The multivariate Gaussian distribution

February 25, 2022

The multivariate Gaussian distribution

Covariance matrices

Let $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)^t$ be an n-dimensional random vector with expectation $\boldsymbol{m}_X = \mathbb{E}(\boldsymbol{X}) = (m_{X_1}, m_{X_2}, \ldots, m_{X_n})^t$.

Definition

The covariance matrix of $\boldsymbol{X}$ is the square $n \times n$ matrix $\boldsymbol{K}_X = (k_{ij})$ defined as

$$\boldsymbol{K}_X = \mathbb{E}\left((\boldsymbol{X} - \boldsymbol{m}_X)(\boldsymbol{X} - \boldsymbol{m}_X)^t\right)$$

Therefore

- *If $i \neq j$, then $k_{ij} = \mathbb{E}\left((X_i - m_{X_i})(X_j - m_{X_j})\right) = \mathrm{Cov}(X_i, X_j)$.*
- *The diagonal entries are $k_{ii} = \mathbb{E}\left((X_i - m_{X_i})^2\right) = \mathrm{Var}(X_i)$.*

Covariance matrices

Proposition

The covariance matrix $\boldsymbol{K}_X$ is:

- *symmetric, $k_{ij} = \mathrm{Cov}(X_i, X_j) = \mathrm{Cov}(X_j, X_i) = k_{ji}$.*
- *positive-semidefinite; that is to say, $\boldsymbol{z}^t \boldsymbol{K}_X \boldsymbol{z} \geqslant 0$ for all $\boldsymbol{z} = (z_1, z_2, \ldots, z_n)^t \in \mathbb{R}^n$.*

Proof: Let $Y = z_1 X_1 + \cdots + z_n X_n = \boldsymbol{z}^t \boldsymbol{X}$. Notice that

$$m_Y = \mathbb{E}\left(\sum_{i=1}^n z_i X_i\right) = \sum_{i=1}^n z_i m_{X_i} = \boldsymbol{z}^t \boldsymbol{m}_X$$

and hence

$$Y - m_Y = \boldsymbol{z}^t (\boldsymbol{X} - \boldsymbol{m}_X)$$

Covariance matrices

Therefore,

$$\begin{aligned} \mathbf{z}^t \mathbf{K}_X \mathbf{z} &= \mathbf{z}^t \,\mathbb{E}\left((\mathbf{X} - \mathbf{m}_X)(\mathbf{X} - \mathbf{m}_X)^t\right) \mathbf{z} \\ &= \mathbb{E}\left(\mathbf{z}^t (\mathbf{X} - \mathbf{m}_X)(\mathbf{X} - \mathbf{m}_X)^t \mathbf{z}\right) \\ &= \mathbb{E}\left((Y - m_Y)(Y - m_Y)^t\right) \\ &= \mathbb{E}\left((Y - m_Y)^2\right) = \mathrm{Var}(Y) \geqslant 0 \end{aligned}$$

Covariance matrices

Let us say that the "centered" random variables

$$X_1 - m_{X_1},\ X_2 - m_{X_2},\ \ldots,\ X_n - m_{X_n}$$

are linearly independent (in the Linear Algebra sense) if the equality

$$\sum_{i=0}^{n} z_i\,(X_i - m_{X_i}) = 0 \quad \text{(with probability 1)}$$

implies $z_1 = z_2 = \cdots = z_n = 0$.

Covariance matrices

Theorem

The random variables

$$X_1 - m_{X_1},\ X_2 - m_{X_2},\ \ldots,\ X_n - m_{X_n}$$

are linearly independent if and only if $\mathbf{K}_X$ is positive-definite; that is, if and only if

$$\mathbf{z}^t \mathbf{K}_X \mathbf{z} > 0 \quad \text{for all } \mathbf{z} \neq 0.$$

Covariance matrices

Proof: Let $Y = \mathbf{z}^t \mathbf{X}$ and observe that

$$\begin{aligned} &\mathbf{z}^t \mathbf{K}_X \mathbf{z} = 0 \text{ for some } \mathbf{z} \neq 0 \\ &\iff \mathrm{Var}(Y) = 0 \text{ for some } \mathbf{z} \neq 0 \\ &\iff Y - m_Y = 0 \ \text{(with probability 1) for some } \mathbf{z} \neq 0 \\ &\iff \sum_{i=1}^{n} z_i\,(X_i - m_{X_i}) = 0 \ \text{(with probability 1)} \\ &\qquad\qquad \text{for some } \mathbf{z} = (z_1, z_2, \ldots, z_n)^t \neq 0 \\ &\iff X_1 - m_{X_1}, X_2 - m_{X_2}, \ldots, X_n - m_{X_n} \\ &\qquad\qquad \text{are not linealy independent.} \end{aligned}$$

Linear transformations

Theorem

Let $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)^t$ be an n-dimensional random vector, let $\boldsymbol{A}$ be an $m \times n$ real matrix, let $\boldsymbol{b}$ be a constant real $m \times 1$ vector, and let $\boldsymbol{Y} = (Y_1, Y_2, \ldots, Y_m)^t$ be the m-dimensional random vector defined as $\boldsymbol{Y} = \boldsymbol{A}\,\boldsymbol{X} + \boldsymbol{b}$.

