Stochastic Processes and Financial Mathematics (37363)

Chapter 2

Multivariate Gaussian random variables

Alex Novikov and Scott Alexander

School of Mathematical and Physical Sciences, UTS

Autumn 2025

Chapter outline

Topics:

- Constructing Gaussian vector RVs
- Affine-linear transform of Gaussian vector RVs
- Theorem on normal correlation (TNC)
  - vector RVs
  - scalar RVs

Constructing Gaussian vector RVs

We begin by providing two (equivalent) definitions of joint Gaussian random vectors.

The first definition is based on scalar Gaussian RVs introduced in Chapter 1.

Definition 1 (Gaussian random vector 1)

We say that the random vector $\boldsymbol{X} = (X_1, \ldots, X_n)^T$ has a joint Gaussian (normal) distribution if it has the representation

$$\boldsymbol{X} = \boldsymbol{m} + \boldsymbol{L}\boldsymbol{\xi}$$

where $\boldsymbol{m} \in \mathbb{R}^n$ is a non-random vector, $\boldsymbol{L} \in \mathbb{R}^{n\times k}$ is a non-random matrix, $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_k)^T$, $\xi_i \sim N(0,1)$ and $\xi_1, \ldots, \xi_k$ are independent.

Terminology. The vector $\boldsymbol{X} = \boldsymbol{m} + \boldsymbol{L}\boldsymbol{\xi}$ is called an affine-linear transformation of the RV $\boldsymbol{\xi}$.

Constructing Gaussian vector RVs

Notation.
The joint Gaussian vector RV $\boldsymbol{X}$ has expectation

$$E[\boldsymbol{X}] = E[\boldsymbol{m} + \boldsymbol{L}\boldsymbol{\xi}] = \boldsymbol{m} + \boldsymbol{L}E[\boldsymbol{\xi}] = \boldsymbol{m}$$

and autocovariance

$$\begin{aligned}\boldsymbol{Q} := \operatorname{cov}(\boldsymbol{X}, \boldsymbol{X}) &= E[(\boldsymbol{X} - \boldsymbol{m})(\boldsymbol{X} - \boldsymbol{m})^T] = E[\boldsymbol{L}\boldsymbol{\xi}(\boldsymbol{L}\boldsymbol{\xi})^T] \\ &= \boldsymbol{L}E[\boldsymbol{\xi}\boldsymbol{\xi}^T]\boldsymbol{L}^T = \boldsymbol{L}\boldsymbol{I}_k\boldsymbol{L}^T = \boldsymbol{L}\boldsymbol{L}^T\end{aligned}$$

and we write $\boldsymbol{X} \sim N(\boldsymbol{m}, \boldsymbol{Q})$.

Caution.
In most applications of this definition we will have $k \geq n$.

However, there may be cases where $k < n$.

In this case the $n \times n$ autocovariance matrix $\boldsymbol{Q}$ will have rank k and will not be invertible.

Constructing Gaussian vector RVs

The second definition is based on characteristic and moment generating functions also introduced in Chapter 1.

Definition 2 (Gaussian random vector 2)

We say that the random vector $\boldsymbol{X} = (X_1, \ldots, X_n)^T$ has a joint Gaussian (normal) distribution if for all $\boldsymbol{u} \in \mathbb{R}^n$ the CF of $\boldsymbol{X}$ has the form

$$\varphi_{\boldsymbol{X}}(\boldsymbol{u}) := E[e^{i\boldsymbol{u}\cdot\boldsymbol{X}}] = e^{i\boldsymbol{m}\cdot\boldsymbol{u} - \frac{1}{2}\boldsymbol{u}\cdot\boldsymbol{Q}\boldsymbol{u}}$$

or equivalently if the MGF of $\boldsymbol{X}$ has the form

$$M_{\boldsymbol{X}}(\boldsymbol{u}) := E[e^{\boldsymbol{u}\cdot\boldsymbol{X}}] = e^{\boldsymbol{m}\cdot\boldsymbol{u} + \frac{1}{2}\boldsymbol{u}\cdot\boldsymbol{Q}\boldsymbol{u}}$$

where $\boldsymbol{m} = E[\boldsymbol{X}] \in \mathbb{R}^n$ and $\boldsymbol{Q} = \mathrm{cov}(\boldsymbol{X}, \boldsymbol{X}) \in \mathbb{R}^{n\times n}$ and $\boldsymbol{z} \cdot \boldsymbol{Q}\boldsymbol{z} \geq 0$ for all $\boldsymbol{z} \in \mathbb{R}^n$.

Notation.
$\boldsymbol{X} \sim N(\boldsymbol{m}, \boldsymbol{Q})$.

Constructing Gaussian vector RVs

The following example illustrates how both definitions can be used to show that a vector of independent, standard normal RVs is joint Gaussian.

Example.
Let $\boldsymbol{\xi} = (\xi_1, \dots, \xi_n)^T$, $\xi_i \sim N(0,1)$ and $\xi_1, \dots, \xi_n$ are independent. We have

$$E[\xi_i] = 0 \text{ for all } i,$$

$$\text{var}(\xi_i) = \text{cov}(\xi_i, \xi_i) = 1 \text{ for all } i$$

and by independence

$$\text{cov}(\xi_i, \xi_j) = 0 \text{ for all } i \neq j$$

so

$$\boldsymbol{m} = \mathbf{0} \quad \text{and} \quad \boldsymbol{Q} = \boldsymbol{I}_n$$

where $\mathbf{0}$ is the zero vector in $\mathbb{R}^n$ and $\boldsymbol{I_n}$ is the identity matrix in $\mathbb{R}^{n \times n}$.

So by Definition 1, $\boldsymbol{\xi} \sim N(\mathbf{0}, \boldsymbol{I}_n)$. To see this take $\boldsymbol{m} = \mathbf{0}$ and $\boldsymbol{L} = \boldsymbol{I}_n$ in this definition.

