Exam 2

Jeremy Morris

March 23, 2006

4.2 Consider a bivariate normal population with $\mu_1 = 0, \mu_2 = 2, \sigma_{11} = 2, \sigma_{22} = 1$ and $\rho_{12} = .5$.

(a) Write out the bivariate normal density.

The multivariate normal density is defined by the following equation.

$$f(\boldsymbol{x}) = \frac{1}{(2\pi)^{p/2}|\boldsymbol{\Sigma}|^{1/2}} e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})}$$

With $\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix}$ and $\sigma_{12} = \rho_{12}\sqrt{\sigma_{11}}\sqrt{\sigma_{22}}$. For this example, we have the following definitions for $\boldsymbol{\Sigma}$, and $|\boldsymbol{\Sigma}|$.

$$\begin{aligned}
\boldsymbol{\Sigma} &= \begin{pmatrix} 2 & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & 1 \end{pmatrix} \\
|\boldsymbol{\Sigma}| &= (2)(1) - (\sqrt{2}/2)^2 = \frac{7}{4} \\
\boldsymbol{\Sigma}^{-1} &= \frac{4}{7}\begin{pmatrix} 1 & \frac{-\sqrt{2}}{2} \\ \frac{-\sqrt{2}}{2} & 2 \end{pmatrix}
\end{aligned}$$

To complete the definition of the density function we will derive the squared generalized distance expression $(\boldsymbol{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})$.

$$(\boldsymbol{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu}) = \begin{pmatrix} x_1 & x_2 - 2 \end{pmatrix} \begin{pmatrix} \frac{4}{7} & -\frac{2\sqrt{2}}{7} \\ -\frac{2\sqrt{2}}{7} & \frac{8}{7} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 - 2 \end{pmatrix}$$

$$\begin{pmatrix} x_1 & x_2 - 2 \end{pmatrix} \begin{pmatrix} \frac{4}{7}x_1 - (x_2-2)\frac{2\sqrt{2}}{7} \\ -\frac{2\sqrt{2}}{7}x_1 + (x_2-2)\frac{8}{7} \end{pmatrix}$$

$$\frac{1}{7}\begin{pmatrix} x_1 & x_2 - 2 \end{pmatrix} \begin{pmatrix} 4x_1 - 2\sqrt{2}x_2 + 4\sqrt{2} \\ -2\sqrt{2}x_1 + 8x_2 - 16 \end{pmatrix}$$

$$\frac{1}{7}x_1(4x_1 - 2\sqrt{2}x_2 + 4\sqrt{2}) + \frac{1}{7}(x_2-2)(-2\sqrt{2}x_1 + 8x_2 - 16)$$

$$\varphi(x_1, x_2) = \frac{1}{7}(4x_1^2 + 8\sqrt{2}x_1 - 4\sqrt{2}x_1x_2 - 32x_2 + 8x_2^2 + 32) \tag{1}$$

Then the bivariate normal distribution function is defined as follows.

$$f(x_1, x_2) = \frac{1}{\sqrt{7}\pi} e^{-\frac{1}{2}\varphi(x_1,x_2)}$$

(b) Write out the squared generalized distance expression $(\boldsymbol{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})$ as a function of x_1 and x_2.

Derived in part (a) equation 1.

(c) Determine (and sketch) the constant-density contour that contains 50% of the probability.

To determine the constant-density contour, we need to calculate the eigenvalues and eigenvectors of $\boldsymbol{\Sigma}$. This is done by solving for the roots of the equation $|\boldsymbol{\Sigma} - \lambda I|$.

$$\begin{aligned}
(2-\lambda)(1-\lambda) - 1/2 &= 0 \\
2 - 2\lambda - \lambda + \lambda^2 - 1/2 &= 0 \\
\lambda^2 - 3\lambda + 3/2 &= 0
\end{aligned}$$

Using the quadratic equation, we get the eigenvalues

$$\begin{aligned}
\lambda_1 &= \frac{3+\sqrt{3}}{2} \\
\lambda_2 &= \frac{3-\sqrt{3}}{2}
\end{aligned}$$

Then we find the eigenvector for λ_1.

