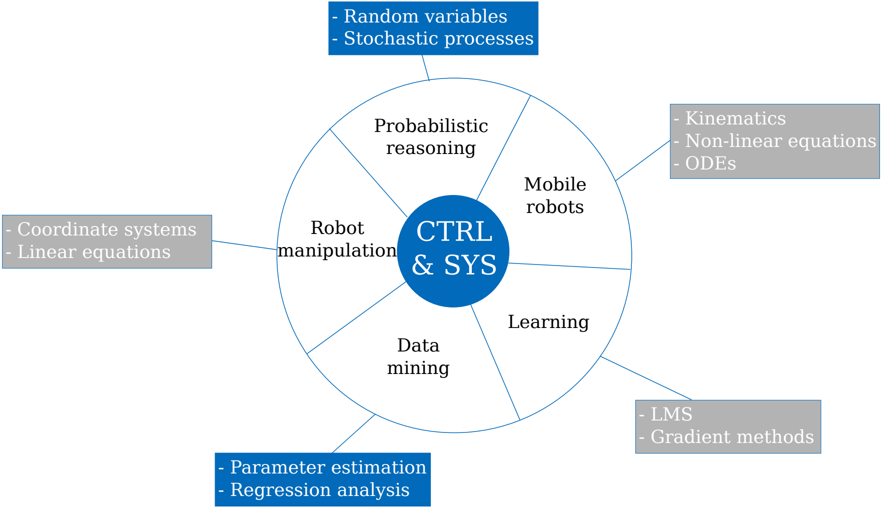
Sensor Models



Today's Topics

Bayes' theorem and sensor models

Random variables

- Discrete and continuous random variables
- Probability distributions and densities
- Expected value, variance
- Relation to measurement processes

$\rightarrow$ Statistical independence

- Joint distributions
- Correlation and covariance

Good to recall

- what a random variable is
- what a density (PDF) and a cumulative distribution function (CDF) is
- how you can compute the expected value and the variance of a random variable / distribution
- how you can approximate the expected value and the variance via the empirical mean and the empirical variance
- nice website: https://analystprep.com/study-notes/frm/part-1/quantitative-analysis/multivariate-random-variables/

Bayes' theorem and sensor models

Random variables

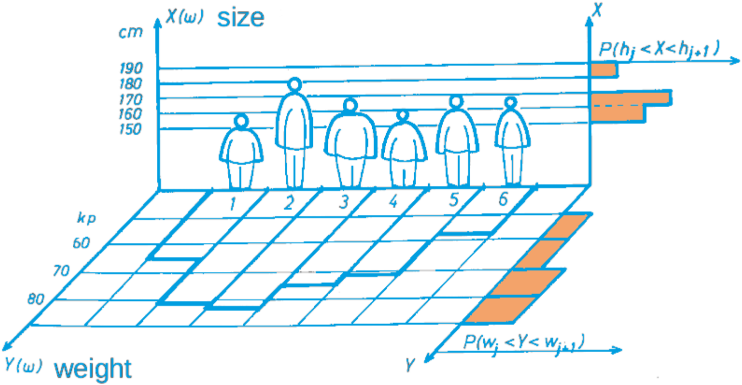
- Discrete and continuous random variables
- Probability distributions and densities
- Expected value, variance
- Relation to measurement processes

$\rightarrow$ Statistical independence

- $\rightarrow$ Joint distributions
- $\rightarrow$ Correlation and covariance
- $\rightarrow$ Examples leading towards Kalman filter

Statistical Independence

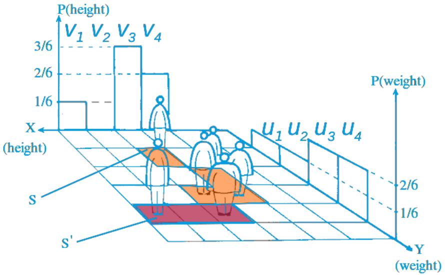
Random variables often depend on one another (e.g. weight and height)



Joint distribution example (H. Haken, Synergetik, Springer Verlag)

To specify such dependencies, we need to define (in)dependent probabilities

Joint Distributions



(H. Haken, Synergetik, Springer Verlag)

