Random Vectors

Mohadeseh Shafiei Kafraj[1]

[1]Gatsby Computational Neuroscience Unit
University College London

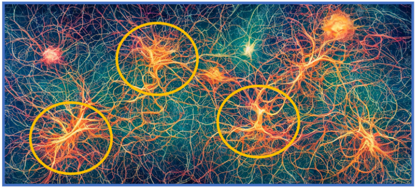
Gatsby Bridging Programme 2023

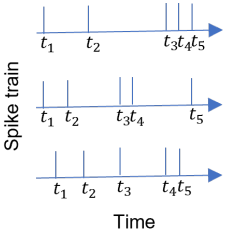
Table of Contents

GATSBY

Random Vectors: When are they useful?

This slide requires a change:





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

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

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

GATSBY

PDF and CDF

$X = (X_1, ..., X_n)^T$: A random vector

$F_x(x)$: *Cumulative* Distribution Function(*CDF*)

$f_x(x)$: Probability *Density* function (*pdf*)

PDF and CDF

By definition, Cumulative Distribution Function(CDF) is:
$F_x(x) = P[X_1 \leq x_1, ..., X_n \leq x_n]$

$x = (x_1, ..., x_n)$ we get:

$F_x(x) = P[X \leq x]$

we associate the events:
$X \leq \infty$ with the certain event, $F_x(\infty) = 1$,and

$X \leq -\infty$ with the impossible event, $F_x(-\infty) = 0$.

PDF and CDF

The probability *density* function (pdf) is defined as:

$$f_x(x) = \frac{\partial^n F_x(x)}{\partial x_1 ... \partial x_n}$$

Equivalently we could have defined it as:

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

Therefore,

$$f_x(x) \Delta x_1 ... \Delta x_n \simeq P[x_1 < X_1 \leq x_1 + \Delta x_1, ..., x_n < X_n \leq x_1 + \Delta x_n]$$

PDF and CDF

pdf is defined as:

$$f_x(x) = \frac{\partial^n F_x(x)}{\partial x_1 \ldots \partial x_n}$$

if we integrate the equation, we obtain:

$$F_x(x) = \int_{-\infty}^{x_1} \ldots \int_{-\infty}^{x_n} f_x(x')dx_1' \ldots dx_n' = \int_{-\infty}^{x} f_x(x')dx'$$

more generally:

$P[B] = \int_{x \in B} f_x(x')dx'$, where $B \subset R^N$

PDF and CDF

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

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

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

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

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

mixture: a linear combination

PDF and CDF

Joint distribution of *two* random vectors:

$X = (X_1, ..., X_n).T$

$Y = (Y_1, ..., Y_M).T$

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

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

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

PDF and CDF: Here's a fun example!

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

where in this example,

$X = [x, y]$, therefore,

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

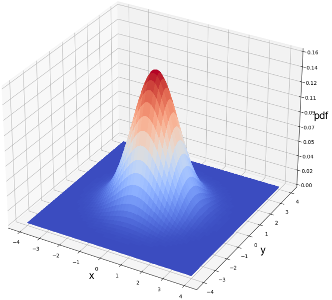
Step 1: For this interesting distribution, implement step 1 to see what pdf and CDF look like, for different mean vector and covariance matrices!

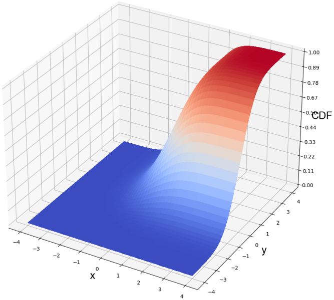
PDF and CDF: 2D

Step1 :

mean, $\mu = [0, 0]$,

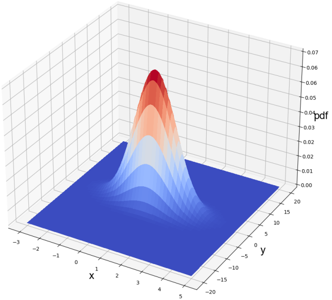
covariance matrix $\Sigma = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

