Lecture 11

ECE 278 Mathematics for MS Comp Exam

Multivariate gaussian probability density function

- A ***multivariate gaussian probability density function*** is a joint probability density function for a block $\mathbf{x}$ of real random variables with components $\underline{x}_i$ given by

$$f_{\underline{\mathbf{x}}}(\mathbf{x}) \quad = \quad \frac{1}{\sqrt{(2\pi)^N \det \mathbb{C}}} e^{-\frac{1}{2}(\mathbf{x}-\langle \mathbf{x} \rangle)^T \mathbb{C}^{-1}(\mathbf{x}-\langle \mathbf{x} \rangle)}. \qquad (1)$$

- The matrix $\mathbb{C}$ is the ***real covariance matrix*** defined for any multivariate probability density function as

$$\mathbb{C} \quad = \quad \left\langle \left(\underline{\mathbf{x}} - \langle \mathbf{x} \rangle \right) \left(\underline{\mathbf{x}} - \langle \mathbf{x} \rangle \right)^T \right\rangle . \qquad (2)$$

- The square symmetric matrix $\mathbb{C}$ has a determinant $\det \mathbb{C}$.

- The diagonal matrix element C_{ii} is the variance of the random variable $\underline{x}_i$.

Correlation and Independence

- The off-diagonal matrix element C_{ij} is the covariance of the two random variables $\underline{x}_i$ and $\underline{x}_j$. These two elements are uncorrelated if C_{ij} equals zero.
- In general, this need not be a strong statement, but for jointly gaussian random variables, it means that they are independent.
- It is possible to have a joint probability density function such that each marginal density function is a gaussian probability density function, yet the joint probability density function is not jointly gaussian, and not given by (1).
    - asked on this years MS Comp exam.
- This means that knowing that each marginal probability density function is gaussian is not sufficient to infer that the joint probability density function is jointly gaussian.

Bivariate Gaussian Random Variables

- A zero-mean bivariate gaussian random variable consists of two random, zero-mean gaussian components $\underline{x}$ and $\underline{y}$, which may be correlated.
- The covariance matrix given in (2) is

$$\mathbb{C} \quad = \quad \begin{bmatrix} \sigma_x^2 & \rho_{xy}\sigma_x\sigma_y \\ \rho_{xy}\sigma_x\sigma_y & \sigma_y^2 \end{bmatrix}, \tag{3}$$

where

$$\rho_{xy} \quad \doteq \quad \left\langle \underline{xy} \right\rangle / \sigma_x\sigma_y, \tag{4}$$

is defined as the ***correlation coefficient***.

Bivariate gaussian random variable-2

- An example of a two-dimensional joint gaussian probability density function is shown in plan view in Figure 1.

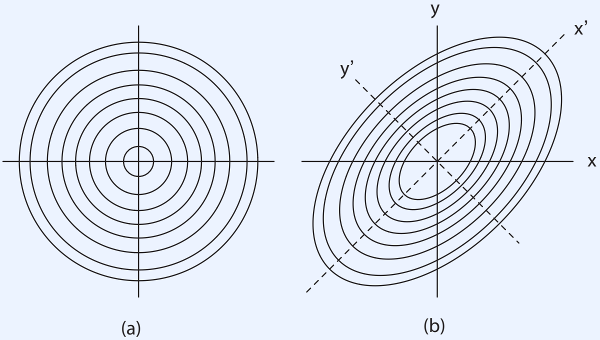


Figure: Contours of the joint gaussian probability density function $f_{\underline{x},\underline{y}}(x,y)$ as a function of the correlation coefficient ρ_{xy}: (a) $\rho_{xy} = 0$, (b) $\rho_{xy} = 0.5$.

Bivariate gaussian random variable-3

- If $\sigma_x = \sigma_y = \sigma$, then (1) reduces to

$$f_{\underline{x},\underline{y}}(x,y) \quad = \quad \frac{1}{2\pi\sigma^2\sqrt{1-\rho_{xy}^2}} \exp\left(-\frac{x^2 - 2\rho_{xy}xy + y^2}{2\sigma^2(1-\rho_{xy}^2)}\right). \qquad (5)$$

Moreover, if $\rho_{xy} = 0$, then $f_{\underline{x},\underline{y}}(x,y)$ is a product distribution in the chosen coordinate system.

