Random Vectors

Mohadeseh Shafiei Kafraj[1]

[1]Gatsby Computational Neuroscience Unit
University College London

Gatsby Bridging Programme 2023

GATSBY

Table of Contents

Random Vectors: When are they useful?

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

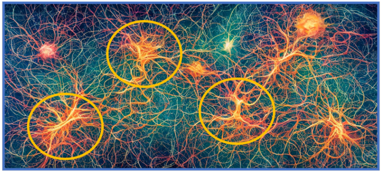
$$= f\,(height, weight, blood\ pressure, red - blood\ cell\ count, red - blood\ cell\ cpund, heart\ rate$$

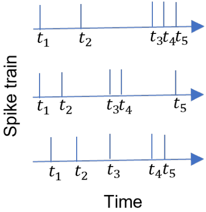
All of these variables are dependent on each other!

We need to capture this dependency!

We need to study them 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 exploring the timing of the first five spikes after a stimulus is presented.
Our hypothesis is that the timing of each spike is influenced by its predecessors, suggesting that these timings are not independent.
Therefore! It's important that we capture these dependencies!

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, $F_{\boldsymbol{X}}(\infty) = 1$,and

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

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 $B \subset R^N$

PDF and CDF

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

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

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

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

GATSBY

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

GATSBY

PDF and CDF: Here's a fun example!

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

where in this example,

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

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

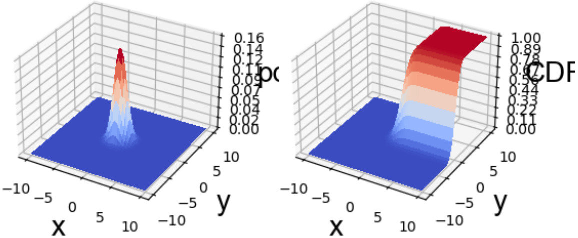
Step 1: For this interesting distribution, implement step 1 to see how pdf and CDF look like, for different mean vectors and covariance matrices!

PDF and CDF: 2D

Step1 :

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

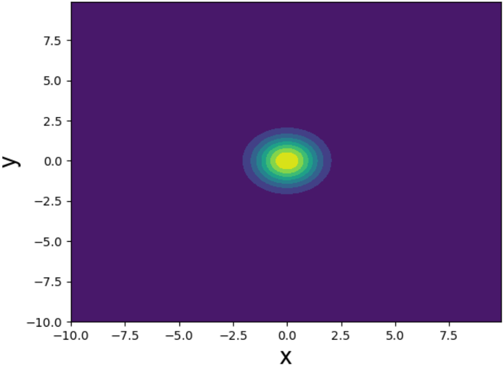
covariance matrix $\boldsymbol{\Sigma} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$



PDF and CDF: 2D

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

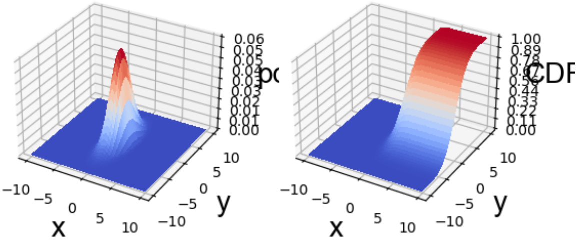
covariance matrix $\boldsymbol{\Sigma} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$



PDF and CDF: 2D

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

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

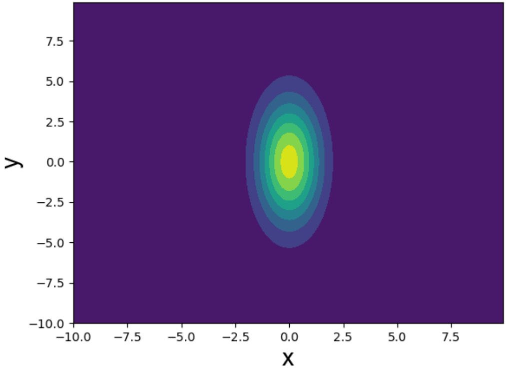


PDF and CDF: 2D

Step1 :

