Honours Multivariate Analysis

Lecture 3 - The Multivariate Normal Distribution

Stefan S. Britz

Department of Statistical Sciences
University of Cape Town

Course Outline

1. Introduction, Examples of Multivariate Data †
2. TOOLS
    1. Visualization and Summary Statistics †
    2. Singular Value Decomposition, Eigenvalue Decomposition and Spectral Decomposition revisited †
    3. The Multivariate Normal Distribution †
    4. Multivariate Maximum Likelihood Estimation †
    5. Multivariate Inference †
3. EXPLORATORY ANALYSIS
    1. Principal Component Analysis §
    2. Factor Analysis §
    3. Correspondence Analysis §
4. CONFIRMATORY ANALYSIS
    1. For grouped Multivariate Data:
        1. Manova †
        2. Discriminant Analysis §
    2. Regression
        1. Multivariate Regression §
        2. Canonical Correlation Analysis §

† Mr Stefan Britz

§ Mr Miguel Rodo

The Univariate Normal Distribution

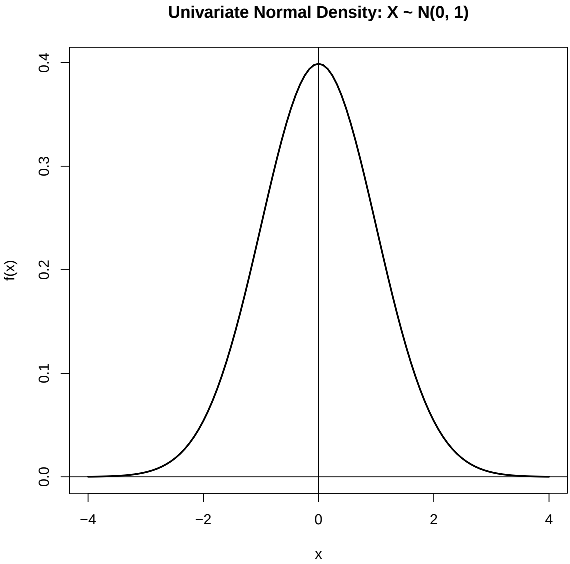
The univariate normal probability density function for a random variable $X \sim N(\mu, \sigma^2)$ is

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \ -\infty < x < \infty$$

where

- $\dfrac{1}{\sigma\sqrt{2\pi}}$ is a normalizing constant
- $\dfrac{(x-\mu)^2}{\sigma^2} = \underbrace{(x-\mu)\sigma^{-2}(x-\mu)}_{\text{square of distance from } x \text{ to } \mu \text{ in std.dev. units}}$

The Univariate Normal Distribution



The Multivariate Normal Distribution

Consider now $\boldsymbol{X}_{p\times 1}$, with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. The multivariate generalized distance is

$$(\boldsymbol{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})$$

and the multivariate version of the normalizing constant is

$$(2\pi)^{-\frac{p}{2}}|\boldsymbol{\Sigma}|^{-\frac{1}{2}}.$$

The Multivariate Normal Distribution

Consider now $\boldsymbol{X}_{p\times 1}$, with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. The multivariate generalized distance is

$$(\boldsymbol{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})$$

and the multivariate version of the normalizing constant is

$$(2\pi)^{-\frac{p}{2}}|\boldsymbol{\Sigma}|^{-\frac{1}{2}}.$$

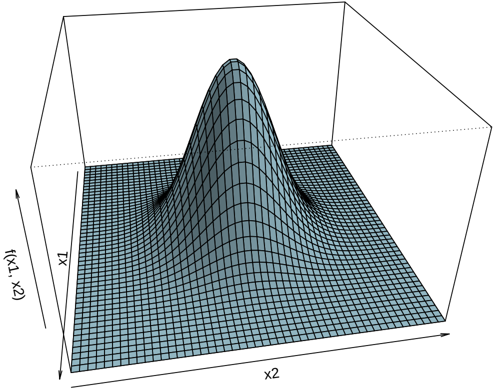
Thus the multivariate normal density function is

$$f(\boldsymbol{x}) = \frac{1}{(2\pi)^{\frac{p}{2}}|\boldsymbol{\Sigma}|^{\frac{1}{2}}} \exp\left[-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right], \ \forall \ \boldsymbol{x}$$

and we write $\boldsymbol{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

The Bivariate Normal Distribution

When $p = 2$, we have the bivariate normal distribution.



See `Lecture3.R` for more examples of plots, and this interactive illustration.

