Random Vectors

Mohadeseh Shafiei Kafraj[1]

[1]Gatsby Computational Neuroscience Unit
University College London

Gatsby Bridging Programme 2023

Table of Contents

Random Vectors: When are they useful?

$$\textit{The probability that a child chosen at random is healthy}$$

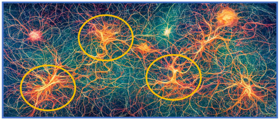
$$= \boldsymbol{f}\,(height, weight, blood\ pressure, red - blood\ cell\ count, heart\ rate)$$

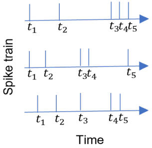
All of these variables are dependent on each other!

We need to capture this dependency!

*We need to study these variables together as a *Random Vector*!*

Random Vectors: When are they useful?





$$T_1 = \left(t_1^1, t_2^1, t_3^1, t_4^1, t_5^1\right)^T$$

$$T_2 = \left(t_1^2, t_2^2, t_3^2, t_4^2, t_5^2\right)^T$$

$$T_3 = \left(t_1^3, t_2^3, t_3^3, t_4^3, t_5^3\right)^T$$

We're interested in the timing of the first five spikes after a stimulus is presented.

The timing of each spike is dependent on the previous ones.

We need to study $\boldsymbol{t_1, \dots, t_5}$ *together as a random vector,* $\boldsymbol{T}$*!*

PDF and CDF

$\boldsymbol{X} = (X_1, ..., X_n)^T$: A random vector

$F_{\boldsymbol{X}}(\boldsymbol{x})$: *Cumulative* Distribution Function(*CDF*)

$f_{\boldsymbol{X}}(\boldsymbol{x})$: Probability *Density* function (*pdf*)

PDF and CDF

By definition, Cumulative Distribution Function(CDF) is:

$$F_{\boldsymbol{X}}(\boldsymbol{x}) = P[X_1 \leq x_1, ..., X_n \leq x_n]$$

$\boldsymbol{x} = (x_1, ..., x_n)$ we get:

$$F_{\boldsymbol{X}}(\boldsymbol{x}) = P[\boldsymbol{X} \leq \boldsymbol{x}]$$

we associate the events:
$\boldsymbol{X} \leq \infty$ with the certain event, $F_{\boldsymbol{X}}(\infty) = 1$, (Axiom 2 of probability theory! $P(\Omega = 1)$

$\boldsymbol{X} \leq -\infty$ with the impossible event, $F_{\boldsymbol{X}}(-\infty) = 0.(P(\emptyset) = 0)$

PDF and CDF

The probability *density* function (pdf) is defined as (if the derivative exists!):

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{\partial^n F_{\boldsymbol{X}}(\boldsymbol{x})}{\partial x_1 \ldots \partial x_n}$$

Equivalently we could have defined it as:

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \lim_{\Delta x_1 \to 0, \ldots, \Delta x_n \to 0} \frac{P[x_1 < X_1 \leq x_1 + \Delta x_1, \ldots, x_n < X_n \leq x_1 + \Delta x_n]}{\Delta x_1 \ldots \Delta x_n}$$

Therefore,

$$f_{\boldsymbol{X}}(\boldsymbol{x}) \Delta x_1 \ldots \Delta x_n \simeq P[x_1 < X_1 \leq x_1 + \Delta x_1, \ldots, x_n < X_n \leq x_1 + \Delta x_n]$$

PDF and CDF

pdf is defined as:

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{\partial^n F_{\boldsymbol{X}}(\boldsymbol{x})}{\partial x_1 \ldots \partial x_n}$$

if we integrate the equation, we obtain:

$$F_{\boldsymbol{X}}(\boldsymbol{x}) = \int_{-\infty}^{x_1} \ldots \int_{-\infty}^{x_n} f_{\boldsymbol{X}'}(\boldsymbol{x}') dx_1' \ldots dx_n' = \int_{-\infty}^{\boldsymbol{x}} f_{\boldsymbol{X}'}(\boldsymbol{x}') d\boldsymbol{x}'$$

more generally:

$P[B] = \int_{\boldsymbol{x} \in B} f_{\boldsymbol{X}}(\boldsymbol{x}) d\boldsymbol{x}$, where $B \subset R^N$

Exercise 1

Example

Let $f_{\mathbf{X}}(\mathbf{x})$ be given as

$$f_{\mathbf{X}}(\mathbf{x}) = Ke^{-\mathbf{x}^T \boldsymbol{\Lambda}} u(\mathbf{x}),$$

where $\boldsymbol{\Lambda} = (\lambda_1, \ldots, \lambda_n)^T$ with $\lambda_i > 0$ all $i, \mathbf{x} = (x_1, \ldots, x_n)^T$, $u(\mathbf{x}) = 1$ if $x_i \geq 0$ $i = 1, \ldots, n$ and zero otherwise, and K is a constant to be determined. What value of K will enable $f_{\mathbf{X}}(\mathbf{x})$ to be a pdf?