Joint probability

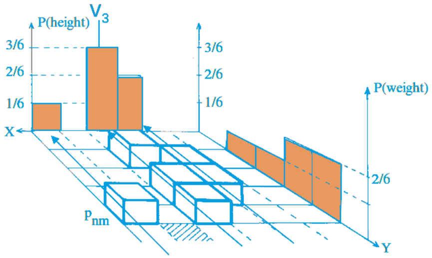
$$P_{\text{joint}}(X = u_m, Y = v_n) = p_{nm}$$

where $X(\omega) = u_m$ and $Y(\omega) = v_n$

$S = \{(u, v) | u \in X(\omega), v \in Y(\omega)\}$ is a subset of R^2

$S' \subseteq S : P_{\text{joint}}((X, Y) \in S') = P_{\text{joint}}(\{\omega | (X(\omega), Y(\omega)) \in S'\})$

Marginalisation



(H. Haken, Synergetik, Springer Verlag)

$$P\left(X = v_m\right) = \sum_n P_{joint}\left(X = v_m, Y = u_n\right)$$
$$= \sum_n p_{mn}$$

The joint probability distribution P_{joint} lives on top of the product of the ranges of the two random variables X and Y

How do we get back the one-dimensional probability distributions P_{height} and P_{weight}?

First, fix one height $X = v_m$, then build $P_{\text{height}}(X = v_m)$ by summation regardless of the value of Y; then repeat this process for all v_m

Marginalisation

The reduction of the joint distribution to a lower dimensional one via a summation process is called **marginalisation**

Joint CDF and PDF

Joint CDF

$$F_2(x, y) = P_{\text{joint}}(X \leq x \text{ and } Y \leq y)$$

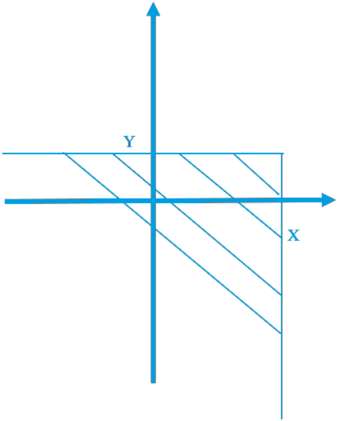
Joint PDF

$$f_2(x, y) = \frac{\partial^2 F_2(x, y)}{\partial x \partial y}$$

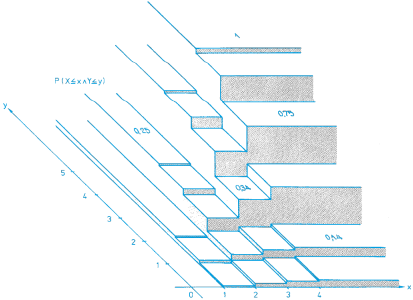
Marginalisation

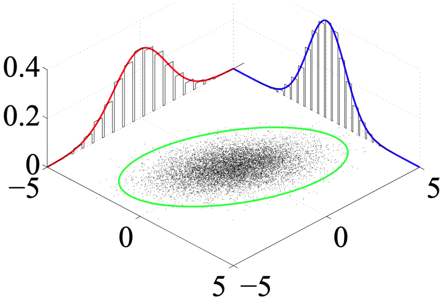
$$f_X(x) = \int_{-\infty}^{\infty} f_2(x, y) dy$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_2(x, y) dx$$



Joint CDF and PDF





$$
\begin{aligned}
F_2(x,y) &= \\
P(X \le x, Y \le y) &= \\
\sum_{\substack{t_1 \in W(X)\\ t_1 \le x}} \sum_{\substack{t_2 \in W(Y)\\ t_2 \le y}} f_2(t_1,t_2)
\end{aligned}
$$

Independence

Independence

A random variable X is independent of a random variable Y if and only if

$$\begin{aligned} F_2(x,y) &= P_{\text{joint}}\left(X \leq x \text{ and } Y \leq y\right) \\ &= P_{\text{joint}}\left(X \leq x, Y \leq y\right) \\ &= P_X(X \leq x) P_Y(Y \leq y) \\ &= F_X(x) F_Y(y) \\ \implies & f_2(x,y) = f_X(x) f_Y(y) \end{aligned}$$

In the case of independent variables X and Y, the joint CDF F_2 **factors** into two separate CDFs: F_X and F_Y.
If X, Y independent, discrete $\implies$ the joint $P_{joint}(X,Y)$ table becomes a **multiplication table**.