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

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

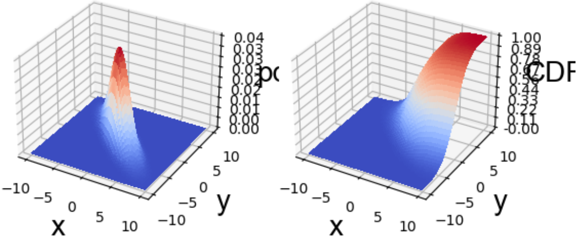


PDF and CDF: 2D

Step1 :

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

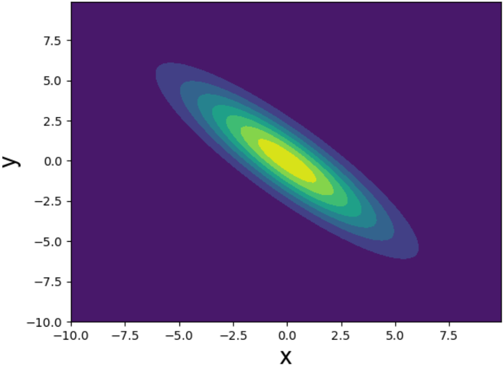
covariance matrix $\boldsymbol{\Sigma} = \begin{bmatrix} 9 & -8 \\ -8 & 9 \end{bmatrix}$



PDF and CDF: 2D

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

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



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

Step 2: Implement step 2 to see what pdf and CDF look like, for different mean vector and covariance matrices!

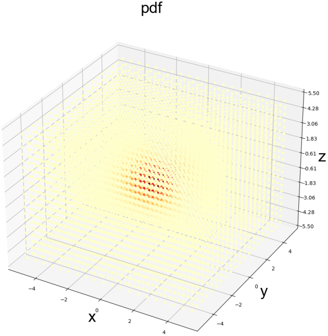
GATSBY

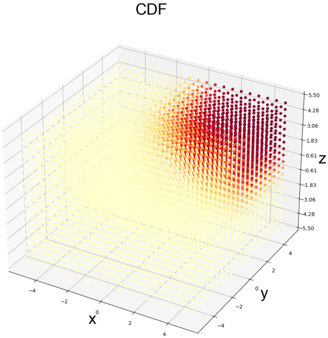
PDF and CDF: 3D

Step2 :

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

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



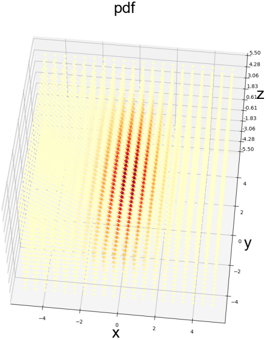


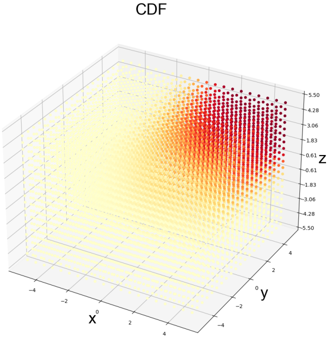
PDF and CDF: 3D

Step2 :

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

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



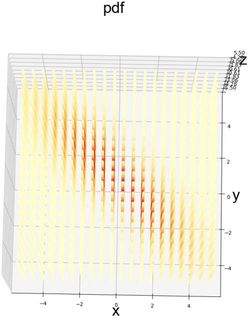


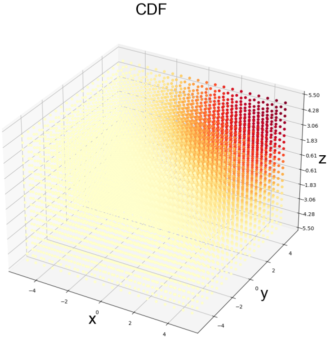
PDF and CDF: 3D

Step2 :

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

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





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:

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

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

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

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

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

Define

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

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

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 many interesting properties! we will discuss it!

Correlation matrix

The correlation matrix R is defined by:

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

$$\boldsymbol{R} = \boldsymbol{K} + \boldsymbol{\mu}\boldsymbol{\mu}^T \text{ , and}$$

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

GATSBY

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

$\boldsymbol{X}$, 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!

