Statistics 351 (Fall 2008)

A Geometric Description of the Bivariate Normal Density Function

Recall that if $\mathbf{X} = (X_1, X_2)' \in \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Lambda})$ with

$$\boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \quad \text{and} \quad \boldsymbol{\Lambda} = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$$

where $\sigma_1 > 0$, $\sigma_2 > 0$, and $-1 < \rho < 1$ so that $\det[\boldsymbol{\Lambda}] > 0$, then the density function of $\mathbf{X}$ is

$$f_{\mathbf{X}}(x_1, x_2)$$

$$= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{ -\frac{1}{2(1-\rho^2)} \left(\left(\frac{x_1 - \mu_1}{\sigma_1} \right)^2 - 2\rho \frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \left(\frac{x_2-\mu_2}{\sigma_2} \right)^2 \right) \right\}. \tag{$*$}$$

Example. Suppose that X_1, X_2 are independent $\mathcal{N}(0, 1)$ random variables so that the joint density of $\mathbf{X} = (X_1, X_2)'$ is

$$f_{\mathbf{X}}(x_1, x_2) = \frac{1}{2\pi} \exp\left\{ -\frac{1}{2}(x_1^2 + x_2^2) \right\}, \quad -\infty < x_1, x_2 < \infty.$$

We can visualize the graph of $f_{\mathbf{X}}(x_1, x_2)$ as a "sombrero" above the x_1, x_2-plane centred at $(0, 0)$. That is, consider the usual one-dimensional bell-curve and imagine rotating it around the x_2 coordinate axis to create the "sombrero."

Formally, the level curves (or contour lines) are found when the function $f_{\mathbf{X}}(x_1, x_2)$ is constant. Setting $f_{\mathbf{X}}(x_1, x_2) = C$ gives

$$C = \frac{1}{2\pi} \exp\left\{ -\frac{1}{2}(x_1^2 + x_2^2) \right\}$$

and so

$$-2\log(2\pi C) = x_1^2 + x_2^2.$$

Note that the maximal value of the function $f_{\mathbf{X}}(x_1, x_2)$ occurs when $(x_1, x_2) = (0, 0)$ and so we have the restriction that $0 < C \le 1/(2\pi)$. Observe that $-2\log(2\pi C) \ge 0$ for this range of C so that $-2\log(2\pi C)$ is just another positive constant, say K. Thus, we see that

$$x_1^2 + x_2^2 = K$$

which describes the equation of a circle of radius $\sqrt{K}$ centred at the origin, and explains our description of the density $f_{\mathbf{X}}(x_1, x_2)$ as a "sombrero."

Example. If, instead, $\mathbf{X} = (X_1, X_2)'$ where X_1 and X_2 are independent with $X_1 \in \mathcal{N}(\mu_1, \sigma_1^2)$ and $X_2 \in \mathcal{N}(\mu_2, \sigma_2^2)$, then the level curves of the density $f_{\mathbf{X}}(x_1, x_2)$ are of the form

$$\left(\frac{x_1 - \mu_1}{\sigma_1} \right)^2 + \left(\frac{x_2 - \mu_2}{\sigma_2} \right)^2 = K$$

which describes an ellipse centred at the point (μ_1, μ_2) whose major and minor axes are parallel to the x_1, x_2 coordinate axes.

Example. Finally, suppose that $\mathbf{X}$ is bivariate normal with density given by $(*)$ so that the level curves of $f_{\mathbf{X}}(x_1, x_2)$ are of the form

$$\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 - 2\rho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2 = K.$$

If we assume that $\rho \neq 0$, then it is not so obvious what the level curves are.

In order to see more clearly what is happening, consider the density function in the form

$$f_{\mathbf{X}}(x_1, x_2) = \frac{1}{2\pi}\frac{1}{\sqrt{\det[\mathbf{\Lambda}]}}\exp\left\{-\frac{1}{2}(\overline{x}-\boldsymbol{\mu})'\mathbf{\Lambda}^{-1}(\overline{x}-\boldsymbol{\mu})\right\}.$$

Since $\mathbf{\Lambda}$ is a covariance matrix, we know that its eigenvalues λ_1, λ_2 are real. Furthermore, it can be diagonalized as

$$\mathbf{\Lambda} = CDC'$$

where $D = \operatorname{diag}(\lambda_1, \lambda_2)$ and C is an orthogonal matrix whose columns are the normalized eigenvectors $\overline{v}_1$, $\overline{v}_2$ of $\mathbf{\Lambda}$. Thus,

$$\mathbf{\Lambda}^{-1} = (CDC')^{-1} = (C')^{-1}D^{-1}C^{-1} = (C^{-1})'D^{-1}C^{-1} = (C')'D^{-1}C' = CD^{-1}C'$$

using the fact that $C^{-1} = C'$ since C is orthogonal. (This also tells us that the eigenvectors of $\mathbf{\Lambda}$ are necessarily the same as the eigenvectors of $\mathbf{\Lambda}^{-1}$.) If we now consider the quadratic form $Q(\overline{x})$ based on $\mathbf{\Lambda}^{-1}$, then we see that

