T Q P_T x$

With $Y = \gamma Q_{m\perp} X$ the problem is:

$H_0 \sim N(\mu\gamma Q_{m\perp} x, \gamma^2 I)$; $\mu \leq 0$
$H_1 \sim N(\mu\gamma Q_{m\perp} x, \gamma^2 I)$; $\mu > 0$

A sufficient statistic is $t = \frac{m^T y}{\|m\| \|y\|}$. Then our test becomes

$$\phi(t) = \begin{cases} 1 & t > \nu \quad : H_1 \\ 0 & t \leq \nu \quad : H_0 \end{cases}$$