The Bivariate Normal Distribution

When $p = 2$,

- $\mu_1 = \mathrm{E}(X_1)$, $\mu_2 = \mathrm{E}(X_2)$,
- $\sigma_{11} = \mathrm{Var}(X_1)$, $\sigma_{22} = \mathrm{Var}(X_2)$, $\sigma_{12} = \sigma_{21} = \mathrm{Cov}(X_1, X_2)$
- $\rho_{12} = \dfrac{\sigma_{12}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{22}}} = \mathrm{Cor}(X_1, X_2)$

Using $\mathbf{\Sigma} = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{bmatrix}$, we can calculate

$$\mathbf{\Sigma}^{-1} = \frac{1}{\sigma_{11}\sigma_{22} - \sigma_{12}^2} \begin{bmatrix} \sigma_{22} & -\sigma_{12} \\ -\sigma_{12} & \sigma_{11} \end{bmatrix} = \frac{1}{\sigma_{11}\sigma_{22}(1 - \rho_{12}^2)} \begin{bmatrix} \sigma_{22} & -\sigma_{12} \\ -\sigma_{12} & \sigma_{11} \end{bmatrix},$$

since $\sigma_{12}^2 = \sigma_{11}\sigma_{22}\rho_{12}^2$. Therefore, we have

The Bivariate Normal Distribution

$$
\begin{aligned}
&(\boldsymbol{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu}) \\
&= \frac{1}{\sigma_{11}\sigma_{22}(1-\rho_{12}^2)} \times \\
&\begin{bmatrix} x_1-\mu_1 & x_2-\mu_2 \end{bmatrix} \begin{bmatrix} \sigma_{22} & -\rho_{12}\sqrt{\sigma_{11}}\sqrt{\sigma_{22}} \\ -\rho_{12}\sqrt{\sigma_{11}}\sqrt{\sigma_{22}} & \sigma_{11} \end{bmatrix} \begin{bmatrix} x_1-\mu_1 \\ x_2-\mu_2 \end{bmatrix} \\
&= \frac{\sigma_{22}(x_1-\mu_1)^2+\sigma_{11}(x_2-\mu_2)^2-2\rho_{12}\sqrt{\sigma_{11}}\sqrt{\sigma_{22}}(x_1-\mu_1)(x_2-\mu_2)}{\sigma_{11}\sigma_{22}(1-\rho_{12}^2)} \\
&= \frac{1}{1-\rho_{12}^2} \times \left[\left(\frac{x_1-\mu_1}{\sqrt{\sigma_{11}}}\right)^2+\left(\frac{x_2-\mu_2}{\sqrt{\sigma_{22}}}\right)^2-2\rho_{12}\left(\frac{x_1-\mu_1}{\sqrt{\sigma_{11}}}\right)\left(\frac{x_2-\mu_2}{\sqrt{\sigma_{22}}}\right)\right]
\end{aligned}
$$

The Bivariate Normal Distribution

We can write the normalizing constant as

$$\frac{1}{(2\pi)^{\frac{p}{2}}|\mathbf{\Sigma}|^{1/2}} = \frac{1}{2\pi\sqrt{|\mathbf{\Sigma}|}} = \frac{1}{2\pi\sqrt{\sigma_{11}\sigma_{22}(1-\rho_{12}^2)}}$$

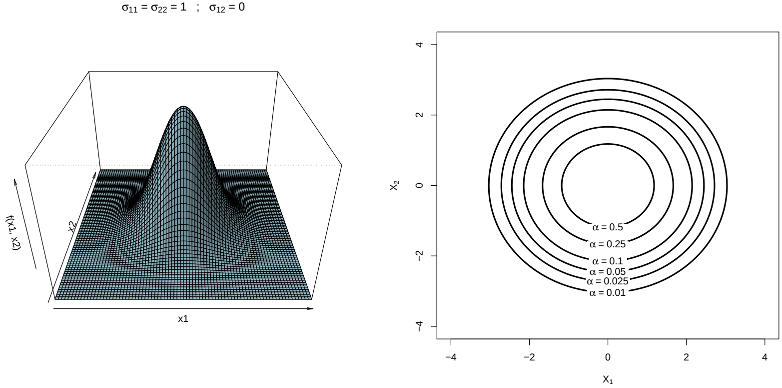
Thus

$$\begin{aligned} f(x_1, x_2) = & \frac{1}{2\pi\sqrt{\sigma_{11}\sigma_{22}(1-\rho_{12}^2)}} \times \\ & \boldsymbol{e}^{-\frac{1}{2(1-\rho_{12}^2)} \times \left[\left(\frac{x_1-\mu_1}{\sqrt{\sigma_{11}}}\right)^2 + \left(\frac{x_2-\mu_2}{\sqrt{\sigma_{22}}}\right)^2 - 2\rho_{12}\left(\frac{x_1-\mu_1}{\sqrt{\sigma_{11}}}\right)\left(\frac{x_2-\mu_2}{\sqrt{\sigma_{22}}}\right)\right]} \end{aligned}$$

If X_1 and X_2 are uncorrelated, $\rho_{12} = 0$ and $f(x_1, x_2) = f(x_1) \times f(x_2)$.

