Mathematical Methods for Automation Engineering M
– Jointly Distributed Random Variables –

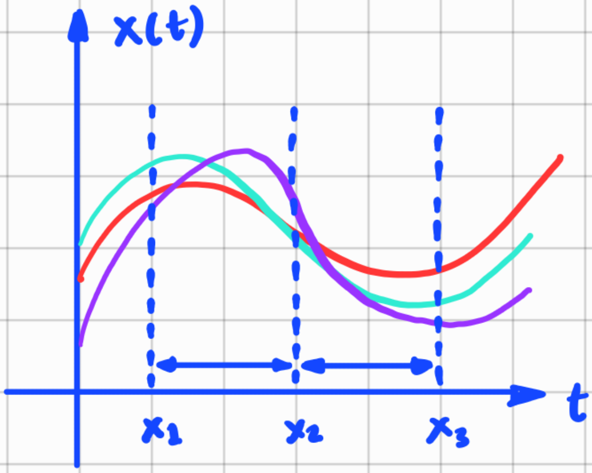
Andrea Mentrelli

Department of Mathematics &
Alma Mater Research Center on Applied Mathematics, AM^2
University of Bologna

andrea.mentrelli@unibo.it

A.Y. 2024/25

A random variable $X(t)$ not necessarily time-variant, can be discrete or continuous



Jointly Distributed Random Variables

Let us consider a random experiment with a sample space Ω

Definition (Multidimensional (jointly distributed) random variables)

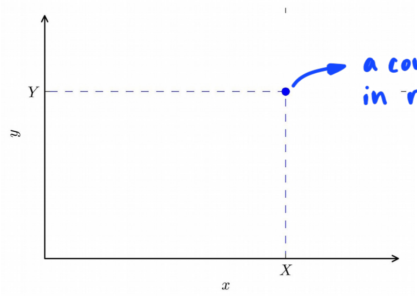
A *m-dimensional random variable* $(X_1, X_2, \ldots X_m)$ is a random variable that associates to an event $E \subset \Omega$ an ordered set of m real numbers $(x_1, x_2, \ldots x_m) \in \mathbb{R}^m$

- $\boxed{X_1, X_2, \ldots X_m}$ are called $\boxed{\textbf{components}}$ of the random variable, and are themselves random variables
- multidimensional (**jointly distributed**) random variables can be **discrete**, **continuous**, or **mixed**
- For two-dimensional, or **double random variables** $\boxed{(m = 2)}$ we shall use the symbols $\boxed{(X_1,\ X_2) \text{ or } (X,\ Y)}$

Double random variables – Geometric interpretation

random variable with two components

The components of a double random variable (X, Y) can be interpreted as the coordinates of a **point** of the Cartesian plane Oxy



a component in random variables

Some examples:

- coordinates (X, Y) of a target ✓
- weight and height (W, H) of a person ✓
- diameter and length (Φ, L) of screws ✓

dependent on each other

Jointly Distributed Random Variables

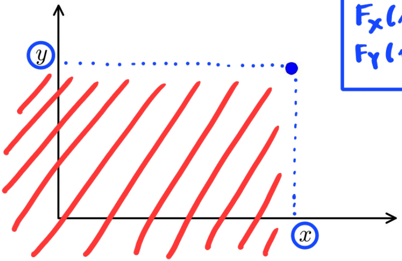
Definition (Joint cumulative distribution function)

We call *joint cumulative distribution function* $\boxed{\mathcal{F}_{XY}(x,y)}$ of the double random variable (X,Y), the probability that the inequalities $X \leqslant x$ and $Y \leqslant y$ are *simultaneously* verified, namely

$$\boxed{\mathcal{F}_{XY}(x,y) = \mathcal{P}(X \leqslant x,\ Y \leqslant y)}$$

Straight generalization

$$\boxed{\begin{aligned} F_X(x) &= P(X \leq x) \\ F_Y(y) &= P(Y \leq y) \end{aligned}}$$



Jointly Distributed Random Variables

Let the components $\boxed{X \text{ and } Y \text{ be discrete random variables}}$ taking on the values $\{x_1, x_2, \ldots x_n\}$ and $\{y_1, y_2, \ldots y_m\}$, respectively

$p_i = P(X = x_i)$

Definition (Joint probability mass function)

