An Introduction to signal Detection and Estimation

BUPT homework

HW-1

1st problem

2. Suppose Y is a random variable that, under hypothesis H_0, has pdf

$$p_0(y) = \begin{cases} (2/3)(y+1), & 0 \le y \le 1 \\ 0, & \text{otherwise.} \end{cases}$$

and, under hypothesis H_1, has pdf

$$p_1(y) = \begin{cases} 1, & 0 \le y \le 1 \\ 0, & \text{otherwise.} \end{cases}$$

(a) Find the Bayes rule and minimum Bayes risk for testing H_0 versus H_1 with uniform costs and equal priors.

(b) Find the minimax rule and minimax risk for uniform costs.

(c) Find the Neyman-Pearson rule and the corresponding detection probability for false-alarm probability $\alpha \in (0,1)$.

solution (a):

Firstly, according to $(II.B.10)$, write the likelihood ratio as:

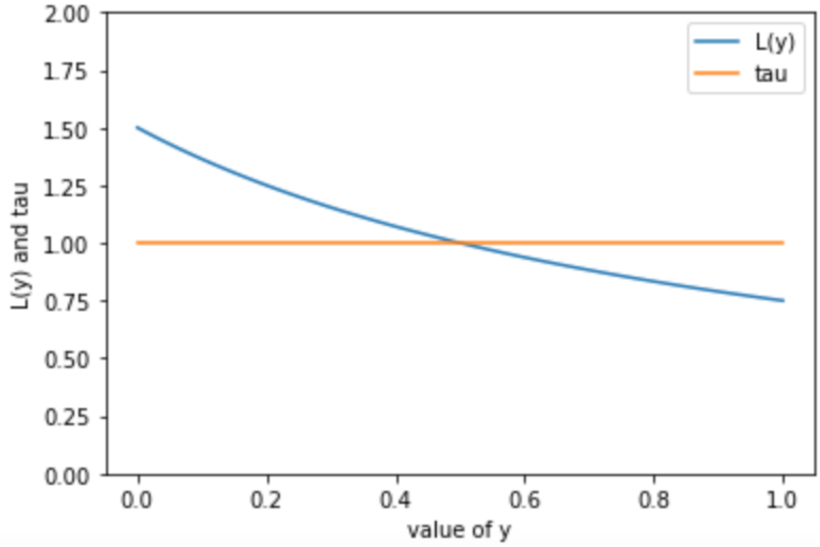
$$\begin{aligned} L(y) &= \frac{p_1(y)}{p_0(y)} \\ &= \begin{cases} \frac{3}{2(y+1)}, & 0 \le y \le 1 \\ meaninglesss, & \text{otherwise} \end{cases} \end{aligned}$$

according to $(II.B.12)$ The bayes rule can be writen as:

$$\delta_B(y) = \begin{cases} 1 & \text{if } L(y) \ge \tau \\ 0 & \text{if } L(y) < \tau \end{cases}$$

where $\tau = 1$ under the uniform costs and equal priors,

The plot of $L(y)$ and τ is drawing as :



Thus the bayes rule is rewritten to the form as:

$$\delta_B(y) = \begin{cases} 1, \text{ if } \frac{3}{2(y+1)} \geq 1 \\ 0 & \text{if } \frac{3}{2(y+1)} < 1 \end{cases}$$
$$= \begin{cases} 1 & \text{if } y \leq 0.5 \\ 0 & \text{if } y > 0.5 \end{cases}$$

The bayes risk is given by $(II.B.14)$, and because the uniform costs and equal priors, the Bayes risk can be written as:

$$r(\delta_B) = \pi_0 P_0(\Gamma_1) + \pi_1 P_1(\Gamma_0)$$
$$= 0.5 \int_0^{0.5} p_0(y) + 0.5 \int_{0.5}^{1} p_1(y)$$

Where $p(y) = 0.5 \cdot (p_1(y) + p_0(y)) = \frac{2y+5}{6}$, the risk is rewritten as:

$$r(\delta_B) = 0.5 \cdot \left[\int_0^{0.5} 2/3 \cdot (y+1) + \int_{0.5}^{1} 1 \right]$$

```
fun1 = @(y) (2/3)./(y+1);
q1 = integral(fun1, 0, 0.5)

fun2 = @(y) y.^0
q2 = integral(fun2, 0.5, 1)

fanal = 0.5*(q1+q2)
```

var the matlab code shown above, the Bayes risk is calculated as:

$$r(\delta_B) = 0.3852$$

Solution(b):

The function V Is given by :

$$V(\pi_0)=\begin{cases}\pi_0, & 0\le\pi_0<3/7\\ \pi_0\left(\int_0^{\tau'}\frac{2}{3}(y+1)\cdot dy\right)+(1-\pi_0)\left(\int_{\tau'}^{1}1\cdot dy\right), & 3/7\le\pi_0<3/5\\ 1-\pi_0, & 3/5\le\pi_0<1\end{cases}$$

Where $\tau'=\frac{3-5\pi_0}{2\pi_0}$ Is the threotle. Since $V(0)=0, V(1)=0$, the least favorable prior π_L

is in the interior $(0,1)$ in this case. And it is unnecessary to consider teh case that $P_0(L(Y)=\tau)=P_1(L(Y)=\tau)=0$ since $L(Y)$ is a continuous random variable. So an equalizer rule is found by solving:

$$\int_0^{\tau'}\frac{2}{3}(y+1)\cdot dy=\int_{\tau'}^{1}1\cdot dy$$

Then, $\tau_L'=\frac{-5+\sqrt{37}}{2}$, so $\pi_L'=\frac{3}{2\tau_0+5}=\frac{3}{\sqrt{37}}$, so the minimax rule is:

$$\delta_{\pi_L}(y)\quad=\begin{cases}1, & \text{if } y\le\frac{3}{\sqrt{37}}\\ 0, & \text{if } y>\frac{3}{\sqrt{37}}\end{cases}$$

Then the minimax risk for the uniform costs is:

$$r_u(\delta_\pi)=0.5\cdot\int_0^{\frac{3}{\sqrt{37}}}2/3\cdot(y+1)+0.5\cdot\int_{\frac{3}{\sqrt{37}}}^{1}1\cdot dy=0.4099$$

```
fun1 = @(y) 1/3.*(y+1);
q1 = integral(fun1, 0, 3/sqrt(37))

fun2 = @(y) 0.5*y.^0;
q2 = integral(fun2, 3/sqrt(37), 1)

fanal = q1+q1
```

Solution (c):

First, write the false-alarm probability as:

$$\begin{aligned}P_0(p_1(Y)>\eta\cdot p_0(Y))&=P_0(L(Y)>\eta)\\&=P_0(Y<\eta')\\&=\frac{1}{3}\eta'^2+\frac{2}{3}\eta'\end{aligned}$$

Where $\eta'=\frac{3-2\eta}{2\eta}$ and get the η_0' by sovling $P_0(p_1(Y)>\eta\cdot p_0(Y))=\alpha$:

$$\eta_0'=\sqrt{1+3\alpha}-1$$

Then the Neyman-Pearson rule is given by:

$$\tilde{\delta}_{NP}(y) = \begin{cases} 1 & \text{if } y \geq \eta_0' \\ 0 & \text{if } y < \eta_0' \end{cases}$$

Where $\eta_0^{'} = \sqrt{1+3\alpha} - 1$.

2rd problem

4. Repeat Exercise 2 for the situation in which p_0 and p_1 are given instead by

$$p_0(y) = \begin{cases} e^{-y}, & y \geq 0 \\ 0, & y < 0 \end{cases}$$

and

$$p_1(y) = \begin{cases} \sqrt{2/\pi} e^{-y^2/2}, & y \geq 0 \\ 0, & y < 0. \end{cases}$$

For part (a) consider arbitrary priors.

solution:

Firstly, written the likelihood ratio as:

$$\begin{aligned} L(y) &= \frac{p_1(y)}{p_0(y)}, \quad y \geq 0 \\ &= \sqrt{2/\pi} e^{-y^2/2+y} \end{aligned}$$

Then the Threshold $\tau = \frac{\pi_0}{\pi_1}$, so the Γ_1 Can be found by:

$$\begin{aligned} & L(y) \geq \tau \\ \Rightarrow & \sqrt{2/\pi} e^{-y^2/2+y} \geq \frac{\pi_0}{(1-\pi_0)} \\ \Rightarrow & y^2 - 2y + 2 \cdot ln\left(\sqrt{\frac{\pi}{2}} \cdot \frac{\pi_0}{1-\pi_0}\right) \leq 0 \end{aligned}$$

Then the final bayes rule under $\forall \pi_0 \in [0, 1]$ can be obtained as follows:

$$\delta_B(y) = \begin{cases} 1, & y \in [1-\sqrt{1-c}, 1+\sqrt{1-c}] \\ 0, & y \in [0, 1-\sqrt{1-c}) \cup (1-\sqrt{1-c}, +\infty) \end{cases} \quad when \quad \pi_0 \in [0, \frac{1}{\sqrt{\frac{\pi}{2e}}+1})$$

Where $c = 2 \cdot ln(\sqrt{\frac{\pi}{2}} \cdot \frac{\pi_0}{1-\pi_0})$

$$\delta_B(y) = 0 \quad when \quad \pi_0 \in [\frac{1}{\sqrt{\frac{\pi}{2e}}+1}, +\infty)$$

3rd problem

7. (a) Consider the hypothesis pair

$$H_0 : Y = N$$
$$\text{versus}$$
$$H_1 : Y = N + S$$

where N and S are independent random variables each having pdf

$$p(x) = \begin{cases} e^{-x}, & x \geq 0 \\ 0, & x < 0. \end{cases}$$

Find the likelihood ratio between H_0 and H_1.

(b) Find the threshold and detection probability for α-level Neyman-Pearson testing in (a).

(c) Consider the hypothesis pair

$$H_0 : Y_k = N_k, \quad k = 1, \ldots, n$$
$$\text{versus}$$
$$H_1 : Y_k = N_k + S, \quad k = 1, \ldots, n$$

where $n > 1$ and $N_1, \ldots, N_n$, and S are independent random variables each having the pdf given in (a). Find the likelihood ratio.