Bivariate gaussian random variable-4

- For this case, the bivariate gaussian density function is called a ***circularly-symmetric*** density function with the bivariate gaussian random variable called a ***circularly-symmetric gaussian random variable***.

- The joint gaussian probability density function $f_{\underline{x},\underline{y}}(x,y)$, now also including a nonzero mean for each component, can then be written as

$$f_{\underline{x},\underline{y}}(x,y) \quad = \quad \left[\frac{1}{\sqrt{2\pi}\sigma}e^{-(x-\langle\underline{x}\rangle)^2/2\sigma^2}\right]\left[\frac{1}{\sqrt{2\pi}\sigma}e^{-(y-\langle\underline{y}\rangle)^2/2\sigma^2}\right], \tag{6}$$

  where the probability density function of each component is a 1-D gaussian

- Therefore, uncorrelated jointly gaussian random variables are independent.

Multivariate gaussian random variable-again

- For a set of N independent real gaussian random variables with $\mathbb{C} = \sigma^2 \mathbb{I}_N$, the joint probability density function is

$$f_{\underline{\mathbf{x}}}(\mathbf{x}) = \frac{1}{\sqrt{(2\pi\sigma^2)^N}} e^{-(\mathbf{x}-\langle \mathbf{x} \rangle)^2/2\sigma^2}, \tag{7}$$

which factors as a product of the N single variable gaussian densities.

Diagonalization of a Covariance Matrix

- A real covariance matrix is a symmetric matrix and can be diagonalized by a change of basis.
- Therefore, any multivariate gaussian probability density function has basis for which the probability density function expressed in this basis is a product distribution.
- The resulting marginal gaussian random variables in this basis are independent, but need not have the same mean and variance.
- For example, consider the two-dimensional gaussian probability density function given in (5) with diagonal elements $\sigma_x^2 = \sigma_y^2 = \sigma^2$ and off-diagonal elements $\sigma^2 \rho_{xy}$.

UCSD Jacobs | Electrical and Computer Engineering

Diagonalization of a Covariance Matrix-2

- Define a new basis (x', y') that is a rotation of the original basis (x, y).
- The components in the new basis for this example can be expressed by a unitary transformation $\mathbb{R}$ of components in the original basis as given by

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \mathbb{R} \begin{bmatrix} x \\ y \end{bmatrix}.$$

- The matrix $\mathbb{R}$ is generated from the normalized eigenvectors of the covariance matrix $\mathbb{C}$ given in (3) and satisfies the matrix equation

$$\mathbb{R}^T \mathbb{C} \mathbb{R} = \mathbb{D},$$

where $\mathbb{D}$ is a diagonal matrix with diagonal elements given by the eigenvalues of $\mathbb{C}$.

Diagonalization of a Covariance Matrix-3

- Using these eigenvalues, the variances of the uncorrelated gaussian random variables in this new basis are $\sigma_{x'}^2 = \sigma^2(1+\rho_{xy})$ and $\sigma_{y'}^2 = \sigma^2(1-\rho_{xy})$, which can be equal only if $\rho_{xy} = 0$.
- Using the normalized eigenvectors of $\mathbb{C}$, the components of the new basis are $x' = \frac{1}{\sqrt{2}}(x+y)$ and $y' = \frac{1}{\sqrt{2}}(x-y)$.
- The joint gaussian probability density function in the new basis is a product distribution given by

$$\begin{aligned} f(x',y') &= \left[\frac{1}{\sqrt{2\pi\sigma^2(1+\rho_{xy})}}e^{-x'^2/2\sigma^2(1+\rho_{xy})}\right]\times \\ &\quad \left[\frac{1}{\sqrt{2\pi\sigma^2(1-\rho_{xy})}}e^{-y'^2/2\sigma^2(1-\rho_{xy})}\right], \end{aligned}$$

which is written to show that each marginal probability density function in the new basis is an independent gaussian probability density function with each distribution having a different variance.