PDF and CDF

constraint: $(P[B] \neq 0)$

mixture *CDF*: $F_{\boldsymbol{X}}(\boldsymbol{x}) = \sum_{i=1}^{n} F_{\boldsymbol{X}|B_i}(\boldsymbol{x}|B_i)P[B_i]$

Consider $B_i, i = 1, ..., n$: n disjoint and exhaustive events, with $P[B_i] > 0$, $\cup B_i = \Omega$. $B_i B_j = \emptyset$

conditional *CDF*: $F_{\boldsymbol{X}|B}(\boldsymbol{x}|B) = P[X \leq x|B] = \frac{P[X \leq x, B]}{P[B]}$

conditional *pdf*: $f_{\boldsymbol{X}|B}(\boldsymbol{x}|B) = \frac{\partial^n F_{\boldsymbol{X}|B}(\boldsymbol{x}|B)}{\partial x_1 \dots \partial x_n}$

mixture *pdf*: $f_{\boldsymbol{X}}(\boldsymbol{x}) = \sum_{i=1}^{n} f_{\boldsymbol{X}|B_i}(\boldsymbol{x}|B_i)P[B_i]$

mixture: a linear combination

PDF and CDF

Joint distribution of *two* random vectors:

$\boldsymbol{X} = (X_1, ..., X_n).T$

$\boldsymbol{Y} = (Y_1, ..., Y_M).T$

$F_{\boldsymbol{XY}}(\boldsymbol{x}, \boldsymbol{y}) = P[\boldsymbol{X} \leq \boldsymbol{x}, \boldsymbol{Y} \leq \boldsymbol{y}]$

joint density: $f_{\boldsymbol{XY}}(\boldsymbol{x}, \boldsymbol{y}) = \frac{\partial^{n+m} F_{\boldsymbol{X}Y}(\boldsymbol{x},\boldsymbol{y})}{\partial x_1 ... \partial x_n \partial y_1 ... \partial y_m}$

marginal density: $f_{\boldsymbol{X}}(\boldsymbol{x}) = \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} f_{\boldsymbol{XY}}(\boldsymbol{x}, \boldsymbol{y}) dy_1 ... dy_n$

Expectation Vector

The expectation of the vector $\boldsymbol{X} = (X_1, ..., X_n)^T$ is a vector $\boldsymbol{\mu} = (\mu_1, ..\mu_n)^T$ whose elements are given by:

$\mu_i = \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} x_i f_{\boldsymbol{X}}(x_1, ..., x_n) dx_1 ... dx_n.$

$f_{X_i}(x_i) = \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} f_{\boldsymbol{X}}(\boldsymbol{x}) dx_1 ... dx_{i-1} dx_{i+1} ... dx_n$

$\mu_i = \int_{-\infty}^{\infty} x_i f_{X_i}(x_i) dx_i$

Exercise 2

Example

For $-\infty < x_i < \infty$, $i = 1, 2, \ldots, n$, let

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{(2\pi)^{n/2}\sigma_1 \ldots \sigma_n} \exp\left\{-\frac{1}{2}\left(\sum_{i=1}^{n}\left(\frac{x_i}{\sigma_i}\right)^2\right)\right\}.$$

Show that all the marginal pdf's are Gaussian.

Reminder:

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} exp(-\frac{1}{2}(\frac{x-\mu}{\sigma})^2)$$

Eigenvalues and Eigenvectors: Reminder

*** λ_is are the solutions of this equation:

$det(\boldsymbol{M} - \lambda \boldsymbol{I}) = 0,$

$\boldsymbol{I}$ is identity matrix

*** Characteristic equation:

$\boldsymbol{M}\phi_i = \lambda_i \phi_i$
equivalently, $\quad (\boldsymbol{M} - \lambda_i \boldsymbol{I})\phi_i = \mathbf{0} \quad \phi_i \neq \mathbf{0}$

eigenvectors are normalized if: $\phi_i^T \phi_i = ||\phi_i||^2 = 1$

Exercise 3: Reminder

Example

find the Eigenvalues and Eigenvectors of this matrix!

Note that it is real and symmetric!

$$\boldsymbol{M} = \begin{bmatrix} 4 & 2 \\ 2 & 4 \end{bmatrix}$$

GATSBY

Exercise 3: Reminder

$$\boldsymbol{M} = \begin{bmatrix} 4 & 2 \\ 2 & 4 \end{bmatrix}$$

Eigenvalues: $det(M - \lambda I) = 0$

$$det \begin{bmatrix} 4-\lambda & 2 \\ 2 & 4-\lambda \end{bmatrix} = (4-\lambda)^2 - 4 = 0$$

Eigenvalues: $\lambda_1 = 6, \lambda_2 = 2$

The eigenvector associated with λ_1:

$(M - 6I)\phi_1 = 0$, implies that, $\phi_1 = \frac{1}{\sqrt{2}}(1,1)^T$

The eigenvector associated with λ_2:

$(M - 2I)\phi_2 = 0$, implies that, $\phi_2 = \frac{1}{\sqrt{2}}(1,-1)^T$

Covariance Matrix

The covariance matrix associated with a real random vector X is:

$\boldsymbol{K} = E[(\boldsymbol{X} - \boldsymbol{\mu})(\boldsymbol{X} - \boldsymbol{\mu})^T]$, Define $\boldsymbol{K}_{ij} = E[(X_i - \mu_i)(X_j - \mu_j)^T]$

Particularly: $\sigma_i^2 = K_{ii}$, so we can write K as: $\begin{bmatrix} \sigma_1^2 & \dots & K_{1n} \\ \dots & \sigma_i^2 & \dots \\ K_{n1} & \dots & \sigma_n^2 \end{bmatrix}$

1- if $\boldsymbol{X}$ is real, all the elements of $\boldsymbol{K}$ are *real*.