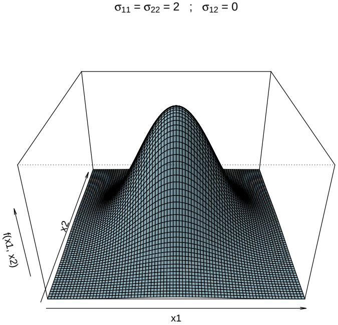
Bivariate Contours

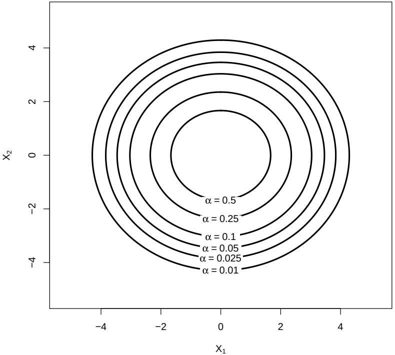
The shape of a bivariate distribution and the relationship between the variables can also be explored via a contour plot.



Bivariate Contours

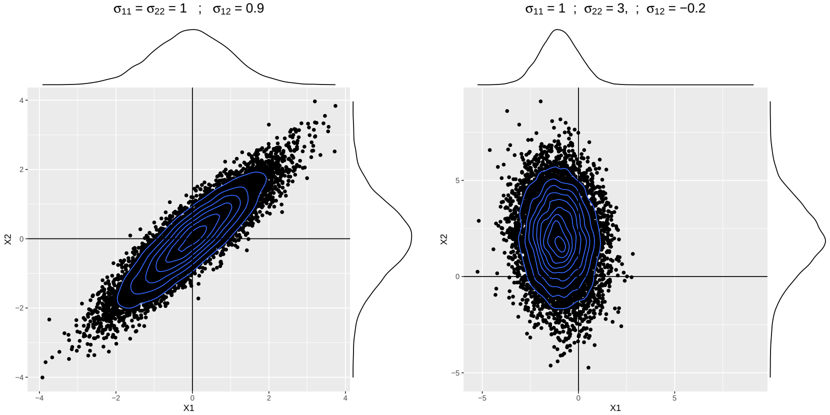
The shape of a bivariate distribution and the relationship between the variables can also be explored via a contour plot.





Bivariate Contours

We can also draw contours around data simulated from bivariate normal distributions.



Again, see `Lecture3.R`.

Homework exercise 3.1

Johnson & Wichern exercise 4.2

Consider a bivariate normal population with

$$\mu_1 = 0,\ \mu_2 = 2,\ \sigma_{11} = 2,\ \sigma_{22} = 1,\ \rho_{12} = 0.5$$

1. Write out the bivariate normal density.
2. 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.

Graphical interpretation of bivariate normal density

- The multivariate normal density is constant on surfaces where the square of the distance from $\boldsymbol{x}$ to the mean, $\boldsymbol{\mu}$ is constant.

Graphical interpretation of bivariate normal density

- The multivariate normal density is constant on surfaces where the square of the distance from $\boldsymbol{x}$ to the mean, $\boldsymbol{\mu}$ is constant.

- These constant probability contours are defined as all $\boldsymbol{x}$ such that $(\boldsymbol{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu}) = c^2$.

Graphical interpretation of bivariate normal density

- The multivariate normal density is constant on surfaces where the square of the distance from x to the mean, $\boldsymbol{\mu}$ is constant.

- These constant probability contours are defined as all x such that $(\boldsymbol{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu}) = c^2$.

- The axes of each ellipsoid of constant density are in the direction of the eigenvectors of $\boldsymbol{\Sigma}^{-1}$ and their lengths are proportional to the reciprocals of the square roots of the eigenvalues of $\boldsymbol{\Sigma}^{-1}$.

Graphical interpretation of bivariate normal density

- The multivariate normal density is constant on surfaces where the square of the distance from $\boldsymbol{x}$ to the mean, $\boldsymbol{\mu}$ is constant.

- These constant probability contours are defined as all $\boldsymbol{x}$ such that $(\boldsymbol{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu}) = c^2$.

- The axes of each ellipsoid of constant density are in the direction of the eigenvectors of $\boldsymbol{\Sigma}^{-1}$ and their lengths are proportional to the reciprocals of the square roots of the eigenvalues of $\boldsymbol{\Sigma}^{-1}$.