(d) Find the threshold for α-level Neyman-Pearson testing in (c).

solution （a）：

write the pdf of H_0 and H_1:

$$p_0(y) = p(y) = e^{-y}, \quad y \in (0, +\infty)$$

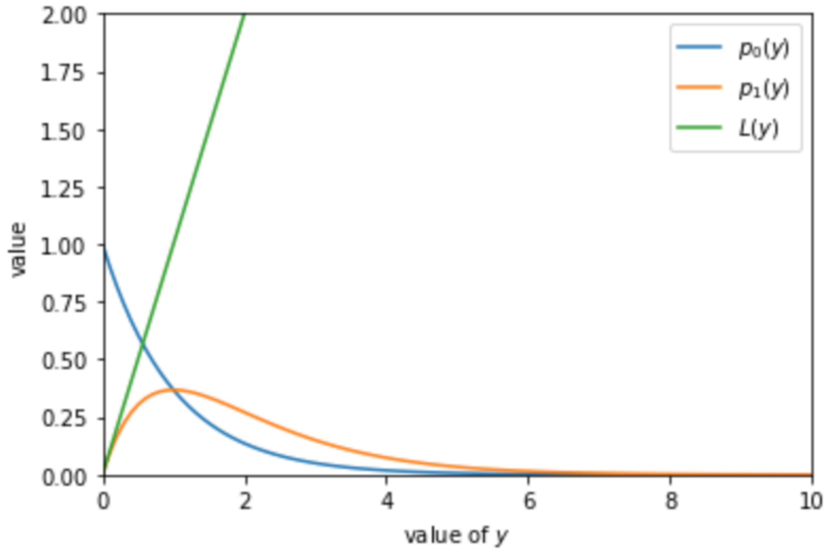
And for H_1

$$p_1(y) = \int_{-\infty}^{\infty} p(y-s)p(y)ds = \int_0^y e^{s-y}e^{-s}ds = ye^{-y}, \quad y \in (0, +\infty)$$

Thus, the likelihood ratio is:

$$L(y) = \frac{p_1(y)}{p_0(y)} = y, \quad y > 0$$

The plot of the pdf and the likelihood ratio is :



Solution (b):

First, write the false-alarm probability as:

$$\begin{aligned} P_0(p_1(Y) > \eta \cdot p_0(Y)) &= P_0(L(Y) > \eta) \\ &= P_0(Y > \eta) \\ &= \int_0^{\eta} e^{-y} \cdot dy = e^{-\eta} \end{aligned}$$

Then the threshold η Can be obtained by:

$$\begin{aligned} P_F(\delta_{NP}) &\quad = e^{-\eta} = \alpha \\ \Longrightarrow &\quad \eta = -ln(\alpha) \end{aligned}$$

So the detection probability is :

$$P_D\left(\delta_{NP}\right) = P_1(Y > \eta) = \int_{\eta}^{\infty} ye^{-y}dy = (\eta + 1)e^{-\eta} = \alpha(1 - \ln \alpha), \quad 0 < \alpha < 1$$

Solution (c):

Similar as (a), first write the pdf:

$$p_0(y) = \prod_{k=1}^{n} p\left(y_k\right) = \prod_{k=1}^{n} e^{-y_k}, \quad 0 < \min\left\{y_1, y_2, \ldots, y_n\right\}$$

And for H_1:

$$\begin{aligned} p_1(y) &= \int_{-\infty}^{\infty}\left[\prod_{k=1}^{n} p\left(y_k - s\right)\right] p(s) ds = \int_0^{\min\{y_1, y_2, \ldots, y_n\}}\left[\prod_{k=1}^{n} e^{s - y_k}\right] e^{-s} ds \\ &= \frac{p_0(y)}{n-1}\left[e^{(n-1)\min\{y_1, y_2, \ldots, y_n\}} - 1\right], \quad 0 < \min\left\{y_1, y_2, \ldots, y_n\right\} \end{aligned}$$

Thus, the likelihood ratio can be written as:

$$L(y)=\frac{1}{n-1}\left[e^{(n-1)\min\{y_1,y_2,...,y_n\}}-1\right],\quad 0<\min\{y_1,y_2,\ldots,y_n\}$$

Solution (d):

Firstly, write the false alarm probability as:

$$\begin{aligned}P_F(\delta_{NP}) =&P_0(L(Y)>\eta)=P_0\left(\min\{Y_1,Y_2,\ldots,Y_n\}>\eta'\equiv\frac{\log((n-1)\eta+1)}{n-1}\right)\\ &=P_0\left(\bigcap_{k=1}^{n}(Y_k>\eta')\right)=\prod_{k=1}^{n}P_0(Y_k>\eta')=\prod_{k=1}^{n}e^{-\eta'}=e^{-n\eta'}\end{aligned}$$

Where the $\eta'=-\frac{1}{n}\log\alpha$, and the threshold η is:

$$\eta=\frac{e^{(n-1)\eta'}-1}{n-1}=\frac{\alpha^{-(n-1)/n}-1}{n-1}$$

4th problem

15. Consider the composite hypothesis testing problem:

$$H_0: Y \text{ has density } p_0(y)=\tfrac{1}{2}e^{-|y|},\quad y\in\mathbb{R}$$

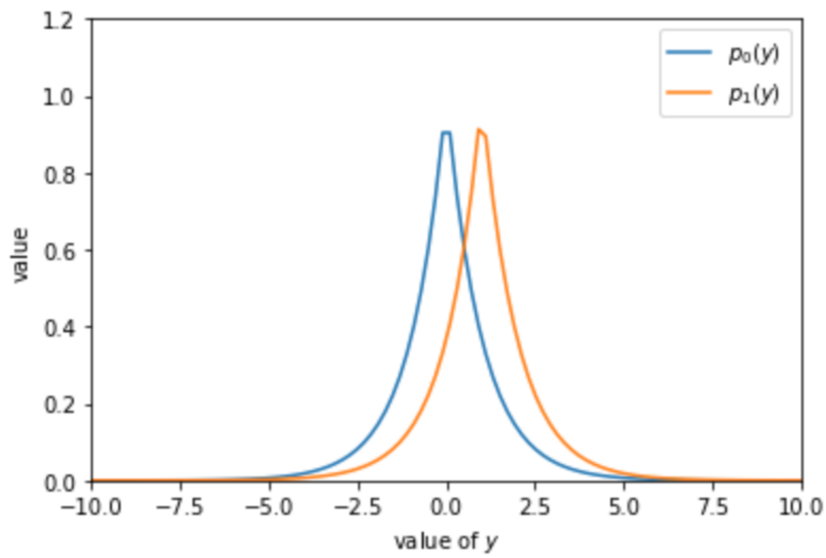
versus

$$H_1: Y \text{ has density } p_\theta(y)=\tfrac{1}{2}e^{-|y-\theta|}, y\in\mathbb{R},\theta>0.$$

(a) Describe the locally most powerful α-level test and derive its power function.

(b) Does a uniformly most powerful test exist? If so, find it and derive its power function. If not, find the generalized likelihood ratio test for H_0 versus H_1.

The pdf under $H_0,\ H_1$



Solution (a):

According to $(\mathrm{II.E.31})$, the α-level $locally\ most\ powerful$ (LMP) test is :

$$\tilde{\delta}_{lo}(y)=\begin{cases}1 & \text{if } \left.\frac{\partial p_\theta(y)}{\partial\theta}\right|_{\theta=0}>\eta p_0(y)\\ \gamma, & \text{if } \left.\frac{\partial p_\theta(y)}{\partial\theta}\right|_{\theta=0}=\eta p_0(y)\\ 0 & \text{if } \left.\frac{\partial p_\theta(y)}{\partial\theta}\right|_{\theta=0}<\eta p_0(y)\end{cases}$$

And :

$$\left.\frac{\partial p_\theta(y)}{\partial\theta}\right|_{\theta=0}=\begin{cases}\frac{1}{2}e^{-y}, & y>0\\ -\frac{1}{2}e^{-y}, & y<0\end{cases}$$

$$\Longrightarrow \frac{\left.\frac{\partial p_\theta(y)}{\partial\theta}\right|_{\theta=0}}{p_0(y)}=\begin{cases}1, & y>0\\ 0, & y<0\end{cases}=sgn(y)$$

Thus,

$$\tilde{\delta}_{lo}(y)=\begin{cases}1 & \text{if } \operatorname{sgn}(y)>\eta\\ \gamma, & \text{if } \operatorname{sgn}(y)=\eta\\ 0 & \text{if } \operatorname{sgn}(y)<\eta\end{cases}$$

Since $\alpha=E_0[\tilde{\delta}_{lo}(y)]$ The randomization γ is :

$$\gamma=\frac{\alpha-P_0(\operatorname{sgn}(Y)>\eta)}{P_0(\operatorname{sgn}(Y)=\eta)}$$

this implies the range of η, i.e., $\eta\in\{-1,1\}$, notice that $\gamma>0$, and $P_0(\operatorname{sgn}(Y)>\eta)=\begin{cases}0, & \eta=1\\ 1/2, & \eta=-1\end{cases}$ and $P_0(\operatorname{sgn}(Y)=\eta)=1/2, \forall\eta\in\{-1,1\}$, so the randomization is :

$$\gamma=\begin{cases}2\alpha & \text{if } 0<\alpha<1/2\\ 2\alpha-1 & \text{if } 1/2\le\alpha<1\end{cases}$$

The LMP rules is:

$$\delta_{lo}(y)=\begin{cases}2\alpha & \text{if } y>0\\ 0 & \text{if } y\le 0\text{'}\end{cases}\quad \text{for } 0<\alpha<1/2$$

$$\delta_{lo}(y)=\begin{cases}1 & \text{if } y\ge 0\\ 2\alpha-1 & \text{if } y<0\end{cases}\quad \text{for } 1/2\le\alpha<1$$

The power function is the detection probability :

$$P_D\left(\tilde{\delta}_{lo};\theta\right)=P_\theta(\operatorname{sgn}(Y)>\eta)+\gamma P_\theta(\operatorname{sgn}(Y)=\eta)$$

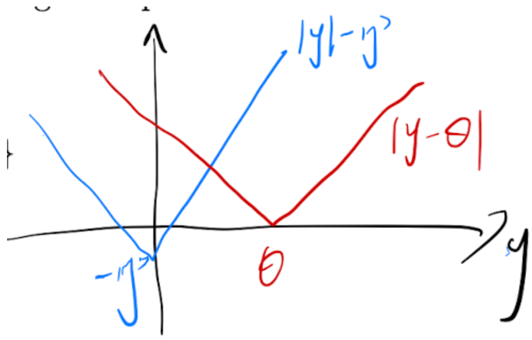
$$=\begin{cases}2\alpha\int_0^\infty\frac{1}{2}e^{-|y-\theta|}dy & \text{if } 0<\alpha<1/2\\ \int_0^\infty\frac{1}{2}e^{-|y-\theta|}dy+(2\alpha-1)\int_{-\infty}^0\frac{1}{2}e^{-|y-\theta|}dy & \text{if } 1/2\le\alpha<1\end{cases}$$

$$=\begin{cases}\alpha\left(2-e^{-\theta}\right) & \text{if } 0<\alpha<1/2\\ 1+(\alpha-1)e^{-\theta} & \text{if } 1/2\le\alpha<1\end{cases}$$