2- $K_{ij} = K_{ji}$, the covariance matrix is *real symmetric*!

3- Real symmetric matrices have interesting properties! we will discuss it!

4- Covariance matrix is square!

5- Covariance matrix is positive semi-definite!

6- All eigenvalues are real and non-negative!

Properties of Covariance Matrices

Covariance matrixes are **real symmetric!**

let $\boldsymbol{M}$ be any $n \times n$ **real symmetric!**, the quadratic form associated with $\boldsymbol{M}$ is the **scalar** $q(z)$ defined by:

$q(\boldsymbol{z}) = \boldsymbol{z}^T \boldsymbol{M}\boldsymbol{z}, \boldsymbol{z}$ is any column vector!

if $q(\boldsymbol{z}) \geq 0:$ positive semidefinite

if $q(\boldsymbol{z}) > 0:$ positive definite

Theorem: positive definite matrix

Theorem

A real symmetric matrix $\boldsymbol{M}$ is positive definite if and only if all its eigenvalues are positive!

If a real symmetric matrix is positive definite its eigenvalues are positive!

Proof: if $\boldsymbol{M}$ is positive definite, thus for any vector $\boldsymbol{x} \neq 0$, $\boldsymbol{x}^T \boldsymbol{M} \boldsymbol{x} > 0$

choose x be an eigenvector, ϕ_i, so,

$\phi_i^T \boldsymbol{M} \phi_i > 0$,

$\phi_i^T \boldsymbol{M} \phi_i = \lambda_i$

$\lambda_i > 0$. Done!

Theorem: positive definite matrix

If a real symmetric matrix has positive eigenvalues it is positive definite!

we should show that for any vector $x \neq 0$,

$\boldsymbol{x}^T \boldsymbol{M} \boldsymbol{x} > 0$

let $\boldsymbol{x} = \boldsymbol{U}\boldsymbol{y}$

$\boldsymbol{x}^T \boldsymbol{M} \boldsymbol{x} = (\boldsymbol{U}\boldsymbol{Y})^T \boldsymbol{M}(\boldsymbol{U}\boldsymbol{y}) = \boldsymbol{y}^T \boldsymbol{U}^T \boldsymbol{M}\boldsymbol{U}\boldsymbol{y} =$

$\boldsymbol{y}^T \boldsymbol{\Lambda} \boldsymbol{y} = \sum_{i=1}^{n} \lambda_i y_i^2 > 0$.Done!

Exercise 4

Example

Explain why none of the following matrices can be covariance matrices associated with real random vectors.

$$\begin{bmatrix} 2 & -4 & 0 \\ -4 & 3 & 1 \\ 0 & 1 & 2 \end{bmatrix} \quad \begin{bmatrix} 4 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & -2 \end{bmatrix} \quad \begin{bmatrix} 6 & 1+j & 2 \\ 1-j & 5 & -1 \\ 2 & -1 & 6 \end{bmatrix} \quad \begin{bmatrix} 4 & 6 & 2 \\ 6 & 9 & 3 \\ 9 & 12 & 16 \end{bmatrix}$$

(a) (b) (c) (d)

Matrix Similarity

Covariance matrixes are **real symmetric!**

Definition: Two $n \times n$ matrix A, and B are similar if there exist an $n \times n$ matrix T with $det(T) \neq 0$ such that

$$T^{-1}AT = B$$

Theorem

Theorem 1: An $n \times n$ matrix M is similar to a diagonal matrix if and only if M has n linearly independent eigenvectors.

Theorem

Theorem 2: Let M be a real symmetric matrix. Then M has n mutually orthogonal unit eigenvectors $\phi_1, ..., \phi_n$.

Matrix Similarity

Theorem

Theorem 3: Let M be a real symmetric matrix. Matrix M is similar to matrix Λ:

$$U^{-1}MU = \Lambda$$

$$\Lambda = \begin{bmatrix} \lambda_1 & \dots & 0 \\ 0 & \dots & \lambda_n \end{bmatrix}$$

$$U = \begin{bmatrix} \vdots & \vdots & \vdots \\ \phi_1 & \dots & \phi_n \\ \vdots & \vdots & \vdots \end{bmatrix}$$

Exercise 5

Example

The correlation matrix R is defined by:

$$\boldsymbol{R} = E[\boldsymbol{X}\boldsymbol{X}^T]$$

Let's show that:

$$\boldsymbol{K} = \boldsymbol{R} - \boldsymbol{\mu}\boldsymbol{\mu}^T,$$

where $\boldsymbol{K}$ is the covariance matrix.

Exercise 6

Example

Find the covariance and expectation of vector $\boldsymbol{Y}$, under the transformation, $\boldsymbol{Y} = \boldsymbol{A}^T \boldsymbol{X} + \boldsymbol{B}$, based on the mean and covariance of random vector $\boldsymbol{X}$.