$$\boldsymbol{\Sigma}\boldsymbol{e}_1 - \lambda_1\boldsymbol{e}_1 = 0$$

$$\Rightarrow \begin{aligned}
(2 - \tfrac{3}{2} - \tfrac{\sqrt{3}}{2})x_1 + \tfrac{\sqrt{2}}{2}x_2 &= 0 \\
\tfrac{\sqrt{2}}{2}x_1 + (2 - \tfrac{3}{2} - \tfrac{\sqrt{3}}{2})x_1 &= 0
\end{aligned}$$

$$\Rightarrow \begin{pmatrix} 1-\sqrt{3} & \sqrt{2} \\ \sqrt{2} & -1-\sqrt{3} \end{pmatrix} \boldsymbol{e}_1 = \boldsymbol{0}$$

$$\Rightarrow \boldsymbol{e}_1 = \begin{pmatrix} \frac{-\sqrt{2}}{1-\sqrt{3}} \\ 1 \end{pmatrix}$$

And the eigenvector for λ_2.

$$\boldsymbol{\Sigma}\boldsymbol{e}_2 - \lambda_2\boldsymbol{e}_2 = 0$$

$$\Rightarrow \begin{aligned}
(2 - \tfrac{3}{2} + \tfrac{\sqrt{3}}{2})x_1 + \tfrac{\sqrt{2}}{2}x_2 &= 0 \\
\tfrac{\sqrt{2}}{2}x_1 + (2 - \tfrac{3}{2} + \tfrac{\sqrt{3}}{2})x_1 &= 0
\end{aligned}$$

$$\Rightarrow \begin{pmatrix} 1+\sqrt{3} & \sqrt{2} \\ \sqrt{2} & -1+\sqrt{3} \end{pmatrix} \boldsymbol{e}_1 = \boldsymbol{0}$$

$$\Rightarrow \boldsymbol{e}_2 = \begin{pmatrix} \frac{\sqrt{2}}{1+\sqrt{3}} \\ -1 \end{pmatrix}$$

4.6 Let $\boldsymbol{X}$ be distributed as $N_3(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where $\boldsymbol{\mu} = [1, -1, 2]$ and

$$\boldsymbol{\Sigma} = \begin{bmatrix} 4 & 0 & -1 \\ 0 & 5 & 0 \\ -1 & 0 & 2 \end{bmatrix}$$

Which of the following random variables are independent? Explain.

(a) X_1 *and* X_2.
X_1 and X_2 are independent because $cov(X_1, X_2) = cov(X_2, X_1) = 0$.

(b) X_1 *and* X_3.
X_1 and X_3 are not independent because $cov(X_1, X_3) = cov(X_3, X_1) = -1$.

(c) X_2 *and* X_3.
X_2 and X_3 are independent because $cov(X_2, X_3) = cov(X_3, X_2) = 0$.

(d) (X_1, X_3) *and* X_2.
To determine the answer, we need to rearrange the covariance matrix and partition it. The new covariance matrix follows.

$$\boldsymbol{\Sigma} = \left[\begin{array}{cc|c} 4 & -1 & 0 \\ -1 & 2 & 0 \\ \hline 0 & 0 & 5 \end{array}\right]$$

We conclude that (X_1, X_3) and X_2 are independent.

(e) X_1 *and* $X_1 + 3X_2 - 2X_3$.
We have the multivariate normal distribution

$$\boldsymbol{AX} = \begin{bmatrix} X_1 \\ X_1 + 3X_2 - 2X_3 \end{bmatrix}$$

Where

$$\begin{aligned} \boldsymbol{A} &= \begin{bmatrix} 1 & 0 & 0 \\ 1 & 3 & -2 \end{bmatrix} \\ \boldsymbol{X} &= \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} \end{aligned}$$

And $\boldsymbol{AX}$ has the distribution $N_2(\boldsymbol{A\mu}, \boldsymbol{A\Sigma A'})$. Here we show the matrix $\boldsymbol{A\Sigma A'}$ to determine independence.

$$\begin{aligned} \boldsymbol{A\Sigma A'} &= \begin{bmatrix} 1 & 0 & 0 \\ 1 & 3 & -2 \end{bmatrix} \begin{bmatrix} 4 & 0 & -1 \\ 0 & 5 & 0 \\ -1 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 3 \\ 0 & -2 \end{bmatrix} \\ &= \begin{bmatrix} 4 & 0 & -1 \\ 6 & 15 & -5 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 3 \\ 0 & -2 \end{bmatrix} \\ &= \begin{bmatrix} 4 & 6 \\ 6 & 61 \end{bmatrix} \end{aligned}$$

And we conclude that (X_1, X_3) and X_2 are not independent.