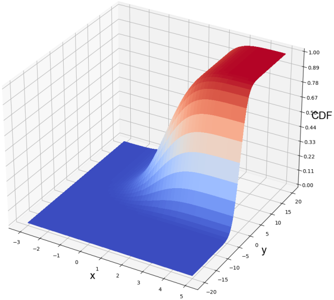




PDF and CDF: 2D

Step1 :
mean, $\mu = [1, 0]$,
covariance matrix $\Sigma = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}$



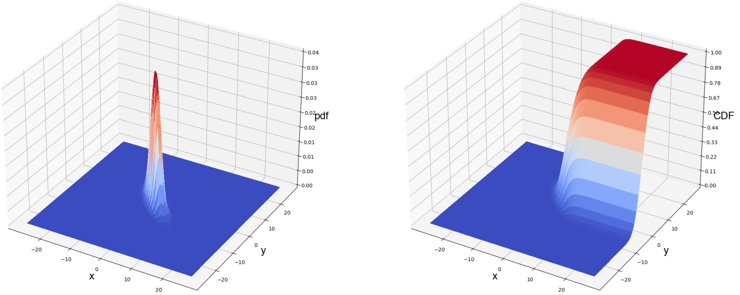


PDF and CDF: 2D

Step1 :

mean, $\mu = [0, 0]$,

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



PDF and CDF: 3D

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

where in this example,

$X = [x, y, z]$, therefore,

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

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

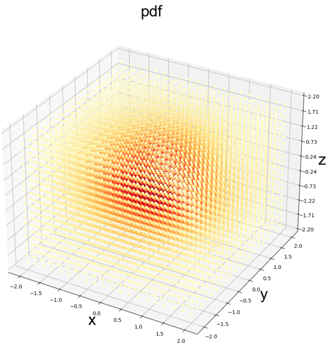
GATSBY

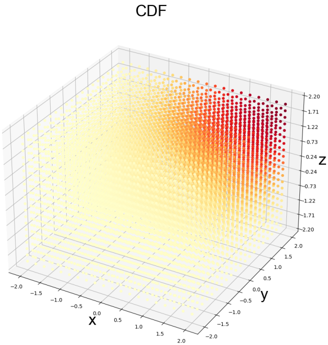
PDF and CDF: 3D

Step2 :

mean, $\mu = [0, 0, 0]$,

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



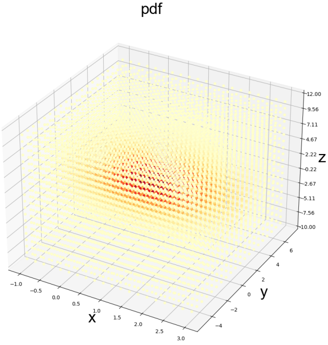


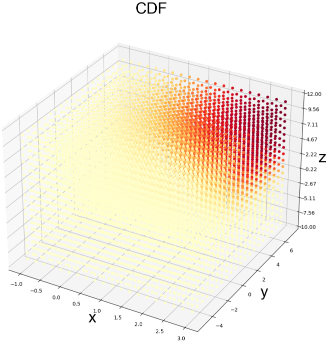
PDF and CDF: 3D

Step2 :

mean, $\mu = [1, 1, 1]$,

covariance matrix $\Sigma = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix}$



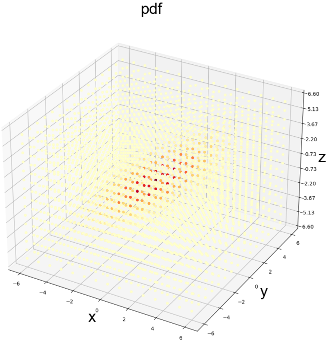


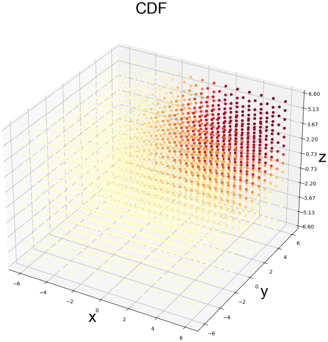
PDF and CDF: 3D

Step2 :

mean, $\mu = [0, 0, 0]$,

covariance matrix $\Sigma = \begin{bmatrix} 3 & 2 & -2 \\ 2 & 3 & 0 \\ -2 & 0 & 3 \end{bmatrix}$