- Therefore, the axes are $\pm c\sqrt{\lambda_i}\boldsymbol{e}_i$, where the eigendecomposition of $\boldsymbol{\Sigma}$ is $\boldsymbol{\Sigma}\boldsymbol{e}_i = \lambda_i\boldsymbol{e}_i$ for $i = 1, 2, \ldots, p$.

Bivariate normal density example

Assume $\sigma_{11} = \sigma_{22}$. Thus the characteristic equation $|\boldsymbol{\Sigma} - \lambda \boldsymbol{I}| = 0$ becomes

$$
\begin{aligned}
0 &= \begin{vmatrix} \sigma_{11} - \lambda & \sigma_{12} \\ \sigma_{12} & \sigma_{11-\lambda} \end{vmatrix} \\
&= (\sigma_{11} - \lambda)^2 - \sigma_{12}^2 \\
&= (\lambda - \sigma_{11} - \sigma_{12})(\lambda - \sigma_{11} + \sigma_{12})
\end{aligned}
$$

$\therefore$ the eigenvalues are $\lambda_1 = \sigma_{11} + \sigma_{12}$ and $\lambda_2 = \sigma_{11} - \sigma_{12}$.

Bivariate normal density example

Assume $\sigma_{11} = \sigma_{22}$. Thus the characteristic equation $|\boldsymbol{\Sigma} - \lambda \boldsymbol{I}| = 0$ becomes

$$
\begin{aligned}
0 &= \begin{vmatrix} \sigma_{11} - \lambda & \sigma_{12} \\ \sigma_{12} & \sigma_{11-\lambda} \end{vmatrix} \\
&= (\sigma_{11} - \lambda)^2 - \sigma_{12}^2 \\
&= (\lambda - \sigma_{11} - \sigma_{12})(\lambda - \sigma_{11} + \sigma_{12})
\end{aligned}
$$

$\therefore$ the eigenvalues are $\lambda_1 = \sigma_{11} + \sigma_{12}$ and $\lambda_2 = \sigma_{11} - \sigma_{12}$.
The first eigenvector, $\boldsymbol{e}_1$, is then determined from

$$
\begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{11} \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \end{bmatrix} = (\sigma_{11} + \sigma_{12}) \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}
$$

which gives us

$$
\begin{aligned}
\sigma_{11} e_1 + \sigma_{12} e_2 &= (\sigma_{11} + \sigma_{12}) e_1 \\
\sigma_{12} e_1 + \sigma_{11} e_2 &= (\sigma_{11} + \sigma_{12}) e_2
\end{aligned}
$$

$\implies e_1 = e_2$.

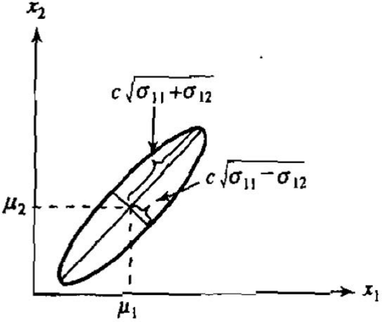
Bivariate normal density example

After normalization we have that the first eigenvalue-eigenvector pair is

$$\lambda_1 = \sigma_{11} + \sigma_{12}, \quad \boldsymbol{e}_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}'$$

Likewise, the second pair is

$$\lambda_2 = \sigma_{11} - \sigma_{12}, \quad \boldsymbol{e}_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}'$$



The Probability Content of the Ellipsoids of Constant Density

Let $\boldsymbol{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with $|\boldsymbol{\Sigma}| > 0$. Then

1. $(\boldsymbol{X} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{X} - \boldsymbol{\mu})$ is distributed as χ_p^2.

2. The $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ distribution assigns a probability of $1 - \alpha$ to the solid ellipsoid

$$\{\boldsymbol{x} : (\boldsymbol{x} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x} - \boldsymbol{\mu}) \leq \chi_p^2(\alpha)\},$$

   where $\chi_p^2(\alpha)$ denotes the upper (100α)th percentile of the χ_p^2 distribution.

   That is,

$$\Pr\left[(\boldsymbol{x} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x} - \boldsymbol{\mu}) \leq \chi_p^2(\alpha)\right] = 1 - \alpha.$$

The Probability Content of the Ellipsoids of Constant Density

Let $\boldsymbol{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with $|\boldsymbol{\Sigma}| > 0$. Then

1. $(\boldsymbol{X} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{X} - \boldsymbol{\mu})$ is distributed as χ^2_p.
2. The $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ distribution assigns a probability of $1 - \alpha$ to the solid ellipsoid

$$\{\boldsymbol{x} : (\boldsymbol{x} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x} - \boldsymbol{\mu}) \leq \chi^2_p(\alpha)\},$$

where $\chi^2_p(\alpha)$ denotes the upper (100α)th percentile of the χ^2_p distribution.