Final answer:

$\boldsymbol{K}_Y = \boldsymbol{A}^T \boldsymbol{K}_X \boldsymbol{A}$

$\boldsymbol{\mu}_Y = \boldsymbol{A}^T \boldsymbol{\mu}_x + \boldsymbol{B}$

Definitions

Consider real n-dimensional random vectors $\boldsymbol{X}$, $\boldsymbol{Y}$ with respective mean vectors $\boldsymbol{\mu}_x, and \boldsymbol{\mu}_y$:

$\boldsymbol{X}$, and $\boldsymbol{Y}$ are *uncorrelated* if:

$$Cov(\boldsymbol{X}, \boldsymbol{Y}) = E[(\boldsymbol{X} - \boldsymbol{\mu}_x)(\boldsymbol{Y} - \boldsymbol{\mu}_y)^T] = 0$$

Let's show that uncorrelatedness implies:

$$E[\boldsymbol{X}\boldsymbol{Y}^T] = \boldsymbol{\mu}_x \boldsymbol{\mu}_y^T$$

Definitions

Consider real n-dimensional random vectors $\boldsymbol{X}$, $\boldsymbol{Y}$ with respective mean vectors $\boldsymbol{\mu}_x, and \boldsymbol{\mu}_y$:

$\boldsymbol{X}$, and $\boldsymbol{Y}$ are *uncorrelated* if:
$Cov(\boldsymbol{X}, \boldsymbol{Y}) = E[(\boldsymbol{X} - \boldsymbol{\mu}_x)(\boldsymbol{Y} - \boldsymbol{\mu}_y)] = 0$, equivalently, $E[\boldsymbol{X}\boldsymbol{Y}^T] = \boldsymbol{\mu}_x \boldsymbol{\mu}_y^T$

$\boldsymbol{X}$, and $\boldsymbol{Y}$ are *orthogonal* if:
$E[\boldsymbol{X}\boldsymbol{Y}^T] = 0$

$\boldsymbol{X}$, and $\boldsymbol{Y}$ are *independent* if:
$f_{\boldsymbol{XY}}(\boldsymbol{x}, \boldsymbol{y}) = f_{\boldsymbol{X}}(\boldsymbol{x}) f_{\boldsymbol{Y}}(\boldsymbol{y})$

*Note: Independence implies uncorrelatedness! But the converse is not *generally* true!

Exercise 7

Prove that Independence implies uncorrelatedness!

Exercise 8

Example

Let $\mathbf{X}_i, i = 1, \ldots, n$ be n mutually orthogonal random vectors. Show that

$$E\left[\left\|\sum_{i=1}^{n} \mathbf{X}_i\right\|^2\right] = \sum_{i=1}^{n} \mathbf{E}\left[\|\mathbf{X}_i\|^2\right].$$

(*Hint*: Use the definition $\|\mathbf{X}\|^2 \triangleq \mathbf{X}^T\mathbf{X}$.)

GATSBY

Exercise 9

Example

Let $\mathbf{X}_i, i = 1, \ldots, n$ be n mutually uncorrelated random vectors with means $\boldsymbol{\mu}_i \triangleq E[\mathbf{X}_i]$. Show that

$$E\left[\left\|\sum_{i=1}^{n}(\mathbf{X}_i - \boldsymbol{\mu}_i)\right\|^2\right] = \sum_{i=1}^{n} E\left[\|\mathbf{X}_i - \boldsymbol{\mu}_i\|^2\right].$$

The Multivariate Gaussian Distribution

We already know that if X is a (scalar) Gaussian random variable, with mean $\boldsymbol{\mu}$, and variable σ^2, its pdf is

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} exp(-\frac{1}{2}(\frac{x-\mu}{\sigma})^2)$$

Now consider a random vector $\boldsymbol{X} = (X_1, ..., X_n)^T$ with independent components, the the pdf is:

$f_{\boldsymbol{X}}(x_1, ...x_n) = \prod_{i=1}^{n} f_{X_i}(x_i) = \frac{1}{(2\pi)^{n/2}\sigma_1...\sigma_n} exp[-\frac{1}{2}\Sigma_{i=1}^{n}(\frac{x_i-\mu_i}{\sigma_i})^2]$, which has the compact form as:

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{1}{(2\pi)^{n/2}[det(\boldsymbol{K})]^{1/2}} exp[-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu})^T \boldsymbol{K}^{-1}(\boldsymbol{x} - \boldsymbol{\mu})]$$

The Multivariate Gaussian Distribution

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{1}{(2\pi)^{n/2}[det(\boldsymbol{K})]^{1/2}} exp[-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T \boldsymbol{K}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})]$$

$$\boldsymbol{K} = \begin{bmatrix} \sigma_1^2 & & 0 \\ & \dots & \\ 0 & & \sigma_n^2 \end{bmatrix}$$

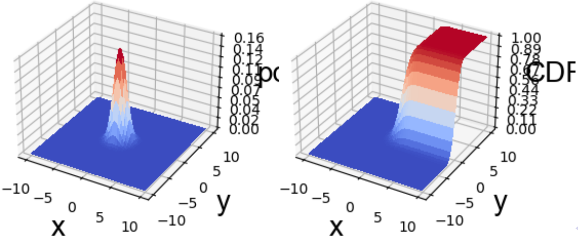
How about if $\boldsymbol{K}$ is a covariance matrix that isn't diagonal? In this case, the pdf is called a multivariate Gaussian pdf.