GATSBY

Expectation vector, Covariance matrix: 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:

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

Is this distribution familiar to you?

let $\boldsymbol{\mu}_Y$, and $\boldsymbol{\mu}_Z$ be the mean vectors of vector $\boldsymbol{Y}$, and $\boldsymbol{Z}$ respectively,

1- compute $\boldsymbol{\mu}_Y$, and $\boldsymbol{\mu}_Z$!, then $\boldsymbol{\mu}_Y \boldsymbol{\mu}_Z^T$

Expectation vector, Covariance matrix: Example!

$\boldsymbol{\mu}_X = (\mu_1, \mu_2, \mu_3, \mu_4)$: 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, $\boldsymbol{\mu}_Y = \boldsymbol{\mu}_Z = (0, 0)^T$ $\boldsymbol{\mu}_Y \boldsymbol{\mu}_Z^T = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

Expectation vector, Covariance matrix: Example!

$\boldsymbol{Y} = (X_1, X_2)^T$
$\boldsymbol{Z} = (X_3, X_4)^T$,
$\boldsymbol{X} = (X_1, X_2, X_3, X_4)^T$

According to the previous part, $\boldsymbol{\mu}_Y \boldsymbol{\mu}_Z^T = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

2- Compute $E(\boldsymbol{Y}\boldsymbol{Z}^T)$!

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

Expectation vector, Covariance matrix: Example!

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

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

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

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

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

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

Expectation vector, Covariance matrix: Example!

$\boldsymbol{Y} = (X_1, X_2)^T$
$\boldsymbol{Z} = (X_3, X_4)^T,$
$\boldsymbol{X} = (X_1, X_2, X_3, X_4)^T$

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

3- Are $\boldsymbol{Y}$, and $\boldsymbol{Z}$ independent?

Expectation vector, Covariance matrix: Example!

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

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

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

Expectation vector, Covariance matrix: Example!

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

*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:
$\boldsymbol{R} = E[\boldsymbol{X}\boldsymbol{X}^T]$
$\boldsymbol{R} = \boldsymbol{R} + \boldsymbol{\mu}\boldsymbol{\mu}^T$ $\boldsymbol{K} = \boldsymbol{R} - \boldsymbol{\mu}\mu^T$

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

GATSBY

Expectation vector, Covariance matrix: Example!

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

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

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,

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

Expectation vector, Covariance matrix: Example!

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

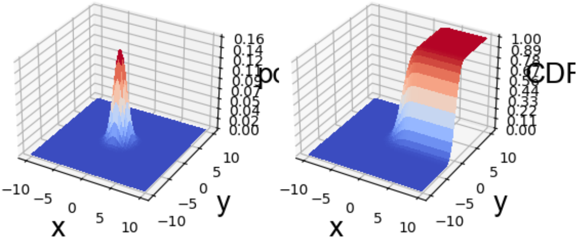
$\frac{1}{\sqrt{2*\pi}} \int_{-\infty}^{\infty} x_i^2 exp(\frac{-1}{2} x_i^2).1 = \frac{\sqrt{2\pi}}{\sqrt{2\pi}}.1 = 1$

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

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

Expectation vector, Covariance matrix: Example!

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



Properties of Covariance Matrices

Covariance matrixes are **real symmetric!**

let $\boldsymbol{M}$ be any $n \times nr.s.matrix$, 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

GATSBY

Eigenvalues and Eigenvectors

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

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

Eigenvalues and Eigenvectors: example

find the Eigenvalues and Eigenvectors of this matrix!

Note that it is real and symmetric!

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

Eigenvalues and Eigenvectors: example

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

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, $\phi_2 = \frac{1}{\sqrt{2}}(1, -1)^T$

Eigenvalues and Eigenvectors

Definition: Two $n \times no$ 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: An $n \times n$ matrix M is similar to a diagonal matrix if and only if M has n linearly independent eigenvectors.

