Massachusetts Institute of Technology
Department of Electrical Engineering & Computer Science
6.041/6.431: Probabilistic Systems Analysis
(Spring 2005)

Problem Set 7G Solutions
Due: April 6, 2005

G1. (a) Since K is positive definite and symmetric, its inverse K^{-1} is also positive definite and symmetric. A symmetric matrix can be diagonalized. Let D be the diagonalized matrix of K^{-1}, where D is a 2×2 matrix composed of the eigenvalues of K^{-1}, λ_1 and λ_2, each with corresponding eigenvectors, $\underline{P_1}$ and $\underline{P_2}$. $\underline{P_1}$ is the column vector,$(p_{11}, p_{21})^{\mathrm{t}}$, and $\underline{P_2}$ is $(p_{12}, p_{22})^{\mathrm{t}}$.
A change of variables can be made where

$$(\underline{x} - \underline{\mu}) = P(\underline{y} - \underline{\mu}),$$

where P is a 2×2 matrix composed of the two column eigenvectors $\underline{P_1}$ and $\underline{P_2}$ and $\underline{y}$ is the column vector $(y_1, y_2)^{\mathrm{t}}$. Since $(\underline{x} - \underline{\mu})$ equals $P(\underline{y} - \underline{\mu})$, x_1 can be expressed as a linear combination of y_1 and y_2, $x_1 = ay_1 + by_2$ and likewise, x_2 can be expressed as $x_2 = cy_1 + dy_2$.

$$(\underline{x} - \underline{\mu})^{\mathrm{t}} K^{-1} (\underline{x} - \underline{\mu}) = (\underline{y} - \underline{\mu})^{\mathrm{t}} D (\underline{y} - \underline{\mu}) = \lambda_1 (y_1 - \mu_1)^2 + \lambda_2 (y_2 - \mu_2)^2$$

The PDF of X_1 and X_2 now becomes:

$$f_{(Y_1,Y_2)}(y_1, y_2) = \frac{1}{2\pi\sqrt{|K|}} e^{-(\lambda_1(y_1-\mu_1)^2 + \lambda_2(y_2-\mu_2)^2)/2},$$

Let $\lambda_1 = 1/\sigma_1^2$ and $\lambda_2 = 1/\sigma_2^2$. Then $\sqrt{|K|} = \sigma_1\sigma_2$ and the PDF can be written as follows:

$$f_{(Y_1,Y_2)}(y_1, y_2) = \frac{1}{2\pi\sigma_1\sigma_2} e^{-(\frac{(y_1-\mu_1)^2}{2\sigma_1^2} + \frac{(y_2-\mu_2)^2}{2\sigma_2^2})}.$$

$f_{(Y_1,Y_2)}(y_1, y_2)$ can now be separated into the product of the two marginal PDF's of Y_1 and Y_2.

$$f_{Y_1}(y_1) = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-(\frac{(y_1-\mu_1)^2}{2\sigma_1^2})}$$

$$f_{Y_2}(y_2) = \frac{1}{\sqrt{2\pi}\sigma_2} e^{-(\frac{(y_2-\mu_2)^2}{2\sigma_2^2})}$$

Therefore, Y_1 and Y_2 are both independent normal random variables, where $X1 = aY_1 + bY_2$ and $X_2 = cY_1 + dY_2$. (In this problem Y_1 is U and Y_2 is V).

(b)

$$\begin{aligned} M_Y(s) &= \mathbf{E}[e^{s(\alpha X_1 + \beta X_2 + \gamma)}] \\ &= \mathbf{E}[e^{s(\alpha(aU+bV) + \beta(cU+dV) + \gamma)}] \\ &= e^{s\gamma} E[e^{sU(\alpha a + \beta c)} e^{sU(\alpha b + \beta d)}] \end{aligned}$$

Let $c_1 = \alpha a + \beta c$ and let $c_2 = \alpha b + \beta d$

$$\begin{aligned} M_Y(s) &= e^{s\gamma} \mathbf{E}[e^{sc_1 U} e^{sc_2 V}] \\ &= e^{s\gamma} M_U(sc_1) M_V(sc_2) \end{aligned}$$

Since $U \sim \mathcal{N}(\mu_U, \sigma_U^2)$ and $V \sim \mathcal{N}(\mu_V, \sigma_V^2)$, the transforms of U and V are the following:

$$M_U(s) = e^{\frac{\sigma_U^2 s^2}{2} + \mu_U s}, \quad M_V(s) = e^{\frac{\sigma_V^2 s^2}{2} + \mu_V s}$$

Finally

$$M_Y(s) = e^{\frac{(\sigma_U^2 c_1^2 + \sigma_V^2 c_2^2)}{2} s^2 + (c_1\mu_U + c_2\mu_V + \gamma)s},$$

which implies that Y is Gaussian with mean equal to $(c_1\mu_U + c_2\mu_V + \gamma)$ and variance $(\sigma_U^2 c_1^2 + \sigma_V^2 c_2^2)$.