Then

$$\boldsymbol{m}_Y = \boldsymbol{A}\,\boldsymbol{m}_X + \boldsymbol{b}, \qquad \boldsymbol{K}_Y = \boldsymbol{A}\,\boldsymbol{K}_X\,\boldsymbol{A}^t$$

Observe that if $\boldsymbol{A} = (a_{ij})$ and $\boldsymbol{b} = (b_i)$, then

$$Y_i = a_{i1}X_1 + \cdots + a_{in}X_n + b_i, \quad 1 \leqslant i \leqslant m$$

Linear transformations

Proof: By the linearity of the expectation operator we have

$$\boldsymbol{m}_Y = \mathbb{E}(\boldsymbol{Y}) = \mathbb{E}(\boldsymbol{AX} + \boldsymbol{b}) = \boldsymbol{A}\,\mathbb{E}(\boldsymbol{X}) + \boldsymbol{b} = \boldsymbol{Am}_X + \boldsymbol{b}$$

Analogously,

$$\begin{aligned}\boldsymbol{K}_Y &= \mathbb{E}\left((\boldsymbol{Y} - \boldsymbol{m}_Y)(\boldsymbol{Y} - \boldsymbol{m}_Y)^t\right) \\ &= \mathbb{E}\left(\boldsymbol{A}\,(\boldsymbol{X} - \boldsymbol{m}_X)(\boldsymbol{X} - \boldsymbol{m}_X)^t\,\boldsymbol{A}^t\right) \\ &= \boldsymbol{A}\,\mathbb{E}\left((\boldsymbol{X} - \boldsymbol{m}_X)(\boldsymbol{X} - \boldsymbol{m}_X)^t\right)\boldsymbol{A}^t = \boldsymbol{AK}_X\boldsymbol{A}^t\end{aligned}$$

Gaussian characteristic functions

Let X_1, X_2, …, X_n be independent Gaussian random variables, $X_i \sim \mathrm{N}(m_{X_i}, \sigma^2_{X_i})$. Their joint characteristic function is

$$\begin{aligned}M_X(\omega_1, \omega_2, \ldots, \omega_n) &= M_{X_1}(\omega_1)M_{X_2}(\omega_2)\cdots M_{X_n}(\omega_n) \\ &= \prod_{i=1}^{n} \exp\left(i\omega_i m_{X_i} - \frac{1}{2}\sigma^2_{X_i}\omega_i^2\right) \\ &= \exp\left(\sum_{i=1}^{n}\left(i\omega_i m_{X_i} - \frac{1}{2}\sigma^2_{X_i}\omega_i^2\right)\right) \\ &= \exp\left(i\boldsymbol{\omega}^t\boldsymbol{m}_X - \frac{1}{2}\boldsymbol{\omega}^t\boldsymbol{K}_X\boldsymbol{\omega}\right)\end{aligned}$$

Gaussian characteristic functions

- $\boldsymbol{\omega} = (\omega_1, \omega_2, \ldots, \omega_n)^t$ and $\boldsymbol{m}_X = (m_{X_1}, \ldots, m_{X_n})$ is the expectation vector.
- Moreover,

$$\boldsymbol{K}_X = \begin{pmatrix} \sigma^2_{X_1} & 0 & . & . & . & 0 \\ 0 & \sigma^2_{X_2} & 0 & . & . & 0 \\ & & . & & & \\ & & & . & & \\ 0 & . & . & . & 0 & \sigma^2_{X_n} \end{pmatrix}$$

is the covariance matrix.

The matrix $\boldsymbol{K}_X$ is diagonal, because the random variables X_i, $1 \leqslant i \leqslant n$, are independent and hence $\mathrm{Cov}(X_i, X_j) = 0$ if $i \neq j$.

Gaussian random vectors

Definition

An n-dimensional random vector $\boldsymbol{X} = (X_1, X_2, \dots, X_n)^t$ is Gaussian if its characteristic function has the form

$$M_X(\omega_1, \omega_2, \dots, \omega_n) = \exp\left(i\boldsymbol{\omega}^t \boldsymbol{m} - \frac{1}{2}\boldsymbol{\omega}^t \boldsymbol{K}\boldsymbol{\omega}\right),$$

where $\boldsymbol{\omega}^t = (\omega_1, \omega_2, \dots, \omega_n)$, $\boldsymbol{m}$ is a column $n \times 1$ vector of real numbers, and $\boldsymbol{K}$ is a symmetric positive-semidefinite $n \times n$ matrix.