Constructing Gaussian vector RVs

We can also obtain this directly via

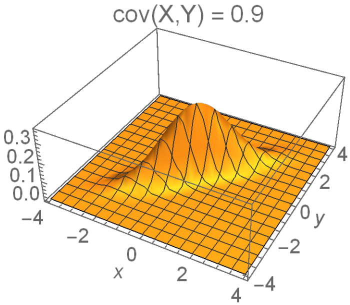
$$
\begin{aligned}
\varphi_{\boldsymbol{\xi}}(\boldsymbol{u}) =& E[e^{i\boldsymbol{u}\cdot\boldsymbol{\xi}}] = E\left[e^{\sum_{k=1}^{n} iu_k\xi_k}\right] = E\left[\prod_{k=1}^{n} e^{iu_k\xi_k}\right] \\
=& \prod_{k=1}^{n} E[e^{iu_k\xi_k}] \quad \text{(by independence)} \\
=& \prod_{k=1}^{n} e^{-\frac{1}{2}u_k^2} \quad \text{(CF of } N(0,1) \text{ RV -- proof as exercise)} \\
=& e^{-\frac{1}{2}\sum_{k=1}^{n} u_k^2} = e^{-\frac{1}{2}\boldsymbol{u}\cdot\boldsymbol{u}} = e^{-\frac{1}{2}|\boldsymbol{u}|^2}. \qquad (1)
\end{aligned}
$$

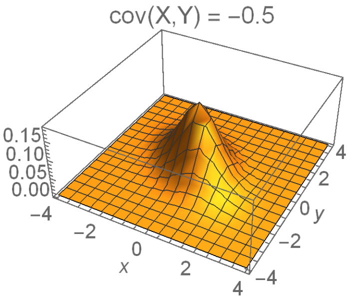
So by Definition 2, $\boldsymbol{\xi} \sim N(\mathbf{0}, \boldsymbol{I}_n)$. To see this take $\boldsymbol{m} = \mathbf{0}$ and $\boldsymbol{Q} = \boldsymbol{I}_n$ in this definition.

Constructing Gaussian vector RVs

Below is Mathematica code and plots of two 2D Gaussian RV PDFs.

```
ρ = 0.9;
g1 = Plot3D[PDF[MultinormalDistribution[{0, 0}, {{1, ρ}, {ρ, 1}}], {x, y}], {x, -4, 4},
    {y, -4, 4}, PlotRange → All, AxesLabel → {x, y}, PlotLabel → "cov(X,Y) = 0.9"];
ρ = -0.5;
g2 = Plot3D[PDF[MultinormalDistribution[{0, 0}, {{1, ρ}, {ρ, 1}}], {x, y}], {x, -4, 4},
    {y, -4, 4}, PlotRange → All, AxesLabel → {x, y}, PlotLabel → "cov(X,Y) = -0.5"];
GraphicsGrid[{{g1, g2}}]
```





Constructing Gaussian vector RVs

Comments. Note that according to these definitions any deterministic vector $\boldsymbol{x}$ has a Gaussian distribution $\boldsymbol{x} \sim N(\boldsymbol{x}, \mathbf{0})$.

More generally, if $\boldsymbol{Q} > 0$ (that is $\boldsymbol{z} \cdot \boldsymbol{Q}\boldsymbol{z} > 0$ for any $\boldsymbol{z} \neq \mathbf{0}$) then $\det(\boldsymbol{Q}) > 0$.

This in turn means the inverse matrix $\boldsymbol{Q}^{-1}$ exists allowing the PDF for $\boldsymbol{X} \sim N(\boldsymbol{m}, \boldsymbol{Q})$ to be defined as

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{1}{(2\pi)^{n/2}\sqrt{\det(\boldsymbol{Q})}} e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{m})\cdot \boldsymbol{Q}^{-1}(\boldsymbol{x}-\boldsymbol{m})}.$$

Definition 1 is a kind of direct probabilistic approach based on use of a fundamental RV ($\xi_i \sim N(0,1)$ and independent). It is used in simulations.

Definition 2 is a kind of analytical approach based on analytical tools (the CF and MGF).

Constructing Gaussian vector RVs

The following theorem shows the equivalence of the two previous definitions.

Theorem 1 (constructing multivariate Gaussian RVs)

Definitions 1 and 2 are equivalent with $\boldsymbol{m} = E[\boldsymbol{X}]$ and $\boldsymbol{Q} = \text{cov}(\boldsymbol{X}, \boldsymbol{X})$.

Proof.

Part 1. Showing Definition 1 implies Definition 2.

We have $\boldsymbol{\xi} = (\xi_1, ..., \xi_n)^T \sim N(\boldsymbol{0}, \boldsymbol{I}_n)$ so

$$
\begin{aligned}
E[e^{i\boldsymbol{u}\cdot\boldsymbol{X}}] &= E[e^{i\boldsymbol{u}\cdot(\boldsymbol{m}+\boldsymbol{L}\boldsymbol{\xi})}] = e^{i\boldsymbol{m}\cdot\boldsymbol{u}}E[e^{i\boldsymbol{u}\cdot\boldsymbol{L}\boldsymbol{\xi}}] \quad \text{(non-random terms)} \\
&= e^{i\boldsymbol{m}\cdot\boldsymbol{u}}E[e^{i\boldsymbol{L}^T\boldsymbol{u}\cdot\boldsymbol{\xi}}] = e^{i\boldsymbol{m}\cdot\boldsymbol{u}}e^{-\frac{1}{2}(\boldsymbol{L}^T\boldsymbol{u})\cdot\boldsymbol{L}^T\boldsymbol{u}} \\
&\quad \text{(replace } \boldsymbol{u} \text{ with } \boldsymbol{L}^T\boldsymbol{u} \text{ in equation (1) to see)} \\
&= e^{i\boldsymbol{m}\cdot\boldsymbol{u}}e^{-\frac{1}{2}(\boldsymbol{L}^T\boldsymbol{u})^T\boldsymbol{L}^T\boldsymbol{u}} = e^{i\boldsymbol{m}\cdot\boldsymbol{u}-\frac{1}{2}\boldsymbol{u}^T\boldsymbol{L}\boldsymbol{L}^T\boldsymbol{u}} = e^{i\boldsymbol{m}\cdot\boldsymbol{u}-\frac{1}{2}\boldsymbol{u}\cdot(\boldsymbol{L}\boldsymbol{L}^T)\boldsymbol{u}}
\end{aligned}
$$

and we can conclude that Definition 2 holds with $\boldsymbol{Q} = \boldsymbol{L}\boldsymbol{L}^T$.

Constructing Gaussian vector RVs

Part 2. Showing Definition 2 implies Definition 1.

We have $\boldsymbol{X} \sim N(\boldsymbol{m}, \boldsymbol{Q})$ and assume that $\boldsymbol{z} \cdot \boldsymbol{Q}\boldsymbol{z} > 0$ for all $\boldsymbol{z} \neq \boldsymbol{0}$ (shorthand $\boldsymbol{Q} > 0$) so that $\det(\boldsymbol{Q}) > 0$.

We need to find $\boldsymbol{L}$ such that $\boldsymbol{X} = \boldsymbol{m} + \boldsymbol{L}\boldsymbol{\xi}$ where $\boldsymbol{\xi} \sim N(\boldsymbol{0}, \boldsymbol{I}_n)$.