We call *joint probability mass function* p_{ij} of a double discrete random variable (X, Y), the probability that the pair (X, Y) takes on the value (x_i, y_j), namely

$$\boxed{p_{ij} = \mathcal{P}\left(X = x_i, Y = y_j\right)}$$

Clearly

$$0 \leqslant p_{ij} \leqslant 1 \qquad \boxed{\sum_{i=1}^{n} \sum_{j=1}^{m} p_{ij} = 1}$$

Jointly Distributed Random Variables

Definition (Joint probability density function)

Let $\mathcal{F}_{XY}(x, y)$ be the joint cumulative distribution function of a double continuous random variable which admits second order partial derivatives. We call *joint probability density function* the function defined as follows

with two variables

$$f_{XY}(x, y) = \frac{\partial^2 \mathcal{F}_{XY}(x, y)}{\partial x \partial y}$$

$$\boxed{f_X = \frac{dF_X}{dX}}$$

it is seen that

$$\mathcal{P}(x < X \leqslant x + dx,\ y < Y \leqslant y + dy) = f_{XY}(x, y)\ dx\, dy$$

- **1-D random variables**: The integral of $f_X(x)$ over an interval $I \in \mathbb{R}$ gives the probability that X takes on values in I
- **2-D random variables**: The integral of $f_{XY}(x, y)$ over a domain $\mathcal{D} \in \mathbb{R}^2$ gives the probability that (X, Y) takes on values in $\mathcal{D}$

Jointly Distributed Random Variables

About the **joint probability density/cumulative distribution functions**

- Given $\mathcal{D} \in \mathbb{R}^2$

$$\mathcal{P}\left((X, Y) \in \mathcal{D}\right) = \int_{\mathcal{D}} f_{XY}\left(\xi, \eta\right)\, d\xi\, d\eta$$

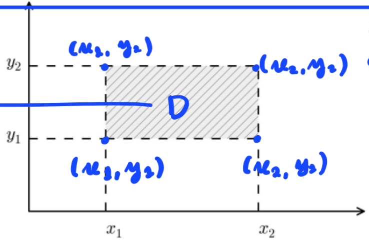
in one dimentional case:
$P(x_1 \le X \le x_2)$
$= F(x_2) - F(x_1)$
$= \int_{x_1}^{x_2} f_X(x)\, dx$

- If $\mathcal{D} = [x_1,\, x_2] \times [y_1,\, y_2]$

$$D = \mathcal{P}\left(x_1 < X \leqslant x_2,\; y_1 < Y \leqslant y_2\right) = \int_{x_1}^{x_2} \int_{y_1}^{y_2} f_{XY}\left(\xi, \eta\right)\, d\xi\, d\eta$$

Joint density function

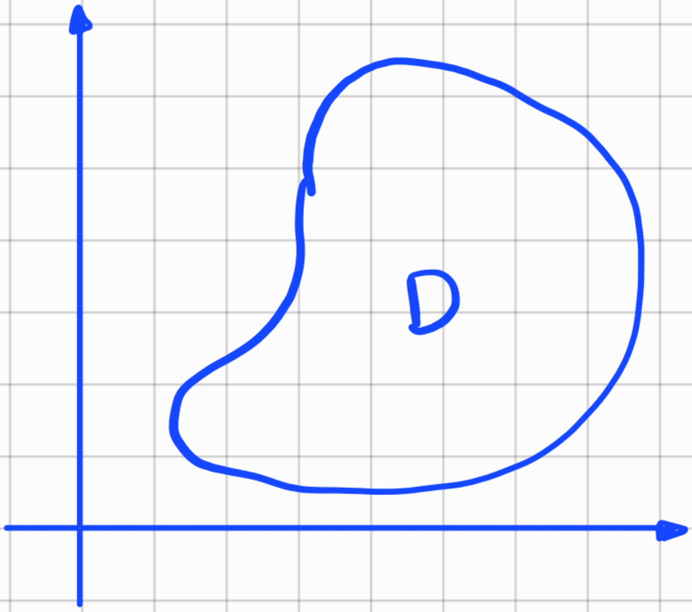
$$= \mathcal{F}_{XY}\left(x_2, y_2\right) - \mathcal{F}_{XY}\left(x_1, y_2\right) - \left(\mathcal{F}_{XY}\left(x_2, y_1\right) - \mathcal{F}_{XY}\left(x_1, y_1\right)\right)$$