Solution (b):

the NP critical region is

$$\Gamma_\theta = \left\{ \frac{p_\theta(y)}{p_0(y)} > \eta \right\} = \{|y| - |y-\theta| > \eta'\}$$

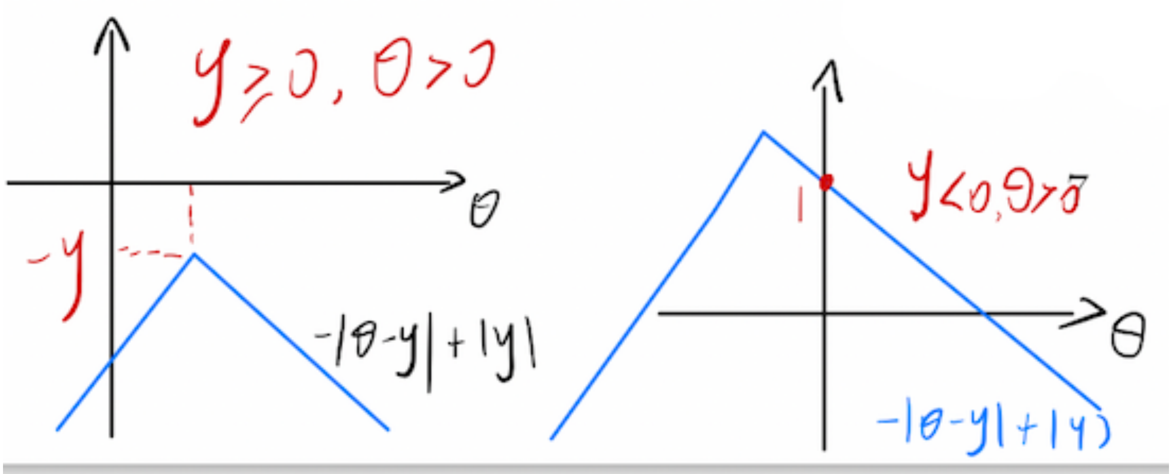


According to the figure shown above, the NP critical region is:

$$\Gamma_\theta = \begin{cases} (-\infty, \infty) & \text{if } \eta' < -\theta \\ \left(\left(\frac{\eta'+\theta}{2}\right), \infty\right) & \text{if } -\theta \le \eta' \le \theta \\ \phi & \text{if } \eta' > \theta \end{cases}$$

Obviously, the critical region is relevant to the parameter θ, so the UMP test doesn't exist.

The generalized likelihood ratio test uses this statistic:



$$\sup_{\theta>0} e^{|y|-|y-\theta|} = \exp\left\{\sup_{\theta>0}(|y| - |y-\theta|)\right\}$$
$$= \begin{cases} 1 & \text{if } y < 0 \\ e^y & \text{if } y \ge 0 \end{cases}$$

5th problem

16. In Section B, we formulated and solved the binary Bayesian hypothesis-testing problem. Generalize this formulation and solution to M hypotheses for $M > 2$.

Similar as $(\mathrm{II.B.6})$, the overall cost function can be written as:

$$r(\delta)=\sum_{j=0}^{M-1}\sum_{i=0}^{M-1}\pi_jC_{ij}P_j\left(\Gamma_i\right)=\sum_{i=0}^{M-1}\left[\sum_{j=0}^{M-1}\pi_jC_{ij}P_j\left(\Gamma_i\right)\right]$$
$$=\sum_{i=0}^{M-1}\left[\sum_{j=0}^{M-1}\pi_jC_{ij}\int_{\Gamma_i}p_j(y)\mu(dy)\right]=\sum_{i=0}^{M-1}\int_{\Gamma_i}\left[\sum_{j=0}^{M-1}\pi_jC_{ij}p_j(y)\right]\mu(dy)$$

in order to minimize the cost $r(\delta)$, each region Γ_i should enpower the $\left[\sum_{j=0}^{M-1}\pi_jC_{ij}p_j(y)\right]$ to be the smallest for $i\in \text{Indexs of all hypothesis}$ then the bayes rule decision regions are defined as:

$$\Gamma_i=\left\{y\in\Gamma \mid \sum_{j=0}^{M-1}\pi_jC_{ij}p_j(y)=\min_{0\le k\le M-1}\sum_{j=0}^{M-1}\pi_jC_{kj}p_j(y)\right\}$$

6th problem

19. Consider the following pair of hypotheses concerning a sequence $Y_1, Y_2, \ldots, Y_n$ of independent random variables

$$H_0: Y_k \sim \mathcal{N}(\mu_0, \sigma_0^2), \quad k=1,2,\ldots,n$$

versus

$$H_1: Y_k \sim \mathcal{N}(\mu_1, \sigma_1^2), \quad k=1,2,\ldots,n$$

where μ_0, μ_1, σ_0^2, and σ_1^2 are known constants.

(a) Show that the likelihood ratio can be expressed as a function of the parameters μ_0, μ_1, σ_0^2, and σ_1^2, and the quantities $\sum_{k=1}^n Y_k^2$ and $\sum_{k=1}^n Y_k$.

(b) Describe the Neyman-Pearson test for the two cases $(\mu_0 = \mu_1, \sigma_1^2 > \sigma_0^2)$ and $(\sigma_0^2=\sigma_1^2, \mu_1 > \mu_0)$.

(c) Find the threshold and ROC's for the case $\mu_0=\mu_1, \sigma_1^2>\sigma_0^2$ with $n=1$.

---

solution (a):

the pdf for $H_0,\ H_1$ are:

$$p_1(y)=\prod_{k=1}^{n}\frac{1}{\sqrt{2\pi\sigma_0}}e^{-(y_k-\mu_0)^2/2\sigma_0^2}$$

and

$$p_0(y)=\prod_{k=1}^{n}\frac{1}{\sqrt{2\pi\sigma_1}}e^{-(y_k-\mu_1)^2/2\sigma_1^2}$$

So the likehood ratio is:

$$L(y) = \frac{\prod_{k=1}^{n} \frac{1}{\sqrt{2\pi}\sigma_1} e^{-(y_k-\mu_1)^2/2\sigma_1^2}}{\prod_{k=1}^{n} \frac{1}{\sqrt{2\pi}\sigma_0} e^{-(y_k-\mu_0)^2/2\sigma_0^2}}$$
$$= \left(\frac{\sigma_0}{\sigma_1}\right)^n e^{\frac{n}{2}\left(\frac{\mu_0^2}{\sigma_0^2}-\frac{\mu_1^2}{\sigma_1^2}\right)} e^{\left(\frac{1}{2\sigma_0^2}-\frac{1}{2\sigma_1^2}\right)\sum_{k=1}^{n} y_k^2} e^{\left(\frac{\mu_1}{\sigma_1^2}-\frac{\mu_0}{\sigma_0^2}\right)\sum_{k=1}^{n} y_k}$$

Solution (b):

1. if $\left(\mu_0 = \mu_1, \sigma_1^2 > \sigma_0^2\right)$

in this case, the likehood ratio become:

$$L(y) = \left(\frac{\sigma_0}{\sigma_1}\right)^n e^{\left(\frac{1}{2\sigma_0^2}-\frac{1}{2\sigma_1^2}\right)\cdot\sum_{k=1}^{n}(y_k-\mu)^2}$$

So the Neyman-Pearson test operates by comparing $\sum_{k=1}^{n}(y_k-\mu)^2$ to a threshold

2. if $\left(\sigma_0^2 = \sigma_1^2, \mu_1 > \mu_0\right)$

in this case, the likehood ratio become:

$$L(y) = exp\left(2(\mu_1-\mu_0)\cdot\sum_{k=1}^{n} y_k + n\cdot(\mu_0-\mu_1)\right)$$

So the Neyman-Pearson test operates by comparing $\sum_{k=1}^{n} y_k$ to a threshold

Solution (c):

in this case, the likehood ratio become:

$$L(y) = \left(\frac{\sigma_0}{\sigma_1}\right)^n e^{\left(\frac{1}{2\sigma_0^2}-\frac{1}{2\sigma_1^2}\right)\cdot\sum_{k=1}^{n}(y_k-\mu)^2}$$
$$= \frac{\sigma_0}{\sigma_1}\cdot e^{\left(\frac{1}{2\sigma_0^2}-\frac{1}{2\sigma_1^2}\right)\cdot(y_1-\mu)^2}$$

The form of the NP test is:

$$\delta_{NP}(y) = \begin{cases} 1 & \text{if } (y_1-\mu)^2 \geq \eta' \\ 0 & \text{if } (y_1-\mu)^2 < \eta' \end{cases}$$

write the false-alarm probability as:

$$\begin{aligned} P_0(p_1(Y) > \eta\cdot p_0(Y)) &= P_0(L(Y) > \eta) \\ &= P_0((y_1-\mu)^2 > \eta^{'}) \\ &= 2\left[1-\Phi\left(\frac{\sqrt{\eta'}}{\sigma_0}\right)\right] \end{aligned}$$

since $P_0(p_1(Y) > \eta\cdot_0(Y)) = P_F(\delta_{NP}) = \alpha$, so for the α the threshold η is set as :

$$\eta' = \left[\sigma_0 \Phi^{-1}\left(1-\frac{\alpha}{2}\right)\right]^2$$

and the detection probability is :

$$\begin{aligned}
P_D\left(\delta_{NP}\right) = P_1(\Gamma_1) &= P_1\left(Y_1 - \mu < -\sqrt{\eta'} \cup Y_1 - \mu > \sqrt{\eta'}\right) \\
&= 1 - P_1\left(-\sqrt{\eta'} \le Y_1 - \mu \le \sqrt{\eta'}\right) = 2\left[1 - \Phi\left(\frac{\sqrt{\eta'}}{\sigma_1}\right)\right] \\
&= 2\left[1 - \Phi\left(\frac{\sigma_0}{\sigma_1}\Phi^{-1}\left(1-\frac{\alpha}{2}\right)\right)\right], \quad 0 < \alpha < 1
\end{aligned}$$

HW-2

1

1. Show that the filter with impulse response

$$\tilde{h}_k = \begin{cases} \tilde{s}_{n-k}, & 0 \le k \le n-1 \\ 0, & \text{otherwise} \end{cases}$$

with $\underline{\tilde{s}} = \mathbf{\Sigma}_N^{-1}\underline{s}$ has maximum output signal-to-noise ratio at time n among all linear filters, when the input signal is $\underline{s} = (s_1, \ldots, s_n)^T$ and the input noise has zero mean and covariance $\mathbf{\Sigma}_N$.

let $\underline{h}_n = [h_{n,1}, \ldots, h_{n,l}, \ldots, h_{n,n}]^T$ be the vector of the impulse response of a general discrete-time linear filter at time n .