There are many ways to find such an $\boldsymbol{L}$, such as the Cholesky decomposition

$$\boldsymbol{Q} = \boldsymbol{H}^T \boldsymbol{H},$$

which exists because $\boldsymbol{Q} > 0$. Here $\boldsymbol{H}$ is an upper triangular matrix and $\boldsymbol{H}^T$ is a lower triangular matrix.

Constructing Gaussian vector RVs

Also because $\boldsymbol{Q} > 0$ the inverse $\boldsymbol{Q}^{-1}$ exists and

$$\begin{aligned}\det(\boldsymbol{Q}) &= \det(\boldsymbol{H}^T\boldsymbol{H}) = \det(\boldsymbol{H}^T)\det(\boldsymbol{H}) = (\det(\boldsymbol{H}))^2 \\ &\qquad \text{(det. of triangular matrix is product of diag. components)} \\ &> 0\end{aligned}$$

so $\det(\boldsymbol{H}^T) \neq 0$ showing $(\boldsymbol{H}^T)^{-1}$ also exists.

Taking $\boldsymbol{L} = \boldsymbol{H}^T$ we have

$$\boldsymbol{X} = \boldsymbol{m} + \boldsymbol{H}^T\boldsymbol{\xi}$$

or

$$\boldsymbol{\xi} = (\boldsymbol{H}^T)^{-1}(\boldsymbol{X} - \boldsymbol{m}) = (\boldsymbol{H}^{-1})^T(\boldsymbol{X} - \boldsymbol{m})$$

by the transpose and inverse properties of real matrices.

Constructing Gaussian vector RVs

Therefore

$$\begin{aligned}
E[e^{i\boldsymbol{u}\cdot\boldsymbol{\xi}}] &= E[e^{i\boldsymbol{u}\cdot(\boldsymbol{H}^{-1})^T(\boldsymbol{X}-\boldsymbol{m})}] = E[e^{i(\boldsymbol{H}^{-1}\boldsymbol{u})\cdot(\boldsymbol{X}-\boldsymbol{m})}] \\
&= e^{-i(\boldsymbol{H}^{-1}\boldsymbol{u})\cdot\boldsymbol{m}}E[e^{i(\boldsymbol{H}^{-1}\boldsymbol{u})\cdot\boldsymbol{X}}] \quad \text{(non-random terms)} \\
&= e^{-i(\boldsymbol{H}^{-1}\boldsymbol{u})\cdot\boldsymbol{m}}e^{i\boldsymbol{m}\cdot\boldsymbol{H}^{-1}\boldsymbol{u}-\frac{1}{2}(\boldsymbol{H}^{-1}\boldsymbol{u})\cdot\boldsymbol{Q}\boldsymbol{H}^{-1}\boldsymbol{u}} \quad \text{(by Definition 2)} \\
&= e^{-\frac{1}{2}\boldsymbol{u}^T(\boldsymbol{H}^{-1})^T\boldsymbol{Q}\boldsymbol{H}^{-1}\boldsymbol{u}} = e^{-\frac{1}{2}\boldsymbol{u}^T(\boldsymbol{H}^T)^{-1}\boldsymbol{Q}\boldsymbol{H}^{-1}\boldsymbol{u}} = e^{-\frac{1}{2}\boldsymbol{u}^T\boldsymbol{I}_n\boldsymbol{u}} \\
&= e^{-\frac{1}{2}\boldsymbol{u}\cdot\boldsymbol{u}}
\end{aligned}$$

as

$$\boldsymbol{Q} = \boldsymbol{H}^T\boldsymbol{H} \;\Rightarrow\; (\boldsymbol{H}^T)^{-1}\boldsymbol{Q}\boldsymbol{H}^{-1} = (\boldsymbol{H}^T)^{-1}\boldsymbol{H}^T\boldsymbol{H}\boldsymbol{H}^{-1} = \boldsymbol{I}_n.$$

So we see $\boldsymbol{\xi} \sim N(\boldsymbol{0}, \boldsymbol{I}_n)$, or $\xi_i \sim N(0,1)$ for all i with

$$\text{cov}(\xi_i, \xi_j) = 0 \text{ for all } i \neq j$$

implying their independence.

Constructing Gaussian vector RVs

Below is some Mathematica code and output used to obtain and verify Cholesky decomposition.

```
QQ = {{1, 2}, {2, 5}};
HH = CholeskyDecomposition[QQ];
HH // MatrixForm
```

$$\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$$

```
Transpose[HH].HH // MatrixForm
```

$$\begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}$$

Constructing Gaussian vector RVs

Exercise.
Let $\boldsymbol{\xi} \sim N(\boldsymbol{m}, \boldsymbol{Q})$ where

$$\boldsymbol{m} = \begin{pmatrix} m_1 \\ m_2 \end{pmatrix}, \quad \boldsymbol{Q} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}, \quad -1 \leq \rho \leq 1.$$

Find $E[e^{i\boldsymbol{u}\cdot\boldsymbol{\xi}}]$ and show that $\rho = 0$ implies ξ_1 and ξ_2 are independent.

Solution. By Definition 2

$$\begin{aligned} E[e^{i\boldsymbol{u}\cdot\boldsymbol{\xi}}] &= e^{i\boldsymbol{m}\cdot\boldsymbol{u} - \frac{1}{2}\boldsymbol{u}\cdot\boldsymbol{Q}\boldsymbol{u}} \\ &= e^{i(m_1u_1+m_2u_2)-\frac{1}{2}(\sigma_1^2u_1^2+2\rho\sigma_1\sigma_2u_1u_2+\sigma_2^2u_2^2)} \\ &= e^{i(m_1u_1+m_2u_2)-\frac{1}{2}(\sigma_1^2u_1^2+\sigma_2^2u_2^2)} \quad \text{(taking } \rho = 0) \\ &= e^{im_1u_1-\frac{1}{2}\sigma_1^2u_1^2}e^{im_2u_2-\frac{1}{2}\sigma_2^2u_2^2} = E[e^{iu_1\xi_1}]E[e^{iu_2\xi_2}] \end{aligned}$$

showing ξ_1 and ξ_2 are independent.

Affine-linear transform of Gaussian vector RVs

The next theorem is a major result that show a linear affine transformation applied to a joint Gaussian (normal) random vector produces another joint Gaussian random vector.