interval in two dimentional

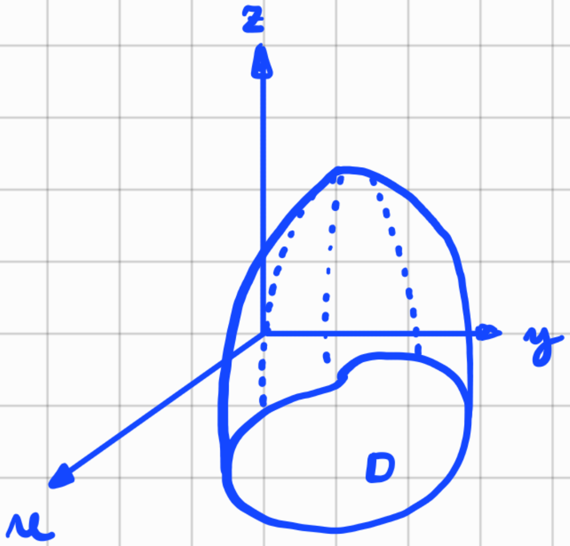
- When $x_1, y_1 \to -\infty$ we recover $\mathcal{P}\left(X \leqslant x_2,\; Y \leqslant y_2\right) = \mathcal{F}_{XY}\left(x_2, y_2\right)$

In general case for two-dimentional case:



$$\boxed{P((X,Y)\in D) = \int_D f_{XY}(\xi,\eta)\,d\xi\,d\eta}$$

Or even in three dimentional:



Jointly Distributed Random Variables

Definition (Marginal cumulative distribution function)

We call *marginal cumulative distribution functions* the functions defined as follows

if we take two independent variables, and let one of the variable goes to infinite

$$\mathcal{F}_X(x) = \mathcal{F}_{XY}(x, \infty) \qquad \mathcal{F}_Y(y) = \mathcal{F}_{XY}(\infty, y)$$

$\lim_{y \to \infty} (x, y) = F_X(x)$

which coincide with the cumulative distribution functions of the components of the random variable

For a continuous double random variable

$$\boxed{\begin{aligned} \mathcal{F}_X(x) = \mathcal{F}_{XY}(x, \infty) &= \int_{-\infty}^{x} \int_{-\infty}^{\infty} f_{XY}(\xi, \eta)\, d\xi\, d\eta \\ \mathcal{F}_Y(y) = \mathcal{F}_{XY}(\infty, y) &= \int_{-\infty}^{\infty} \int_{-\infty}^{y} f_{XY}(\xi, \eta)\, d\xi\, d\eta \end{aligned}}$$

Jointly Distributed Random Variables

Definition (Marginal probability mass/density function)

Given a discrete double random variable, we call *marginal probability mass function* the probability defined as follows

$$\boxed{\begin{aligned} p(x_i) &= \mathcal{P}(X = x_i) = \sum_{j=1}^{m} p(x_i, y_j) \\ p(y_j) &= \mathcal{P}(Y = y_j) = \sum_{i=1}^{n} p(x_i, y_j) \end{aligned}}$$

For a continuous double random variable, the *marginal probability density function* is defined as follows

$$\boxed{f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, \eta)\, d\eta \qquad f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(\xi, y)\, d\xi}$$

Jointly Distributed Random Variables

The notion of **independence** is crucial in the theory of probability

Theorem (Independence)

Two random variables X and Y are independent if for all x, y

$$\boxed{\begin{array}{c} f_{XY}(x,y) = f_X(x)\, f_Y(y) \\ \mathcal{F}_{XY}(x,y) = \mathcal{F}_X(x)\, \mathcal{F}_Y(y) \end{array}}$$

From the previous theorem it follows that, when the components X and Y are independent

can be written as a product of two components

$$\begin{aligned} \mathcal{P}(x_1 < X \leqslant x_2,\, y_1 < Y \leqslant y_2) &= \mathcal{P}(x_1 < X \leqslant x_2)\, \mathcal{P}(y_1 < Y \leqslant y_2) \\ &= [\mathcal{F}_X(x_2) - \mathcal{F}_X(x_1)]\,[\mathcal{F}_Y(y_2) - \mathcal{F}_Y(y_1)] \end{aligned}$$

$$\iint f_{XY}(x,y)\, dx = \iint f_X(x) f_Y(y)\, dx\, dy = \int f_X(x)\, dx \int f_Y(y)\, dy$$

Jointly Distributed Random Variables

Example (1)

The joint cumulative distribution function of a double random variable (X, Y) is $(\alpha, \beta > 0)$

$$\mathcal{F}_{XY}(x, y) = \begin{cases} \left(1 - e^{-\alpha x}\right)\left(1 - e^{-\beta y}\right) & x, y \geqslant 0 \\ 0 & \text{elsewhere} \end{cases}$$