We write $\boldsymbol{X} \sim \mathrm{N}(\boldsymbol{m}, \boldsymbol{K})$ or, equivalently, we say that the random variables X_1, X_2, …, X_n are jointly Gaussian.

Marginal distributions

If $\boldsymbol{m} = (m_j)$ and $\boldsymbol{K} = (k_{ij})$, then

$$M_{X_j}(\omega) = M_X(0, \dots, \omega, \dots, 0) = \exp\left(i\, m_j\, \omega - \frac{1}{2} k_{jj}\, \omega^2\right)$$

- It can be shown that $k_{jj} \geqslant 0$ for all j, $1 \leqslant j \leqslant n$, because $\boldsymbol{K}$ is positive-semidefinite.

If $k_{jj} > 0$, then the component X_j is a 1-dimensional Gaussian random variable with parameters

$$\mathbb{E}(X_j) = m_j, \quad \mathrm{Var}(X_j) = k_{jj}$$

Marginal distributions

- If $k_{jj} = 0$ for some j, then

$$M_{X_j}(\omega) = e^{i m_j \omega}$$

Hence $\mathbb{P}(X_j = m_j) = 1$ and so $X_j = m_j$ (with probability 1).

We can regard the "constant" variable X_j as a "degenerate" $\mathrm{N}(m_j, 0)$ random variable.

Marginal distributions

Moreover,

$$m_{11;X_r X_s} = \frac{1}{i^2} \left. \frac{\partial^2 M_X(\omega_1, \omega_2, \dots, \omega_n)}{\partial \omega_r\, \partial \omega_s} \right|_{(0,0,\dots,0)} = k_{rs} + m_r m_s$$

Therefore

$$\mathrm{Cov}(X_r, X_s) = k_{rs}$$

- So we conclude that $\boldsymbol{K}$ is the covariance matrix of $\boldsymbol{X}$.

Eigenvalues of the covariance matrix

Since $\boldsymbol{K}_X$ is a symmetric matrix, it can be converted into a diagonal matrix by means of an orthogonal transformation.

There exists an orthogonal matrix $\boldsymbol{C}$ (i.e., $\boldsymbol{C}\boldsymbol{C}^t = \boldsymbol{C}^t\boldsymbol{C} = \boldsymbol{I}$) such that

$$\boldsymbol{C}\boldsymbol{K}_X\boldsymbol{C}^t = \boldsymbol{D} = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n),$$

where the numbers $\lambda_i \in \mathbb{R}$ are the eigenvalues of $\boldsymbol{K}_X$.

Equivalently,

$$\boldsymbol{K}_X = \boldsymbol{C}^t\boldsymbol{D}\boldsymbol{C}$$

Eigenvalues of the covariance matrix

Hence,

$$0 \leqslant \boldsymbol{z}^t\boldsymbol{K}_X\boldsymbol{z} = \boldsymbol{z}^t(\boldsymbol{C}^t\boldsymbol{D}\boldsymbol{C})\boldsymbol{z} = (\boldsymbol{C}\boldsymbol{z})^t\boldsymbol{D}\,(\boldsymbol{C}\boldsymbol{z})$$
$$= \boldsymbol{y}^t\boldsymbol{D}\boldsymbol{y} = \sum_{i=1}^{n} \lambda_i y_i^2,$$

where $\boldsymbol{y} = \boldsymbol{C}\boldsymbol{z} \in \mathbb{R}^n$ can be arbitrarily chosen, because $\boldsymbol{z} = \boldsymbol{C}^t\boldsymbol{y}$.

We deduce that all the eigenvalues of $\boldsymbol{K}_X$ are nonnegative,

$$\lambda_1 \geqslant 0, \lambda_2 \geqslant 0, \ldots, \lambda_n \geqslant 0$$

Linear independence of the components

Proposition

Let $\boldsymbol{X}$ be a Gaussian random vector. Then the covariance matrix $\boldsymbol{K}_X$ is positive-definite if and only if any of the following statements hold:

- $\lambda_i > 0$ *for all* $1 \leqslant i \leqslant n$.
- *The variables* $X_i - m_{X_i}$, $1 \leqslant i \leqslant n$, *are linearly independent.*
- $\det(\boldsymbol{K}_X) > 0$.
- $\boldsymbol{K}_X$ *has an inverse.*
- *The multivariate density* $f_X(x_1, x_2, \ldots, x_n)$ *exists.*

Linear independence of the components

(Remark)

Gaussian random vectors are also defined if $\boldsymbol{K}_X$ is not invertible (i.e., if $\det(K_X) = 0$).