That is,

$$\Pr\left[(\boldsymbol{x} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x} - \boldsymbol{\mu}) \leq \chi^2_p(\alpha)\right] = 1 - \alpha.$$

Homework exercise 3.2

1. Prove the above by using the spectral decomposition of the covariance matrix.
2. Determine (and sketch) the constant-density contour that contains 90% of the probability for the examples in exercise 3.1.

Properties of the Multivariate Normal Distribution 1

Linear combinations of the components of $\boldsymbol{X}$ are normally distributed.

If $\boldsymbol{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ then ANY linear combination

$$\boldsymbol{a}'\boldsymbol{X} = a_1 X_1 + a_2 X_2 + \ldots + a_p X_p \sim N(\boldsymbol{a}'\boldsymbol{\mu}, \boldsymbol{a}'\boldsymbol{\Sigma}\boldsymbol{a})$$

and vice-versa, if $\boldsymbol{a}'\boldsymbol{X} \sim N(\boldsymbol{a}'\boldsymbol{\mu}, \boldsymbol{a}'\boldsymbol{\Sigma}\boldsymbol{a})$ for every $\boldsymbol{a}$, then $\boldsymbol{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

Also, $\boldsymbol{X} + \boldsymbol{d} \sim N_p(\boldsymbol{\mu} + \boldsymbol{d}, \boldsymbol{\Sigma})$ where $\boldsymbol{d}$ is a vector of constants.

Linear combinations of the components of $\boldsymbol{X}$ are normally distributed

Example:

Let $\boldsymbol{a}' = [1, 0, \dots, 0]$, then

$$a'X = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{bmatrix} = X_1$$

$$\boldsymbol{a}'\boldsymbol{\mu} = \mu_1$$

$$\boldsymbol{a}'\boldsymbol{\Sigma}\boldsymbol{a} = \sigma_{11}$$

$$\rightarrow \boldsymbol{a}'\boldsymbol{X} \sim N(\mu_1, \sigma_{11})$$

Homework exercise 3.3

Johnson & Wichern exercise 4.4(a)

Given $\boldsymbol{X} \sim N_3(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with $\boldsymbol{\mu}' = [2,\ -3,\ 1]$ and $\boldsymbol{\Sigma} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 2 & 2 \end{bmatrix}$,

find the distribution of $3X_1 - 2X_2 + X_3$.

Properties of the Multivariate Normal Distribution 2

If $\boldsymbol{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ then q linear combinations, then

$$\underbrace{\boldsymbol{A}}_{q \times p} \underbrace{\boldsymbol{X}}_{p \times 1} = \begin{bmatrix} a_{11}X_1 + \ldots + a_{1p}X_p \\ a_{21}X_1 + \ldots + a_{2p}X_p \\ \vdots \\ a_{q1}X_1 + \ldots + a_{qp}X_p \end{bmatrix} \sim N_q(\boldsymbol{A\mu}, \boldsymbol{A\Sigma A}')$$

q linear combinations of $\boldsymbol{X}$

Example:

If $\boldsymbol{X} \sim N_3(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and $\boldsymbol{A} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix}$, then

$$\boldsymbol{AX} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} = \begin{bmatrix} X_1 - X_2 \\ X_2 - X_3 \end{bmatrix}$$

$$\boldsymbol{A}\mu = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} = \begin{bmatrix} \mu_1 - \mu_2 \\ \mu_2 - \mu_3 \end{bmatrix}$$

$$\boldsymbol{A\Sigma A}' = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_{22} & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_{33} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -1 & 1 \\ 0 & -1 \end{bmatrix}$$

Properties of the Multivariate Normal Distribution 3

All subsets of the components of $\boldsymbol{X}$ have a (multivariate) normal distribution.
If we partition $\boldsymbol{X}$ as

$$\boldsymbol{X}_{p\times 1} = \left[\begin{array}{c} \underbrace{\boldsymbol{X}_1}_{q\times 1} \\ \hline \underbrace{\boldsymbol{X}_2}_{(p-q)\times 1} \end{array}\right]$$

then

$$\boldsymbol{\mu}_{p\times 1} = \left[\begin{array}{c} \underbrace{\boldsymbol{\mu}_1}_{q\times 1} \\ \hline \underbrace{\boldsymbol{\mu}_2}_{(p-q)\times 1} \end{array}\right] \quad \text{and} \quad \boldsymbol{\Sigma}_{p\times p} = \left[\begin{array}{c|c} \underbrace{\boldsymbol{\Sigma}_{11}}_{q\times q} & \underbrace{\boldsymbol{\Sigma}_{12}}_{q\times(p-q)} \\ \hline \underbrace{\boldsymbol{\Sigma}_{21}}_{(p-q)\times q} & \underbrace{\boldsymbol{\Sigma}_{22}}_{(p-q)\times(p-q)} \end{array}\right]$$

yielding $\boldsymbol{X}_1 \sim N_q(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11})$ and $\boldsymbol{X}_2 \sim N_{(p-q)}(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_{22})$.

