UNIVERSITY OF TORONTO
Department of Electrical and Computer Engineering
ECE1639F Fall 2008
Solutions to Problem Set 1

1. Problem 1.1

(i)

$$\begin{aligned} f_X(x) &= \alpha e^{-\alpha x} \quad x \geq 0 \\ f_Y(y) &= \beta e^{-\beta y} \quad y \geq 0 \end{aligned}$$

$$\begin{aligned} z &= X+Y \\ \Rightarrow F_Z(z) &= \int\int_{x+y\leq z} f_X(x)f_Y(y)dxdy \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{z-x} f_X(x)f_Y(y)dxdy \end{aligned}$$

$$\begin{aligned} f_Z(z) &= \int_{-\infty}^{\infty} f_X(x)f_Y(z-x)dx \\ &= \int_0^z \alpha\beta e^{-\alpha x}e^{-\beta(z-x)}dx \\ &= \alpha\beta e^{-\beta z}\int_0^z e^{(\beta-\alpha)x}dx = \alpha\beta e^{-\beta z}(e^{(\beta-\alpha)z}-1) = \frac{\alpha\beta}{\beta-\alpha}(e^{-\alpha z}-e^{-\beta z}) \quad z \geq 0 \end{aligned}$$

To find $E(X|Z)$, first find $f_{X,Z}(x,y)$.
But $\begin{bmatrix} X \\ Z \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} X \\ Y \end{bmatrix}$ or $\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} X \\ Z \end{bmatrix}$
Thus

$$f_{X,Z}(x,z) = f_{X,Y}(x,z-x) = f_X(x)f_Y(z-x) = \alpha\beta e^{-\alpha x}e^{-\beta(z-x)} \quad z \geq x \geq 0$$

$$\begin{aligned} f_{X|Z}(x|z) &= \frac{\alpha\beta e^{-\alpha x}e^{-\beta(z-x)}}{\frac{\alpha\beta}{\beta-\alpha}(e^{-\alpha z}-e^{-\beta z})} \quad z \geq x \geq 0 \\ &= (\beta-\alpha)\tfrac{e^{-\alpha x}e^{-\beta(z-x)}}{e^{-\alpha z}-e^{-\beta z}} \quad z \geq x \geq 0 \end{aligned}$$

To verify that it is a probability density function, we compute

$$\begin{aligned} \int_0^z f_{X|Z}(x|z)dx &= \int_0^z (\beta-\alpha)\frac{e^{-\alpha x}e^{-\beta(z-x)}}{e^{-\alpha z}-e^{-\beta z}}dx \\ &= \int_0^z \frac{(\beta-\alpha)e^{-\beta z}}{(e^{-\alpha z}-e^{-\beta z})}\int_0^z e^{-(\alpha-\beta)x}dx \\ &= \frac{e^{-\beta z}}{(e^{-\alpha z}-e^{-\beta z})}(e^{-(\alpha-\beta)z}-1) \\ &= 1 \end{aligned}$$

$$
\begin{aligned}
E(X|Z=z) &= \int_0^z x(\beta-\alpha)\frac{e^{-\alpha x}e^{-\beta(z-x)}}{e^{-\alpha z}-e^{-\beta z}}dx = \frac{(\beta-\alpha)e^{-\beta z}}{(e^{-\alpha z}-e^{-\beta z})}\int_0^z xe^{-(\alpha-\beta)x}dx \\
&= \frac{e^{-\beta z}}{(e^{-\alpha z}-e^{-\beta z})}[ze^{-(\alpha-\beta)z} - \int_0^z e^{-(\alpha-\beta)x}dx] \\
&= \frac{e^{-\beta z}}{(e^{-\alpha z}-e^{-\beta z})}[ze^{-(\alpha-\beta)z} - \frac{1}{(\alpha-\beta)}(1-e^{-(\alpha-\beta)z})] \\
&= \frac{ze^{-\alpha z}}{e^{-\alpha z}-e^{-\beta z}} + \frac{1}{\alpha-\beta} \qquad (1)
\end{aligned}
$$