Expectation Vector

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

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

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

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

Covariance Matrix

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

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

Define

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

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 X is real, all the elements of 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:

$R = E[XX^T]$

$R = R + \mu\mu^T$, and

$K = R - \mu\mu^T$

Definitions

Consider real n-dimensional random vectors X, Y with respective mean vectors $\mu_x, and \mu_y$:

X, and Y are *uncorrelated* if:
$E[XY^T] = \mu_x \mu_y^T$

X, and Y are *orthogonal* if:
$E[XY^T] = 0$

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

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

Expectation vector, Covariance matrix: Example!

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

and the joint PDF as:

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

Is this distribution familiar to you?

let μ_Y, and μ_Z be the mean vectors of vector Y, and Z respectively,

1- compute μ_Y, and μ_Z!, then $\mu_Y \mu_Z^T$

Expectation vector, Covariance matrix: Example!

$\mu_X = (\mu_1, \mu_2, \mu_3, \mu_4)$: the expectation of X vector, therefore

$\mu_Y = (\mu_1, \mu_2), \mu_Z = (\mu_3, \mu_4)$

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

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

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

Therefore, $\mu_Y = \mu_Z = (0,0)^T \ \mu_Y \mu_Z^T = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

Expectation vector, Covariance matrix: Example!

$Y = (X_1, X_2)^T$
$Z = (X_3, X_4)^T$,
$X = (X_1, X_2, X_3, X_4)^T$

According to the previous part, $\mu_Y \mu_Z^T = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

2- Compute $E(YZ^T)$!

Are Y, and Z orthogonal? Are Y, and Z uncorrelated?

Expectation vector, Covariance matrix: Example!

$$E(YZ^T) = \begin{bmatrix} EX_1X_3] & E[X_1X_4] \\ E[X_2X_3] & E[X_2X_4] \end{bmatrix}$$

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

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

therefore,

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

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

Expectation vector, Covariance matrix: Example!

$Y = (X_1, X_2)^T$,
$Z = (X_3, X_4)^T$,
$X = (X_1, X_2, X_3, X_4)^T$

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

3- Are Y, and Z independent?

Expectation vector, Covariance matrix: Example!

$f_{YZ}(yz) = f_X(x) = \frac{1}{4\pi^2} exp(-\frac{1}{2}x^T x) =$

$\frac{1}{2\pi} exp(-\frac{1}{2}(x_1^2 + x_2^2)).\frac{1}{2\pi} exp(-\frac{1}{2}(x_3^2 + x_3^2)) = f_Y(y) f_Z(z)$

Therefore, Y, and Z are independent!

Expectation vector, Covariance matrix: Example!

Compute the correlation matrix, R, and covariance matrix, K for the joint vector, X!

*reminder:

$K = E[(X - \mu)(X - \mu)^T]$
Define
$K_{ij} = E[(X_i - \mu_i)(X_j - \mu_j)]$
The correlation matrix R is defined by:
$R = E[XX^T]$
$R = R + \mu\mu^T \; K = R - \mu\mu^T$

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

Expectation vector, Covariance matrix: Example!

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

$K = R - \mu\mu^T = R$

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

Expectation vector, Covariance matrix: Example!