In such cases, the Gaussian vector $\boldsymbol{X}$ has not a density and either one has

- $X_j = m_j$ *(with probability 1) for some* j, $1 \leqslant j \leqslant n$,

or, more generally,

- *the centered variables* $X_i - m_{X_i}$, $1 \leqslant i \leqslant n$, *are not linearly independent.*

Uncorrelation and independence

Theorem

If the random variables X_1, X_2, ..., X_n are jointly Gaussian and pairwise uncorrelated, then they are jointly independent.

Proof:

$$\mathrm{Cov}(X_i, X_j) = 0 \Longrightarrow \boldsymbol{K}_X = \mathrm{diag}(\sigma^2_{X_1}, \sigma^2_{X_2}, \ldots, \sigma^2_{X_n})$$

Uncorrelation and independence

Therefore

$$\begin{aligned} M_X(\omega_1, \omega_2, \ldots, \omega_n) &= \exp\left(i\boldsymbol{\omega}^t \boldsymbol{m}_X - \frac{1}{2}\boldsymbol{\omega}^t \boldsymbol{K}_X \boldsymbol{\omega}\right) \\ &= \exp\left(\sum_{k=1}^{n}\left(i\omega_k m_{X_k} - \frac{1}{2}\sigma^2_{X_k}\omega_k^2\right)\right) \\ &= \prod_{k=1}^{n} \exp\left(i\omega_k m_{X_k} - \frac{1}{2}\sigma^2_{X_k}\omega_k^2\right) \\ &= M_{X_1}(\omega_1) M_{X_2}(\omega_2) \cdots M_{X_n}(\omega_n) \end{aligned}$$

Linear combinations

Theorem

Let $\boldsymbol{X}$ be an n-dimensional Gaussian random vector, let $\boldsymbol{A}$ be an $m \times n$ real matrix, let $\boldsymbol{b}$ be a constant real $m \times 1$ vector, and let $\boldsymbol{Y} = \boldsymbol{A}\boldsymbol{X} + \boldsymbol{b}$.

Then, $\boldsymbol{Y}$ is an m-dimensional Gaussian random vector with

$$\boldsymbol{m}_Y = \boldsymbol{A}\,\boldsymbol{m}_X + \boldsymbol{b}, \qquad \boldsymbol{K}_Y = \boldsymbol{A}\boldsymbol{K}_X\boldsymbol{A}^t$$

- If $m \leqslant n$, $\boldsymbol{A}$ has full rang m, and $\boldsymbol{X}$ has a probability density $f_X(x_1, \ldots, x_n)$, then the Gaussian random vector $\boldsymbol{Y}$ also has a density $f_Y(y_1, \ldots, y_m)$.

Linear combinations

Proof: It only remains to be proved that $\boldsymbol{Y}$ is Gaussian.

$$\begin{aligned} M_Y(\omega_1, \omega_2, \ldots, \omega_n) &= \mathbb{E}\left(e^{i\boldsymbol{\omega}^t \boldsymbol{Y}}\right) \\ &= \mathbb{E}\left(e^{i\boldsymbol{\omega}^t(\boldsymbol{A}\boldsymbol{X}+\boldsymbol{b})}\right) = e^{i\boldsymbol{\omega}^t\boldsymbol{b}}\, M_X(\boldsymbol{\omega}^t\boldsymbol{A}) \\ &= \exp\left(i\boldsymbol{\omega}^t\boldsymbol{b}\right) \cdot \exp\left(i\,(\boldsymbol{\omega}^t\boldsymbol{A})\boldsymbol{m}_X) - \frac{1}{2}(\boldsymbol{\omega}^t\boldsymbol{A})\boldsymbol{K}_X(\boldsymbol{\omega}^t\boldsymbol{A})^t\right) \\ &= \exp\left(i\boldsymbol{\omega}^t\boldsymbol{b}\right) \cdot \exp\left(i\,\boldsymbol{\omega}^t(\boldsymbol{A}\boldsymbol{m}_X) - \frac{1}{2}\,\boldsymbol{\omega}^t(\boldsymbol{A}\boldsymbol{K}_X\boldsymbol{A}^t)\boldsymbol{\omega}\right) \\ &= \exp\left(i\,\boldsymbol{\omega}^t(\boldsymbol{A}\boldsymbol{m}_X + \boldsymbol{b}) - \frac{1}{2}\,\boldsymbol{\omega}^t(\boldsymbol{A}\boldsymbol{K}_X\boldsymbol{A}^t)\boldsymbol{\omega}\right) \\ &= \exp\left(i\,\boldsymbol{\omega}^t\boldsymbol{m}_Y - \frac{1}{2}\boldsymbol{\omega}^t\boldsymbol{K}_Y\boldsymbol{\omega}\right) \end{aligned}$$

Linear combinations

Theorem

The n-dimensional random vector $\boldsymbol{X} = (X_1, \dots, X_n)^t$ is Gaussian if and only if the 1-dimensional random variable

$$Y = a_1 X_1 + \cdots + a_n X_n = \boldsymbol{a}^t \boldsymbol{X}$$

is Gaussian for all $\boldsymbol{a} = (a_1, a_2, \dots a_n)^t \in \mathbb{R}^n$.