Properties of the Multivariate Normal Distribution 3

All subsets of the components of $\boldsymbol{X}$ have a (multivariate) normal distribution. If we partition $\boldsymbol{X}$ as

$$\boldsymbol{X}_{p\times 1} = \left[\begin{array}{c} \underbrace{\boldsymbol{X}_1}_{q\times 1} \\ \hline \underbrace{\boldsymbol{X}_2}_{(p-q)\times 1} \end{array}\right]$$

then

$$\boldsymbol{\mu}_{p\times 1} = \left[\begin{array}{c} \underbrace{\boldsymbol{\mu}_1}_{q\times 1} \\ \hline \underbrace{\boldsymbol{\mu}_2}_{(p-q)\times 1} \end{array}\right] \quad \text{and} \quad \boldsymbol{\Sigma}_{p\times p} = \left[\begin{array}{c|c} \underbrace{\boldsymbol{\Sigma}_{11}}_{q\times q} & \underbrace{\boldsymbol{\Sigma}_{12}}_{q\times(p-q)} \\ \hline \underbrace{\boldsymbol{\Sigma}_{21}}_{(p-q)\times q} & \underbrace{\boldsymbol{\Sigma}_{22}}_{(p-q)\times(p-q)} \end{array}\right]$$

yielding $\boldsymbol{X}_1 \sim N_q(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11})$ and $\boldsymbol{X}_2 \sim N_{(p-q)}(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_{22})$.

The result follows from defining $\boldsymbol{A} = \left[\begin{array}{c|c} \underbrace{\boldsymbol{I}}_{q\times q} & \underbrace{\boldsymbol{0}}_{q\times(p-q)} \end{array}\right]$ and applying property 2.

Properties of the Multivariate Normal Distribution 4

Zero covariance implies that the corresponding components are independently distributed.

There are 3 results to consider:

Properties of the Multivariate Normal Distribution 4

Zero covariance implies that the corresponding components are independently distributed.

There are 3 results to consider:

1. If $\underbrace{\boldsymbol{X}_1}_{q_1 \times 1}$ and $\underbrace{\boldsymbol{X}_2}_{q_2 \times 1}$ are independent, then $\mathrm{Cov}(\boldsymbol{X}_1, \boldsymbol{X}_2) = \underbrace{\mathbf{0}}_{q_1 \times q_2}$.

Properties of the Multivariate Normal Distribution 4

Zero covariance implies that the corresponding components are independently distributed.

There are 3 results to consider:

1. If $\underbrace{\boldsymbol{X}_1}_{q_1 \times 1}$ and $\underbrace{\boldsymbol{X}_2}_{q_2 \times 1}$ are independent, then $\mathrm{Cov}(\boldsymbol{X}_1, \boldsymbol{X}_2) = \underbrace{\mathbf{0}}_{q_1 \times q_2}$.
2. If $\left[\begin{array}{c} \boldsymbol{X}_1 \\ \hline \boldsymbol{X}_2 \end{array}\right] \sim N_{q_1+q_2}\left(\left[\begin{array}{c} \boldsymbol{\mu}_1 \\ \hline \boldsymbol{\mu}_2 \end{array}\right], \left[\begin{array}{c|c} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \hline \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{array}\right]\right)$ then $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ are independent if and only if $\boldsymbol{\Sigma}_{12} = \mathbf{0}$.

Properties of the Multivariate Normal Distribution 4

Zero covariance implies that the corresponding components are independently distributed.

There are 3 results to consider:

1. If $\underbrace{\boldsymbol{X}_1}_{q_1 \times 1}$ and $\underbrace{\boldsymbol{X}_2}_{q_2 \times 1}$ are independent, then $\mathrm{Cov}(\boldsymbol{X}_1, \boldsymbol{X}_2) = \underbrace{\mathbf{0}}_{q_1 \times q_2}$.

2. If $\left[\begin{array}{c} \boldsymbol{X}_1 \\ \hline \boldsymbol{X}_2 \end{array}\right] \sim N_{q_1+q_2}\left(\left[\begin{array}{c} \boldsymbol{\mu}_1 \\ \hline \boldsymbol{\mu}_2 \end{array}\right], \left[\begin{array}{c|c} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \hline \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{array}\right]\right)$ then $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ are independent if and only if $\boldsymbol{\Sigma}_{12} = \mathbf{0}$.