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

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

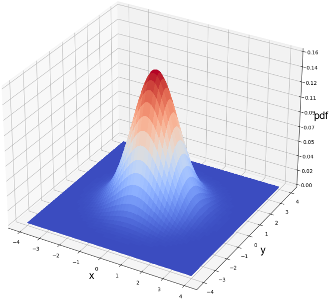
So, $K = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

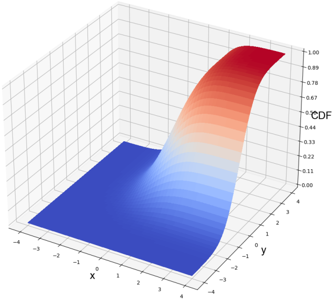
*Reminder $\sigma_i^2 = K_{ii}$, so we can write K as: $\begin{bmatrix} \sigma_1^2 & ... & K_1 n \\ ... & K_{ii}^2 & ... \\ K_{n1} & ... & \sigma_n^2 \end{bmatrix}$

GATSBY

Expectation vector, Covariance matrix: Example!

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





Properties of Covariance Matrices

Covariance matrixes are real symmetric!

let M be any $n \times nr.s.matrix$, the quadratic form associated with M is the scalar $q(z)$ defined by:

$q(z) = z^T Mz, z$ is any column vector!

if $Z^T Mz \geq 0$: positive semidefinite

if $Z^T Mz > 0$: positive definite

Eigenvalues and Eigenvectors

Characteristic equation: $M\phi = \lambda\phi, \phi \neq \mathbf{0}$

eigenvalue: λ

λs are the solutions of this equation: $det(M - \lambda I) = 0$, I is identity matrix

eigenvector: $\phi = (\phi_1, ..., \phi_n)^T$

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

Eigenvalues and Eigenvectors: example

find the Eigenvalues and Eigenvectors of this matrix!

Note that it is real and symmetric!

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

Eigenvalues and Eigenvectors: example

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

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:

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

The eigenvector associated with λ_2:

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

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

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

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

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

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

$$U^T U = I, (U^T = U^{-1}), so :$$

$$U^T MU = \Lambda$$

Example: Decorrelation of random vectors!

$X = (X_1, X_2, X_3)^T$: a random vector

$$K_X = \begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & 0 \\ 1 & 0 & 2 \end{bmatrix}$$

solve $det(K - \lambda I) = 0$ to find eigenvalues, and then use the characteristic function, $M\phi = \lambda\phi$ to find eigenvectors!

Example: Decorrelation of random vectors!

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

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

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

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

GATSBY

Example: Decorrelation of random vectors!

Find the new vector Y under the transformation of $Y = U^T$, where U is defined as:

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

GATSBY

Example: Decorrelation of random vectors!

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

What is the covariance of the transformed vector Y?

Example: Decorrelation of random vectors!

$$Reminder : Cov(X) = K_X = E[XX^T] - E[X]E[X]^T$$

$$Reminder : Cov(Y) = K_Y = E[YY^T] - E[Y]E[Y]^T$$

$$K_Y = E[U^T XX^T U] - E[U^T X]E[X^T U] =$$

$$U^T E[XX^T]U - U^T E[X]E[X]^T U = U^T(E[XX^T] - E[X]E[X]^T)U =$$

$$K_Y = U^T K_x U$$

Example: Decorrelation of random vectors!

Form characteristic function we have that,

$K_X U = U\Lambda$

which implies that

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

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

therefore, $U^{-1} K_X U = U^T K_X U = \Lambda$

Example: Decorrelation of random vectors!

Finally! $K_Y = U^T K_x U = \Lambda$

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

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

components of Y are Uncorrelated!

Theorem: positive definite matrix

A real symmetric matrix M is positive definite if and only if all its eigenvalues are positive!
proof:

First: if M is positive definite, thus for any vector $x \neq 0$,

$x^T M x > 0$

choose x be an eigenvector, ϕ_i, so,

$\phi_i^T M \phi_i > 0$,

$\phi_i^T M \phi_i = \lambda_i$

$\lambda_i > 0$!

Theorem: positive definite matrix

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

$x^T M x > 0$

let $x = Uy$

$x^T Mx {=} (UY)^T M(Uy) = y^T U^T MUy =$

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

The Multidimensional Gaussian Law

This is ...

Distribution of the Sample Mean

This is ...

Conditional Gaussian distributions

This is ...

Marginal Gaussian distributions

This is ...