Coding exercise, PDF and CDF: 2D

In this exercise,

$\boldsymbol{X} = [x, y]$

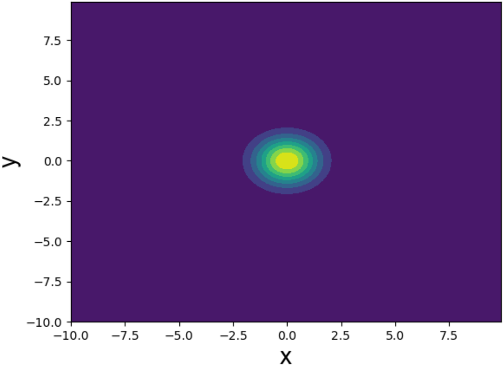
$\boldsymbol{\mu} = [0, 0]$, covariance matrix $\boldsymbol{\Sigma} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$



PDF and CDF: 2D

mean, $\boldsymbol{\mu} = [0, 0]$,

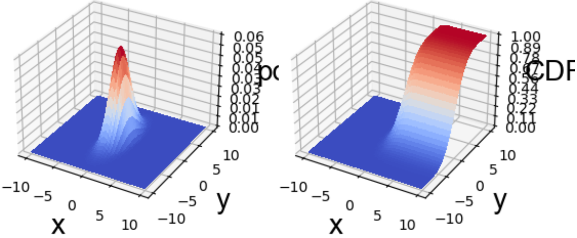
covariance matrix $\boldsymbol{\Sigma} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$



PDF and CDF: 2D

mean, $\boldsymbol{\mu} = [0, 0]$,

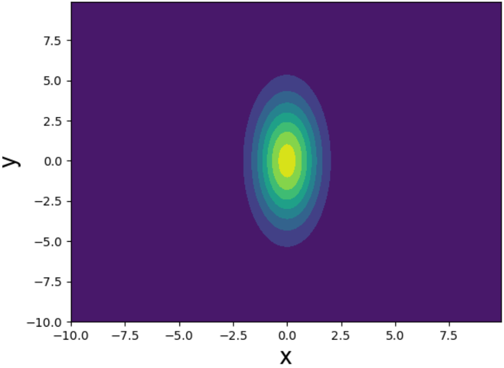
covariance matrix $\boldsymbol{\Sigma} = \begin{bmatrix} 1 & 0 \\ 0 & 7 \end{bmatrix}$



PDF and CDF: 2D

mean, $\boldsymbol{\mu} = [0, 0]$,

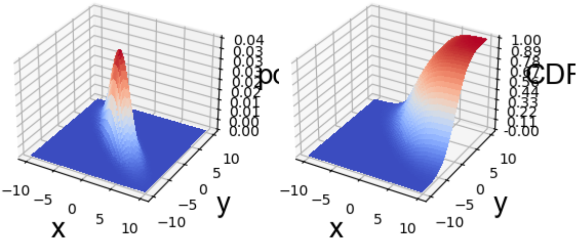
covariance matrix $\boldsymbol{\Sigma} = \begin{bmatrix} 1 & 0 \\ 0 & 7 \end{bmatrix}$



Step1 :

mean, $\boldsymbol{\mu} = [0, 0]$,

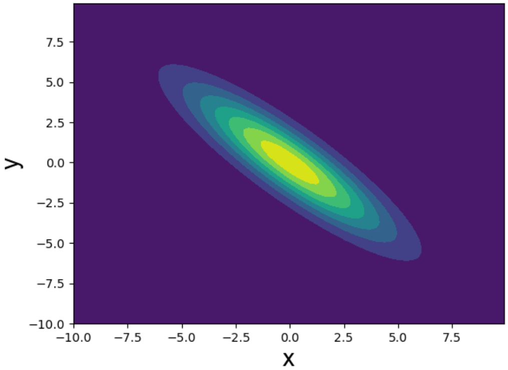
covariance matrix $\boldsymbol{\Sigma} = \begin{bmatrix} 9 & -8 \\ -8 & 9 \end{bmatrix}$



PDF and CDF: 2D

mean, $\boldsymbol{\mu} = [0, 0]$,

covariance matrix $\boldsymbol{\Sigma} = \begin{bmatrix} 9 & -8 \\ -8 & 9 \end{bmatrix}$



PDF and CDF: 3D

pdf: $f_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{exp(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu}))}{\sqrt{2\pi^k|\Sigma|}}$,

where in this example,

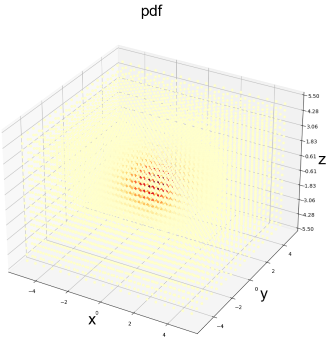
$\boldsymbol{X} = [x, y, z]$, therefore,