The SNR is :

$$SNR = \frac{\left|\sum_{l=1}^{n} h_{n,l} s_l\right|^2}{E\left\{\left(\sum_{l=1}^{n} h_{n,l} N_l\right)^2\right\}} = \frac{\left|\underline{h}_n^T \underline{s}\right|^2}{\underline{h}_n^T \mathbf{\Sigma}_N \underline{h}_n}$$

Since $\mathbf{\Sigma}_N = \mathbf{\Sigma}_N^{1/2}\mathbf{\Sigma}_N^{1/2}$, the SNR can be further wirtten as:

$$SNR = \frac{\left|\left(\mathbf{\Sigma}_N^{1/2}\underline{h}_n\right)^T \mathbf{\Sigma}_N^{-1/2}\underline{s}\right|^2}{\left\|\mathbf{\Sigma}_N^{1/2}\underline{h}_n\right\|^2}$$

Then according to the Schwarz Inequality $\left|\underline{x}^T\underline{y}\right| \le \|x\|\|y\|$, we have:

$$SNR \le \left\|\mathbf{\Sigma}_N^{1/2}\underline{s}\right\|^2$$

And notice the if and only if $\mathbf{\Sigma}_N^{1/2}\underline{h}_n = \lambda\mathbf{\Sigma}_N^{-1/2}\underline{s}$, the $SNR = \left\|\mathbf{\Sigma}_N^{1/2}\underline{s}\right\|^2$, so the max ratio is obtained if and only if $\underline{h}_n = \lambda\mathbf{\Sigma}_N^{-1}\underline{s}$, when $\lambda = 1$, $\underline{h}_n = \underline{\tilde{s}}$, then $\tilde{h}_k = \underline{h}_{n,n-k} = \tilde{s}_{n-k}, \quad 0 \le k \le n-1$

2

3. Consider the M-ary decision problem: ($\Gamma = \mathbb{R}^n$)

$$\begin{aligned} H_0 : \underline{Y} &= \underline{N} + \underline{s}_0 \\ H_1 : \underline{Y} &= \underline{N} + \underline{s}_1 \\ &\vdots \\ H_{M-1} : \underline{Y} &= \underline{N} + \underline{s}_{M-1}, \end{aligned}$$

where $\underline{s}_0, \underline{s}_1, \ldots, \underline{s}_{M-1}$ are known signals with equal energies, $\| \underline{s}_0 \|^2 = \| \underline{s}_1 \|^2 = \cdots = \| \underline{s}_{M-1} \|^2$.

(a) Assuming $\underline{N} \sim \mathcal{N}(\underline{0}, \sigma^2 \mathbf{I})$, find the decision rule achieving minimum error probability when all hypotheses are equally likely.

(b) Assuming further that the signals are orthogonal, show that the minimum error probability is given by

$$P_e = 1 - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} [\Phi(x)]^{M-1} e^{-(x-d)^2/2} dx$$

where $d^2 = \| \underline{s}_0 \|^2 / \sigma^2$.

(a)

Follow the conclusion of the 5-th problem in HW-1, the critical regions are:

$$\Gamma_i = \left\{ y \in \Gamma \mid \sum_{j=0}^{M-1} \pi_j C_{ij} p_j(y) = \min_{0 \le k \le M-1} \sum_{j=0}^{M-1} \pi_j C_{kj} p_j(y) \right\}$$

$$\Gamma_i = \left\{ y \in \Gamma \mid \sum_{j=0}^{M-1} p_j(y) - p_i(y) = \min_{0 \le k \le M-1} \left[\sum_{j=0}^{M-1} p_j(y) - p_i(y) \right] \right\}$$

$$\Longrightarrow \Gamma_i = \left\{ \underline{y} \in R^n \mid p_i(\underline{y}) = \max_{0 \le k \le M-1} p_k(\underline{y}) \right\}$$

Since p_k Follows distribution $N\left(\underline{s}_k, \sigma^2 \mathbf{I}\right)$, this reduce to :

$$\Gamma_i = \left\{ \underline{y} \in R^n \mid \left\| \underline{y} - \underline{s}_i \right\|^2 = \min_{0 \le k \le M-1} \left\| y - s_k \right\|^2 \right\}$$

$$= \left\{ \underline{y} \in R^n \mid \underline{s}_i^T \underline{y} = \max_{0 \le k \le M-1} \underline{s}_k^T \underline{y} \right\}$$

(b)

since all hypotheses are equally likely, the error probability is denoted by:

$$P_e = \frac{1}{M}\sum_{k=0}^{M} P_k\left(\Gamma_k^c\right)$$

Note that :

$$\underline{y} = \max_{0\le k\le M-1} \underline{s}_k^T \underline{y} \Leftrightarrow \max_{0\le l\neq k\le M-1} \underline{s}_l^T \underline{Y} < \underline{s}_k^T \underline{Y}$$

, Therefore:

$$P_k\left(\Gamma_k^c\right) = 1 - P_k\left(\Gamma_k\right) = 1 - P_k\left(\max_{0\le l\neq k\le M-1} \underline{s}_l^T \underline{Y} < \underline{s}_k^T \underline{Y}\right)$$

Notice that $\underline{s}_1^T\underline{Y}, \underline{s}_2^T\underline{Y}, \cdots, \underline{s}_n^T\underline{Y}$ are i.i.d Gaussian random variable with the same variances $\sigma^2\|s_1\|^2$. Consider $z = \underline{s}_k^T\underline{Y}$, thus:

$$\begin{aligned} & P_k\left(\max_{0\le l\neq k\le M-1} \underline{s}_l^T \underline{Y} < \underline{s}_k^T \underline{Y}\right) \\ = & \frac{1}{\sqrt{2\pi}\sigma\left\|s_1\right\|}\int_{-\infty}^{\infty} P_k\left(\max_{0\le l\neq k\le M-1} \underline{s}_l^T \underline{Y} < z\right) e^{-\left(z-\|s_1\|^2\right)/2\sigma^2\|s_1\|^2} dz \end{aligned}$$

Where:

$$\begin{aligned} P_k\left(\max_{0\le l\neq k\le M-1} \underline{s}_L^T \underline{Y} < z\right) &= P_k\left(\bigcap_{0\le l\neq k\le M-1}\left\{\underline{s}_l^T \underline{Y} < z\right\}\right) \\ &= \prod_{0\le l\neq k\le M-1} P_k\left(\underline{s}_l^T \underline{Y} < z\right) \\ &= \left[\Phi\left(\frac{z}{\sigma\left\|s_1\right\|}\right)\right]^{M-1} \end{aligned}$$

setting $x = z/\sigma\left\|s_1\right\|$ and combining the above Eqs, the P_e is denoted by:

$$1 - P_k\left(\Gamma_k\right) = \frac{1}{\sqrt{2\pi}} \quad \int_{-\infty}^{\infty}[\Phi(x)]^{M-1} e^{-(x-d)\frac{2}{2}} dx, k = 0, \cdots, M-1$$

3

6. Suppose $\underline{Y} \sim \mathcal{N}(\underline{\mu}, \mathbf{\Sigma})$. For each $k \geq 2$, define $\hat{Y}_k = E\{Y_k | Y_1, \ldots, Y_{k-1}\}$ and $\hat{\sigma}^2_{Y_k} = \text{Var}(Y_k | Y_1, \ldots, Y_{k-1})$. Also define $\hat{Y}_1 = E\{Y_1\}$ and $\hat{\sigma}^2_{Y_1} = \text{Var}(Y_1)$. Define a sequence $I_1, I_2, \ldots, I_n$ by

$$I_k = (Y_k - \hat{Y}_k)/\hat{\sigma}_{Y_k}.$$

Show that $\underline{I} \sim \mathcal{N}(0, \mathbf{I})$, and thus that the above scheme provides whitening of $\underline{Y}$.

First, notice that $\underline{I}$ is a linear transformation of $\underline{Y}$, and is Gaussian. We only need to prove that $E\{\underline{I}\} = \underline{0}$ and $\text{cov}(\underline{I}) = \mathbf{I}$. We have:

$$E\{I_k\} = \frac{E\{Y_k\} - E\left\{\hat{Y}_k\right\}}{\hat{\sigma}_k}$$

$E\left\{\hat{Y}_k\right\}$ is an iterated expectation of Y_k; so $E\{Y_k\} = E\left\{\hat{Y}_k\right\}$And thus $E\{I_k\} = 0, k = 1, \cdots, n$.

Next need to prove that $\text{cov}(\underline{I}) = \mathbf{I}$, :

$$\text{Var}(I_k) = E\left\{I_k^2\right\} = \frac{E\left\{\left(Y_k - \hat{Y}_k\right)^2\right\}}{\hat{\sigma}^2_{Y_k}} = \frac{\hat{\sigma}^2_{Y_k}}{\hat{\sigma}^2_{Y_k}} = 1$$

$\text{for } l < k, \text{ we have :}$

$$\begin{aligned} \text{cov}(I_k, I_l) &= E\{I_k I_l\} \\ &= \frac{E\left\{\left(Y_k - \hat{Y}_k\right)\left(Y_l - \hat{Y}_l\right)\right\}}{\hat{\sigma}_{Y_k}\hat{\sigma}_{Y_l}} \end{aligned}$$

Then:

$$\begin{aligned} &E\left\{\left(Y_k - \hat{Y}_k\right)\left(Y_l - \hat{Y}_l\right)\right\} = E\left\{E\left\{\left(Y_k - \hat{Y}_k\right)\left(Y_l - \hat{Y}_l\right) \mid Y_1, \cdots, Y_{k-1}\right\}\right\} \\ &= E\left\{\left(E\{Y_k \mid Y_1, \cdots, Y_{k-1}\} - \hat{Y}_k\right)\left(Y_l - \hat{Y}_l\right)\right\} = E\left\{\left(\hat{Y}_k - \hat{Y}_k\right)\left(Y_l - \hat{Y}_l\right)\right\} = 0 \end{aligned}$$

So fo all l, we all have $\text{cov}(I_k, I_l) = 0$

In a nut shell, $E\{\underline{I}\} = \underline{0}$ and $\text{cov}(\underline{I}) = \mathbf{I}$, so we have the discussion proved.