Linear combinations

Proof: If $\boldsymbol{X}$ is Gaussian so is Y.

Reciprocally, suppose that for all $\boldsymbol{\omega}^t = (\omega_1, \dots, \omega_n) \in \mathbb{R}^n$, the random variable $Y = \boldsymbol{\omega}^t \boldsymbol{X}$ is Gaussian.

Since

$$m_Y = \boldsymbol{\omega}^t \boldsymbol{m}_X, \qquad \sigma_Y^2 = \boldsymbol{\omega}^t \boldsymbol{K}_X \boldsymbol{\omega}$$

we have

$$\begin{aligned} M_X(\omega_1, \dots, \omega_n) &= \mathbb{E}\left(e^{i\boldsymbol{\omega}^t \boldsymbol{X}}\right) = \mathbb{E}\left(e^{iY}\right) \\ &= M_Y(1) = \exp\left(i m_Y - \frac{1}{2}\sigma_Y^2\right) \\ &= \exp\left(i\boldsymbol{\omega}^t \boldsymbol{m}_X - \frac{1}{2}\boldsymbol{\omega}^t \boldsymbol{K}_X \boldsymbol{\omega}\right) \end{aligned}$$

Linear combinations

Let $\boldsymbol{K}$ be a positive-semidefinite symmetric matrix and set:

$$\begin{aligned} \boldsymbol{C}\boldsymbol{K}\boldsymbol{C}^t &= \boldsymbol{D} = \operatorname{diag}(\lambda_1, \dots, \lambda_n), \quad \lambda_i \geqslant 0 \\ \boldsymbol{D}^{1/2} &= \operatorname{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n}) \\ \boldsymbol{K}^{1/2} &= \boldsymbol{C}^t \boldsymbol{D}^{1/2} \boldsymbol{C} \end{aligned}$$

Then $\boldsymbol{D}^{1/2}\boldsymbol{D}^{1/2} = \boldsymbol{D}$, the matrix $\boldsymbol{K}^{1/2}$ is symmetric, and

$$\begin{aligned} \boldsymbol{K}^{1/2}\boldsymbol{K}^{1/2} &= (\boldsymbol{C}^t \boldsymbol{D}^{1/2} \boldsymbol{C})(\boldsymbol{C}^t \boldsymbol{D}^{1/2} \boldsymbol{C}) \\ &= \boldsymbol{C}^t \boldsymbol{D}^{1/2} (\boldsymbol{C}\boldsymbol{C}^t) \boldsymbol{D}^{1/2} \boldsymbol{C} \\ &= \boldsymbol{C}^t (\boldsymbol{D}^{1/2}\boldsymbol{D}^{1/2}) \boldsymbol{C} = \boldsymbol{C}^t \boldsymbol{D} \boldsymbol{C} = \boldsymbol{K} \end{aligned}$$

Linear combinations

Proposition

Let $\boldsymbol{Z}$ be a random vector whose components $Z_1, \dots, Z_n$ are independent $\mathrm{N}(0,1)$-distributed random variables, let $\boldsymbol{K}$ be a $n \times n$ positive-semidefinite symmetric matrix, let $\boldsymbol{m}$ be a $n \times 1$ real constant vector, and set $\boldsymbol{X} = \boldsymbol{K}^{1/2}\boldsymbol{Z} + \boldsymbol{m}$.

Then $\boldsymbol{X} \sim \mathrm{N}(\boldsymbol{m}, \boldsymbol{K})$.