Theorem 2 (affine-linear transform of Gaussian RVs)

Let $\boldsymbol{X} = (X_1, \ldots, X_k)^T \sim N(\boldsymbol{m}, \boldsymbol{Q})$ and let $\boldsymbol{Y} = (Y_1, \ldots, Y_n)^T$ be an affine-linear transformation of $\boldsymbol{X}$, i.e.

$$\boldsymbol{Y} = \boldsymbol{b} + \boldsymbol{A}\boldsymbol{X}$$

where $\boldsymbol{b} \in \mathbb{R}^n$ and $\boldsymbol{A} \in \mathbb{R}^{n \times k}$ are deterministic.

Then $\boldsymbol{Y} \sim N(\boldsymbol{A}\boldsymbol{m} + \boldsymbol{b}, \boldsymbol{A}\boldsymbol{Q}\boldsymbol{A}^T)$.

Affine-linear transform of Gaussian vector RVs

Proof.
Method 1. The CF of $\boldsymbol{Y}$ is given by

$$
\begin{aligned}
E[e^{i\boldsymbol{u}\cdot\boldsymbol{Y}}] &= E[e^{i\boldsymbol{u}\cdot(\boldsymbol{b}+\boldsymbol{A}\boldsymbol{X})}] = e^{i\boldsymbol{b}\cdot\boldsymbol{u}}E[e^{i\boldsymbol{u}\cdot\boldsymbol{A}\boldsymbol{X}}] \quad \text{(non-random terms)} \\
&= e^{i\boldsymbol{b}\cdot\boldsymbol{u}}E[e^{i(\boldsymbol{A}^T\boldsymbol{u})\cdot\boldsymbol{X}}] \\
&= e^{i\boldsymbol{b}\cdot\boldsymbol{u}}e^{i\boldsymbol{m}\cdot(\boldsymbol{A}^T\boldsymbol{u})-\frac{1}{2}(\boldsymbol{A}^T\boldsymbol{u})\cdot\boldsymbol{Q}\boldsymbol{A}^T\boldsymbol{u}} \quad \text{(replace } \boldsymbol{u} \text{ with } \boldsymbol{A}^T\boldsymbol{u} \text{ in Def. 2)} \\
&= e^{i\boldsymbol{b}\cdot\boldsymbol{u}}e^{i(\boldsymbol{A}\boldsymbol{m})\cdot\boldsymbol{u}-\frac{1}{2}\boldsymbol{u}\cdot\boldsymbol{A}\boldsymbol{Q}\boldsymbol{A}^T\boldsymbol{u}} = e^{i(\boldsymbol{b}+\boldsymbol{A}\boldsymbol{m})\cdot\boldsymbol{u}-\frac{1}{2}\boldsymbol{u}\cdot\boldsymbol{A}\boldsymbol{Q}\boldsymbol{A}^T\boldsymbol{u}}
\end{aligned}
$$

which by Definition 2 gives $\boldsymbol{Y} \sim N(\boldsymbol{b}+\boldsymbol{A}\boldsymbol{m}, \boldsymbol{A}\boldsymbol{Q}\boldsymbol{A}^T)$.

Affine-linear transform of Gaussian vector RVs

Method 2. Note that for suitable deterministic $\boldsymbol{L} \in \mathbb{R}^{k \times k}$ we can re-write $\boldsymbol{Y}$ as

$$\boldsymbol{Y} = \boldsymbol{b} + \boldsymbol{A}\boldsymbol{X} = \boldsymbol{b} + \boldsymbol{A}(\boldsymbol{m} + \boldsymbol{L}\boldsymbol{\xi}) = \boldsymbol{b} + \boldsymbol{A}\boldsymbol{m} + \boldsymbol{A}\boldsymbol{L}\boldsymbol{\xi}$$

where $\boldsymbol{\xi} = (\xi_1, \dots, \xi_k)^T \sim N(\boldsymbol{0}, \boldsymbol{I})$, which by Definition 1 shows that $\boldsymbol{Y}$ is Gaussian (replace $\boldsymbol{m}$ with $\boldsymbol{b} + \boldsymbol{A}\boldsymbol{m}$ and $\boldsymbol{L}$ with $\boldsymbol{A}\boldsymbol{L}$ to see).

The expectation of $\boldsymbol{Y}$ is

$$\begin{aligned} E[\boldsymbol{Y}] &= E[\boldsymbol{b} + \boldsymbol{A}\boldsymbol{m} + \boldsymbol{A}\boldsymbol{L}\boldsymbol{\xi}] \\ &= \boldsymbol{b} + \boldsymbol{A}\boldsymbol{m} + \boldsymbol{A}\boldsymbol{L}E[\boldsymbol{\xi}] \quad \text{(non-random terms)} \\ &= \boldsymbol{b} + \boldsymbol{A}\boldsymbol{m}. \end{aligned}$$

Affine-linear transform of Gaussian vector RVs

The autocovariance of $\boldsymbol{Y}$ is

$$\begin{aligned}
\operatorname{cov}(\boldsymbol{Y}, \boldsymbol{Y}) &= E\left[(\boldsymbol{Y} - E[\boldsymbol{Y}])(\boldsymbol{Y} - E[\boldsymbol{Y}])^T\right] \\
&= E\left[\boldsymbol{AL}\xi(\boldsymbol{AL}\xi)^T\right] = E\left[\boldsymbol{AL}\xi\xi^T\boldsymbol{L}^T\boldsymbol{A}^T\right] \\
&= \boldsymbol{AL}E\left[\xi\xi^T\right]\boldsymbol{L}^T\boldsymbol{A}^T \quad \text{(non-random terms)} \\
&= \boldsymbol{AL}E\left[(\xi - E[\xi])(\xi - E[\xi])^T\right]\boldsymbol{L}^T\boldsymbol{A}^T \quad (E[\xi] = \mathbf{0}) \\
&= \boldsymbol{AL}\operatorname{cov}(\xi, \xi)\boldsymbol{L}^T\boldsymbol{A}^T = \boldsymbol{ALI}_n\boldsymbol{L}^T\boldsymbol{A}^T = \boldsymbol{ALL}^T\boldsymbol{A}^T \\
&= \boldsymbol{AQA}^T \quad \text{(see proof (Part 1) of Theorem1)}
\end{aligned}$$

which gives $\boldsymbol{Y} \sim N(\boldsymbol{b} + \boldsymbol{Am}, \boldsymbol{AQA}^T)$.

Theorem on normal correlation (TNC) – vector RVs

Theorem 3 (theorem on normal correlation – vector version)

Let random vectors $\boldsymbol{\theta} = (\theta_1, ..., \theta_m)^T$ and $\boldsymbol{\xi} = (\xi_1, .., \xi_n)^T$ have a joint Gaussian distribution. Then the following properties hold.

1. If $\mathrm{cov}(\boldsymbol{\theta}, \boldsymbol{\xi}) = \mathbf{0}$ then $\boldsymbol{\theta}$ and $\boldsymbol{\xi}$ are independent.