4

7. Consider the hypothesis pair

$$H_0 : Y_k = N_k, \qquad k = 1, \dots, n$$

versus

$$H_1 : Y_k = N_k + \Theta S_k, \quad k = 1, \dots, n$$

where $\underline{N} \sim \mathcal{N}(\underline{0}, \Sigma)$, $\underline{s}$ is known, and Θ is a random variable independent of $\underline{N}$.

(a) Find the α-level Neyman-Pearson detector and ROCs assuming that Θ is a discrete random variable taking the values $+1$ and -1 with equal probabilities (i.e., $P(\Theta = +1) = P(\Theta = -1) = 1/2$).

(b) Suppose that $\Theta \sim \mathcal{N}(0, \sigma_\theta^2)$. Assuming $\Sigma = \sigma^2 \mathbf{I}$, show that the likelihood ratio is of the form

$$L(\underline{y}) = k_1 e^{k_2 \|\underline{s}^T \underline{y}\|^2}$$

where k_1 and k_2 are positive constants. Find k_2.

(a)

First write the likelihood ratio, according to $(\mathrm{III.B.63})$:

$$L(\underline{y}) = \int_\Lambda \exp\left\{\left[\underline{s}^T(\theta)\underline{y} - \frac{1}{2}\|\underline{s}(\theta)\|^2\right]/\sigma^2\right\} w(\theta)\mu(d\theta)$$

So we have $L(y) = P(\Theta = -1) \cdot L(y, \Theta = -1) + P(\Theta = 1) \cdot L(y, \Theta = 1)$, then:

$$\begin{aligned} L(y) =& P(\Theta = -1) \cdot L(y, \Theta = -1) + P(\Theta = 1) \cdot L(y, \Theta = 1) \\ =& \frac{1}{2} e^{\underline{S}^T \mathbf{\Sigma}^{-1} \underline{y} - d^2/2} + \frac{1}{2} e^{-\underline{s}^T \mathbf{\Sigma}^{-1} \underline{y} - d^2/2} \\ =& e^{-d^2/2} \cosh \underline{s}^T \mathbf{\Sigma}^{-1} \underline{y} \end{aligned}$$

Then according to the defination of $T(y)$, we have:

$$T(\underline{y}) \equiv \left|\underline{s}^T \mathbf{\Sigma}^{-1} \underline{y}\right|$$

According to $(\mathrm{III.B.27})$, $d^2 = \underline{s}^T \mathbf{\Sigma}^{-1} \underline{s}$.

the Neyman-Pearson test is of the form:

$$\tilde{\delta}_{NP}(\underline{y}) = \begin{cases} 1 & \text{if } T(\underline{y}) > \eta \\ \gamma, & \text{if } T(\underline{y}) = \eta \\ 0 & \text{if } T(\underline{y}) < \eta \end{cases}$$

To set the threshold η ,

$P_0(T(\underline{Y})>\eta)=1-P\left(-\eta \leq \underline{s}^T \boldsymbol{\Sigma}^{-1} \underline{N} \leq \eta\right)=1-\Phi(\eta / d)+\Phi(-\eta / d)=2[1-\Phi(\eta / d)]$

Where $\eta=d \Phi^{-1}(1-\alpha / 2)$

Then the detection probability is:

$$
\begin{gathered}
P_D\left(\tilde{\delta}_{NP}\right)=\frac{1}{2} P_1(T(\underline{Y})>\eta \mid \Theta=+1)+\frac{1}{2} P_1(T(\underline{Y})>\eta \mid \Theta=-1) \\
=\frac{1}{2}\left[1-P\left(-\eta \leq-d^2+\underline{s}^T \boldsymbol{\Sigma}^{-1} \underline{N} \leq \eta\right)\right]+\frac{1}{2}\left[1-P\left(-\eta \leq+d^2+\underline{s}^T \boldsymbol{\Sigma}^{-1} \underline{N} \leq \eta\right)\right] \\
=2-\Phi\left(\Phi^{-1}(1-\alpha / 2)+d\right)-\Phi\left(\Phi^{-1}(1-\alpha / 2)-d\right)
\end{gathered}
$$

(b)

according to (III.B.63) :

$$
\begin{gathered}
L(\underline{y})=\int_{\Lambda} \exp \left\{\left[\underline{s}^T(\theta) \underline{y}-\frac{1}{2}\|\underline{s}(\theta)\|^2\right] / \sigma^2\right\} w(\theta) \mu(d \theta) \\
=k_1 e^{k_2 |\underline{s}^T \underline{y}} \frac{1}{\sqrt{2 \pi v}} \int_{-\infty}^{\infty} e^{-(\theta-\mu)^2 / 2 v} d \theta \\
=k_1 e^{k_2 |\underline{s}^T \underline{y}}
\end{gathered}
$$

Where \$
v^{2}=\frac{\sigma_{\theta}^{2} n \bar{s}^{2}}{\sigma_{\theta}^{2} \sigma^{2}+n \bar{s}^{2}}, \
\mu=\frac{v^{2} \underline{s}^{T} \underline{y}}{2} \
k_{1}=\frac{v}{\sigma_{\theta}} \
k_{2}=\frac{v^{2}}{4}
\$

5

13. Consider the model

$$Y_k = \theta^{1/2} s_k R_k + N_k, \quad k = 1, \ldots, n$$

where $s_1, s_2, \ldots, s_n$ is a known signal sequence, $\theta \geq 0$ is a constant, and $R_1, R_2, \ldots, R_n, N_1, N_2, \ldots, N_n$ are i.i.d. $\mathcal{N}(0,1)$ random variables

(a) Consider the hypothesis pair

$$H_0 : \theta = 0$$
$$\text{versus}$$
$$H_1 : \theta = A$$

where A is a known positive constant. Describe the structure of the Neyman-Pearson detector.

(b) Consider now the hypothesis pair

$$H_0 : \theta = 0$$
$$\text{versus}$$
$$H_1 : \theta > 0.$$

Under what conditions on $s_1, s_2, \ldots, s_n$ does a UMP test exist?

(c) For the hypothesis pair of part (b) with $s_1, s_2, \ldots, s_n$ general, is there a *locally* optimum detector? If so, find it. If not, describe the generalized likelihood ratio test.

(a)

the hypothesis canb be transferred into:

$$\begin{aligned} &H_0 : N_k \\ &\text{versus} \\ &H_1 : A^{1/2} s_k R_k + N_k \end{aligned}$$

According to (III.B.111) :

$$T(\underline{y}) = \frac{1}{n} \underline{y}^T \boldsymbol{\Sigma}_S \underline{y}$$

Where $\boldsymbol{\Sigma}_\mathbf{S} = \operatorname{diag}\left\{As_1^2, As_2^2, \ldots, As_n^2\right\}$, then:

$$T(\underline{y}) = \sum_{k=1}^{n} \frac{As_k^2}{As_k^2+1} y_k^2$$

(b)

No, it does not exit

(c)

an LMP test can be based on the statistic:

$$T_{lo}(\underline{y}) = \sum_{k=1}^{n} s_k^2 y_k^2$$

6

15. Consider the problem of Example III.B.5 in which the amplitude sequence $a_1, a_2, \ldots, a_n$ is given by

$$a_k = Ab_k, \quad k = 1, 2, \ldots, n,$$

where $\sum_{k=1}^{n} b_k^2 = n$, and A is a positive random variable, independent of the phase Θ, having the Rayleigh density with parameter A_0; i.e.,

$$p_A(a) = (a/A_0^2)\exp\{-a^2/2A_0^2\}, \quad a \geq 0.$$

Find the Neyman-Pearson detector, including the threshold for size α, and derive an expression for the ROC's.

Under the condition that $A = a$, let $L_a(y)$ denote the likelihood ratio, then the unconditioned ratio is:

$$L(\underline{y}) = \int_0^{\infty} L_a(\underline{y}) p_A(a) da = \int_0^{\infty} e^{-na^2/4\sigma^2} I_0\left(a^2\hat{r}/\sigma^2\right) p_A(a) da$$

Where $\hat{r} \equiv r/A$, $r = \sqrt{y_c^2 + y_s^2}$, Note that:

$$\hat{r} = \sqrt{\left(\sum_{k=1}^{n} b_k \cos\left((k-1)\omega_c T_s\right) y_k\right)^2 + \left(\sum_{k=1}^{n} b_k \sin\left((k-1)\omega_c T_s\right) y_k\right)^2}$$

To get the NP rules,

$$\frac{\partial L(\underline{y})}{\partial \hat{r}} = \frac{1}{\sigma^2} \int_0^{\infty} e^{-na^2/4\sigma^2} a^2 I_0'\left(a^2\hat{r}/\sigma^2\right) p_A(a) da > 0$$

so the NP test is of the form:

$$\tilde{\delta}_{NP}(\underline{y}) = \begin{cases} 1 & \text{if } \hat{r} > \tau' \\ \gamma, & \text{if } \hat{r} = \tau' \\ 0 & \text{if } \hat{r} < \tau' \end{cases}$$

to set the τ', let $P_0\left(\hat{R} > \tau'\right) = \alpha$. According to (III.B.72), we have:

$$P_0\left(\hat{R} > \tau'\right) = e^{-(\tau')^2/n\sigma^2}$$

Then according to (III.B.72), $P_1\left(\hat{R} > \tau' \mid A = a\right) = Q(b, \tau_0)$, where $b^2 = na^2/2\sigma^2$ and $\tau_0 = \sqrt{2/n}\tau'/\sigma = \sqrt{-2\log\alpha}$. Thus, the detection probability is :

$$\begin{aligned} P_D = \int_0^\infty Q\left(\frac{a}{\sigma}\sqrt{n/2}, \tau_0\right)p_A(a)da &= \int_0^\infty \int_{\tau_0}^\infty xe^{-(x^2+na^2/2\sigma^2)/2} I_0\left(x\frac{a}{\sigma}\sqrt{n/2}\right)\frac{a}{A_0^2}e^{-a^2/2A_0^2}dxda \\ &= \int_{\tau_0}^\infty xe^{-x^2/2}\int_0^\infty \frac{a}{A_0^2}e^{-a^2/2a_0^2}I_0\left(x\frac{a}{\sigma}\sqrt{n/2}\right)dadx \end{aligned}$$

Where $a_0 = \sqrt{\frac{2A_0^2\sigma^2}{nA_0^2+2\sigma^2}}$. Further let $y = a/a_0$, the P_D Becomes:

$$P_D = \frac{a_0^2}{A_0^2}\int_{\tau_0}^\infty xe^{-x^2(1-b_0^2)/2}\int_0^\infty ye^{-(y^2+b_0^2x^2)/2}I_0(b_0xy)dydx = \frac{a_0^2}{A_0^2}\int_{\tau_0}^\infty xe^{-x^2(1-b_0^2)/2}Q(b_0x, 0)dx$$

Where $b_0^2 = na_0^2/2\sigma^2$.