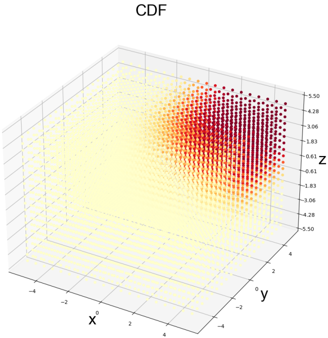
CDF: $= F_{[x,y,z]}([x, y, z]) = \int_{-\infty}^{x}\int_{-\infty}^{y}\int_{-\infty}^{z} f_{\boldsymbol{X}}(x', y', z')dx'dy'dz'$

PDF and CDF: 3D

mean, $\boldsymbol{\mu} = [0, 0, 0]$,

covariance matrix $\boldsymbol{\Sigma} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$



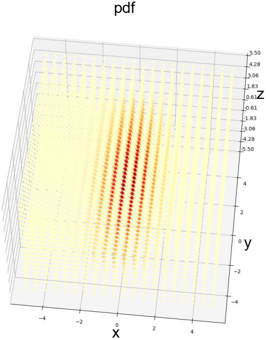


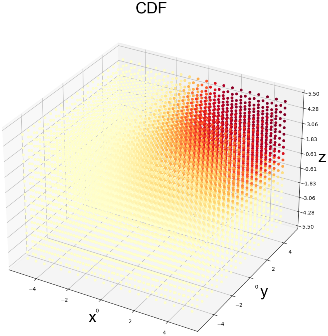
PDF and CDF: 3D

Step2 :

mean, $\boldsymbol{\mu} = [0, 0, 0]$,

covariance matrix $\boldsymbol{\Sigma} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 1 \end{bmatrix}$



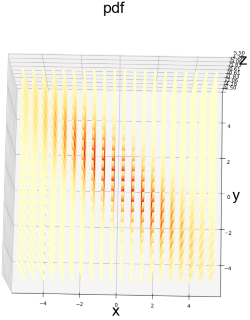


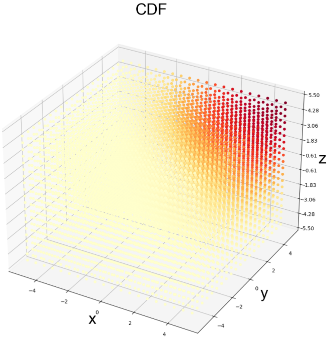
PDF and CDF: 3D

Step2 :

mean, $\boldsymbol{\mu} = [0, 0, 0]$,

covariance matrix $\boldsymbol{\Sigma} = \begin{bmatrix} 6 & -5 & 0 \\ -5 & 6 & 0 \\ 0 & 0 & 1 \end{bmatrix}$





Exercise 10

Example

For vectors
$\boldsymbol{Y} = (X_1, X_2)^T$,
$\boldsymbol{Z} = (X_3, X_4)^T$,
we write their joint vector as
$\boldsymbol{X} = (X_1, X_2, X_3, X_4)^T$

and the joint PDF as:

$f_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{1}{4\pi^2} exp(-\frac{1}{2}\boldsymbol{x}^T\boldsymbol{x})$

are $\boldsymbol{Y}, and \boldsymbol{Z}$ orthogonal? uncorrelated? independent?

GATSBY

Exercise 10

$\boldsymbol{\mu}_X = (\mu_1, \mu_2, \mu_3, \mu_4)$: the expectation of X vector, therefore

$\boldsymbol{\mu}_Y = (\mu_1, \mu_2), \boldsymbol{\mu}_Z = (\mu_3, \mu_4)$

$$\mu_1 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 f_x(x_1, x_2, x_3, x_4) dx_1 dx_2 dx_3 dx_4$$

$$\mu_1 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 \frac{1}{4\pi^2} exp(-\frac{1}{2}x^T x) dx_1 dx_2 dx_3 dx_4$$

$$\mu_1 = \frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 exp(-\frac{1}{2}(x_1^2+x_2^2+x_3^2+x_4^2)) dx_1 dx_2 dx_3 dx_4 = 0$$

Same for $\boldsymbol{\mu}_2, \boldsymbol{\mu}_3, \boldsymbol{\mu}_4$, Therefore, $\boldsymbol{\mu}_Y = \boldsymbol{\mu}_Z = (0, 0)^T$ $\boldsymbol{\mu}_Y \boldsymbol{\mu}_Z^T = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

Exercises 10

$$E(\boldsymbol{YZ}^T) = \begin{bmatrix} E[X_1X_3] & E[X_1X_4] \\ E[X_2X_3] & E[X_2X_4] \end{bmatrix}$$

$E[X_1X_3] =$
$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1x_3 exp(-\frac{1}{2}(x_1^2 + x_2^2 + x_3^2 + x_4^2))dx_1dx_2dx_3dx_4 =$

$\int_{-\infty}^{\infty} x_1 exp(-\frac{1}{2}x_1^2)dx_1 . \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_3 exp(-\frac{1}{2}(x_2^2 + x_3^2 + x_4^2))dx_2dx_3dx_4 =$
0

The same computation is true for other joint expectations.therefore,

$E(\boldsymbol{YZ}^T) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$: $\boldsymbol{Y}$, and $\boldsymbol{Z}$ are orthogonal!