2. If $\boldsymbol{u} \cdot \mathrm{cov}(\boldsymbol{\xi}, \boldsymbol{\xi})\boldsymbol{u} > 0$ for any $\boldsymbol{u} \neq \mathbf{0}$ then the RV

$$\boldsymbol{\eta} := \boldsymbol{\theta} - E[\boldsymbol{\theta}] - \mathrm{cov}(\boldsymbol{\theta}, \boldsymbol{\xi})\,\mathrm{cov}(\boldsymbol{\xi}, \boldsymbol{\xi})^{-1}(\boldsymbol{\xi} - E[\boldsymbol{\xi}])$$

and $\boldsymbol{\xi}$ are independent.

3. If $\boldsymbol{u} \cdot \mathrm{cov}(\boldsymbol{\xi}, \boldsymbol{\xi})\boldsymbol{u} > 0$ for any $\boldsymbol{u} \neq \mathbf{0}$ then

$$E[\boldsymbol{\theta}|\boldsymbol{\xi}] = E[\boldsymbol{\theta}] + \mathrm{cov}(\boldsymbol{\theta}, \boldsymbol{\xi})\,\mathrm{cov}(\boldsymbol{\xi}, \boldsymbol{\xi})^{-1}(\boldsymbol{\xi} - E[\boldsymbol{\xi}]).$$

Theorem on normal correlation (TNC) – vector RVs

Theorem 3 (cont.)

4. If $\boldsymbol{u} \cdot \mathrm{cov}(\boldsymbol{\xi}, \boldsymbol{\xi})\boldsymbol{u} > 0$ for any $\boldsymbol{u} \neq \mathbf{0}$, then the conditional autocovariance

$$\begin{aligned}
\mathrm{cov}(\boldsymbol{\theta}, \boldsymbol{\theta}|\boldsymbol{\xi}) &= E\big[(\boldsymbol{\theta} - E[\boldsymbol{\theta}|\boldsymbol{\xi}])(\boldsymbol{\theta} - E[\boldsymbol{\theta}|\boldsymbol{\xi}])^T|\boldsymbol{\xi}\big] \\
&= E\big[(\boldsymbol{\theta} - E[\boldsymbol{\theta}|\boldsymbol{\xi}])(\boldsymbol{\theta} - E[\boldsymbol{\theta}|\boldsymbol{\xi}])^T\big] \\
&= \mathrm{cov}(\boldsymbol{\theta}, \boldsymbol{\theta}) - \mathrm{cov}(\boldsymbol{\theta}, \boldsymbol{\xi})\,\mathrm{cov}(\boldsymbol{\xi}, \boldsymbol{\xi})^{-1}\,\mathrm{cov}(\boldsymbol{\theta}, \boldsymbol{\xi})^T.
\end{aligned}$$

Note to Part 1.
We know that for any RVs, independence implies zero covariance.

TNC Part 1 works the other way, stating that for joint Gaussian RVs, zero covariance implies independence.

This is convenient, as verifying zero covariance is easier than verifying independence directly.

Theorem on normal correlation (TNC) – vector RVs

Note to Part 2 and 3.
If we combine these results we have

$$\boldsymbol{\eta} := \boldsymbol{\theta} - E[\boldsymbol{\theta}|\boldsymbol{\xi}] \quad \text{and} \quad \boldsymbol{\xi}$$

are independent.

In the supporting notes to this chapter, we show that ordinary least squares regression is a particular application of Part 3.

In this context we can interpret this result as saying that the noise terms $\boldsymbol{\eta}$ and the predictors (covariates) $\boldsymbol{\xi}$ are independent.

However, in this subject the main use we have for this result is in the proof of Part 4, with the second line as stated following immediately.

Theorem on normal correlation (TNC) – vector RVs

Note to Part 3 and 4.
Although we don't prove this, the conditional distribution

$$\boldsymbol{\theta}|\{\boldsymbol{\xi} = \boldsymbol{x}\} \sim N(\boldsymbol{\mu}(\boldsymbol{x}), \boldsymbol{\sigma}^2(\boldsymbol{x}))$$

where

$$\boldsymbol{\mu}(\boldsymbol{x}) := E[\boldsymbol{\theta}|\boldsymbol{\xi} = \boldsymbol{x}]$$

and

$$\boldsymbol{\sigma}^2(\boldsymbol{x}) := \mathrm{cov}(\boldsymbol{\theta}, \boldsymbol{\theta}|\boldsymbol{\xi} = \boldsymbol{x}).$$

are given by Parts 3 and 4 respectively.

Theorem on normal correlation (TNC) – vector RVs

Proof.
Part 1. If we can show the joint CF of $\boldsymbol{\theta}, \boldsymbol{\xi}$ is the product of the marginal CFs, i.e.

$$E[e^{i(\boldsymbol{z}\cdot\boldsymbol{\theta}+\boldsymbol{u}\cdot\boldsymbol{\xi})}] = E[e^{i\boldsymbol{z}\cdot\boldsymbol{\theta}}]E[e^{i\boldsymbol{u}\cdot\boldsymbol{\xi}}],$$

then by the one-to-one relationship between CFs and distributions we will have shown that $\boldsymbol{\theta}$ and $\boldsymbol{\xi}$ are independent.

Proceeding, first note that because $\boldsymbol{\theta}, \boldsymbol{\xi}$ are joint Gaussian then for $\boldsymbol{z} \in \mathbb{R}^m$ and $\boldsymbol{u} \in \mathbb{R}^n$

$$\boldsymbol{z}\cdot\boldsymbol{\theta} + \boldsymbol{u}\cdot\boldsymbol{\xi},$$

are also are joint Gaussian by Theorem 2.