Proof: Notice that $\mathbb{E}(\boldsymbol{X}) = \boldsymbol{K}^{1/2}\,\mathbb{E}(\boldsymbol{Z}) + \boldsymbol{m} = \boldsymbol{m}$. It only remains to be proved that the covariance matrix of $\boldsymbol{X}$ is $\boldsymbol{K}$. Indeed,

$$\boldsymbol{K}_X = \boldsymbol{K}^{1/2}\, \boldsymbol{K}_Z\, (\boldsymbol{K}^{1/2})^t = \boldsymbol{K}^{1/2}\, \boldsymbol{I}\, (\boldsymbol{K}^{1/2})^t = \boldsymbol{K}^{1/2}\boldsymbol{K}^{1/2} = \boldsymbol{K}$$

The multivariate Gaussian density

Let $\boldsymbol{Z} \sim \mathrm{N}(0, \boldsymbol{I})$ and $\boldsymbol{X} = \boldsymbol{K}^{1/2}\boldsymbol{Z} + \boldsymbol{m}$, where $\boldsymbol{K}$ is an $n \times n$ positive-definite symmetric matrix (hence $\det(\boldsymbol{K}) > 0$).

We know that $\boldsymbol{X} \sim \mathrm{N}(\boldsymbol{m}, \boldsymbol{K})$. Let us find $f_X(x_1, x_2, \ldots, x_n)$.

We have

$$f_Z(z_1, z_2, \ldots, z_n) = f_{Z_1}(z_1) f_{Z_2}(z_2) \cdots f_{Z_n}(z_n)$$
$$= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}}\, e^{-\frac{1}{2} z_i^2} = \frac{1}{(2\pi)^{n/2}}\, e^{-\frac{1}{2}\boldsymbol{z}^t \boldsymbol{z}}$$

The multivariate Gaussian density

Let $\boldsymbol{K}^{-1/2} = (\boldsymbol{K}^{1/2})^{-1}$. The matrix $\boldsymbol{K}^{-1/2}$ is symmetric and, moreover,

$$\boldsymbol{K}^{-1/2}\boldsymbol{K}^{-1/2} = \boldsymbol{K}^{-1}$$

- The equation $\boldsymbol{x} = \boldsymbol{K}^{1/2}\boldsymbol{z} + \boldsymbol{m}$ has a unique solution

$$\boldsymbol{z} = \boldsymbol{K}^{-1/2}(\boldsymbol{x} - \boldsymbol{m})$$

- The jacobian of the transformation is

$$J(z_1, z_2, \ldots, z_n) = \det(\boldsymbol{K}^{1/2}) = \sqrt{\det(\boldsymbol{K})}$$

The multivariate Gaussian density

In this way, by the transformation theorem we have

$$f_X(x_1, x_2, \ldots, x_n) = \left. \frac{f_Z(z_1, z_2, \ldots, z_n)}{|J(z_1, z_2, \ldots, z_n)|} \right|_{\boldsymbol{z} = \boldsymbol{K}^{-1/2}(\boldsymbol{x} - \boldsymbol{m})}$$
$$= \frac{1}{(2\pi)^{n/2}\sqrt{\det(\boldsymbol{K})}} e^{-\frac{1}{2}\left(\boldsymbol{K}^{-1/2}(\boldsymbol{x}-\boldsymbol{m})\right)^t \left(\boldsymbol{K}^{-1/2}(\boldsymbol{x}-\boldsymbol{m})\right)}$$
$$= \frac{1}{(2\pi)^{n/2}\sqrt{\det(\boldsymbol{K})}} e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{m})^t \left(\boldsymbol{K}^{-1/2}\right)^t \boldsymbol{K}^{-1/2}(\boldsymbol{x}-\boldsymbol{m})}$$
$$= \frac{1}{(2\pi)^{n/2}\sqrt{\det(\boldsymbol{K})}} e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{m})^t \boldsymbol{K}^{-1}(\boldsymbol{x}-\boldsymbol{m})}$$

The multivariate Gaussian density

Proposition

Let $\boldsymbol{X} \sim \mathrm{N}(\boldsymbol{m}, \boldsymbol{K})$ where $\det(\boldsymbol{K}) > 0$. Then $\boldsymbol{X}$ has a density f_X such that

$$f_X(x_1, x_2, \ldots, x_n) = \frac{1}{(2\pi)^{n/2}\sqrt{\det(\boldsymbol{K})}} \cdot \exp\left(-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{m})^t \boldsymbol{K}^{-1}(\boldsymbol{x} - \boldsymbol{m})\right)$$

The multivariate Gaussian density

For instance, for $n = 2$ we obtain:

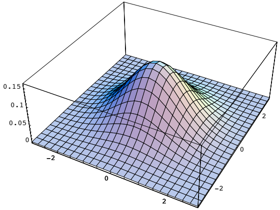
$$f_{XY}(x,y) = \frac{1}{2\pi\,\sqrt{1-\rho^2}\,\sigma_X\sigma_Y}\,\exp\left(-\frac{1}{2}\cdot\frac{1}{1-\rho^2}\cdot a(x,y)\right),$$

where

$$a(x,y) = \left(\frac{x-m_X}{\sigma_X}\right)^2 - 2\rho\,\frac{x-m_X}{\sigma_X}\cdot\frac{y-m_Y}{\sigma_Y} + \left(\frac{y-m_Y}{\sigma_Y}\right)^2$$

