Communication and Detection Theory

Signal and Information Processing Laboratory
Institut für Signal- und Informationsverarbeitung

Spring Semester 2017

Prof. Dr. A. Lapidoth

Model Answers to Exercise 12 of May 16, 2017

http://www.isi.ee.ethz.ch/teaching/courses/cdt

Problem 1 ***Hypothesis Testing with Two Observations***

(i) The problem of guessing H based on $\mathbf{Y}_1$ is of the kind addressed in Section 20.14 (with σ^2 being 1, J being 2, and the prior being uniform). Using the results of that section we conclude that

$$\langle \mathbf{Y}_1, \boldsymbol{\mu} \rangle_{\mathrm{E}}$$

forms a one-dimensional sufficient statistic and that an optimal rule is to guess "$H = 0$" when $\langle \mathbf{Y}_1, \boldsymbol{\mu} \rangle_{\mathrm{E}}$ is positive and to guess "$H = 1$" otherwise.

(ii) To find an optimal rule for guessing H based on Y_2, we compute the likelihood ratio. We first consider the case $\alpha \geq 0$:

$$\begin{aligned}
\frac{f_{Y_2|H=0}(y_2)}{f_{Y_2|H=1}(y_2)} &= \frac{e^{-(y_2-\alpha)}\,\mathrm{I}\{y_2 - \alpha \geq 0\}}{e^{-(y_2+\alpha)}\,\mathrm{I}\{y_2 + \alpha \geq 0\}} \\
&= \frac{e^{-(y_2-\alpha)}\,\mathrm{I}\{y_2 \geq \alpha\}}{e^{-(y_2+\alpha)}\,\mathrm{I}\{y_2 \geq -\alpha\}} \\
&= \begin{cases} \frac{0}{0} & \text{if } -\infty < y_2 < -\alpha, \\ 0 & \text{if } -\alpha \leq y_2 < \alpha, \\ e^{2\alpha} & \text{if } \alpha \leq y_2 < \infty, \end{cases} \qquad \alpha \geq 0, \\
&= \begin{cases} 1 & \text{if } -\infty < y_2 < -\alpha, \\ 0 & \text{if } -\alpha \leq y_2 < \alpha, \\ e^{2\alpha} & \text{if } \alpha \leq y_2 < \infty, \end{cases} \qquad \alpha \geq 0, \qquad (1)
\end{aligned}$$

where we have used the convention of (20.39). Similarly for $\alpha \leq 0$:

$$\begin{aligned}
\frac{f_{Y_2|H=0}(y_2)}{f_{Y_2|H=1}(y_2)} &= \frac{e^{-(y_2-\alpha)}\,\mathrm{I}\{y_2 - \alpha \geq 0\}}{e^{-(y_2+\alpha)}\,\mathrm{I}\{y_2 + \alpha \geq 0\}} \\
&= \frac{e^{-(y_2-\alpha)}\,\mathrm{I}\{y_2 \geq \alpha\}}{e^{-(y_2+\alpha)}\,\mathrm{I}\{y_2 \geq -\alpha\}} \\
&= \begin{cases} 1 & \text{if } -\infty < y_2 < \alpha, \\ \infty & \text{if } \alpha \leq y_2 < -\alpha, \\ e^{2\alpha} & \text{if } -\alpha \leq y_2 < \infty, \end{cases} \qquad \alpha \leq 0. \qquad (2)
\end{aligned}$$

An optimal rule for the case $\alpha > 0$ is to guess "$H = 1$" if $-\alpha \le y_2 < \alpha$, and to guess "$H = 0$" otherwise (because $e^{2\alpha}$ is greater than one when α is positive). An optimal rule for the case $\alpha \le 0$ is to guess "$H = 0$" if $\alpha \le y_2 < -\alpha$ and to guess "$H = 1$" otherwise.

(iii) A two dimensional sufficient statistic is the pair

$$\langle \mathbf{Y}_1, \boldsymbol{\mu} \rangle_{\mathrm{E}}, Y_2$$

because conditional on H the observations $\mathbf{Y}_1$ and Y_2 are independent; see Proposition 22.4.6.