$$Q(\overline{x}) = \overline{x}'\mathbf{\Lambda}^{-1}\overline{x} = \overline{x}'CD^{-1}C'\overline{x} = (\overline{x}'C)D^{-1}(C'\overline{x}) = \overline{y}'D^{-1}\overline{y}$$

where $\overline{y} = C'\overline{x}$. Since $D^{-1} = \operatorname{diag}(\lambda_1^{-1}, \lambda_2^{-1})$, we conclude

$$Q(\overline{x}) = \frac{y_1^2}{\lambda_1} + \frac{y_2^2}{\lambda_2}$$

which describes the equation of an ellipse centred at $(0,0)$ whose major and minor axes are parallel to the y_1, y_2 coordinate axes. (Notice that the major axis will be the one corresponding to the largest eigenvalue and that the minor axis will be the one corresponding to the smallest eigenvalue.) However, C is an orthogonal matrix so that its columns are orthonormal. Thus, the action of C applied to any vector in $\mathbb{R}^2$ is simply to rotate that vector. In particular, the angle between any two vectors is preserved. Formally, if $\overline{e}_1 = (1,0)'$ and $\overline{e}_2 = (0,1)'$ denote the basis vectors for the y_1, y_2 coordinate axes, then since the columns of C are the eigenvectors of $\mathbf{\Lambda}$ we see that $C\overline{e}_1 = \overline{v}_1$ and $C\overline{e}_2 = \overline{v}_2$ where $\overline{v}_1$ and $\overline{v}_2$ are the eigenvectors of $\mathbf{\Lambda}$. Since $C^{-1} = C'$ this says that $C'\overline{v}_1 = \overline{e}_1$ and $C'\overline{v}_2 = \overline{e}_2$. Furthermore, since C is an angle-preserving rotation of the plane, so too is its inverse C'. Thus, we conclude that C' transforms the x_1, x_2-plane to the y_1, y_2-plane sending the orthonormal basis $(\overline{v}_1, \overline{v}_2)$ to the orthonormal basis $(\overline{e}_1, \overline{e}_2)$. Hence, the quadratic form

$$Q(\overline{x}) = \overline{x}'\mathbf{\Lambda}^{-1}\overline{x}$$

describes an ellipse whose major and minor axes are parallel to the eigenvectors of $\mathbf{\Lambda}$. If we consider the quadratic form

$$(\overline{x}-\boldsymbol{\mu})'\mathbf{\Lambda}^{-1}(\overline{x}-\boldsymbol{\mu})$$

instead, then we see that this describes an ellipse centred at (μ_1, μ_2) whose major and minor axes are parallel to the $\overline{v}_1$, $\overline{v}_2$ coordinate axes. That is, the effect of subtracting $\boldsymbol{\mu}$ is to translate the entire picture so that it is centred at (μ_1, μ_2) and that the major and minor axes are the $\overline{v}_1 + \boldsymbol{\mu}$, $\overline{v}_2 + \boldsymbol{\mu}$ coordinate axes

Example. Suppose that the random vector $\mathbf{X} = (X_1, X_2)'$ has a multivariate normal distribution with mean vector $\boldsymbol{\mu} = \overline{0}$ and covariance matrix

$$\mathbf{\Lambda} = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$$

where $0 < \rho < 1$. The level curves of $f_{\mathbf{X}}(x_1, x_2)$ are of the form

$$x_1^2 - 2\rho x_1 x_2 + x_2^2 = K.$$

In order to sketch the level curves, we will diagonalize $\mathbf{\Lambda}$. The characteristic equation for $\mathbf{\Lambda}$ is

$$0 = \det[\mathbf{\Lambda} - \lambda I] = (1-\lambda)^2 - \rho^2 = \lambda^2 - 2\lambda + (1-\rho^2) = (\lambda - 1 - \rho)(\lambda - 1 + \rho)$$

which implies the eigenvalues of $\mathbf{\Lambda}$ are

$$\lambda_1 = 1 + \rho \quad \text{and} \quad \lambda_2 = 1 - \rho.$$

The corresponding normalized eigenvectors are

$$v_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} \quad \text{and} \quad v_2 = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}.$$

Thus, the level curves are ellipses rotated $45°$. That is, the orthogonal matrix C which diagonalizes $\mathbf{\Lambda}$ is

$$C = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} \cos(\pi/4) & -\sin(\pi/4) \\ \sin(\pi/4) & \cos(\pi/4) \end{bmatrix}$$

which is the *rotation matrix* that rotates the standard basis vectors counterclockwise by $\pi/4 = 45°$. Since $0 < \rho < 1$ so that $\lambda_1 > \lambda_2$, we see that the major axis is parallel v_1 and the minor axis is parallel to v_2.

Remark. An alternative way to see that the level curves are ellipses rotated by $45°$ is to write

$$x_1^2 - 2\rho x_1 x_2 + x_2^2 = (1+\rho)\left(\frac{x_1}{\sqrt{2}} - \frac{x_2}{\sqrt{2}}\right)^2 + (1-\rho)\left(\frac{x_1}{\sqrt{2}} + \frac{x_2}{\sqrt{2}}\right)^2.$$