4.9 Refer to Exercise 4.8, but modify the construction by replacing the break point 1 by c so that

$$X_2 = \begin{cases} -X_1 & \text{if } -c \leq X_1 \leq c \\ X_1 & \text{elsewhere} \end{cases}$$

Show that c can be chosen so that $\text{Cov}(X_1, X_2) = 0$, but that the two random variables are not independent.

First, we note (from exercise 4.8) that $X_1 \sim N(0,1)$. For $c = 0$, we have

$$\begin{aligned} \text{Cov}(X_1, X_2) &= E[X_1(X_1)] - E[X_1]E[X_1] \\ &= E[(X_1)^2] = \text{var}(X_1) + E[X_1]^2 \\ &= 1 \end{aligned}$$

For c very large we have

$$\begin{aligned} \text{Cov}(X_1, X_2) &= E[X_1(-X_1)] + E[X_1]E[X_1] \\ &= -E[(X_1)^2] = -(\text{var}(X_1) + E[X_1]^2) \\ &= -1 \end{aligned}$$

Since the covariance is a smooth function of c, then by the mean value theorem, $\text{Cov}(X_1, X_2) = 0$ at some point, but that the two random variables are not independent.

4.10 Show each of the following.

(a) $\begin{vmatrix} \boldsymbol{A} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{B} \end{vmatrix} = |\boldsymbol{A}||\boldsymbol{B}|$

Factor $\begin{vmatrix} \boldsymbol{A} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{B} \end{vmatrix}$ so that we have the following.

$$\begin{vmatrix} \boldsymbol{A} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{B} \end{vmatrix} = \begin{vmatrix} \boldsymbol{A} & \mathbf{0} \\ \mathbf{0}' & \boldsymbol{I} \end{vmatrix} \begin{vmatrix} \boldsymbol{I} & \mathbf{0} \\ \mathbf{0}' & \boldsymbol{B} \end{vmatrix}$$

Then take the determinants to get

$$\begin{aligned} \begin{vmatrix} \boldsymbol{A} & \mathbf{0} \\ \mathbf{0}' & \boldsymbol{I} \end{vmatrix} \begin{vmatrix} \boldsymbol{I} & \mathbf{0} \\ \mathbf{0}' & \boldsymbol{B} \end{vmatrix} &= (|\boldsymbol{A}\boldsymbol{I}| - |\mathbf{0}\mathbf{0}'|)(|\boldsymbol{I}\boldsymbol{B}| - |\mathbf{0}\mathbf{0}'|) \\ &= |\boldsymbol{A}||\boldsymbol{B}| \end{aligned}$$

(b) $\begin{vmatrix} \boldsymbol{A} & \boldsymbol{C} \\ \mathbf{0}' & \boldsymbol{B} \end{vmatrix} = |\boldsymbol{A}||\boldsymbol{B}|$ for $|\boldsymbol{A}| \neq 0$.

As in part (a), factor the determinant so that we have the following.

$$\begin{vmatrix} \boldsymbol{A} & \boldsymbol{C} \\ \mathbf{0}' & \boldsymbol{B} \end{vmatrix} = \begin{vmatrix} \boldsymbol{A} & \mathbf{0} \\ \mathbf{0}' & \boldsymbol{B} \end{vmatrix} \begin{vmatrix} \boldsymbol{I} & \boldsymbol{A}^{-1}\boldsymbol{C} \\ \mathbf{0}' & \boldsymbol{I} \end{vmatrix}$$

Then, by part (a), we get $\begin{vmatrix} \boldsymbol{A} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{B} \end{vmatrix} = |\boldsymbol{A}||\boldsymbol{B}|$, and $\begin{vmatrix} \boldsymbol{I} & \boldsymbol{A}^{-1}\boldsymbol{C} \\ \mathbf{0}' & \boldsymbol{I} \end{vmatrix} = 1$.