(iv) Define

$$T_1 = \langle \mathbf{Y}_1, \boldsymbol{\mu} \rangle_{\mathrm{E}}$$

so

$$f_{T_1|H=0}(t_1) = \frac{1}{\sqrt{2\pi \|\boldsymbol{\mu}\|^2}} e^{-\frac{(t_1 - \|\boldsymbol{\mu}\|^2)^2}{2\|\boldsymbol{\mu}\|^2}}$$

and

$$f_{T_1|H=1}(t_1) = \frac{1}{\sqrt{2\pi \|\boldsymbol{\mu}\|^2}} e^{-\frac{(t_1 + \|\boldsymbol{\mu}\|^2)^2}{2\|\boldsymbol{\mu}\|^2}}.$$

Consequently,

$$\frac{f_{T_1|H=0}(t_1)}{f_{T_1|H=1}(t_1)} = e^{2t_1}, \quad t_1 \in \mathbb{R}.$$

By the conditional independence

$$\begin{aligned}\frac{f_{T_1,Y_2|H=0}(t_1, y_2)}{f_{T_1,Y_2|H=1}(t_1, y_2)} &= \frac{f_{T_1|H=0}(t_1)}{f_{T_1|H=1}(t_1)} \frac{f_{Y_2|H=0}(y_2)}{f_{Y_2|H=1}(y_2)} \\ &= e^{2t_1} \frac{f_{Y_2|H=0}(y_2)}{f_{Y_2|H=1}(y_2)},\end{aligned}$$

which can be computed using (1) and (2). For the case $\alpha > 0$ an optimal decision rule is to guess "$H = 1$" if $-\alpha \le y_2 < \alpha$ or $e^{2(\alpha + t_1)} < 1$, i.e., to guess "$H = 1$" if $-\alpha \le y_2 < \alpha$ or $\alpha + t_1 < 0$. For the case $\alpha \le 0$ an optimal rule is to guess "$H = 0$" if $\alpha \le y_2 < -\alpha$ or $e^{2(\alpha+t_1)} > 1$, i.e., to guess "$H = 0$ if $\alpha \le y_2 < -\alpha$ or $\alpha + t_1 > 0$.

Problem 2 — *Optimality Does Not Imply Sufficiency*

(i) We shall establish the sufficiency by showing that for every nondegenerate prior (π_0, π_1) the *a posteriori* distribution of H given Y_1, Y_2, and Θ is computable from $(Y_1 + Y_2, \Theta, \pi_0, \pi_1)$ (Proposition 20.12.4).

$$\begin{aligned}&\Pr[H = 0 \,|\, Y_1 = y_1, Y_2 = y_2, \Theta = \theta] \\ &= \frac{\pi_0 f_{Y_1,Y_2|H=0,\Theta=\theta}(y_1, y_2)}{\pi_0 f_{Y_1,Y_2|H=0,\Theta=\theta}(y_1, y_2) + \pi_1 f_{Y_1,Y_2|H=1,\Theta=\theta}(y_1, y_2)} \\ &= \frac{\pi_0 \frac{1}{2\pi\theta^2} \exp\Bigl(-\frac{(y_1-1)^2}{2\theta^2} - \frac{(y_2-1)^2}{2\theta^2}\Bigr)}{\pi_0 \frac{1}{2\pi\theta^2} \exp\Bigl(-\frac{(y_1-1)^2}{2\theta^2} - \frac{(y_2-1)^2}{2\theta^2}\Bigr) + \pi_1 \frac{1}{2\pi\theta^2} \exp\Bigl(-\frac{(y_1+1)^2}{2\theta^2} - \frac{(y_2+1)^2}{2\theta^2}\Bigr)} \\ &= \frac{\pi_0 \exp\Bigl(\frac{y_1+y_2}{\theta^2}\Bigr)}{\pi_0 \exp\Bigl(\frac{y_1+y_2}{\theta^2}\Bigr) + \pi_1 \exp\Bigl(-\frac{y_1+y_2}{\theta^2}\Bigr)}. \qquad (3)\end{aligned}$$

Similarly,

$$\Pr[H=1 \,|\, Y_1=y_1, Y_2=y_2, \Theta=\theta] = \frac{\pi_1 \exp\Bigl(-\frac{y_1+y_2}{\theta^2}\Bigr)}{\pi_1 \exp\Bigl(-\frac{y_1+y_2}{\theta^2}\Bigr) + \pi_0 \exp\Bigl(\frac{y_1+y_2}{\theta^2}\Bigr)}, \tag{4}$$

thus demonstrating that the a posteriori distribution of H is computable from (y_1+y_2, θ) and the prior (π_0, π_1).