1. find the marginal cumulative distribution functions of X and Y
2. determine whether X and Y are independent or not

1. The marginal cumulative distribution functions are

$$\mathcal{F}_X(x) = \mathcal{F}_{XY}(x, \infty) = 1 - e^{-\alpha x} \quad (x \geqslant 0)$$
$$\mathcal{F}_Y(y) = \mathcal{F}_{XY}(\infty, y) = 1 - e^{-\beta y} \quad (y \geqslant 0)$$

2. $\mathcal{F}_{XY}(x, y) = \mathcal{F}_X(x)\, \mathcal{F}_Y(y) \quad \Rightarrow$ the components are independent

Example (1)

The joint cumulative distribution function of a double random variable (X, Y) is $(\alpha, \beta > 0)$

$$\mathcal{F}_{XY}(x, y) = \begin{cases} \left(1 - e^{-\alpha x}\right)\left(1 - e^{-\beta y}\right) & x, y \geqslant 0 \\ 0 & \text{elsewhere} \end{cases}$$

1. find the marginal cumulative distribution functions of X and Y
2. determine whether X and Y are independent or not

1) The marginal cum. functions are:

$$F_X(x) = \lim_{y \to \infty} (1 - e^{-\alpha x})(1 - e^{-\beta y}) = 1 - e^{-\alpha x}$$

$$F_Y(y) = \lim_{x \to \infty} (1 - e^{-\alpha x})(1 - e^{-\beta y}) = 1 - e^{-\beta y}$$

Jointly Distributed Random Variables

Numerical characteristics of double random variables

Definition (Initial moment of order (k, s))

Given a double random variable (X, Y), we call *initial moment of order* (k, s), denoted by the symbol $\alpha_{k,s}$, the expected value of the product of the power k and s of the components X and Y

$$\alpha_{k,s} = E\left(X^k Y^s\right)$$

$\alpha_k = \sum x_i^k p_i$

generalization from one dimentional case

- discrete case

$$\boxed{\alpha_{k,s} = \sum_i \sum_j x_i^k y_j^s p_{ij}}$$

- continuous case

$\alpha_{1,0} = \iint x f_{XY}(x,y)\, dx dy$

$\alpha_{0,1} = \iint y f_{XY}(x,y)\, dx dy$

$$\boxed{\alpha_{k,s} = \int_{\mathbb{R}^2} x^k y^s f_{XY}(x, y)\ dxdy}$$

Jointly Distributed Random Variables

The **initial moments of the first order** $\alpha_{1,0}$ and $\alpha_{0,1}$ correspond to the expected value of the components X and Y

$$\alpha_{1,0} \equiv E\left(X^1 Y^0\right) = E\left(X\right) \equiv \mu_X \qquad \alpha_{0,1} \equiv E\left(X^0 Y^1\right) = E\left(Y\right) \equiv \mu_Y$$

The expected values μ_X and μ_Y represent the coordinates on the Cartesian plane of the point around which the random points (X, Y) are *dispersed*

Definition (Centered random variables)

Given a double random variable $\boxed{(X, Y)}$ with expected values (μ_X, μ_Y), we call *centered double random variable* the random variable $\boxed{(X_c, Y_c)}$ whose components are defined as follows

$$\boxed{X_c = X - \mu_X} \qquad \boxed{Y_c = Y - \mu_Y}$$

Definition (Centered moment of order (k, s))

Given a double random variable (X, Y), we call *centered moment of order* (k, s), and we denote it with the symbol $\mu_{k,s}$, the expected value of the *product* of the powers k and s, respectively, of the centered components X_c and Y_c

$$\mu_{k,s} = E\left(X_c^k Y_c^s\right)$$

$\mu_k = \sum (x_i - \mu_x)^k p_i$

- discrete case

$k, s \geq 0$

$$\boxed{\mu_{k,s} = \sum_i \sum_j (x_i - \mu_X)^k (y_j - \mu_Y)^s p_{ij}}$$

$\mu_{2,0}$, $\mu_{0,2}$, $\mu_{1,1}$ ← three possibilities of moment order of 2

- continuous case

$\mu_{2,0} = \sigma_x^2$

$\mu_{0,2} = \sigma_y^2$

$$\boxed{\mu_{k,s} = \int_{\mathbb{R}^2} (x - \mu_X)^k (y - \mu_Y)^s f_{XY}(x, y)\, dxdy}$$