Since $Q(b, 0) = 1, \ \forall b$, and $1 - b_0^2 = a_0^2/A_0^2$

$$P_D = \frac{a_0^2}{A_0^2}\int_{\tau_0}^\infty xe^{-x^2(1-b_0^2)/2}dx = e^{-\tau_0^2(1-b_0^2)/2} = \exp\left(-\frac{\tau_0^2}{2\left(1+\frac{nA_0^2}{2\sigma^2}\right)}\right) = \alpha^{x_0}$$

Where $x_0 = \frac{1}{1+\frac{nA_0^2}{2\sigma^2}}$

7

16. Find the $\underline{\hat{S}}$ solving

$$e^{(\underline{\hat{S}}^T\underline{y}-\frac{1}{2}\|\underline{\hat{S}}\|^2)/\sigma^2} = \int_{\mathbf{R}^n} e^{(\underline{s}^T\underline{y}-\frac{1}{2}\|\underline{s}\|^2)/\sigma^2}p_{\underline{S}}(\underline{s})d\underline{s}$$

for the case in which $p_{\underline{S}}$ is the $\mathcal{N}(\underline{0}, \mathbf{\Sigma}_S)$ density.

The right side of the eqution shares the same form of the (III.B.84). Thus the equation is rewritten into:

$$2\underline{\hat{S}}^T\underline{y} - \|\underline{\hat{S}}\|^2 = \underline{y}^T\mathbf{\Sigma_S}\left(\sigma^2\mathbf{I} + \mathbf{\Sigma_S}\right)^{-1}\underline{y} + \sigma^2\sum_{k=1}^n \log\left(\frac{\sigma^2}{\sigma^2+\lambda_k}\right)$$

$\lambda_1, \lambda_2, \ldots, \lambda_n$ are the eigenvalues of $\mathbf{\Sigma}_S$

Rewrite the equation:

$$\begin{aligned}\|\underline{\hat{S}} - \underline{y}\|^2 = \|\underline{y}\|^2 - \underline{y}^T \mathbf{\Sigma_S} \left(\sigma^2 \mathbf{I} + \mathbf{\Sigma_S}\right)^{-1} \underline{y} - \sigma^2 \sum_{k=1}^{n} \log\left(\frac{\sigma^2}{\sigma^2 + \lambda_k}\right) \\ \equiv \sigma^2 \left[\underline{y}^T \left(\sigma^2 \mathbf{I} + \mathbf{\Sigma_S}\right)^{-1} \underline{y} - \sum_{k=1}^{n} \log\left(\frac{\sigma^2}{\sigma^2 + \lambda_k}\right)\right]\end{aligned}$$

and solved by:

$$\underline{\hat{S}} = \underline{y} \pm \frac{\sigma}{\|\underline{v}\|} \left[\underline{y}^T \left(\sigma^2 \mathbf{I} + \mathbf{\Sigma_S}\right)^{-1} \underline{y} - \sum_{k=1}^{n} \log\left(\frac{\sigma^2}{\sigma^2 + \lambda_k}\right)\right]^{1/2} \underline{v}$$

(For any nonzero vector $\underline{v}$)

HW-3

1

13. Suppose we toss a coin n independent times and define an observation sequence

$$Y_k = \begin{cases} 1 & \text{if the } k\text{th outcome is heads} \\ 0 & \text{if the } k\text{th outcome is tails} \end{cases}$$

$k = 1, 2, \ldots, n$. Let $\theta = P(Y_k = 1), k = 1, \ldots, n$.

(a) Find an MVUE of θ.

(b) Find the ML estimate of θ. Find its bias and variance.

(c) Compute the Cramér-Rao lower bound and compare with results from (a) and (b).

(a)

First we have:

$$p_\theta(y) = \theta^{T(\underline{y})} (1 - \theta)^{(n - T(\underline{y}))}$$

Where :

$$T(\underline{y}) = \sum_{k=1}^{n} y_k$$

and notice:

$$\phi = \log\left(\theta/(1-\theta)\right)$$
$$C(\phi) = e^{n\phi}$$

Thus, any unbiased function of T is an MVUE for θ . Also notice that $E_\theta\{T(\underline{Y})\} = n\theta$, an estimate is given by:

$$\hat{\theta}_{MV}(\underline{y}) = \frac{T(\underline{y})}{n} = \frac{1}{n}\sum_{k=1}^{n} y_k$$

(b)

According to $(\mathrm{IV.D.2})$

$$\hat{\theta}_{ML}(y) = \arg\left\{\max_{\theta\in\Lambda} p_\theta(y)\right\}$$
$$= \arg\left\{\max_{0<\theta<1} \theta^{T(\underline{y})}(1-\theta)^{(n-T(\underline{y}))}\right\} = T(\underline{y})/n = \hat{\theta}_{MV}(\underline{y})$$

the bias:

$$E_\theta\left\{\hat{\theta}_{ML}(\underline{Y})\right\} = \theta$$

According to $(\mathrm{IV.D.9})$ the variance:

$$Var(\hat{\theta}_{ML}(\underline{Y})) = \theta(1-\theta)/n$$

(c)

According to $(\mathrm{IV.C.22})$ (condition 5) and $(\mathrm{IV.C.26})$:

$$I_\theta = -E_\theta\left\{\frac{\partial^2}{\partial\theta^2}\log p_\theta(Y)\right\} = -E_\theta\left\{-\frac{T(\underline{Y})}{\theta^2} - \frac{n-T(\underline{Y})}{(1-\theta)^2}\right\}$$
$$= \frac{n}{\theta} + \frac{n}{1-\theta} = \frac{n}{\theta(1-\theta)}$$

So the CRLB is:

$$CRLB = \frac{1}{I_\theta} = \frac{\theta(1-\theta)}{n} = \mathrm{Var}_\theta\left(\hat{\theta}_{ML}(\underline{Y})\right)$$

2

15. Suppose Y is Poisson. Find the ML estimate of its rate. Compute the bias, variance, and Cramér-Rao lower bound.

First we have:

$$p_\theta(y) = \frac{e^{-\theta}\theta^y}{y!}, \quad y \in 0, 1, \ldots.$$

Then:

$$\begin{aligned}\frac{\partial}{\partial\theta}\log p_\theta(y) &= \frac{\partial}{\partial\theta}(-\theta + y\log\theta) \\ &= -1 + \frac{y}{\theta}\end{aligned}$$

Thus:

$$\hat{\theta}_{ML}(y) = y$$

Then the bias:

$$E_\theta\left\{\hat{\theta}_{ML}(Y)\right\} = \mathrm{Var}_\theta\left(\hat{\theta}_{ML}(Y)\right) = \theta$$

Finally according to $(\mathrm{IV.C.22})$ (condition 5) and $(\mathrm{IV.C.26})$:

$$I_\theta = -E_\theta\left\{\frac{\partial^2}{\partial\theta^2}\log p_\theta(Y)\right\} = \frac{E_\theta\{Y\}}{\theta^2} = \frac{1}{\theta}$$

The $CRLB$ is thus:

$$CRLB = \frac{1}{I_\theta} = \theta$$

3

20. Consider the observation model

$$Y_k = \theta^{1/2} s_k R_k + N_k, \quad k = 1, 2, \ldots, n$$

where $s_1, s_2, \ldots, s_n$ is a known signal, $N_1, N_2, \ldots, N_n$, R_1, $R_2, \ldots, R_n$ are i.i.d. $\mathcal{N}(0,1)$ random variables, and $\theta \geq 0$ is an unknown parameter.

(a) Find the likelihood equation for estimating θ from $Y_1, Y_2, \ldots, Y_n$.

(b) Find the Cramér-Rao lower bound on the variance of unbiased estimates of θ.

(c) Suppose $s_1, s_2, \ldots, s_n$ is a sequence of $+1$'s and -1's. Find the MLE of θ explicitly.

(d) Compute the bias and variance of your estimate from (c), and compare the latter with the Cramér-Rao lower bound.

(a)

because Y_k is i.i.d and follows $\mathcal{N}\left(0, 1+\theta s_k^2\right)$, the likelihood equation for estimating θ is :

$$\begin{aligned} \frac{\partial}{\partial \theta} \log p_\theta(\underline{y}) &= \sum_{k=1}^{n} \frac{\partial}{\theta}\left\{-\frac{1}{2}\log\left(1+\theta s_k^2\right) - \frac{y_k^2}{2\left(1+\theta s_k^2\right)}\right\} \\ &= -\frac{1}{2}\sum_{k=1}^{n}\left\{\frac{s_k^2}{1+\theta s_k^2} - \frac{y_k^2 s_k^2}{\left(1+\theta s_k^2\right)^2}\right\} \end{aligned}$$

from which the likelihood equation becomes:

$$\sum_{k=1}^{n} \frac{s_k^2\left(y_k^2 - 1 - \hat{\theta}_{ML}(\underline{y}) s_k^2\right)}{\left(1+\hat{\theta}_{ML}(\underline{y}) s_k^2\right)^2} = 0$$

(b)

Similar as the 2nd problem, write I_θ:

$$\begin{aligned} I_\theta = -E_\theta\left\{\frac{\partial^2}{\partial \theta^2}\log p_\theta(\underline{Y})\right\} &= \sum_{k=1}^{n}\left\{\frac{s_k^4 E_\theta\left\{Y_k^2\right\}}{\left(1+\theta s_k^2\right)^3} - \frac{s_k^4}{2\left(1+\theta_k^2\right)^2}\right\} \\ &= \frac{1}{2}\sum_{k=1}^{n}\frac{s_k^4}{\left(1+\theta s_k^2\right)^2} \end{aligned}$$

So the $CRLB$ is :

$$CRLB = \frac{1}{I_\theta} = \frac{2}{\sum_{k=1}^{n} \frac{s_k^4}{\left(1+\theta s_k^2\right)^2}}$$

(c)

according to the Eq. above:

$$\sum_{k=1}^{n} \frac{s_k^2 \left(y_k^2 - 1 - \hat{\theta}_{ML}(\underline{y}) s_k^2\right)}{\left(1 + \hat{\theta}_{ML}(\underline{y}) s_k^2\right)^2} = 0$$

And $s_k^2 = 1$, then $\hat{\theta}_{ML}(\underline{y})$ Becomes:

$$\hat{\theta}_{ML}(\underline{y}) = \left(\frac{1}{n} \sum_{k=1}^{n} y_k^2\right) - 1$$

(d)

the bias:

$$E_\theta \left\{\hat{\theta}_{ML}(\underline{Y})\right\} = \left(\frac{1}{n} \sum_{k=1}^{n} E_\theta \left\{Y_k^2\right\}\right) - 1 = \theta$$

The variance:

$$\operatorname{Var}_\theta \left(\hat{\theta}_{ML}(\underline{Y})\right) = \frac{1}{n^2} \sum_{k=1}^{n} \operatorname{Var}_\theta \left(Y_k^2\right) = \frac{1}{n^2} \sum_{k=1}^{n} 2(! + \theta)^2 = \frac{2(1+\theta)^2}{n}$$

the MLE is an MVUE in this case.