To verify $E[E(X|Z)] = E(X)$, we calculate

$$
\begin{aligned}
\int_0^\infty E(X|Z=z)f_Z(z)dz &= \int_0^\infty \frac{\alpha\beta e^{-\beta z}}{(\beta-\alpha)}[ze^{-(\alpha-\beta)z} + \frac{1}{\alpha-\beta}e^{-(\alpha-\beta)z} - \frac{1}{\alpha-\beta}]dz \\
&= \int_0^\infty \frac{\alpha\beta z e^{-\alpha z}}{(\beta-\alpha)}dz + \int_0^\infty \frac{\alpha\beta}{(\alpha-\beta)^2}e^{-\beta z}dz - \int_0^\infty \frac{\alpha\beta}{(\alpha-\beta)^2}e^{-\alpha z}dz \\
&= \frac{\frac{\beta}{\alpha}}{\beta-\alpha} + \frac{\alpha}{(\alpha-\beta)^2} - \frac{\beta}{(\alpha-\beta)^2} \\
&= \frac{\frac{\beta}{\alpha}}{\beta-\alpha} + \frac{1}{\alpha-\beta} \\
&= \frac{\frac{\beta}{\alpha}-1}{\beta-\alpha} \\
&= \frac{1}{\alpha} = E(X)
\end{aligned}
$$

(ii) If $\alpha = \beta$, we have

$$f_Z(z) = \int_0^z \alpha^2 e^{-\alpha z}dx = \alpha^2 z e^{-\alpha z}$$

Similarly,

$$f_{X,Z}(x,z) = f_X(x)f_Y(z-x) = \alpha^2 e^{-\alpha z} \quad z \geq x \geq 0$$

so that

$$f_{X|Z}(x|z) = \frac{1}{z} \quad z \geq x \geq 0$$

This yields

$$E(X|Z=z) = \int_0^z x f_{X|Z}(x|z)dx = \frac{z}{2}$$

so that

$$E(X|Z) = \frac{Z}{2}$$

We can also start with (1) and take the limit as $\alpha \to \beta$. First note

$$
\begin{aligned}
\frac{e^{-\beta z}}{(e^{-\alpha z}-e^{-\beta z})} &= \frac{1}{e^{-(\alpha-\beta)z}-1} \\
&= \frac{1}{-(\alpha-\beta)z + h.o.t.}
\end{aligned}
$$

where $h.o.t.$ denotes higher order terms in $(\alpha-\beta)$. Also

$$\begin{aligned} ze^{-(\alpha-\beta)}z - \frac{1}{(\alpha-\beta)}[1-e^{-(\alpha-\beta)z}] &= z[1-(\alpha-\beta)z + h.o.t.] \\ &\quad + \frac{1}{(\alpha-\beta)}[-(\alpha-\beta)z + \frac{(\alpha-\beta)^2 z^2}{2} + h.o.t.] \\ &= -\frac{(\alpha-\beta)z^2}{2} + h.o.t. \end{aligned}$$

Combining, we find that

$$\lim_{\alpha\to\beta} E(X|Z) = \frac{Z}{2}$$

2. Problem 1.4

$$f_X(x) = 1, \quad 0 \le x \le 1$$

Hence $m_x = \frac{1}{2}$. From Example 1.1.3, we also have

$$f_Y(y) = ye^{-y}, \quad y \ge 0$$

Using the results from Example 1.1.2, we get

$$m_Y = \int_0^\infty y^2 e^{-y} dy = 2$$

Next we determine $\text{cov}(X,Y) = EXY - m_x m_y$. Since

$$\begin{aligned} EXY &= \int_0^\infty \int_{e^{-y}}^1 xy f_{X,Y}(x,y) dx dy \\ &= \int_0^\infty \int_{e^{-y}}^1 xy \frac{1}{x} e^{-y} dx dy \\ &= \int_0^\infty (1-e^{-y}) y e^{-y} dy \\ &= 1 - \int_0^\infty y e^{-2y} dy \\ &= 1 - \frac{1}{4} = \frac{3}{4} \end{aligned}$$