Example 1

- A real autocovariance matrix $\mathbb{C}$ of a of a real multivariate gaussian random variable is given by

$$\mathbb{C} = \begin{bmatrix} 1 & 1 \\ 1 & 4 \end{bmatrix}$$

- Find the basis that yields a product distribution
- **Solution**
- The inverse of $\mathbb{C}$ is

$$\mathbb{C}^{-1} = \frac{\mathbf{1}}{\mathbf{3}} \begin{pmatrix} 4 & -1 \\ -1 & 1 \end{pmatrix}$$

  and the determinant is 3.

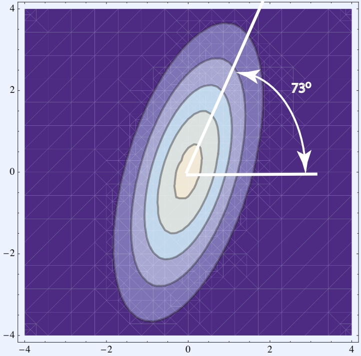
- Using these expressions, the joint distribution is

$$\begin{aligned} f_{\underline{\mathbf{x}}}(\mathbf{x}) &= \frac{1}{(2\pi)^{N/2}\sqrt{\det \mathbb{C}}} e^{-\frac{1}{2}(\mathbf{x}-\langle\mathbf{x}\rangle)^T \mathbb{C}^{-1}(\mathbf{x}-\langle\mathbf{x}\rangle)} \\ &= \frac{1}{6\pi} e^{-\left(4x^2-2xy+y^2\right)/6} \end{aligned}$$

  where $\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}$.

Example 1 -cont.

- A plot of the distribution is below



Example 1-cont.

- To find the the angle θ between the x axis and the x' axis. (This is the rotation of the unitary matrix that diagonalizes $\mathbb{C}$.)
- we need the eigenvalues of the autocovariance matrix, which are

$$\frac{1}{2}\left(5 \pm \sqrt{13}\right)$$

and the eigenvectors are

$$\begin{bmatrix} \frac{1}{2}\left(5+\sqrt{13}\right)-4 \\ 1 \end{bmatrix} \text{ and } \begin{bmatrix} \frac{1}{2}\left(5-\sqrt{13}\right)-4 \\ 1 \end{bmatrix}.$$

- The angle of the major axis that corresponds to the largest eigenvalue with respect to the x-axis is the angle of the corresponding eigenvector

$$\tan^{-1}\left(\frac{y}{x}\right) = \tan^{-1}\left(\frac{1}{\frac{1}{2}\left(5+\sqrt{13}\right)-4}\right) = 73.155^{o}$$

Example 2

- The bivariate gaussian probability density function has the form

$$p_{\underline{x},\underline{y}}(x,y) \quad = \quad Ae^{-(ax^2+2bxy+cy^2)}.$$

- (a) Express the constant A in terms of a, b, and c.
- **Solution**
- Express the joint probability density function in the standard form of a multivariant gaussian distribution

$$f_{\underline{x},\underline{y}}(x,y) = \frac{1}{\sqrt{(2\pi)^N \det \mathbb{C}}} e^{-\frac{1}{2}(\mathbf{x}-\langle \mathbf{x} \rangle)^T \mathbb{C}^{-1}(\mathbf{x}-\langle \mathbf{x} \rangle)},$$

where $\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}$.

- Comparing to the standard form the means $\langle x \rangle$ and $\langle y \rangle$ are both zero because there is no constant term.

Example 2-cont.

- Moreover, the inverse of the covariance matrix $\mathbb{C}^{-1}$ is

$$\mathbb{C}^{-1} = 2 \begin{bmatrix} a & b \\ b & c \end{bmatrix}.$$

- The covariance matrix given by

$$\mathbb{C} = \frac{1}{2(ac - b^2)} \begin{bmatrix} c & -b \\ -b & a \end{bmatrix}.$$

- The determinant of $\mathbb{C}$ is

$$\det \mathbb{C} = \frac{1}{4(ac - b^2)}.$$

Example 2-cont.

- The variances are diagonal terms of the covariance matrix and are given by

$$\sigma_x^2 = \frac{c}{4(ac-b^2)} \qquad \sigma_y^2 = \frac{a}{4(ac-b^2)}.$$

- Using these expressions, the normalization is

$$A = \frac{1}{\sqrt{(2\pi)^2 \det \mathbb{C}}} = \frac{\sqrt{ac-b^2}}{\pi}.$$

Example 2-cont.