So by Definition 2

$$E[e^{i(\boldsymbol{z}\cdot\boldsymbol{\theta}+\boldsymbol{u}\cdot\boldsymbol{\xi})}] = e^{iE[\boldsymbol{z}\cdot\boldsymbol{\theta}+\boldsymbol{u}\cdot\boldsymbol{\xi}]-\frac{1}{2}\operatorname{var}(\boldsymbol{z}\cdot\boldsymbol{\theta}+\boldsymbol{u}\cdot\boldsymbol{\xi})}.$$

Theorem on normal correlation (TNC) – vector RVs

Next note that

$$\operatorname{var}(\boldsymbol{z}\cdot\boldsymbol{\theta}+\boldsymbol{u}\cdot\boldsymbol{\xi})=\operatorname{var}(\boldsymbol{z}\cdot\boldsymbol{\theta})+\operatorname{var}(\boldsymbol{u}\cdot\boldsymbol{\xi})+2\operatorname{cov}(\boldsymbol{z}\cdot\boldsymbol{\theta},\boldsymbol{u}\cdot\boldsymbol{\xi})$$

where

$$\begin{aligned}\operatorname{cov}(\boldsymbol{z}\cdot\boldsymbol{\theta},\boldsymbol{u}\cdot\boldsymbol{\xi})&=\boldsymbol{z}\cdot\operatorname{cov}(\boldsymbol{\theta},\boldsymbol{\xi})\boldsymbol{u}\quad\text{(check this yourselves)}\\&=\boldsymbol{z}\cdot\mathbf{0}\boldsymbol{u}=0.\end{aligned}$$

Therefore

$$\begin{aligned}E[e^{i(\boldsymbol{z}\cdot\boldsymbol{\theta}+\boldsymbol{u}\cdot\boldsymbol{\xi})}]&=e^{iE[\boldsymbol{z}\cdot\boldsymbol{\theta}+\boldsymbol{u}\cdot\boldsymbol{\xi}]-\frac{1}{2}\operatorname{var}(\boldsymbol{z}\cdot\boldsymbol{\theta}+\boldsymbol{u}\cdot\boldsymbol{\xi})}\\&=e^{iE[\boldsymbol{z}\cdot\boldsymbol{\theta}+\boldsymbol{u}\cdot\boldsymbol{\xi}]-\frac{1}{2}(\operatorname{var}(\boldsymbol{z}\cdot\boldsymbol{\theta})+\operatorname{var}(\boldsymbol{u}\cdot\boldsymbol{\xi}))}\\&=e^{iE[\boldsymbol{z}\cdot\boldsymbol{\theta}]-\frac{1}{2}\operatorname{var}(\boldsymbol{z}\cdot\boldsymbol{\theta})}e^{iE[\boldsymbol{u}\cdot\boldsymbol{\xi}]-\frac{1}{2}\operatorname{var}(\boldsymbol{u}\cdot\boldsymbol{\xi})}\\&=E[e^{i\boldsymbol{z}\cdot\boldsymbol{\theta}}]E[e^{i\boldsymbol{u}\cdot\boldsymbol{\xi}}]\end{aligned}$$

and we have the desired result.

Theorem on normal correlation (TNC) – vector RVs

Part 2. Observe that

$$
\begin{aligned}
\operatorname{cov}(\boldsymbol{\eta}, \boldsymbol{\xi}) &= E\big[(\boldsymbol{\eta} - E[\boldsymbol{\eta}])(\boldsymbol{\xi} - E[\boldsymbol{\xi}])^T\big] = E\big[\boldsymbol{\eta}(\boldsymbol{\xi} - E[\boldsymbol{\xi}])^T\big] \\
&= E\big[\big(\boldsymbol{\theta} - E[\boldsymbol{\theta}] - \operatorname{cov}(\boldsymbol{\theta}, \boldsymbol{\xi}) \operatorname{cov}(\boldsymbol{\xi}, \boldsymbol{\xi})^{-1}(\boldsymbol{\xi} - E[\boldsymbol{\xi}])\big)(\boldsymbol{\xi} - E[\boldsymbol{\xi}])^T\big] \\
&= E\big[(\boldsymbol{\theta} - E[\boldsymbol{\theta}])(\boldsymbol{\xi} - E[\boldsymbol{\xi}])^T\big] \\
&\quad - \operatorname{cov}(\boldsymbol{\theta}, \boldsymbol{\xi}) \operatorname{cov}(\boldsymbol{\xi}, \boldsymbol{\xi})^{-1} E\big[(\boldsymbol{\xi} - E[\boldsymbol{\xi}])(\boldsymbol{\xi} - E[\boldsymbol{\xi}])^T\big] \\
&\quad \text{(linearity and non-random terms)} \\
&= \operatorname{cov}(\boldsymbol{\theta}, \boldsymbol{\xi}) - \operatorname{cov}(\boldsymbol{\theta}, \boldsymbol{\xi}) \operatorname{cov}(\boldsymbol{\xi}, \boldsymbol{\xi})^{-1} \operatorname{cov}(\boldsymbol{\xi}, \boldsymbol{\xi}) \\
&= \operatorname{cov}(\boldsymbol{\theta}, \boldsymbol{\xi}) - \operatorname{cov}(\boldsymbol{\theta}, \boldsymbol{\xi}) = \mathbf{0}
\end{aligned}
$$

which by Part 1 TNC shows that $\boldsymbol{\eta}$ and $\boldsymbol{\xi}$ are independent (zero covariance implies independence for joint Gaussian RVs).

Theorem on normal correlation (TNC) – vector RVs

Part 3. The conditional expectation

$$\begin{aligned}
E[\boldsymbol{\eta}|\boldsymbol{\xi}] &= E\left[\boldsymbol{\theta} - E[\boldsymbol{\theta}] - \operatorname{cov}(\boldsymbol{\theta},\boldsymbol{\xi})\operatorname{cov}(\boldsymbol{\xi},\boldsymbol{\xi})^{-1}(\boldsymbol{\xi} - E[\boldsymbol{\xi}])|\boldsymbol{\xi}\right] \\
&= E[\boldsymbol{\theta}|\boldsymbol{\xi}] - E[\boldsymbol{\theta}] - \operatorname{cov}(\boldsymbol{\theta},\boldsymbol{\xi})\operatorname{cov}(\boldsymbol{\xi},\boldsymbol{\xi})^{-1}(E[\boldsymbol{\xi}|\boldsymbol{\xi}] - E[\boldsymbol{\xi}]) \\
&\quad \text{(linearity and non-random terms)} \\
&= E[\boldsymbol{\theta}|\boldsymbol{\xi}] - E[\boldsymbol{\theta}] - \operatorname{cov}(\boldsymbol{\theta},\boldsymbol{\xi})\operatorname{cov}(\boldsymbol{\xi},\boldsymbol{\xi})^{-1}(\boldsymbol{\xi} - E[\boldsymbol{\xi}]) \\
&\quad \text{(see properties of conditional expectation Chapter 1).}
\end{aligned}$$

But from Part 2 of this proof we know that $\boldsymbol{\eta}$ and $\boldsymbol{\xi}$ are independent so

$$E[\boldsymbol{\eta}|\boldsymbol{\xi}] = E[\boldsymbol{\eta}] = \mathbf{0}$$

giving

$$E[\boldsymbol{\theta}|\boldsymbol{\xi}] = E[\boldsymbol{\theta}] + \operatorname{cov}(\boldsymbol{\theta},\boldsymbol{\xi})\operatorname{cov}(\boldsymbol{\xi},\boldsymbol{\xi})^{-1}(\boldsymbol{\xi} - E[\boldsymbol{\xi}]).$$