3. If $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ are independent and are distributed as $N_{q_1}(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11})$ and $N_{q_2}(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_{22})$, respectively, then
$$\left[\begin{array}{c} \boldsymbol{X}_1 \\ \hline \boldsymbol{X}_2 \end{array}\right] \sim N_{q_1+q_2}\left(\left[\begin{array}{c} \boldsymbol{\mu}_1 \\ \hline \boldsymbol{\mu}_2 \end{array}\right], \left[\begin{array}{c|c} \boldsymbol{\Sigma}_{11} & \mathbf{0} \\ \hline \mathbf{0} & \boldsymbol{\Sigma}_{22} \end{array}\right]\right).$$

Zero covariance implies that the corresponding components are independently distributed

Example:

If we let $\underbrace{\boldsymbol{X}}_{3\times 1} \sim N_3(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with $\boldsymbol{\Sigma} = \begin{bmatrix} 4 & 1 & 0 \\ 1 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix}$, then X_1 and X_2 are NOT independent, since $\sigma_{12} = 1$.

BUT if we partition $\boldsymbol{X}$ as follows:

$$\boldsymbol{X} = \left[\begin{array}{c} X_1 \\ X_2 \\ \hline X_3 \end{array}\right], \; \boldsymbol{\Sigma} = \left[\begin{array}{cc|c} 4 & 1 & 0 \\ 1 & 3 & 0 \\ \hline 0 & 0 & 2 \end{array}\right]$$

then

$\boldsymbol{X}_1 = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$ and X_3 have covariance matrix $\boldsymbol{\Sigma}_{12} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ and so $\begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$ and X_3 are independent.

Homework exercise 3.4

Johnson & Wichern exercise 4.3

Let $\boldsymbol{X} \sim N_3(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with $\boldsymbol{\mu}' = \begin{bmatrix} -3 & 1 & 4 \end{bmatrix}$ and $\boldsymbol{\Sigma} = \begin{bmatrix} 1 & -2 & 0 \\ -2 & 5 & 0 \\ 0 & 0 & 2 \end{bmatrix}$

Which of the following pairs of random variables are independent? Explain.

1. X_1 and X_2
2. X_2 and X_3
3. (X_1, X_2) and X_3
4. $\dfrac{X_1 + X_2}{2}$ and X_3

Properties of the Multivariate Normal Distribution 5

The conditional distributions of the components are (multivariate) normally distributed

If $\boldsymbol{X} = \left[\begin{array}{c} \underbrace{\boldsymbol{X}_1}_{q\times 1} \\ \hline \underbrace{\boldsymbol{X}_2}_{(p-q)\times 1} \end{array}\right] \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with $\boldsymbol{\mu} = \left[\begin{array}{c} \boldsymbol{\mu}_1 \\ \hline \boldsymbol{\mu}_2 \end{array}\right]$, and $\boldsymbol{\Sigma} = \left[\begin{array}{c|c} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \hline \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{array}\right]$

then the conditional distribution of $\boldsymbol{X}_1 | \boldsymbol{X}_2 = \boldsymbol{x}_2$ is multivariate normal with mean

$$\mathrm{E}(\boldsymbol{X}_1 | \boldsymbol{X}_2) = \boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\boldsymbol{X}_2 - \boldsymbol{\mu}_2)$$

and covariance matrix

$$\mathrm{Cov}(\boldsymbol{X}_1 | \boldsymbol{X}_2) = \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}.$$

Properties of the Multivariate Normal Distribution 5

The conditional distributions of the components are (multivariate) normally distributed

If $\boldsymbol{X} = \left[\begin{array}{c} \underbrace{\boldsymbol{X}_1}_{q\times 1} \\ \hline \underbrace{\boldsymbol{X}_2}_{(p-q)\times 1} \end{array}\right] \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with $\boldsymbol{\mu} = \left[\begin{array}{c} \boldsymbol{\mu}_1 \\ \hline \boldsymbol{\mu}_2 \end{array}\right]$, and $\boldsymbol{\Sigma} = \left[\begin{array}{c|c} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \hline \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{array}\right]$

then the conditional distribution of $\boldsymbol{X}_1|\boldsymbol{X}_2 = \boldsymbol{x}_2$ is multivariate normal with mean

$$\mathrm{E}(\boldsymbol{X}_1|\boldsymbol{X}_2) = \boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\boldsymbol{X}_2 - \boldsymbol{\mu}_2)$$

and covariance matrix

$$\mathrm{Cov}(\boldsymbol{X}_1|\boldsymbol{X}_2) = \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}.$$

Make sure that you go through the proof for this provided in the notes!