(ii) Let us write

$$f_{\mathbf{Y}|H=0,\Theta=\sigma_i}(y_1,y_2) = h^{(i)}\bigl(\|\mathbf{y}\|^2\bigr)\, g_0^{(i)}(t), \quad i=0,1$$
$$f_{\mathbf{Y}|H=1,\Theta=\sigma_i}(y_1,y_2) = h^{(i)}\bigl(\|\mathbf{y}\|^2\bigr)\, g_1^{(i)}(t), \quad i=0,1$$

where $t = y_1 + y_2$;

$$h^{(i)}(\xi) = \frac{1}{2\pi\sigma_i^2} \exp\Bigl(-\frac{\xi+2}{2\sigma_i^2}\Bigr);$$
$$g_0^{(i)}(t) = \exp\Bigl(\frac{t}{\sigma_i^2}\Bigr);$$
$$g_1^{(i)}(t) = \exp\Bigl(-\frac{t}{\sigma_i^2}\Bigr).$$

When Θ is not observed,

$$\frac{f_{\mathbf{Y}|H=0}(y_1,y_2)}{f_{\mathbf{Y}|H=1}(y_1,y_2)} = \frac{\rho_0 h^{(0)}\bigl(\|\mathbf{y}\|^2\bigr) g_0^{(0)}(t) + \rho_1 h^{(1)}\bigl(\|\mathbf{y}\|^2\bigr) g_0^{(1)}(t)}{\rho_0 h^{(0)}\bigl(\|\mathbf{y}\|^2\bigr) g_1^{(0)}(t) + \rho_1 h^{(1)}\bigl(\|\mathbf{y}\|^2\bigr) g_1^{(1)}(t)},$$

and this — when $\sigma_0^2 \neq \sigma_1^2$ — depends both on $t (= y_1+y_2)$ and $\|\mathbf{y}\|^2 (= y_1^2 + y_2^2)$.

(iii) For a uniform prior we have by (3) that when Θ is observed

$$\begin{aligned}
&\Bigl(\Pr[H=0 \,|\, Y_1=y_1, Y_2=y_2, \Theta=\theta] > \frac{1}{2}\Bigr) \\
\iff &\Bigl(\exp\Bigl(\frac{y_1+y_2}{\theta^2}\Bigr) > \exp\Bigl(-\frac{y_1+y_2}{\theta^2}\Bigr)\Bigr) \\
\iff &\Bigl(\frac{y_1+y_2}{\theta^2} > -\frac{y_1+y_2}{\theta^2}\Bigr) \\
\iff &\bigl(y_1+y_2 > 0\bigr).
\end{aligned}$$

Consequently, for a uniform prior it is optimal to guess "$H=0$" whenever $y_1+y_2>0$. Since this rule is implementable also when Θ is not observed, and since the optimal rule when Θ is not observed cannot outperform the optimal rule when it is, this rule is also optimal when Θ is not observed.

Problem 3 — *Covariance Matrices*

Denote $\mathsf{A} = \begin{pmatrix} 5 & 0 \\ 0 & -1 \end{pmatrix}$, $\mathsf{B} = \begin{pmatrix} 5 & 1 \\ 2 & 2 \end{pmatrix}$, $\mathsf{C} = \begin{pmatrix} 2 & 10 \\ 10 & 1 \end{pmatrix}$, and $\mathsf{D} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$.

A cannot be a covariance matrix because the entry "-1" cannot be the variance of a random variable.

B cannot be a covariance matrix because it is not symmetric.

C cannot be a covariance matrix because for $\boldsymbol{\alpha} = (1,-1)^\mathsf{T}$ we have that $\boldsymbol{\alpha}^\mathsf{T}\mathsf{C}\boldsymbol{\alpha} < 0$ and thus C is

not positive semidefinite.
D is the only possible covariance matrix, since it is symmetric and its two eigenvalues 0 and 2 are both nonnegative.

Problem 4 — *Multivariate Gaussians*

If Z is a Gaussian random variable, then it is also a Gaussian 1-vector; see Section 23.6.1 and in particular (i). And since the vector $(Z, Z)^{\mathsf{T}}$ can be written as a linear transformation of the random 1-vector (Z)

$$(Z, Z)^{\mathsf{T}} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} (Z),$$

it follows that $(Z, Z)^{\mathsf{T}}$ is the result of applying a deterministic linear transformation to a Gaussian vector, and is thus a Gaussian vector (Proposition 23.6.3).
We next turn to the canonical representation of this vector. Denote the vector by $\mathbf{Z}$, so $\mathbf{Z} = (Z, Z)^{\mathsf{T}}$. Its covariance matrix is

$$\mathsf{K}_{\mathbf{ZZ}} = \sigma^2 \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix},$$

whose eigenvalues are 0 and $2\sigma^2$ with corresponding eigenvectors $(1, -1)^{\mathsf{T}}$ and $(1, 1)^{\mathsf{T}}$. Normalizing the eigenvectors to unit norm we obtain that $\mathsf{K}_{\mathbf{ZZ}}\,\mathsf{U} = \mathsf{U}\Lambda$, where

$$\mathsf{U} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}, \quad \Lambda = \sigma^2 \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}.$$