The **centered moments of the second order** $\underline{\mu_{2,0} \text{ and } \mu_{0,2}}$ correspond to the variances of the components X and Y

$$\mu_{2,0} = E\left(X_c^2 Y_c^0\right) = E\left(X_c^2\right) = E\left(\left(X - \mu_X\right)^2\right) \equiv \boxed{\sigma_X^2 \equiv Var\left(X\right)}$$
$$\mu_{0,2} = E\left(X_c^0 Y_c^2\right) = E\left(Y_c^2\right) = E\left(\left(Y - \mu_Y\right)^2\right) \equiv \boxed{\sigma_Y^2 \equiv Var\left(Y\right)}$$

σ_X^2 and σ_Y^2 represent the dispersion of the random points in the direction of the $x-$ and $y-$axis, respectively

Jointly Distributed Random Variables

Definition (Covariance)

Given a double random variable (X, Y), we call *covariance*, denoted by the symbol σ_{XY} or $Cov(XY)$, the centered moment $\mu_{1,1}$

- discrete case

$$\boxed{\sigma_{XY} \equiv Cov(XY) = \sum_i \sum_j (x_i - \mu_X)(y_j - \mu_Y)\, p_{ij}}$$

- continuous case

$$\boxed{\sigma_{XY} \equiv Cov(XY) = \int_{\mathbb{R}^2} (x - \mu_X)(y - \mu_Y)\, f_{XY}(x, y)\; dx\, dy}$$

Note that (this formula will be useful later)

$$\boxed{\sigma_{XY} = E((X - \mu_X)(Y - \mu_Y)) = E(XY) - \mu_X \mu_Y}$$

Jointly Distributed Random Variables

Theorem (Covariance of independent random variables)

The covariance of independent random variables is zero

Proof. (continuous case) Given X and Y (independent)

$$f_{XY}(x,y) = f_X(x)\ f_Y(y)$$

and

$$\begin{aligned}
\sigma_{XY} &= \boxed{\int_{\mathbb{R}^2} (x-\mu_X)(y-\mu_Y)\, f_{XY}(x,y)\ dxdy} \\
&= \int_{-\infty}^{\infty} (x-\mu_X)\, f_X(x)\ dx \int_{-\infty}^{\infty} (y-\mu_Y)\, f_Y(y)\ dy \\
&= E(X_c)\, E(Y_c) = 0
\end{aligned}$$

□

Important! The covariance of independent variables is zero, but the converse is not generally true, i.e. **If the covariance of two random variables is zero, they are not necessarily independent**

Jointly Distributed Random Variables

The covariance σ_{XY} describes how X and Y are *correlated*. Sounds useful, but...

- it is a dimensional quantity (its dimensions are given by the product of the dimensions of the components X and Y)
- it depends on the unit of measures $\Rightarrow$ σ_{XY} is difficult to interpret

Definition (Correlation coefficient)

Given two random variables X and Y, the *correlation coefficient* of X and Y is defined as follows this value is non-dimensional

$$\boxed{r_{XY} = \frac{\sigma_{XY}}{\sqrt{\sigma_X^2 \sigma_Y^2}} \equiv \frac{Cov\,(XY)}{\sqrt{Var\,(X)\,Var\,(Y)}}} = \frac{\sigma_{xy}}{\sigma_x\,\sigma_y}$$

$-1 \leq r_{xy} \leq 1$

- The correlation coefficient r_{XY} is non-dimensional and it does not depend on the unit of measures
- **independence** and **uncorrelatedness** are not the same thing!

Jointly Distributed Random Variables

Some remarks about the correlation coefficient

- r_{XY} is non-dimensional (and independent from the unit of measures)
- it is easily seen that $-1 \leqslant r_{XY} \leqslant 1$
- r_{XY} describes the "intensity" of the **linear relation** between X and Y
  - if $r_{XY} > 0$, X and Y are **positively correlated**
  - if $r_{XY} = 0$, X and Y are **not correlated**
  - if $r_{XY} < 0$, X and Y are **negatively correlated**

More precisely

- **Positive correlation** means that as one variable increases, the other variable *tends to* increase
- If $r_{XY} = \pm 1$, X and Y have an exact linear dependency

in general the taller of the person, it's heavier.

$$\boxed{\begin{array}{lcll} r_{XY} = +1 & \Leftrightarrow & Y = +a\,X + b & (a > 0) \\ r_{XY} = -1 & \Leftrightarrow & Y = -a\,X + b & (a > 0) \end{array}}$$