(c) The multivariate transform of X_1 and X_2 can be written as

$$\begin{aligned}
M_{X_1,X_2}(s_1, s_2) &= \mathbf{E}[e^{s_1 X_1 + s_2 X_2}] \\
&= \mathbf{E}[e^{s_1(aU+bV) + s_2(cU+dV)}] \\
&= \mathbf{E}[e^{(s_1 a + s_2 c)U} + e^{(s_1 b + s_2 d)V}] \\
&= M_U(s)\Big|_{s = s_1 a + s_2 c} * M_V(s)\Big|_{s = s_1 b + s_2 d}
\end{aligned}$$

G2. In this problem, we want to find the linear estimator $g(Y_1, Y_2)$ that minimizes $\mathbf{E}[(X - g(Y_1, Y_2))^2]$. This is an extension of the case of single measurement. Our estimator is of the form

$$g(Y_1, Y_2) = a_1 Y_1 + a_2 Y_2 + b$$

and therefore our goal is to find a_1, a_2 and b that solve

$$\text{minimize } (\mathbf{E}[(X - a_1 Y_1 - a_2 Y_2 - b)^2]) \qquad (*)$$

To achieve $(*)$, we must satisfy

$$\begin{aligned}
\frac{\partial \mathbf{E}[(X - a_1 Y_1 - a_2 Y_2 - b)^2]}{\partial b} &= \mathbf{E}[2(X - a_1 Y_1 - a_2 Y_2 - b)(-1)] &= 0 \\
\frac{\partial \mathbf{E}[(X - a_1 Y_1 - a_2 Y_2 - b)^2]}{\partial a_1} &= \mathbf{E}[2(X - a_1 Y_1 - a_2 Y_2 - b)(-Y_1)] &= 0 \\
\frac{\partial \mathbf{E}[(X - a_1 Y_1 - a_2 Y_2 - b)^2]}{\partial a_2} &= \mathbf{E}[2(X - a_1 Y_1 - a_2 Y_2 - b)(-Y_2)] &= 0
\end{aligned}$$

and $\partial^2 \mathbf{E}[\cdot]/\partial b^2$, $\partial^2 \mathbf{E}[\cdot]/\partial {a_1}^2$, and $\partial^2 \mathbf{E}[\cdot]/\partial {a_2}^2$ must be > 0.

By the linearity of expectation,

$$\begin{aligned}
b &= \mathbf{E}[X] - a_1 \mathbf{E}[Y_1] - a_2 \mathbf{E}[Y_2] && (1) \\
a_1 \mathbf{E}[Y_1^2] &= \mathbf{E}[XY_1] - a_2 \mathbf{E}[Y_1 Y_2] - b\mathbf{E}[Y_1] && (2) \\
a_2 \mathbf{E}[Y_2^2] &= \mathbf{E}[XY_2] - a_1 \mathbf{E}[Y_1 Y_2] - b\mathbf{E}[Y_2]. && (3)
\end{aligned}$$

We now have 3 equations to solve for 3 unknowns.

Consider $(2) - \mathbf{E}[Y_1] \cdot (1)$, we obtain

$$a_1 \mathbf{E}[Y_1^2] - b\mathbf{E}[Y_1] = \mathbf{E}[XY_1] - a_2 \mathbf{E}[Y_1 Y_2] - b\mathbf{E}[Y_1] - \mathbf{E}[X]\mathbf{E}[Y_1] + a_1(\mathbf{E}[Y_1])^2 + a_2 \mathbf{E}[Y_1]\mathbf{E}[Y_2].$$

Arranging algebra and use the fact that $\mathbf{E}[Y_1 Y_2] = \mathbf{E}[Y_1]\mathbf{E}[Y_2]$

$$a_1(\mathbf{E}[Y_1^2] - (\mathbf{E}[Y_1])^2) = \mathbf{E}[XY_1] - \mathbf{E}[X]\mathbf{E}[Y_1].$$

Similarly,

$$a_2(\mathbf{E}[Y_2^2] - (\mathbf{E}[Y_2])^2) = \mathbf{E}[XY_2] - \mathbf{E}[X]\mathbf{E}[Y_2].$$

Therefore,

$$\begin{aligned} a_1 &= (\mathbf{E}[XY_1] - \mathbf{E}[X]\mathbf{E}[Y_1])/\sigma_{Y_1}^2 \\ a_2 &= (\mathbf{E}[XY_2] - \mathbf{E}[X]\mathbf{E}[Y_2])/\sigma_{Y_2}^2 \\ b &= \mathbf{E}[X] - a_1\mathbf{E}[Y_1] - a_2\mathbf{E}[Y_2]. \end{aligned}$$

Writing this expression in the similar term as the case of single measurement:

$$g(Y_1, Y_2) = \mathbf{E}[X] + \frac{\text{Cov}(X,Y_1)}{\sigma_{Y_1}^2}(Y_1 - \mathbf{E}[Y_1]) + \frac{\text{Cov}(X,Y_2)}{\sigma_{Y_2}^2}(Y_2 - \mathbf{E}[Y_2])$$

where we use the fact that $\text{Cov}(X,Y) = \mathbf{E}[(X - \mathbf{E}[X])(Y - \mathbf{E}[Y])] = \mathbf{E}[XY] - \mathbf{E}[X]\mathbf{E}[Y]$.

Note: Convince yourself that the second order condition is satisfied.