Homework exercise 3.5

Johnson & Wichern 4.5(a)

Consider again the bivariate normal population with
$\mu_1 = 0, \mu_2 = 2, \sigma_{11} = 2, \sigma_{22} = 1, \rho_{12} = 0.5.$

Specify the conditional distribution of $X_1 | X_2 = x_2$.

Summary of property 5

1. All conditional distributions are multivariate normal.

2. The conditional mean is of the form

$$\begin{array}{c} \mu_1 + \beta_{1,q+1}(x_{q+1} - \mu_{q+1}) + \ldots + \beta_{1,p}(x_p - \mu_p) \\ \vdots \\ \mu_q + \beta_{q,q+1}(x_{q+1} - \mu_{q+1}) + \ldots + \beta_{q,p}(x_p - \mu_p) \end{array},$$

   where the $\beta' s$ are defined by

$$\mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1} = \begin{bmatrix} \beta_{1,q+1} & \beta_{1,q+2} & \cdots & \beta_{1,p} \\ \beta_{2,q+1} & \beta_{2,q+2} & \cdots & \beta_{2,p} \\ & & \vdots & \\ \beta_{q,q+1} & \beta_{q,q+2} & \cdots & \beta_{q,p} \end{bmatrix}$$

3. The conditional covariance $\mathbf{\Sigma}_{11} - \mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}\mathbf{\Sigma}_{21}$ does not depend upon the values of the conditioning variables.

Linear combinations of random variable vectors

Consider $\underbrace{\boldsymbol{V}_1}_{p\times 1} = c_1\boldsymbol{X}_1 + c_2\boldsymbol{X}_2 + \ldots + c_n\boldsymbol{X}_n = \underbrace{\begin{bmatrix}\boldsymbol{X}_1 & \boldsymbol{X}_2 & \ldots & \boldsymbol{X}_n\end{bmatrix}}_{(p\times n)}\underbrace{\boldsymbol{c}}_{n\times 1}$

where the $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_n$ are mutually independent with each $\boldsymbol{X}_j \sim N_p(\boldsymbol{\mu}_j, \boldsymbol{\Sigma})$.

Then $\boldsymbol{V}_1 \sim N_p\left(\sum_{j=1}^{n} c_j\boldsymbol{\mu}_j, \left(\sum_{j=1}^{n} c_j^2\right)\boldsymbol{\Sigma}\right)$.

Linear combinations of random variable vectors

Consider $\underbrace{\boldsymbol{V}_1}_{p\times 1} = c_1\boldsymbol{X}_1 + c_2\boldsymbol{X}_2 + \ldots + c_n\boldsymbol{X}_n = \underbrace{\begin{bmatrix}\boldsymbol{X}_1 & \boldsymbol{X}_2 & \ldots & \boldsymbol{X}_n\end{bmatrix}}_{(p\times n)}\underbrace{\boldsymbol{c}}_{n\times 1}$

where the $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_n$ are mutually independent with each $\boldsymbol{X}_j \sim N_p(\boldsymbol{\mu}_j, \boldsymbol{\Sigma})$.

Then $\boldsymbol{V}_1$ and $\boldsymbol{V}_2 = b_1\boldsymbol{X}_1 + b_2\boldsymbol{X}_2 + \ldots + b_n\boldsymbol{X}_n$ are jointly multivariate normal with covariance matrix

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

So $\boldsymbol{V}_1$ and $\boldsymbol{V}_2$ are independent if $\boldsymbol{b}'\boldsymbol{c} = \sum_{j=1}^{n} c_j b_j = 0$.

Homework exercise 3.6

Johnson & Wichern exercises 4.16 and 4.17
Let $\boldsymbol{X}1$, $\boldsymbol{X}_2$, $\boldsymbol{X}_3$ and $\boldsymbol{X}_4$ be independent $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ random vectors.

1. Find the marginal distributions of

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

and

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

2. Find the joint density of $\boldsymbol{V}_1$ and $\boldsymbol{V}_2$.

Homework exercise 3.6

Let $\boldsymbol{X}_1$, $\boldsymbol{X}_2$, $\boldsymbol{X}_3$, $\boldsymbol{X}_4$ and $\boldsymbol{X}_5$ be independent and identically distributed random vectors with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$.

3. Find the mean vector and covariance matrices for

$$\frac{1}{5}\boldsymbol{X}_1 + \frac{1}{5}\boldsymbol{X}_2 + \frac{1}{5}\boldsymbol{X}_3 + \frac{1}{5}\boldsymbol{X}_4 + \frac{1}{5}\boldsymbol{X}_5$$

and

$$\boldsymbol{X}_1 - \boldsymbol{X}_2 + \boldsymbol{X}_3 - \boldsymbol{X}_4 + \boldsymbol{X}_5$$

in terms of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$. Also, obtain the covariance between these two linear combinations.