Example of positive correlation: "weight" and "height" of a person

Warning: **correlation does not imply causality**

Jointly Distributed Random Variables

Most of the definitions seen so far are generalizable to the case of **multidimensional random variables with more than two components**

Definition (Cumulative distribution function)

Given a n–dimensional random variable $(X_1, X_2, \dots X_n)$, its *cumulative distribution function* is defined as follows

$$\boxed{\mathcal{F}_{X_1 X_2 \dots X_n}(x_1, x_2, \dots x_n) = \mathcal{P}\left(X_1 \leqslant x_1, X_2 \leqslant x_2, \dots, X_n \leqslant x_n\right)}$$

Definition (Probability density function)

Given a n–dimensional random variable $(X_1, X_2, \dots X_n)$, its *probability density function* is defined as the following mixed partial derivative

$$f_{X_1 X_2 \dots X_n}(x_1, x_2, \dots x_n) = \frac{\partial^n \mathcal{F}_{X_1 X_2 \dots X_n}(x_1, x_2, \dots x_n)}{\partial x_1 \partial x_2 \dots \partial x_n}$$

Jointly Distributed Random Variables

Definition (Marginal CDF/PDF)

Given a n–dimensional random variable $(X_1, X_2, \dots X_n)$, its *marginal cumulative distribution/probability density functions* are defined as follows

$$\mathcal{F}_X(x_k) = \mathcal{F}_{X_1 X_2 \dots X_n}(\infty, \dots, \infty, x_k, \infty, \dots, \infty)$$

$$\boxed{f_X(x_k) = \int_{\mathbb{R}^{n-1}} f_{X_1 X_2 \dots X_n}(x_1, x_2, \dots x_n)\; dx_1 \; \dots dx_{k-1}\, dx_{k+1} \dots dx_n}$$

Also the notion of **independence** is easily extended

Theorem (Independence)

The random variables X_1, X_2, $\dots X_n$ are independent if

$$f_{X_1 X_2 \dots X_n}(x_1, x_2, \dots x_n) = f_{X_1}(x_1)\, f_{X_2}(x_2) \dots f_{X_n}(x_n)$$

or, equivalently,

$$\mathcal{F}_{X_1 X_2 \dots X_n}(x_1, x_2, \dots x_n) = \mathcal{F}_{X_1}(x_1)\, \mathcal{F}_{X_2}(x_2) \dots \mathcal{F}_{X_n}(x_n)$$

The main **statistics** of a n-dimensional random variable $(X_1, X_2, \ldots X_n)$ are

how many orders we can make from n-dimension variables?

- the n **expected values** $E_k \equiv E(X_k)$, for $k = 1, \ldots n$
- the n **variances** $\sigma_k^2 \equiv \sigma_{X_k}^2 \equiv Var(X_k)$, for $k = 1, \ldots n$
- the $n(n-1)$ **covariances** $\sigma_{ij} \equiv \sigma_{X_i X_j} = Cov(X_i X_j)$, for $i, j = 1, \ldots n$ with $i \neq j$

$\sigma_{ij} = \sigma_{ji}$

Note that the **variance** $\sigma_k^2 \equiv \sigma_{X_k}^2 \equiv Var(X_k)$ of the k^{th} random variable X_k is nothing but the **covariance of the random variable X_k with itself**

$$\sigma_{kk} = E\left(\left(X_k - \mu_{X_k}\right)\left(X_k - \mu_{X_k}\right)\right) = E\left(\left(X_k - \mu_{X_k}\right)^2\right) = \sigma_k^2$$

Jointly Distributed Random Variables

The **matrix of the covariances** is defined as follows

$$\|\sigma_{ij}\| = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_{nn} \end{pmatrix}$$

$\sigma_{11} = \sigma_{x_1}^2$

- The matrix is symmetric: $\boxed{\sigma_{ij} = \sigma_{ji}}$
- The elements on the main diagonal are the variances $\underline{\sigma_{kk} \equiv \sigma_k^2}$
- If the random variables are **independent** the matrix is **diagonal**

The **matrix of correlation** is similarly defined

$r_{ii} = \frac{\sigma_{ii}}{\sigma_i \sigma_i} = \frac{\sigma_{ii}}{\sigma_{ii}} = 1$

$$\boxed{r_{ij} = \frac{\sigma_{ij}}{\sigma_i \sigma_j}}$$

- All the elements on the main diagonal are equal to 1