4.10 Show that if $\boldsymbol{A}$ is square,

$$
\begin{aligned}
|\boldsymbol{A}| &= |A_{22}||A_{11} - A_{12}A_{22}^{-1}A_{21}| && \text{for } |A_{22}| \neq 0 \\
&= |A_{11}||A_{22} - A_{21}A_{11}^{-1}A_{12}| && \text{for } |A_{11}| \neq 0
\end{aligned}
$$

Let

$$\boldsymbol{A} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

and factor so that we have the following.

$$
\begin{aligned}
&\begin{pmatrix} I & -A_{12}A_{22}^{-1} \\ 0' & I \end{pmatrix}\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}\begin{pmatrix} I & 0 \\ -A_{22}^{-1}A_{21} & I \end{pmatrix} \\
&= \begin{pmatrix} A_{11} - A_{12}A_{22}^{-1}A_{21} & 0 \\ A_{21} & A_{22} \end{pmatrix}\begin{pmatrix} I & 0 \\ -A_{22}^{-1}A_{21} & I \end{pmatrix} \\
&= \begin{pmatrix} A_{11} - A_{12}A_{22}^{-1}A_{21} & 0 \\ 0 & A_{22} \end{pmatrix}
\end{aligned}
$$

Finally, we have

$$\begin{pmatrix} I & -A_{12}A_{22}^{-1} \\ 0' & I \end{pmatrix}\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}\begin{pmatrix} I & 0 \\ -A_{22}^{-1}A_{21} & I \end{pmatrix} = \begin{pmatrix} A_{11} - A_{12}A_{22}^{-1}A_{21} & 0 \\ 0 & A_{22} \end{pmatrix}$$

Taking the determinant of both sides, we get the following.

$$|I||A||I| = |A_{22}||A_{11} - A_{12}A_{22}^{-1}A_{21}|$$

$$|A| = |A_{22}||A_{11} - A_{12}A_{22}^{-1}A_{21}|$$

For the second relation, we do the same by factoring $\boldsymbol{A}$ so that we have the following.

$$
\begin{aligned}
&\begin{pmatrix} I & 0 \\ -A_{21}A_{11}^{-1} & I \end{pmatrix}\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}\begin{pmatrix} I & -A_{11}^{-1}A_{21} \\ 0' & I \end{pmatrix} \\
&= \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} - A_{21}A_{11}^{-1}A_{12} \end{pmatrix}\begin{pmatrix} I & -A_{11}^{-1}A_{21} \\ 0' & I \end{pmatrix}
\end{aligned}
$$

$$\begin{pmatrix} I & 0 \\ -A_{21}A_{11}^{-1} & I \end{pmatrix}\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}\begin{pmatrix} I & -A_{11}^{-1}A_{21} \\ 0' & I \end{pmatrix} = \begin{pmatrix} A_{11} & 0 \\ 0 & A_{22} - A_{21}A_{11}^{-1}A_{12} \end{pmatrix}$$

Again, taking the determinant of both sides we get the following.

$$|I||A||I| = |A_{11}||A_{22} - A_{21}A_{11}^{-1}A_{12}|$$

$$|A| = |A_{11}||A_{22} - A_{21}A_{11}^{-1}A_{12}|$$

4.16 Let $\boldsymbol{X}_i$, $i = 1, 2, \ldots, 4$, be independent $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ random vectors.

(a) Find the marginal distributions for each of the random vectors

$$\boldsymbol{V}_1 = \frac{1}{4}\boldsymbol{X}_1 - \frac{1}{4}\boldsymbol{X}_2 + \frac{1}{4}\boldsymbol{X}_3 - \frac{1}{4}\boldsymbol{X}_4$$

and

$$\boldsymbol{V}_2 = \frac{1}{4}\boldsymbol{X}_1 + \frac{1}{4}\boldsymbol{X}_2 - \frac{1}{4}\boldsymbol{X}_3 - \frac{1}{4}\boldsymbol{X}_4$$

By result 4.8 in the text, $\boldsymbol{V}_1$ and $\boldsymbol{V}_2$ have the following distribution.

$$N_p\left(\sum_{j=1}^{n} c_j\boldsymbol{\mu}, \left(\sum_{j=1}^{n} c_j^2\right)\boldsymbol{\Sigma}\right)$$

Then we have $\boldsymbol{V}_1 \sim N_p(\mathbf{0}, \frac{1}{4}\boldsymbol{\Sigma})$ and $\boldsymbol{V}_2 \sim N_p(\mathbf{0}, \frac{1}{4}\boldsymbol{\Sigma})$.

(b) Find the joint distribution of the random vectors $\boldsymbol{V}_1$ and $\boldsymbol{V}_2$ defined in (a).

