Multivariate Statistical Analysis

Lecture 05

Fudan University

luoluo@fudan.edu.cn

Outline

Outline

Singular Normal Distributions

In previous section, we focus on non-singular normal normally distributed variate $\mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with $\boldsymbol{\Sigma} \succ \mathbf{0}$ whose density function is

$$n(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{\sqrt{(2\pi)^p \det(\boldsymbol{\Sigma})}} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right).$$

What about the case of singular $\boldsymbol{\Sigma}$?

General Linear Transformation

1. Let $\mathbf{x} \sim \mathcal{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, with $\boldsymbol{\Sigma} \succ \mathbf{0}$. Then

$$\mathbf{y} = \mathbf{C}\mathbf{x}$$

   is distributed according to $\mathcal{N}_p(\mathbf{C}\boldsymbol{\mu}, \mathbf{C}\boldsymbol{\Sigma}\mathbf{C}^\top)$ for non-singular $\mathbf{C} \in \mathbb{R}^{p \times p}$.

2. Let $\mathbf{x} \sim \mathcal{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, with $\boldsymbol{\Sigma} \succ \mathbf{0}$. Then

$$\mathbf{y} = \mathbf{C}\mathbf{x}$$

   is distributed according to $\mathcal{N}_q(\mathbf{C}\boldsymbol{\mu}, \mathbf{C}\boldsymbol{\Sigma}\mathbf{C}^\top)$ for $\mathbf{C} \in \mathbb{R}^{q \times p}$ of rank $q \leq p$.

3. Let $\mathbf{x} \sim \mathcal{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Then

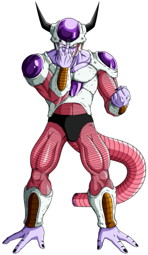
$$\mathbf{y} = \mathbf{C}\mathbf{x}$$

   is distributed according to $\mathcal{N}_q(\mathbf{C}\boldsymbol{\mu}, \mathbf{C}\boldsymbol{\Sigma}\mathbf{C}^\top)$ for any $\mathbf{C} \in \mathbb{R}^{q \times p}$.

5.3×10^5

$c \neq 0$

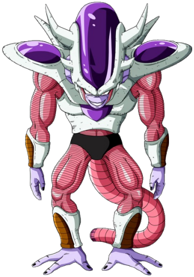
$\sigma^2 > 0$



$> 1.0 \times 10^6$

$\mathbf{C} \in \mathbb{R}^{p \times p}$ is non-singular

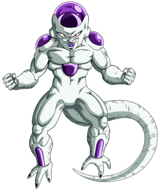
$\boldsymbol{\Sigma} \succ \mathbf{0}$



$2.0 \times 10^6 \sim 3.0 \times 10^6$

$\mathbf{C} \in \mathbb{R}^{q \times p}$ of rank $q \leq p$

$\boldsymbol{\Sigma} \succ \mathbf{0}$



$> 3.0 \times 10^7$

$\mathbf{C} \in \mathbb{R}^{q \times p}$

$\boldsymbol{\Sigma} \succeq \mathbf{0}$

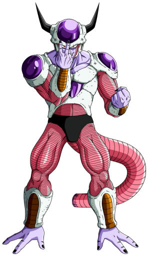
General Linear Transformation

Theorem

Let $\mathbf{x} \sim \mathcal{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$*, with* $\boldsymbol{\Sigma} \succ \mathbf{0}$*. Then*

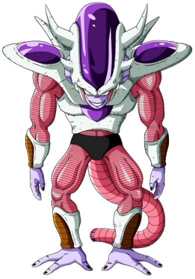
$$\mathbf{z} = \mathbf{D}\mathbf{x}$$

is distributed according to $\mathcal{N}_q(\mathbf{D}\boldsymbol{\mu}, \mathbf{D}\boldsymbol{\Sigma}\mathbf{D}^\top)$ *for* $\mathbf{D} \in \mathbb{R}^{q \times p}$ *of rank* $q \leq p$*.*



non-singular

$$\Longrightarrow$$



full-rank

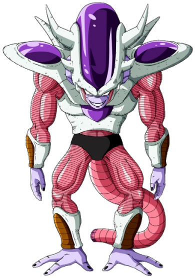
General Linear Transformation

Theorem

Let $\mathbf{x} \sim \mathcal{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. *Then*

$$\mathbf{z} = \mathbf{D}\mathbf{x}$$

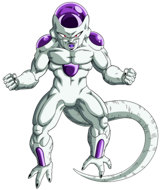
is distributed according to $\mathcal{N}_q(\mathbf{D}\boldsymbol{\mu}, \mathbf{D}\boldsymbol{\Sigma}\mathbf{D}^\top)$ *for any* $\mathbf{D} \in \mathbb{R}^{q\times p}$.



full-rank

understand the singular normal distribution

$\Longrightarrow$



no limitation

Singular Normal Distribution

Singular normal distribution:

1. The mass is concentrated on a given lower dimensional set.
2. The probability associated with any set that does not intersecting the given low-dimensional set is 0.

For example, consider that

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \right).$$

1. Probability of any set that does not intersecting the x_2-axis is 0.
2. The measure of x_2-axis in the space of $\mathbb{R}^2$ is zero.
3. The random vector $\mathbf{x}$ has no density, but its distribution exists.

To be continued...

To be continued...