Theorem

Theorem: Let M be a r.s. matrix with eigenvalues $\lambda_1, .., \lambda_n$. Then M has n mutually orthogonal unit eigenvectors $\phi_1, ..., \phi_n$.

Eigenvalues and Eigenvectors

Theorem

Theorem: Let M be a real symmetric matrix with eigenvalues $\lambda_1, ..., \lambda_n$, matrix M is similar to matrix Λ:

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

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

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

Example: Decorrelation of random vectors!

$\boldsymbol{X} = (X_1, X_2, X_3)^T$: a random vector

$$\boldsymbol{K}_X = \begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & 0 \\ 1 & 0 & 2 \end{bmatrix}$$

solve $det(\boldsymbol{K} - \lambda I) = 0$ to find eigenvalues, and then use this equation, $\boldsymbol{M}\phi_i = \lambda\phi_i$ to find eigenvectors!

Example: Decorrelation of random vectors!

$Reminder : det(\boldsymbol{K} - \lambda_i I) = 0, \boldsymbol{K}\phi_i = \lambda_i \phi_i, i = 1, 2, 3$

$\lambda_1 = 2 : \phi_1 = (0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})^T$

$\lambda_2 = 2 + \sqrt{2} : \phi_2 = (\frac{1}{\sqrt{2}}, \frac{-1}{2}, \frac{1}{2})^T$

$\lambda_3 = 2 - \sqrt{2} : \phi_3 = (\frac{1}{\sqrt{2}}, \frac{1}{2}, \frac{-1}{2})^T$

GATSBY

Example: Decorrelation of random vectors!

Find the new vector Y under the transformation of $\boldsymbol{Y} = \boldsymbol{U}^T \boldsymbol{X}$, where $\boldsymbol{U}$ is defined as:

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

Example: Decorrelation of random vectors!

$$\boldsymbol{Y} = \boldsymbol{U}^T \boldsymbol{X} = \begin{bmatrix} 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{\sqrt{2}} & \frac{1}{2} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}}(X_2 + X_3) \\ \frac{1}{\sqrt{2}}X_1 - \frac{1}{2}X_2 + \frac{1}{2}X_3 \\ \frac{1}{\sqrt{2}}X_1 + \frac{1}{2}X_2 - \frac{1}{2}X_3 \end{bmatrix}$$

What is the covariance of the transformed vector Y?

Example: Decorrelation of random vectors!

$$Reminder : Cov(\boldsymbol{X}) = \boldsymbol{K}_X = E[\boldsymbol{X}\boldsymbol{X}^T] - E[\boldsymbol{X}]E[\boldsymbol{X}]^T$$

$$Reminder : Cov(\boldsymbol{Y}) = \boldsymbol{K}_Y = E[\boldsymbol{Y}\boldsymbol{Y}^T] - E[\boldsymbol{Y}]E[\boldsymbol{Y}]^T$$

$$\boldsymbol{K}_Y = E[\boldsymbol{U}^T\boldsymbol{X}\boldsymbol{X}^T\boldsymbol{U}] - E[\boldsymbol{U}^T\boldsymbol{X}]E[\boldsymbol{X}^T\boldsymbol{U}] =$$

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

$$\boldsymbol{K}_Y = \boldsymbol{U}^T\boldsymbol{K}_x\boldsymbol{U}$$

Example: Decorrelation of random vectors!

Form characteristic function we have that,

$$\boldsymbol{K}_X \boldsymbol{U} = \boldsymbol{U}\boldsymbol{\Lambda}$$

which implies that

$$\boldsymbol{U}^{-1}\boldsymbol{K}_X \boldsymbol{U} = \boldsymbol{U}^{-1}\boldsymbol{U}\boldsymbol{\Lambda} = \boldsymbol{\Lambda}$$

and since all eigenvectors are orthogonal $\boldsymbol{U}^T \boldsymbol{U} = \boldsymbol{I}$ so,
$\boldsymbol{U}^T \boldsymbol{U}\boldsymbol{U}^{-1} = \boldsymbol{U}^T = \boldsymbol{U}^{-1}$

therefore, $\boldsymbol{U}^{-1}\boldsymbol{K}_X \boldsymbol{U} = \boldsymbol{U}^T \boldsymbol{K}_X \boldsymbol{U} = \boldsymbol{\Lambda}$

Example: Decorrelation of random vectors!