$E(\boldsymbol{YZ}^T) = \mu_Y\mu_Z^T$: $\boldsymbol{Y}$, and $\boldsymbol{Z}$ are uncorrelated!

Exercise 10

$$f_{\boldsymbol{Y},\boldsymbol{Z}}(\boldsymbol{yz}) = f_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{1}{4\pi^2} exp(-\frac{1}{2}\boldsymbol{x}^T\boldsymbol{x}) =$$

$$\frac{1}{2\pi} exp(-\frac{1}{2}(x_1^2 + x_2^2)).\frac{1}{2\pi} exp(-\frac{1}{2}(x_3^2 + x_3^2)) = f_{\boldsymbol{Y}}(\boldsymbol{y})f_{\boldsymbol{Z}}(\boldsymbol{z})$$

Therefore, $\boldsymbol{Y}$, and $\boldsymbol{Z}$ are independent!

Exercise 11

Example

For vectors
$\boldsymbol{Y} = (X_1, X_2)^T$,
$\boldsymbol{Z} = (X_3, X_4)^T$,
we write their joint vector as
$\boldsymbol{X} = (X_1, X_2, X_3, X_4)^T$

and the joint PDF as:

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{1}{4\pi^2} exp(-\frac{1}{2}\boldsymbol{x}^T\boldsymbol{x})$$

Compute the correlation matrix, $\boldsymbol{R}$, and covariance matrix, $\boldsymbol{K}$, for the joint vector, $\boldsymbol{X}$!

Exercise 11

*reminder:

$$\boldsymbol{K} = E[(\boldsymbol{X} - \boldsymbol{\mu})(\boldsymbol{X} - \boldsymbol{\mu})^T]$$

Define

$K_{ij} = E[(X_i - \mu_i)(X_j - \mu_j)]$

The correlation matrix R is defined by:

$\boldsymbol{R} = E[\boldsymbol{X}\boldsymbol{X}^T]$

$$\boldsymbol{K} = \boldsymbol{R} - \boldsymbol{\mu}\mu^T$$

*Reminder $\sigma_i^2 = K_{ii}$, so we can write K as: $\begin{bmatrix} \sigma_1^2 & \dots & K_1 n \\ \dots & K_{ii}^2 & \dots \\ K_{n1} & \dots & \sigma_n^2 \end{bmatrix}$

*hint! you will need it! $\int_{-\infty}^{\infty} x^2 exp(-ax^2) = \sqrt{\frac{\pi}{4a^3}}$

Exercise 11

Note that we have shown that $\mu_X = (0,0,0,0,0)^T$, therefore,

$\boldsymbol{K} = \boldsymbol{R}$

and, $K_{ij} = E[(X_i - \mu_i)(X_j - \mu_j)] = E[X_i X_j]$

and, we have shown that if $i \neq j : K_{ij} = E[X_i X_j] = 0$

So, we just need to compute $E[X_i^2]$, and since all variables are independent, and everything is symmetric,

$E[X_1^2] = E[X_2^2] = E[X_3^2] = E[X_4^2]$

Exercise 11

$$E[X_1^2] = \frac{1}{\sqrt{2*\pi}} \int_{-\infty}^{\infty} x_i^2 exp(\frac{-1}{2} x_i^2). \frac{1}{\sqrt{2*\pi^2}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} exp(-\frac{1}{2}(x_2^2 + x_3^2 + x_3^2)) dx_2 dx_3 dx_4 =$$

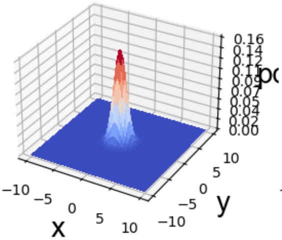
$$\frac{1}{\sqrt{2*\pi}} \int_{-\infty}^{\infty} x_i^2 exp(\frac{-1}{2} x_i^2).1 = \frac{\sqrt{2\pi}}{\sqrt{2\pi}}.1 = 1$$

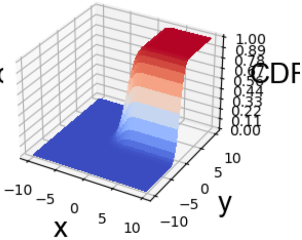
$$\text{So, } K = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

GATSBY

Exercise 11

For both $\boldsymbol{Y}$, and $\boldsymbol{Z}$ vectors, the pdf, is a multivariate Gaussian with:
mean, $\boldsymbol{\mu} = [0, 0]$,
covariance matrix $\boldsymbol{\Sigma} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$



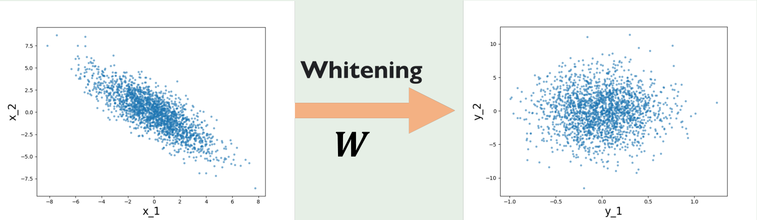


Exercise 12

Example

A random vector $\boldsymbol{X} = (X_1, X_2)$ has the covariance $\boldsymbol{K}_x = \begin{bmatrix} 5 & -4 \\ -4 & 5 \end{bmatrix}$. Find matrix $\boldsymbol{W}$, so that we can generate a new random vector $\boldsymbol{Y} = \boldsymbol{W}\boldsymbol{X}$, whose elements are uncorrelated with the covariance $\boldsymbol{K}_y = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