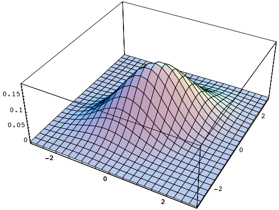
The multivariate Gaussian density

$\sigma_X = \sigma_Y \qquad \rho = 0$



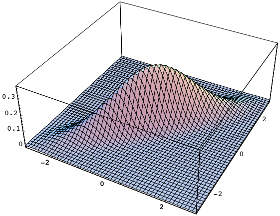
Multidimensional Gaussian density

$\sigma_X = \sigma_Y \qquad \rho = 0.5$

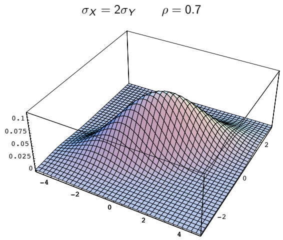


The multivariate Gaussian density

$\sigma_X = \sigma_Y \qquad \rho = 0.9$



The multivariate Gaussian density



Conditional densities

Let X, Y be jointly Gaussian. Then,

$$f_{Y|X}(y|X=x) = \frac{f_{XY}(x,y)}{f_X(x)}$$
$$= \frac{1}{\sqrt{2\pi}\sqrt{1-\rho^2}\,\sigma_Y} \exp\left(-\frac{1}{2}\left(\frac{y-m_{Y|X}}{\sigma_{Y|X}}\right)^2\right)$$

- $m_{Y|X}$ is the expected value of Y given X:

$$m_{Y|X} = \mathbb{E}(Y|X=x) = \rho\,\frac{\sigma_Y}{\sigma_X}(x-m_X) + m_Y$$

- $\sigma^2_{Y|X} = (1-\rho^2)\,\sigma^2_Y$.

Additional results

Theorem

Let $\boldsymbol{X} \sim \mathrm{N}(\boldsymbol{m}, \boldsymbol{K})$ with $\det(\boldsymbol{K}) > 0$. Then the random variable

$$(\boldsymbol{X}-\boldsymbol{m})^t \boldsymbol{K}^{-1}(\boldsymbol{X}-\boldsymbol{m})$$

follows a $\chi^2(n)$-distribution, where n is de dimension of $\boldsymbol{X}$.

For instance, for $n = 2$, the random variable

$$\frac{1}{1-\rho^2}\left(\left(\frac{X_1-m_1}{\sigma_1}\right)^2 - 2\rho\,\frac{X_1-m_1}{\sigma_1}\cdot\frac{X_2-m_2}{\sigma_2} + \left(\frac{X_2-m_2}{\sigma_2}\right)^2\right)$$

is $\chi^2(2)$-distributed.

Additional results

Proof: Consider $\boldsymbol{Y} = \boldsymbol{K}^{-1/2}(\boldsymbol{X}-\boldsymbol{m})$. Then,

$$\mathbb{E}(\boldsymbol{Y}) = \mathbb{E}\left(\boldsymbol{K}^{-1/2}(\boldsymbol{X}-\boldsymbol{m})\right) = \boldsymbol{K}^{-1/2}\mathbb{E}((\boldsymbol{X}-\boldsymbol{m})) = 0$$

$$\begin{aligned}\boldsymbol{K}_Y &= \mathbb{E}\left(\boldsymbol{Y}\boldsymbol{Y}^t\right) = \mathbb{E}\left(\boldsymbol{K}^{-1/2}(\boldsymbol{X}-\boldsymbol{m})(\boldsymbol{X}-\boldsymbol{m})^t(\boldsymbol{K}^{-1/2})^t\right)\\ &= \mathbb{E}\left(\boldsymbol{K}^{-1/2}(\boldsymbol{X}-\boldsymbol{m})(\boldsymbol{X}-\boldsymbol{m})^t\boldsymbol{K}^{-1/2}\right)\\ &= \boldsymbol{K}^{-1/2}\,\mathbb{E}\left((\boldsymbol{X}-\boldsymbol{m})(\boldsymbol{X}-\boldsymbol{m})^t\right)\boldsymbol{K}^{-1/2}\\ &= \boldsymbol{K}^{-1/2}\boldsymbol{K}\boldsymbol{K}^{-1/2} = (\boldsymbol{K}^{-1/2}\boldsymbol{K}^{1/2})(\boldsymbol{K}^{1/2}\boldsymbol{K}^{-1/2}) = \boldsymbol{I}\end{aligned}$$