we get

$$\text{cov}(X,Y) = \frac{3}{4} - \frac{1}{2}2 = -\frac{1}{4}$$

Next, we have

$$EY^2 = \int_0^\infty y^3 e^{-y} dy = 6$$

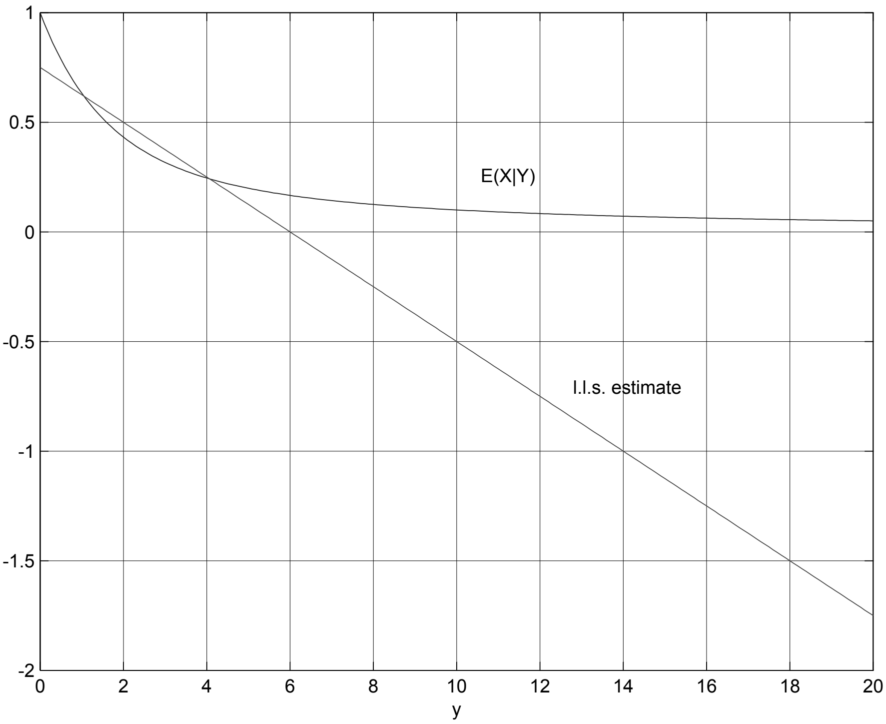
so that

$$\text{cov}(Y) = EY^2 - m_Y^2 = 6 - 4 = 2$$

The linear least squares estimator is given by

$$\hat{X} = \frac{1}{2} - \frac{1}{4}\frac{1}{2}(Y-2) = \frac{3}{4} - \frac{1}{8}Y$$

The plot of the 2 estimators as a function of y is given below.



The linear least squares estimator only gives a reasonable approximation to the conditional mean for $0 \leq y \leq 6$. For larger values, the linear least squares estimator produces a negative estimate, which is quite poor since the probability that X takes negative values is 0!

3. Problem 1.5

(i) $f_{X|Y}(x|y) = \dfrac{|\det \Sigma_Y|^{\frac{1}{2}}}{(2\pi)^{\frac{n}{2}}|\det \Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2}(z-m)^T\Sigma^{-1}(z-m)+\frac{1}{2}(y-m_Y)^T\Sigma_Y^{-1}(y-m_Y)}$

$$J_{X|Y} = (z-m)^T\Sigma^{-1}(z-m) - (y-m_Y)^T\Sigma_Y^{-1}(y-m_Y)$$

(ii) If P is chosen to be $\begin{bmatrix} I & -\Sigma_{XY}\Sigma_Y^{-1} \\ 0 & I \end{bmatrix}$

$$\begin{aligned} P\Sigma P^T &= \begin{bmatrix} I & -\Sigma_{XY}\Sigma_Y^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} \Sigma_X & \Sigma_{XY} \\ \Sigma_{YX} & \Sigma_Y \end{bmatrix} \begin{bmatrix} I & 0 \\ -\Sigma_Y^{-1}\Sigma_{YX} & I \end{bmatrix} \\ &= \begin{bmatrix} I & -\Sigma_{XY}\Sigma_Y^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} \Sigma_X - \Sigma_{XY}\Sigma_Y^{-1}\Sigma_{YX} & \Sigma_{XY} \\ 0 & \Sigma_Y \end{bmatrix} \\ &= \begin{bmatrix} \Sigma_X - \Sigma_{XY}\Sigma_Y^{-1}\Sigma_{YX} & 0 \\ 0 & \Sigma_Y \end{bmatrix} \end{aligned}$$