Answer: $\boldsymbol{W} = \Lambda^{-1/2}\boldsymbol{U}^T$

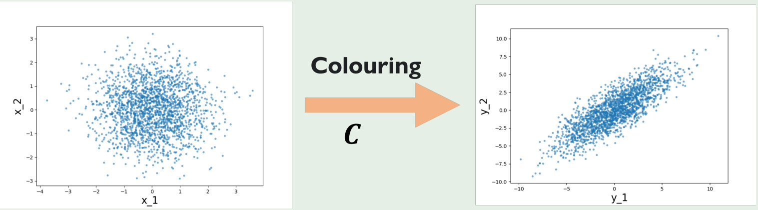


Exercise 13

Example

A random vector $\boldsymbol{X} = (X_1, X_2)$ has the covariance $\boldsymbol{K}_x = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. Find matrix $\boldsymbol{W}$, so that we can generate a new random vector $\boldsymbol{Y} = \boldsymbol{C}\boldsymbol{X}$, whose elements are correlated with the covariance $\boldsymbol{K}_y = \begin{bmatrix} 9 & 7 \\ 7 & 8 \end{bmatrix}$. Answer:

$\boldsymbol{C} = \boldsymbol{U}\boldsymbol{\Lambda}^{1/2}$



Coding exercise

Let's explore how to implement whitening and coloring in Python.

Distribution of the Sample Mean for Multivariate Gaussian Distribution

If $\boldsymbol{X}^1, ..., \boldsymbol{X}^M \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. i.i.d,

then,

$\overline{\boldsymbol{X}} = \frac{1}{M}\Sigma_{m=1}^{M}\boldsymbol{X}^m \sim N(\boldsymbol{\mu}, \frac{1}{M}\boldsymbol{\Sigma})$

*reminder:

$$Cov(\boldsymbol{X}) = \boldsymbol{\Sigma} = E[(\boldsymbol{X} - \boldsymbol{\mu})(\boldsymbol{X} - \boldsymbol{\mu})^T]$$

Distribution of the Sample Mean for Multivariate Gaussian Distribution

If $\boldsymbol{X}^1, ..., \boldsymbol{X}^M \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. i.i.d,

then,

$$\overline{\boldsymbol{X}} = \frac{1}{M}\Sigma_{m=1}^{M}\boldsymbol{X}^m \sim N(\boldsymbol{\mu}, \frac{1}{M}\boldsymbol{\Sigma})$$

*reminder:

$$Cov(\boldsymbol{X}) = \boldsymbol{\Sigma} = E[(\boldsymbol{X} - \boldsymbol{\mu})(\boldsymbol{X} - \boldsymbol{\mu})^T]$$

Proof:

$$E[\overline{\boldsymbol{X}}] = E[\frac{1}{M}\Sigma_{m=1}^{M}\boldsymbol{X}^m] = \frac{1}{M}\Sigma_{m=1}^{M}E[\boldsymbol{X}^m] = \boldsymbol{\mu}$$

GATSBY

$$Cov(\overline{\boldsymbol{X}}) = E[(\frac{1}{M}\Sigma_{m=1}^{M}\boldsymbol{X}^m - \boldsymbol{\mu})(\frac{1}{M}\Sigma_{m=1}^{M}\boldsymbol{X}^m - \boldsymbol{\mu})^T] =$$

$$= \frac{1}{M^2}\Sigma_i^M\Sigma_j^M E[(\boldsymbol{X}^i - \boldsymbol{\mu})(\boldsymbol{X}^j - \mu)^T]$$

if $i \neq j$, $\boldsymbol{X}^i$, X^j are independent, then,

$$E[(\boldsymbol{X}^i - \boldsymbol{\mu})(\boldsymbol{X}^j - \mu)^T] = 0$$

Therefore,

$$\frac{1}{M^2}\Sigma_i^M\Sigma_j^M E[(\boldsymbol{X}^i - \boldsymbol{\mu})(\boldsymbol{X}^j - \mu)^T] =$$
$$\frac{1}{M^2}(\Sigma_i^M E[(\boldsymbol{X}^i - \boldsymbol{\mu})(\boldsymbol{X}^i - \mu)^T] + \Sigma_i^M\Sigma_{j\neq i}^M E[(\boldsymbol{X}^i - \boldsymbol{\mu})(\boldsymbol{X}^j - \mu)^T) =$$
$$\frac{1}{M^2}\Sigma_i^M E[(\boldsymbol{X} - \boldsymbol{\mu})(\boldsymbol{X} - \mu)^T] = \frac{1}{M}E[(\boldsymbol{X} - \boldsymbol{\mu})(\boldsymbol{X} - \mu)^T]$$

$$= \frac{1}{M}Cov(\boldsymbol{X}) = \frac{1}{M}\boldsymbol{\Sigma}$$

Reference

Let's review!

Reference

Probability And Random Processes With Applications To Signal Processing
Stark And Woods

The End