Theorem on normal correlation (TNC) – vector RVs

Part 4. First note that

$$\begin{aligned}\boldsymbol{\theta} - E[\boldsymbol{\theta}|\boldsymbol{\xi}] &= \boldsymbol{\theta} - E[\boldsymbol{\theta}] - \mathrm{cov}(\boldsymbol{\theta}, \boldsymbol{\xi})\,\mathrm{cov}(\boldsymbol{\xi}, \boldsymbol{\xi})^{-1}(\boldsymbol{\xi} - E[\boldsymbol{\xi}]) \\ &= \boldsymbol{\eta}\end{aligned}$$

and $\boldsymbol{\xi}$ are independent by Part 2 of this theorem which gives

$$E\left[(\boldsymbol{\theta} - E[\boldsymbol{\theta}|\boldsymbol{\xi}])(\boldsymbol{\theta} - E[\boldsymbol{\theta}|\boldsymbol{\xi}])^T|\boldsymbol{\xi}\right] = E\left[(\boldsymbol{\theta} - E[\boldsymbol{\theta}|\boldsymbol{\xi}])(\boldsymbol{\theta} - E[\boldsymbol{\theta}|\boldsymbol{\xi}])^T\right]$$

by properties of conditional expectation Chapter 1.

Theorem on normal correlation (TNC) – vector RVs

Then

$$
\begin{aligned}
& E\left[(\boldsymbol{\theta} - E[\boldsymbol{\theta}|\boldsymbol{\xi}])(\boldsymbol{\theta} - E[\boldsymbol{\theta}|\boldsymbol{\xi}])^T\right] \\
&\quad = E\big[(\boldsymbol{\theta} - E[\boldsymbol{\theta}] - \operatorname{cov}(\boldsymbol{\theta},\boldsymbol{\xi})\operatorname{cov}(\boldsymbol{\xi},\boldsymbol{\xi})^{-1}(\boldsymbol{\xi} - E[\boldsymbol{\xi}])) \\
&\qquad\quad (\boldsymbol{\theta} - E[\boldsymbol{\theta}] - \operatorname{cov}(\boldsymbol{\theta},\boldsymbol{\xi})\operatorname{cov}(\boldsymbol{\xi},\boldsymbol{\xi})^{-1}(\boldsymbol{\xi} - E[\boldsymbol{\xi}]))^T\big] \\
&\quad = E\left[(\boldsymbol{\theta} - E[\boldsymbol{\theta}])(\boldsymbol{\theta} - E[\boldsymbol{\theta}])^T\right] \\
&\qquad - E\left[(\boldsymbol{\theta} - E[\boldsymbol{\theta}])(\operatorname{cov}(\boldsymbol{\theta},\boldsymbol{\xi})\operatorname{cov}(\boldsymbol{\xi},\boldsymbol{\xi})^{-1}(\boldsymbol{\xi} - E[\boldsymbol{\xi}]))^T\right] \\
&\qquad - E\left[\operatorname{cov}(\boldsymbol{\theta},\boldsymbol{\xi})\operatorname{cov}(\boldsymbol{\xi},\boldsymbol{\xi})^{-1}(\boldsymbol{\xi} - E[\boldsymbol{\xi}])(\boldsymbol{\theta} - E[\boldsymbol{\theta}])^T\right] \\
&\qquad + E\big[\operatorname{cov}(\boldsymbol{\theta},\boldsymbol{\xi})\operatorname{cov}(\boldsymbol{\xi},\boldsymbol{\xi})^{-1}(\boldsymbol{\xi} - E[\boldsymbol{\xi}]) \\
&\qquad\qquad (\operatorname{cov}(\boldsymbol{\theta},\boldsymbol{\xi})\operatorname{cov}(\boldsymbol{\xi},\boldsymbol{\xi})^{-1}(\boldsymbol{\xi} - E[\boldsymbol{\xi}]))^T\big] \\
&\quad = \operatorname{cov}(\boldsymbol{\theta},\boldsymbol{\theta}) \\
&\qquad - E\left[(\boldsymbol{\theta} - E[\boldsymbol{\theta}])(\boldsymbol{\xi} - E[\boldsymbol{\xi}])^T\right]\operatorname{cov}(\boldsymbol{\xi},\boldsymbol{\xi})^{-1}\operatorname{cov}(\boldsymbol{\theta},\boldsymbol{\xi})^T \\
&\qquad - E\left[(\boldsymbol{\theta} - E[\boldsymbol{\theta}])(\boldsymbol{\xi} - E[\boldsymbol{\xi}])^T\right]\operatorname{cov}(\boldsymbol{\xi},\boldsymbol{\xi})^{-1}\operatorname{cov}(\boldsymbol{\theta},\boldsymbol{\xi})^T \\
&\qquad + \operatorname{cov}(\boldsymbol{\theta},\boldsymbol{\xi})\operatorname{cov}(\boldsymbol{\xi},\boldsymbol{\xi})^{-1}E\left[(\boldsymbol{\xi} - E[\boldsymbol{\xi}])(\boldsymbol{\xi} - E[\boldsymbol{\xi}])^T\right] \\
&\qquad\qquad \operatorname{cov}(\boldsymbol{\xi},\boldsymbol{\xi})^{-1}\operatorname{cov}(\boldsymbol{\theta},\boldsymbol{\xi})^T
\end{aligned}
$$