$$\begin{aligned} P(z-m) &= \begin{bmatrix} I & -\Sigma_{XY}\Sigma_Y^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} x - m_X \\ y - m_Y \end{bmatrix} \\ &= \begin{bmatrix} (x-m_X) - \Sigma_{XY}\Sigma_Y^{-1}(y-m_Y) \\ y - m_Y \end{bmatrix} \end{aligned}$$

(iii)

$$
\begin{aligned}
J_{X|Y} &= \left\{ \begin{bmatrix} x - m_X - \Sigma_{XY}\Sigma_Y^{-1}(y - m_Y) \\ y - m_Y \end{bmatrix}^T \right. \\
& \begin{bmatrix} (\Sigma_X - \Sigma_{XY}\Sigma_Y^{-1}\Sigma_{YX})^{-1}[(x - m_X) - \Sigma_{XY}\Sigma_Y^{-1}(y - m_Y)] \\ \Sigma_Y^{-1}(y - m_Y) \end{bmatrix} \\
& \left. -(y - m_Y)^T\Sigma_Y^{-1}(y - m_Y) \right\} \\
&= (x - m_X - \Sigma_{XY}\Sigma_Y^{-1}(y - m_Y))^T[(\Sigma_X - \Sigma_{XY}\Sigma_Y^{-1}\Sigma_{YX})^{-1}]. \\
& (x - m_X - \Sigma_{XY}\Sigma_Y^{-1}(y - m_Y))
\end{aligned}
$$

Need only show $\dfrac{|\det \Sigma_Y|}{|\det \Sigma|} = \dfrac{1}{|\det(\Sigma_X - \Sigma_{XY}\Sigma_Y^{-1}\Sigma_{YX})|}$.
Note that $\det P = \det P^T = 1$ so that $\det P\Sigma P^T = \det P \det \Sigma \det P^T = \det \Sigma$.
Hence

$$
\begin{aligned}
|\det \Sigma| &= |\det \begin{bmatrix} \Sigma_X - \Sigma_{XY}\Sigma_Y^{-1}\Sigma_{YX} & 0 \\ 0 & \Sigma_Y \end{bmatrix}| \\
&= |\det(\Sigma_X - \Sigma_{XY}\Sigma_Y^{-1}\Sigma_{YX})||\det \Sigma_Y|
\end{aligned}
$$

so that the desired result is obtained. The calculations show that the conditional density $f(x|y)$ has the form of a Gaussian density. The mean value resulting from the conditional density is $E(X|Y)$, which can now be obtained by inspection of $f(x|y)$:

$$E(X|Y) = m_X + \Sigma_{XY}\Sigma_Y^{-1}(Y - m_Y)$$

The conditional covariance is the covariance matrix of the conditional density $f(x|y)$, which again we can obtain by inspection:

$$E\{[X - E(X|Y)][X - E(X|Y)]^T|Y\} = \Sigma_X - \Sigma_{XY}\Sigma_Y^{-1}\Sigma_{YX} = \Sigma_X - \Sigma_{XY}\Sigma_Y^{-1}\Sigma_{XY}^T$$

Since this matrix is completely determined by the given matrix Σ, it is deterministic, does not depend on Y, and is equal to the total expectation $E[X - E(X|Y)][X - E(X|Y)]^T$.

4. From the bivariate Gaussian density, we see that if we write $Z = [X \ Y]^T$,

$$\text{cov}(Z) = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$$

Specializing the results of Problem 1.5, we obtain

$$E(X|Y) = \rho Y$$

The conditional covariance is given by

$$E[(X - E(X|Y))^2|Y] = 1 - \rho^2$$

which as before is independent of Y.