Finally! $\boldsymbol{K}_Y = \boldsymbol{U}^T \boldsymbol{K}_x \boldsymbol{U} = \boldsymbol{\Lambda}$

$$\boldsymbol{\Lambda} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2+\sqrt{2} & 0 \\ 0 & 0 & 2-\sqrt{2} \end{bmatrix}$$

The new covariance matrix of the transformed vector is diagonal, so,

components of $\boldsymbol{Y}$ are Uncorrelated!
A coding example of this type will be shown!

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!

The Multivariate Gaussian Distribution

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

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

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 definite covariance matrix that isn't diagonal? In this case, the pdf is called a multivariate Gaussian pdf.

The Multivariate Gaussian Distribution

Theorem

Let $\boldsymbol{X}$ be an n-dimensional Normal random vector with positive definite covariance matrix $\boldsymbol{K}$ and mean vector $\boldsymbol{\mu}$. Let $\boldsymbol{A}$ be a non-singular linear transformation in n dimensions. then $\boldsymbol{Y} = \boldsymbol{AX}$ is an n-dimensional Normal random vector with covariance matrix $\boldsymbol{Q} = \boldsymbol{AKA}^T$, and mean vector $\beta = \boldsymbol{A\mu}$.

Exercise: The Multivariate Gaussian Distribution

A zero-mean Normal random vector $\boldsymbol{X} = (X_1, X_2)^T$ has covariance matrix $\boldsymbol{K}$ given by

$$\boldsymbol{K}_x = \begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix}$$

find the transformation $\boldsymbol{D}$, $\boldsymbol{Y} = \boldsymbol{DX}$ such that: $\boldsymbol{K}_Y = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

*Reminder: Under the transformation, $\boldsymbol{Y} = \boldsymbol{U}^T\boldsymbol{X}$, the covariance matrix $\boldsymbol{K}_y$ becomes $\boldsymbol{K}_y = \boldsymbol{U}^T\boldsymbol{XU} = \Lambda$.

*Reminder: $\boldsymbol{U} = \boldsymbol{U}^T$

GATSBY

Exercise: The Multivariate Gaussian Distribution

$\boldsymbol{D} = \boldsymbol{ZU}^T$ where $\boldsymbol{Z}$ is the normalizing matrix!

$$Z = \begin{bmatrix} {\lambda_1}^{-1/2} & 0 \\ 0 & {\lambda_2}^{-1/2} \end{bmatrix}$$

Let's check the covariance matrix is indeed an identity matrix!

Is $\boldsymbol{K}_y = \boldsymbol{DK}_x\boldsymbol{D} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$?

Here's a hint!

All you need to do is find the eigenvalues ($det[\boldsymbol{K}_x - \lambda \boldsymbol{I}] = 0$) and eigenvectors $(\boldsymbol{K}_x - \lambda_i \boldsymbol{I})\phi_i = 0$, $\|\phi_i\| = 1$, and substitute them in the equation above!

Exercise: The Multivariate Gaussian Distribution

If we want to generate correlated samples of a random vector $\boldsymbol{X}$, whose covariance matrix $\boldsymbol{K}$ is not diagonal, we can use this transformation:

$\boldsymbol{X} = \boldsymbol{D}^{-1}\boldsymbol{Y}$, where $\boldsymbol{D} = \boldsymbol{Z}\boldsymbol{U}^T$

$\boldsymbol{Y}$ is a Normal random vector with uncorrelated components of unity variance,

$$\boldsymbol{K}_Y = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

It's time to code!

Exercise: Distribution of the Sample Mean

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:

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

Exercise: Distribution of the Sample Mean

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:

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

GATSBY

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

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

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

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

Therefore,

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

$$= \tfrac{1}{M}Cov(\boldsymbol{X}^i) = \tfrac{1}{M}\boldsymbol{\Sigma}$$

GATSBY