4

22. Suppose $\theta > 0$ is a parameter of interest and that given θ, $Y_1, \ldots, Y_n$ is a set of i.i.d. observations with marginal distribution function

$$F_\theta(y) = [F(y)]^{1/\theta}, \quad -\infty < y < \infty,$$

where F is a known distribution function with pdf f.

(a) Show that

$$\hat{\theta}_{MV}(\underline{y}) = -\frac{1}{n}\sum_{k=1}^{n} \log F(y_k)$$

is an MVUE of θ.

(b) Suppose now that θ is replaced by a random variable Θ drawn at random using the prior density

$$w(\theta) = c^m \exp(-c/\theta)/(\Gamma(m)\theta^{m+1}), \quad \theta > 0,$$

where $c > 0$ and $m > 1$ are constants. Use the fact that $E\{\Theta\} = c/(m-1)$ to show that the MMSE estimator of Θ from $Y_1, \ldots, Y_n$ is

$$\hat{\theta}_{\mathrm{MMSE}}(\underline{y}) = \left(c - \sum_{k=1}^{n} \log F(y_k)\right) / \ (m+n-1).$$

(c) Compare $\hat{\theta}_{\mathrm{MV}}$ and $\hat{\theta}_{\mathrm{MMSE}}$ with regard to the role of the prior information.

(a)

Note:

$$p_\theta(\underline{y}) = \exp\left\{\sum_{k=1}^{n} \log F(y_k)/\theta\right\}$$

Also:

$$E_\theta\left\{\sum_{k=1}^{n} \log F(Y_k)\right\} = nE_\theta\{\log F(Y_1)\}$$

$$= \frac{n}{\theta}\int_{-\infty}^{\infty} \log F(Y_1)[F(y_1)]^{(1-\theta)/\theta} f(y_1) dy_1$$

Let $x = \log F(y_1)$, and note $d\log F(y_1) = \frac{f(y_1)}{F(y_1)} dy_1$, $[F(y_1)]^{1/\theta} = \exp\{\log F(y_1)/\theta\}$

. Then:

$$E_\theta\left\{\sum_{k=1}^{n}\log F\left(Y_k\right)\right\}=\frac{n}{\theta}\int_{-\infty}^{0} x e^{x/\theta} dx=-n\theta$$

Thus:

$$E_\theta\left\{\hat{\theta}_{MV}(\underline{Y})\right\}=\theta$$

which implies that $\hat{\theta}_{MV}$ is an MVUE since it is an unbiased function of a complete sufficient statistic.

(b)

It is straightforward to see that $w(\theta \mid y)$ is of the same form as the prior with c replaced by $c-\sum_{k=1}^{n}\log F\left(y_k\right)$, and m replaced by $n+m$. Thus, $E\{\Theta \mid Y\}$is of the same form as the prior with c replaced by $c-\sum_{k=1}^{n}\log F\left(y_k\right)$, and m replaced by $n+m$. So:

$$E\{\Theta \mid Y\}=\frac{c-\sum_{k=1}^{n}\log F\left(Y_k\right)}{m+n-1}$$

(c)

In this example, the prior and posterior distributions have the same form. The only change is that the parameters of that distribution are updated as new data is observed. A prior with this property is said to be a reproducing prior. The prior parameters, c and m, can be thought of as coming from an earlier sample of size m. As n becomes large compared to m, the importance of these prior parameters in the estimate diminishes. Note that $\sum_{k=1}^{n}\log F\left(Y_k\right)$ behaves like $nE\left\{\log F\left(Y_1\right)\right\}$ for large n. Thus, with $n \gg m$, the estimate is appropximately given by the MVUE of Part a. Altenatively, with $m \gg n$, the estimate is approximately the prior mean, $c/(m-1)$. Between these two extremes, there is a balance between prior and observed information.

5

23. Suppose we observe

$$Y_k = A\sin\left(\frac{k\pi}{2} + \Phi\right) + N_k, \quad k = 1, \ldots, n$$

where $\underline{N} \sim \mathcal{N}(\underline{0}, \sigma^2 \mathbf{I})$ and n is even.

(a) Suppose A and Φ are nonrandom with $A \geq 0$ and $\Phi \in [-\pi, \pi]$. Find their ML estimates.

(b) Suppose A and Φ are random and independent with priors

$$w_\Phi(\phi) = \begin{cases} \frac{1}{\pi}, & -\pi \leq \phi \leq \pi \\ 0, & \text{otherwise} \end{cases}$$

$$w_A(a) = \begin{cases} (a/\beta^2)e^{-a^2/2\beta^2} & \text{if } a \geq 0 \\ 0 & \text{if } a < 0 \end{cases}$$

where β is known. Assuming A and Φ are independent of $\underline{N}$, find the MAP estimates of A and Φ.

(c) Under what conditions are the estimates from (a) and (b) approximately equal?

(a)

The log-likelihood function is

$$\log p(\underline{y} \mid A, \phi) = -\frac{1}{2\sigma^2}\sum_{k=1}^{n}\left[y_k - A\sin\left(\frac{k\pi}{2} + \phi\right)\right]^2 - \frac{n}{2}\log\left(2\pi\sigma^2\right).$$

The likelihood equations are thus:

$$\sum_{k=1}^{n}\left[y_k - \hat{A}\sin\left(\frac{k\pi}{2} + \hat{\phi}\right)\right]\sin\left(\frac{k\pi}{2} + \hat{\phi}\right) = 0$$

and

$$\hat{A}\sum_{k=1}^{n}\left[y_k - \hat{A}\sin\left(\frac{k\pi}{2} + \hat{\phi}\right)\right]\cos\left(\frac{k\pi}{2} + \hat{\phi}\right) = 0.$$

These equations are solved by the estimates:

$$\hat{A}_{ML} = \sqrt{y_c^2 + y_s^2}$$

$$\hat{\phi}_{ML} = \tan^{-1}\left(\frac{y_c}{y_s}\right)$$

where

$$y_c = \frac{1}{n}\sum_{k=1}^{n} y_k \cos\left(\frac{k\pi}{2}\right) \equiv \frac{1}{n}\sum_{k=1}^{n/2}(-1)^k y_{2k},$$

$$y_s = \frac{1}{n}\sum_{k=1}^{n} y_k \sin\left(\frac{k\pi}{2}\right) \equiv \frac{1}{n}\sum_{k=1}^{n/2}(-1)^{k+1} y_{2k-1}$$

(b)

Appending the prior to the above problem yields MAP estimates:

$$\hat{\phi}_{MAP} = \hat{\phi}_{ML},$$

$$\hat{A}_{MAP} = \frac{\hat{A}_{ML} + \sqrt{\left(\frac{r}{n}\right)^2 + \frac{2(1+\alpha)\sigma^2}{n}}}{1+\alpha},$$

where $\alpha \equiv \frac{2\sigma^2}{n\beta^2}$.

(c)

Note that, when $\beta \to \infty$ (and the prior "diffuses"), the MAP estimate of A does not approach the MLE of A. However, as $n \to \infty$, the MAP estimate does approach the MLE.

HW-4

1

3. Repeat Exercise 2 for the situation in which each $\underline{s}_k$ is allowed to be a function of the past measurement; i.e., $\underline{s}_k$ can be a function of $\underline{Y}_0^k$. (So, for example, $\{\underline{s}_k\}_{k=0}^{\infty}$ could be a sequence of feedback controls.)

when $\underline{Y}_0^k$ is given the $\underline{S}_k$ is then constant. Then we can write the filter:

$$\underline{\hat{X}}_{t+1|t} = \mathbf{F}_t \underline{\hat{X}}_{t|t} + E\left\{\mathbf{\Gamma}_t \underline{s}_t \mid \underline{Y}_0^t\right\} = \mathbf{F}_t \underline{\hat{X}}_{t|t} + \mathbf{\Gamma}_t \underline{s}_t$$

So that:

$$\mathbf{\Sigma}_{t+1|t} = \mathbf{F}_t \mathbf{\Sigma}_{t|t} \mathbf{F}_t^T + \mathbf{G}_t \mathbf{Q}_t \mathbf{G}_t^T + \operatorname{Cov}\left(\mathbf{\Gamma}_t \underline{s}_t \mid \underline{Y}_0^t\right) = \mathbf{F}_t \mathbf{\Sigma}_{t|t} \mathbf{F}_t^T + \mathbf{G}_t \mathbf{Q}_t \mathbf{G}_t^T$$

Note that the second of these two equations is the same as when there is no measurement feedback. That is to say, the measurement feedback has no effect on the measurement update equations. The conditional statistics of $\underline{X}_t$ and $\underline{Y}_t$ given $\underline{Y}_0^{t-1}$ are Gaussian. Thus, the measurement update is unchanged from the case of no measurement feedback since it depends only on this joint Gaussian property and the linearity of measurement equation.