Also by result 4.8, $\boldsymbol{V}_1$ and $\boldsymbol{V}_2$ are jointly multivariate normal with covariance matrix

$$\begin{bmatrix} \left(\sum_{j=1}^{n} c_j^2\right)\boldsymbol{\Sigma} & (\boldsymbol{b}'\boldsymbol{c})\boldsymbol{\Sigma} \\ (\boldsymbol{b}'\boldsymbol{c})\boldsymbol{\Sigma} & \left(\sum_{j=1}^{n} b_j^2\right)\boldsymbol{\Sigma} \end{bmatrix}$$

With $\boldsymbol{c} = (1/4, -1/4, 1/4, -1/4)'$ and $\boldsymbol{b} = (1/4, 1/4, -1/4, -1/4)'$. So that we have the covariance matrix

$$\begin{bmatrix} \frac{1}{4}\boldsymbol{\Sigma} & \mathbf{0} \\ \mathbf{0} & \frac{1}{4}\boldsymbol{\Sigma} \end{bmatrix}$$

6.11 A likelihood argument provides aditional support for pooling the two independent sample covariance matrices to estimate a common covariance matrix in the case of two normal populations. Give the likelihood function $L(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \boldsymbol{\Sigma})$, for the two independent samples of sizes n_1 and n_2 from $N_p(\boldsymbol{\mu}_1, \boldsymbol{\Sigma})$ and $N_p(\boldsymbol{\mu}_2, \boldsymbol{\Sigma})$ respectively. Show that this likelihood is maximized by the choices $\hat{\boldsymbol{\mu}}_1 = \bar{\boldsymbol{x}}_1, \hat{\boldsymbol{\mu}}_2 = \bar{\boldsymbol{x}}_2$ and

$$\hat{\boldsymbol{\Sigma}} = \frac{1}{n_1 + n_2}\left[(n_1 - 1)\boldsymbol{S}_1 + (n_2 - 1)\boldsymbol{S}_2\right] = \frac{n_1 + n_2 - 1}{n_1 + n_2}\boldsymbol{S}_{\text{pooled}}$$

Here we note the density function for the multivariate normal distribution for reference.

$$f(\boldsymbol{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{p/2}|\boldsymbol{\Sigma}|^{1/2}} e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})}$$

Then we have the likelihood function for the two independent samples as defined below.

$$\begin{aligned} L(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \boldsymbol{\Sigma}) &= \prod_{i=1}^{n_1} f(\boldsymbol{X}_i; \boldsymbol{\mu}_1, \boldsymbol{\Sigma}) \prod_{j=1}^{n_2} f(\boldsymbol{X}_j; \boldsymbol{\mu}_2, \boldsymbol{\Sigma}) \\ &= L(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}) L(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}) \end{aligned}$$

This statement gives us directly that $\hat{\boldsymbol{\mu}}_1 = \bar{\boldsymbol{x}}_1, \hat{\boldsymbol{\mu}}_2 = \bar{\boldsymbol{x}}_2$. Using equation (4-13) in the text, the likelihood function can defined as

$$L(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{Np/2} |\boldsymbol{\Sigma}|^{N/2}} \exp\left(-\frac{1}{2} \ \mathrm{tr} \left[\boldsymbol{\Sigma}^{-1} \left(\sum_{i=1}^{n_1} \Phi_1(\boldsymbol{X}_i) + \sum_{j=1}^{n_2} \Phi_2(\boldsymbol{X}_j) \right) \right] \right)$$

Where

$$\Phi_i(\boldsymbol{x}) = (\boldsymbol{x} - \boldsymbol{\mu}_i)(\boldsymbol{x} - \boldsymbol{\mu}_i)'$$

By result 4.10 in the text, we let $\boldsymbol{B} = \sum_{i=1}^{n_1} \Phi_1(\boldsymbol{x}_i) + \sum_{j=1}^{n_2} \Phi_2(\boldsymbol{x}_j)$, $b = N = n_1 + n_2$ and substitute the MLE's for each mean, then the maximum is reached at

$$\hat{\boldsymbol{\Sigma}} = \frac{1}{2(n_1 + n_2)} \boldsymbol{B}$$

And we get the result

$$\hat{\boldsymbol{\Sigma}} = \frac{1}{n_1 + n_2} \left[(n_1 - 1)\boldsymbol{S}_1 + (n_2 - 1)\boldsymbol{S}_2 \right] = \frac{n_1 + n_2 - 1}{n_1 + n_2} \boldsymbol{S}_{\text{pooled}}$$