The canonical representation is thus

$$\begin{pmatrix} Z \\ Z \end{pmatrix} - \begin{pmatrix} \mathsf{E}[Z] \\ \mathsf{E}[Z] \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & \sqrt{2}\sigma \end{pmatrix} \begin{pmatrix} W^{(1)} \\ W^{(2)} \end{pmatrix}.$$

Problem 5 — *Manipulating Gaussians*

In matrix notation, the vector $(Y, Z)^{\mathsf{T}}$ is $\mathsf{A}\mathbf{W}$, where $\mathbf{W}$ is a standard Gaussian 5-vector and the matrix A is

$$\mathsf{A} = \begin{pmatrix} 3 & 4 & -2 & 1 & -1 \\ 1 & -4 & -2 & 3 & -1 \end{pmatrix}.$$

Consequently, this vector is a centered Gaussian of covariance matrix $\mathsf{A}\mathsf{A}^{\mathsf{T}}$, i.e.

$$\begin{pmatrix} 31 & -5 \\ -5 & 31 \end{pmatrix}.$$

(i) By symmetry it is not difficult to see that the distribution of Y is identical to that of X, so $Y \sim \mathcal{N}(0,1)$. Formally this can be proved by considering the cumulative distribution function:

$$\begin{aligned}
&\Pr\big[Y \leq \xi\big] \\
&\quad = \Pr\big[H=1\big]\,\Pr\big[Y \leq \xi \,\big|\, H=1\big] + \Pr\big[H=-1\big]\,\Pr\big[Y \leq \xi \,\big|\, H=-1\big] \\
&\quad = \Pr\big[H=1\big]\,\Pr\big[X \leq \xi \,\big|\, H=1\big] + \Pr\big[H=-1\big]\,\Pr\big[X \geq -\xi \,\big|\, H=-1\big] \\
&\quad = \frac{1}{2}\Pr\big[X \leq \xi\big] + \frac{1}{2}\Pr\big[X \geq -\xi\big] \\
&\quad = \frac{1}{2}F_X(\xi) + \frac{1}{2} - \frac{1}{2}F_X(-\xi), \quad \xi \in \mathbb{R},
\end{aligned}$$

where the second equality follows from the relation $Y = HX$ and the third because X and H are independent. Taking the derivative of the cummulative distribution function yields that the density of Y is equal to the density of X.

(ii) No, X and Y are *not* correlated. Indeed, since they are both of zero mean, their correlation is $\mathsf{E}[XY]$, which is zero because

$$\begin{aligned}
\mathsf{E}[XY] &= \mathsf{E}\big[HX^2\big] \\
&= \mathsf{E}[H]\,\mathsf{E}\big[X^2\big] \\
&= 0,
\end{aligned}$$

where the second equality follows because H and X are independent, and the third because H is of zero mean.

(iii) Since X and Y are of equal law and both are $\mathcal{N}(0,1)$,

$$\Pr\big[|X| \geq 1\big] = \Pr\big[|Y| \geq 1\big] = 2\,\mathcal{Q}(1).$$

(iv) Since $|X| = |Y|$ (deterministically),

$$\begin{aligned}
\Pr\big[|X| \geq 1,\ |Y| \geq 1\big] &= \Pr\big[|X| \geq 1\big] \\
&= 2\,\mathcal{Q}(1).
\end{aligned}$$

(v) The random variables X and Y are *not independent* because, as seen above,

$$\Pr\big[|X| \geq 1,\ |Y| \geq 1\big] \neq \Pr\big[|X| \geq 1\big]\,\Pr\big[|Y| \geq 1\big].$$

(Had X and Y been independent, then so would have been $|X|$ and $|Y|$, and the above inequality shows that this is not the case.)

(vi) The vector $(X,Y)^{\mathsf{T}}$ is not a Gaussian vector because its components are uncorrelated but not independent. This vector is thus an example of a vector whose components are Gaussian and that is yet not a Gaussian vector.