Additional results

That is, $\mathbf{Y} = \mathbf{K}^{-1/2}(\mathbf{X} - \mathbf{m}) \sim \mathrm{N}(0, \mathbf{I})$. Therefore,

$$(\mathbf{X} - \mathbf{m})^t \mathbf{K}^{-1}(\mathbf{X} - \mathbf{m})$$
$$= (\mathbf{X} - \mathbf{m})^t \mathbf{K}^{-1/2}\mathbf{K}^{-1/2}(\mathbf{X} - \mathbf{m}) = \mathbf{Y}^t\mathbf{Y} = \sum_{i=1}^{n} Y_i^2$$

is a (1-dimensional) random variable following a $\chi^2(n)$-distribution, because the variables Y_i, $1 \leqslant i \leqslant n$, are independent and $\mathrm{N}(0, 1)$-distributed.

Additional results

Theorem

Let $\mathbf{X} \sim \mathrm{N}(\mathbf{m}, \mathbf{K})$ and set $\mathbf{Y} = \mathbf{C}\mathbf{X}$, where $\mathbf{C}$ is an ortogonal matrix such that $\mathbf{C}\mathbf{K}\mathbf{C}^t = \mathbf{D} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$.

Then, $\mathbf{Y} \sim \mathrm{N}(\mathbf{C}\mathbf{m}, \mathbf{D})$.

In particular, the components of $\mathbf{Y}$ are independent and

$$\mathrm{Var}(Y_k) = \lambda_k, \quad k = 1, \ldots, n$$

Remark: It may occur that some eigenvalue is equal to 0, in which case the corresponding component is degenerate.

Additional results

Theorem

Let $\mathbf{X} \sim \mathrm{N}(\mathbf{m}, \sigma^2\mathbf{I})$, where $\sigma^2 > 0$. Let $\mathbf{C}$ be an arbitrary ortogonal matrix, and set $\mathbf{Y} = \mathbf{C}\mathbf{X}$. Then, $\mathbf{Y} \sim \mathrm{N}(\mathbf{C}\mathbf{m}, \sigma^2\mathbf{I})$.

In particular, the components of $\mathbf{Y}$ are independent Gaussian random variables with the same variance σ^2.

Proof:

$$\mathbf{K}_Y = \mathbf{C}\mathbf{K}_X\mathbf{C}^t = \mathbf{C}(\sigma^2\mathbf{I})\mathbf{C}^t = \sigma^2\mathbf{C}\mathbf{C}^t = \sigma^2\mathbf{I}$$

Additional results

Theorem

Let $\mathbf{X} \sim \mathrm{N}(\mathbf{m}, \mathbf{K})$, and suppose that $\mathbf{K}$ can be partitioned (possibly after reordering the components) as follows:

$$\mathbf{K} = \begin{pmatrix} \mathbf{K}_1 & & & \\ & \mathbf{K}_2 & & 0 \\ & & \cdots & \\ & 0 & & \cdots & \\ & & & & \mathbf{K}_p \end{pmatrix}.$$

Then, $\mathbf{X}$ can be partitioned into vectors $\mathbf{X}^{(1)}, \mathbf{X}^{(2)}, \ldots, \mathbf{X}^{(p)}$, where $\mathbf{K}_i$ is the covariance matrix of $\mathbf{X}^{(i)}$, $i = 1, 2, \ldots, p$, and in such a way that the random vectors $\mathbf{X}^{(1)}, \mathbf{X}^{(2)}, \ldots, \mathbf{X}^{(p)}$ are independent.

Example

Let $\boldsymbol{X} = (X_1, X_2, X_3)^t$ be a Gaussian vector with $\boldsymbol{m} = (0,0,0)^t$ and

$$\boldsymbol{K} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 4 \\ 0 & 4 & 9 \end{pmatrix}$$

Then X_1 and (X_2, X_3) are independent.

Example

Indeed,

$$\begin{aligned} M_X(\omega_1,\omega_2,\omega_3) &= \exp\left(i\boldsymbol{\omega}^t\boldsymbol{m} - \frac{1}{2}\boldsymbol{\omega}^t\boldsymbol{K}\boldsymbol{\omega} \right) \\ &= \exp\left(-\frac{1}{2}(\omega_1,\omega_2,\omega_3) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 4 \\ 0 & 4 & 9 \end{pmatrix} \begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix} \right) \\ &= \exp\left(-\frac{1}{2}\omega_1^2 \right) \cdot \exp\left(-\frac{1}{2}(\omega_2,\omega_3) \begin{pmatrix} 2 & 4 \\ 4 & 9 \end{pmatrix} \begin{pmatrix} \omega_2 \\ \omega_3 \end{pmatrix} \right) \\ &= M_{X_1}(\omega_1) \cdot M_{X_2X_3}(\omega_2,\omega_3) \end{aligned}$$