2

4. Suppose we return to the original Kalman-Bucy model, but allow for correlation between the state and measurement noises; i.e., assume everything as before except

$$\operatorname{Cov}(\underline{U}_k, \underline{V}_l) = \begin{cases} \mathbf{C}_k, & k = l \\ \mathbf{0}, & k \neq l \end{cases}$$

where $\mathbf{C}_k$ is a matrix of appropriate dimension. Show that the Kalman predictor is given by

$$\hat{\underline{X}}_{t+1|t} = \mathbf{F}_t \hat{\underline{X}}_{t|t-1} + \mathbf{K}_t(\underline{Y}_t - \mathbf{H}_t \hat{\underline{X}}_{t|t-1})$$

with

$$\hat{\underline{X}}_{0|-1} = \underline{m}_0,$$

where

$$\mathbf{K}_t = (\mathbf{F}_t \mathbf{\Sigma}_{t|t-1} \mathbf{H}_t^T + \mathbf{G}_t \mathbf{C}_t)(\mathbf{H}_t \mathbf{\Sigma}_{t|t-1} \mathbf{H}_t^T + \mathbf{R}_t)^{-1}$$

and

$$\begin{aligned} \mathbf{\Sigma}_{t+1|t} &= \mathbf{F}_t \mathbf{\Sigma}_{t|t-1} \mathbf{F}_t^T \\ &\quad - \mathbf{K}_t(\mathbf{F}_t \mathbf{\Sigma}_{t|t-1} \mathbf{H}_t^T + \mathbf{G}_t \mathbf{C}_t) + \mathbf{G}_t \mathbf{Q}_t \mathbf{G}_t^T \end{aligned}$$

with

$$\mathbf{\Sigma}_{0|-1} = \operatorname{Cov}(\underline{X}_0).$$

Note that although $\underline{U}_t$ and $\underline{V}_t$ are dependent, the Gaussian vector $\underline{U}_t' \equiv \underline{U}_t - \mathbf{C}_t \mathbf{R}_t^{-1} \underline{V}_t$ is independent of $\underline{V}_t$. To take advantage of this, we may add the zero quantity $\mathbf{C}_t \mathbf{R}_t^{-1} \left[\underline{Y}_t - \mathbf{H}_t \underline{X}_t - \underline{V}_t\right]$ to the t^{th} state input, which yields the equivalent state equation

$$\underline{X}_{t+1} = \mathbf{F}_t \underline{X}_t + \mathbf{G}_t \underline{U}_t' + \mathbf{G}_t \mathbf{C}_t \mathbf{R}_t^{-1} \left(\underline{Y}_t - \mathbf{H}_t \underline{X}_t\right)$$

So, we have an equivalent problem with independent state and measurement noises, but with the measurement feedback term $\mathbf{G}_t \mathbf{C}_t \mathbf{R}_t^{-1} \underline{Y}_t$, and with the new state matrix $\left(\mathbf{F}_t - \mathbf{G}_t \mathbf{C}_t \mathbf{R}_t^{-1} \mathbf{H}_t\right)$. We also have a different correlation matrix for the state input, since

$$\operatorname{Cov}\left(\underline{U}_t'\right) = \mathbf{Q}_t - \mathbf{C}_t \mathbf{R}_t^{-1} \mathbf{C}_t^T$$

Applying the result of Exercise 3 and eliminating the measurement update equations yields the given result.

3

6. Consider the standard Kalman-Bucy model with states $\underline{X}_k$ and observations $\underline{Y}_k$. Suppose $0 \leq j \leq t$ and we wish to estimate $\underline{X}_j$ from $\underline{Y}_0^t$. Consider the estimator defined recursively (in t) by

$$\hat{\underline{X}}_{j|t} = \hat{\underline{X}}_{j|t-1} + \mathbf{K}_t^a(\underline{Y}_t - \mathbf{H}_t\hat{\underline{X}}_{t|t-1})$$

where

$$\mathbf{K}_t^a = \mathbf{\Sigma}_{t|t-1}^a\mathbf{H}_t^T[\mathbf{H}_t\mathbf{\Sigma}_{t|t-1}^a\mathbf{H}_t^T + \mathbf{R}_t]^{-1}$$

and

$$\mathbf{\Sigma}_{t+1|t}^a = \mathbf{\Sigma}_{t|t-1}^a[\mathbf{F}_t - \mathbf{K}_t\mathbf{H}_t]^T$$

with

$$\mathbf{\Sigma}_{j|j-1}^a = \mathbf{\Sigma}_{j|j-1}$$

where $\mathbf{H}_t, \hat{\underline{X}}_{t|t-1}, \mathbf{\Sigma}_{t|t-1}, \mathbf{R}_t, \mathbf{F}_t$, and $\mathbf{K}_t$ are as in the one-step prediction problem.

(a) Show that $\mathbf{\Sigma}_{t|t-1}^a = E\{(\underline{X}_j - \hat{\underline{X}}_{j|t-1})\underline{X}_t^T\}$.

(b) Show that $\hat{\underline{X}}_{j|t} = E\{\underline{X}_j|\underline{Y}_0^t\}$.

(a)

a. This result follows by induction on t. We first note that, for $t = j$, the given equality follows by definition. From the state equation we have that, for $k > j$,
$E\left\{\left(\underline{X}_j - \hat{\underline{X}}_{j|k}\right)\underline{X}_{k+1}^T\right\} = E\left\{\left(\underline{X}_j - \hat{\underline{X}}_{j|k}\right)\left(\mathbf{F}_k\underline{X}_k + \mathbf{G}_k\underline{U}_k\right)^T\right\} = E\left\{\left(\underline{X}_j - \hat{\underline{X}}_{j|k}\right)\underline{X}_k^T\right\}\mathbf{F}_k^T$
where we use the fact that $\underline{U}_k$ has zero mean and is independent of both $\underline{X}_j$ and $\hat{\underline{X}}_{j|k}$. Now, on applying the recursion for $\hat{\underline{X}}_{j|k}$ to this equation we have
$E\left\{\left(\underline{X}_j - \hat{\underline{X}}_{j|k}\right)\underline{X}_{k+1}^T\right\} = E\left\{\left(\underline{X}_j - \hat{\underline{X}}_{j|k-1}\right)\underline{X}_k^T\right\}\mathbf{F}_k^T - \mathbf{K}_k^a E\left\{\left(\underline{Y}_k - \mathbf{H}_k\hat{X}_{k|k-1}\right)\underline{X}_k^T\right\}\mathbf{F}_k^T$
We now assume that the given equality is true for $t = k$. From this and the definition of $\mathbf{K}_k^a$, we then have

$$\begin{aligned}
E\left\{\left(\underline{X}_j - \hat{\underline{X}}_{j|k}\right)\underline{X}_{k+1}^T\right\} &= \mathbf{\Sigma}_{k|k-1}^a\left[\mathbf{I} - \mathbf{H}_k^T\left(\mathbf{H}_k\mathbf{\Sigma}_{k|k-1}\mathbf{H}_k^T + \mathbf{R}_k\right)^{-1}E\left\{\left(\underline{Y}_k - \mathbf{H}_k\hat{\underline{X}}_{k|k-1}\right)\underline{X}_k^T\right\}\right] \\
&= \mathbf{\Sigma}_{k|k-1}^a\left[\mathbf{I} - \mathbf{H}_k^T\left(\mathbf{H}_k\mathbf{\Sigma}_{k|k-1}\mathbf{H}_k^T + \mathbf{R}_k\right)^{-1}\mathbf{H}_k E\left\{\left(\underline{X}_k - \hat{X}_{k|k-1}\right)\underline{X}_k^T\right\}\right]\mathbf{F}_k^T \\
&= \mathbf{\Sigma}_{k|k-1}^a\left[\mathbf{I} - \mathbf{H}_k^T\left(\mathbf{H}_k\mathbf{\Sigma}_{k|k-1}\mathbf{H}_k^T + \mathbf{R}_k\right)^{-1}\mathbf{H}_k\mathbf{\Sigma}_{k|k-1}\right]\mathbf{F}_k^T = \mathbf{\Sigma}_{k|k-1}^a\left[\mathbf{I} - \mathbf{H}_k^T\mathbf{K}_k^T\right]\mathbf{F}_k^T
\end{aligned}$$

where the last equality follows by definiton of $\mathbf{K}_k$. Applying the recursion for $\mathbf{\Sigma}_{k+1|k}^a$, we have

$$E\left\{\left(\underline{X}_j - \hat{\underline{X}}_{j|k}\right)\underline{X}_{k+1}^T\right\} = \mathbf{\Sigma}_{k+1|k}^a$$

which shows that the given equation for $t = k + 1$. The induction principle thus gives the desired result.

(b)

Note that $\underline{X}_j$ and $\underline{Y}_t$ are jointly Gaussian conditioned on $\underline{Y}_0^{t-1}$.

$$\operatorname{Cov}\left(\underline{X}_j, \underline{Y}_t \mid \underline{Y}_0^{t-1}\right)\left[\operatorname{Cov}\left(\underline{Y}_t \mid \underline{Y}_0^{t-1}\right)\right]^{-1} = \mathbf{K}_t^a$$

Since $\underline{Y}_t = \mathbf{H}_t \underline{X}_t + \underline{V}_t$, the result from Part a. implies this equality.

4

9. Consider the observation model:

$$Y_k = N_k + \Theta s_k, \quad k = 1, 2, \ldots,$$

where $N_1, N_2, \ldots$ are i.i.d. $\mathcal{N}(0, \sigma^2)$ random variables, $\Theta \sim \mathcal{N}(\mu, \nu^2)$ is independent of $N_1, N_2, \ldots$, and $s_1, s_2, \ldots$ is a known sequence. Let $\hat{\theta}_n$ denote the MMSE estimate of Θ given $Y_1, \ldots, Y_n$. Find recursions for $\hat{\theta}_n$ and for the minimum mean-squared error, $E\{(\hat{\theta}_n - \Theta)^2\}$, by recasting this problem as a Kalman filtering problem.

This is the Kalman-Bucy problem with all dimensions equal to unity, $\mathbf{F}_k \equiv 1, \mathbf{G}_k \equiv \mathbf{Q}_k \equiv 0, \mathbf{H}_k = s_k, \mathbf{R}_k \equiv \sigma^2, \underline{m}_0 \equiv \mu, \mathbf{\Sigma}_0 \equiv v^2, \hat{\underline{X}}_{n+1|n} \equiv \hat{\underline{X}}_{n|n} \equiv \hat{\theta}_n$, and $\mathbf{\Sigma}_{n|n} \equiv \mathbf{\Sigma}_{n|n} \equiv E\left\{\left(\hat{\theta}_n - \Theta\right)^2\right\}$. The desired recursions thus follow by eliminating either set of updates from (V.B.14) - (V.B.16). The resulting estimate is the same as that found in Example IV.B.2.