- (b) Find the marginals, $p_{\underline{x}}(x)$ and $p_{\underline{y}}(y)$, the conditionals $p_{\underline{x}|\underline{y}}(x|y)$ and $p_{\underline{y}|\underline{x}}(y|x)$, the variances $\sigma^2_{\underline{x}}$, $\sigma^2_{\underline{y}}$, and the correlation $\langle \underline{x}\underline{y} \rangle$.
- The means and the variances were derived in part (a). The marginals are given by

$$f_{\underline{x}}(x) = \frac{1}{\sqrt{2\pi\sigma^2_{\underline{x}}}} e^{-x^2/2\sigma^2_x} \qquad f_{\underline{y}}(y) = \frac{1}{\sqrt{2\pi\sigma^2_{\underline{y}}}} e^{-y^2/2\sigma^2_y}$$

- The correlation $\langle \underline{x}\underline{y} \rangle$ is

$$\langle \underline{x}\underline{y} \rangle = \frac{\sqrt{ac-b^2}}{\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy e^{-(ax^2+2bxy+cy^2)} \mathrm{d}x\mathrm{d}y = -\frac{b}{2(ac-b^2)},$$

which is simply the off-diagonal element of the covariance matrix.

Example 2-cont.

- Finally, the conditional distributions are

$$f_{\underline{x}|\underline{y}}(x|y) = \frac{f_{\underline{x},\underline{y}}(x,y)}{f_{\underline{y}}(y)} \qquad p_{\underline{y}|\underline{x}}(y|x) = \frac{f_{\underline{x},\underline{y}}(x,y)}{f_{\underline{x}}(x)}.$$

Example 3-cont.

Consider a joint probability density function $p(x, y)$ is given as

$$p(x,y) = \begin{cases} \frac{1}{2\pi\sigma_x\sigma_y}\exp\left[-\frac{1}{2}\left(\frac{x^2}{2\sigma_x^2}+\frac{y^2}{2\sigma_y^2}\right)\right] & \text{if } xy > 0 \\ 0 & \text{if } xy < 0 \end{cases}.$$

Example 2-cont.

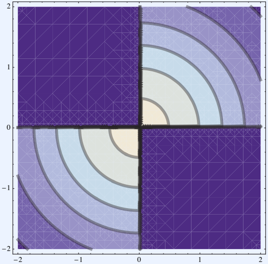
- (a) Show that this function is a valid probability density function.
- **Solution**
- The integral of the joint probability distribution separates to an integral over x and an integral over y. The integral on either x or y is half the value over the whole plane. Therefore

$$\begin{aligned} p_{\underline{x}}(x) &= \frac{1}{2}\int_{\infty}^{\infty} p_{\underline{x}\underline{y}}(x,y)dx \\ &= \frac{1}{4\pi\sigma_x\sigma_y}\exp\left[-\frac{y^2}{4\sigma_y^2}\right]\underbrace{\int_{\infty}^{\infty}\exp\left[-\frac{x^2}{4\sigma_x^2}\right]dx}_{2\sigma_x\sqrt{\pi}} \\ &= \frac{1}{2\sqrt{\pi}\sigma_y}\exp\left[-\frac{y^2}{4\sigma_y^2}\right]. \end{aligned}$$

Including an additional factor of two because of symmetry, this expression integrates to one so that the distribution is a valid probability density function.

Example 3-cont.

- (b) Sketch $p(x, y)$ in plan view and in three dimensions. Is this joint probability density function jointly gaussian?
- **Solution**
- The plot of the function is on the next page for $\sigma_x = \sigma_y = 1$.



-
- It is nonzero in the first and third quadrants when $xy > 0$. It is zero elsewhere. This joint probability density function is not jointly gaussian.

Example 3-cont.

- (c) Find the marginal probability density functions $p_{\underline{x}}(x)$ and $p_{\underline{y}}(y)$ and comment on this result.
- **Solution**
- The marginal distribution for y is of the same form so that each is gaussian. Therefore, knowing that each marginal distribution is a gaussian is not sufficient to infer that the joint distribution is jointly gaussian.