Theorem on normal correlation (TNC) – vector RVs

$$
\begin{aligned}
&= \operatorname{cov}(\boldsymbol{\theta}, \boldsymbol{\theta}) \\
&\quad - \operatorname{cov}(\boldsymbol{\theta}, \boldsymbol{\xi}) \operatorname{cov}(\boldsymbol{\xi}, \boldsymbol{\xi})^{-1} \operatorname{cov}(\boldsymbol{\theta}, \boldsymbol{\xi})^T \\
&\quad - \operatorname{cov}(\boldsymbol{\theta}, \boldsymbol{\xi}) \operatorname{cov}(\boldsymbol{\xi}, \boldsymbol{\xi})^{-1} \operatorname{cov}(\boldsymbol{\theta}, \boldsymbol{\xi})^T \\
&\quad + \operatorname{cov}(\boldsymbol{\theta}, \boldsymbol{\xi}) \operatorname{cov}(\boldsymbol{\xi}, \boldsymbol{\xi})^{-1} \operatorname{cov}(\boldsymbol{\xi}, \boldsymbol{\xi}) \operatorname{cov}(\boldsymbol{\xi}, \boldsymbol{\xi})^{-1} \operatorname{cov}(\boldsymbol{\theta}, \boldsymbol{\xi})^T \\
&= \operatorname{cov}(\boldsymbol{\theta}, \boldsymbol{\theta}) \\
&\quad - \operatorname{cov}(\boldsymbol{\theta}, \boldsymbol{\xi}) \operatorname{cov}(\boldsymbol{\xi}, \boldsymbol{\xi})^{-1} \operatorname{cov}(\boldsymbol{\theta}, \boldsymbol{\xi})^T \\
&\quad - \operatorname{cov}(\boldsymbol{\theta}, \boldsymbol{\xi}) \operatorname{cov}(\boldsymbol{\xi}, \boldsymbol{\xi})^{-1} \operatorname{cov}(\boldsymbol{\theta}, \boldsymbol{\xi})^T \\
&\quad + \operatorname{cov}(\boldsymbol{\theta}, \boldsymbol{\xi}) \operatorname{cov}(\boldsymbol{\xi}, \boldsymbol{\xi})^{-1} \operatorname{cov}(\boldsymbol{\theta}, \boldsymbol{\xi})^T \\
&= \operatorname{cov}(\boldsymbol{\theta}, \boldsymbol{\theta}) - \operatorname{cov}(\boldsymbol{\theta}, \boldsymbol{\xi}) \operatorname{cov}(\boldsymbol{\xi}, \boldsymbol{\xi})^{-1} \operatorname{cov}(\boldsymbol{\theta}, \boldsymbol{\xi})^T
\end{aligned}
$$

using linearity of expectation, removing non-random terms from expectation, definition of covariance and standard properties from linear algebra.

Theorem on normal correlation (TNC) – scalar RVs

Theorem 4 (theorem on normal correlation – scalar version)

Let $(\theta, \xi)^T$ have a joint Gaussian distribution. Then the following properties hold.

1. If $\text{cov}(\theta, \xi) = 0$ then θ and ξ are independent.

2. If $\text{var}(\xi) > 0$ then the RV

$$\eta := \theta - E[\theta] - \frac{\text{cov}(\theta, \xi)}{\text{var}(\xi)}(\xi - E[\xi])$$

and ξ are independent.

3. If $\text{var}(\xi) > 0$ then

$$E[\theta|\xi] = E[\theta] + \frac{\text{cov}(\theta, \xi)}{\text{var}(\xi)}(\xi - E[\xi]).$$

Theorem on normal correlation (TNC) – scalar RVs

Theorem 4 (cont.)

4. If $\text{var}(\xi) > 0$ then the autocovariance

$$\begin{aligned}\text{cov}(\theta, \theta|\xi) &= E\left[(\theta - E[\theta|\xi])^2|\xi\right] \\ &= E\left[(\theta - E[\theta|\xi])^2\right] \\ &= \text{var}(\theta) - \frac{\text{cov}(\theta, \xi)^2}{\text{var}(\xi)}.\end{aligned}$$

Theorem on normal correlation (TNC) – scalar RVs

Example.
Let

$$\begin{pmatrix} \theta \\ \xi \end{pmatrix} = N \sim \left(\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix} \right),$$

i.e. $\theta \sim N(1, 5)$, $\xi \sim N(-1, 1)$, $\text{cov}(\theta, \xi) = 2$ and (θ, ξ) are joint Gaussian.

Then

$$\begin{aligned} E[\theta|\xi] &= E[\theta] + \frac{\text{cov}(\theta, \xi)}{\text{var}(\xi)}(\xi - E[\xi]) \\ &= 1 + \frac{2}{1}(\xi + 1) = 2\xi + 3 \end{aligned}$$

and

$$\begin{aligned} E[(\theta - E(\theta|\xi))^2] &= \text{var}(\theta) - \frac{\text{cov}(\theta, \xi)^2}{\text{var}(\xi)} \\ &= 5 - \frac{2^2}{1} = 5 - 4 = 1. \end{aligned}$$

Theorem on normal correlation (TNC) – scalar RVs

The TNC relies on the assumption that the random vectors $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_m)^T$ and $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_n)^T$ follow a joint Gaussian distribution.

The next example demonstrates that this is essential.

Example.
Consider two independent RVs $\xi \sim N(0,1)$ and $\zeta \sim N(0,1)$ so

$$E[e^{u\xi}] = E[e^{u\zeta}] = e^{\frac{1}{2}u^2}.$$

Consider now the pair of RVs θ and ξ where

$$\theta = |\xi| sgn(\zeta) = \begin{cases} |\xi| & \text{if } \zeta \geq 0, \\ -|\xi| & \text{if } \zeta < 0. \end{cases}$$

We shall show that

$$\theta \sim N(0,1) \quad \text{and} \quad \text{cov}(\theta, \xi) = 0$$

despite θ, ξ being dependent (we show this at the end).

Theorem on normal correlation (TNC) – scalar RVs

The MGF of θ for all $u \in \mathbb{R}$ is

$$\begin{aligned} E[e^{u\theta}] &= E[e^{u|\xi| sgn(\zeta)}] \\ &= E\left[E[e^{u|\xi| sgn(\zeta)}|\xi]\right] \quad \text{(law of iterated conditioning)} \\ &= E\left[\frac{1}{2}e^{u|\xi|} + \frac{1}{2}e^{-u|\xi|}\right] = E[\cosh(u|\xi|)] = E[\cosh(u\xi)] \\ &= \frac{1}{2}E[e^{u\xi} + e^{-u\xi}] = e^{\frac{1}{2}u^2} \end{aligned}$$

showing $\theta \sim N(0,1)$.

But the covariance

$$\begin{aligned} \text{cov}(\theta, \xi) &= E[\theta\xi] - E[\theta]E[\xi] = E[|\xi| sgn(\zeta)\xi] - 0 \\ &= E[|\xi|\xi]E[sgn(\zeta)] \quad (\xi, \zeta \text{ are independent}) \\ &= E[|\xi|\xi] sgn(E[\zeta]) = 0. \end{aligned}$$

Theorem on normal correlation (TNC) – scalar RVs

Finally note that

$$|\xi| = |\theta|$$

and therefore

$$\text{cov}(|\theta|, |\xi|) = \text{var}(|\xi|) = E[\xi^2] - (E[|\xi|])^2 = 1 - \frac{2}{\pi} > 0$$

and we see there is some dependence between θ and ξ.

But Part 1 of Proposition 4 (TNC) says that θ and ξ are independent if $\text{cov}(\theta, \xi) = 0$.

So while θ and ξ may be scalar normal RVs, they cannot be joint-normal.

References I