Independence Example 1

	a)	b)	c)
$P(X=0,Y=0)$	$\frac{1}{4}$	$\frac{1}{2}$	1
$P(X=1,Y=0)$	$\frac{1}{4}$	0	0
$P(X=0,Y=1)$	$\frac{1}{4}$	0	0
$P(X=1,Y=1)$	$\frac{1}{4}$	$\frac{1}{2}$	0

Independence Example 1

	a)	b)	c)
$P(X=0,Y=0)$	$\frac{1}{4}$	$\frac{1}{2}$	1
$P(X=1,Y=0)$	$\frac{1}{4}$	0	0
$P(X=0,Y=1)$	$\frac{1}{4}$	0	0
$P(X=1,Y=1)$	$\frac{1}{4}$	$\frac{1}{2}$	0

a) The random variables are independent (e.g. tossing two coins)

Independence Example 1

	a)	b)	c)
$P(X=0, Y=0)$	$\frac{1}{4}$	$\frac{1}{2}$	1
$P(X=1, Y=0)$	$\frac{1}{4}$	0	0
$P(X=0, Y=1)$	$\frac{1}{4}$	0	0
$P(X=1, Y=1)$	$\frac{1}{4}$	$\frac{1}{2}$	0

a) The random variables are independent (e.g. tossing two coins)

b) The variables aren't independent; they would be if the joint probabilities were products of the marginal probabilities

Independence Example 1

	a)	b)	c)
$P(X=0, Y=0)$	$\frac{1}{4}$	$\frac{1}{2}$	1
$P(X=1, Y=0)$	$\frac{1}{4}$	0	0
$P(X=0, Y=1)$	$\frac{1}{4}$	0	0
$P(X=1, Y=1)$	$\frac{1}{4}$	$\frac{1}{2}$	0

a) The random variables are independent (e.g. tossing two coins)

b) The variables aren't independent; they would be if the joint probabilities were products of the marginal probabilities

c) The variables are independent

Independence Example 2

Throwing a single die, two random variables:

$$X(w) = 1 \text{ if even, } -1 \text{ otherwise}$$

1	2	3	4	5	6
-1	1	-1	1	-1	1

$$Y(w) = 1 \text{ if prime, } -1 \text{ otherwise}$$

1	2	3	4	5	6
-1	1	1	-1	1	-1

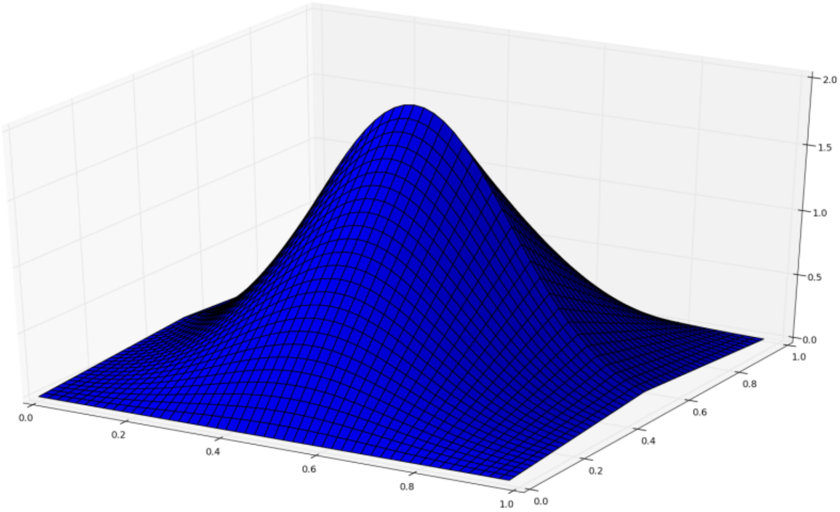
$P(X = -1, Y = -1) = \frac{1}{6}$

$P(X = -1)P(Y = -1) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$

Independence Example 3

$X \backslash Y$	-1	0	1	Σ
-1	$\frac{1}{45}$	$\frac{2}{45}$	$\frac{3}{45}$	$\frac{6}{45}$
0	$\frac{4}{45}$	$\frac{5}{45}$	$\frac{6}{45}$	$\frac{15}{45}$
1	$\frac{7}{45}$	$\frac{8}{45}$	$\frac{9}{45}$	$\frac{24}{45}$
Σ	$\frac{12}{45}$	$\frac{15}{45}$	$\frac{18}{45}$	1

Independence is a Very Strong Attribute



$$\frac{1}{0.2\sqrt{2\pi}} \exp\left(-0.5 \cdot \left(\frac{x-0.5}{0.2}\right)^2\right)(1 - 2\mathsf{abs}(y-0.5))$$

Check: Independence and Expectation

$$
\begin{aligned}
E[XY] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x \cdot y) f_2(x,y)dxdy \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x \cdot y) f_X(x) f_Y(y)dxdy \\
&= \int_{-\infty}^{\infty} x f_X(x)dx \int_{-\infty}^{\infty} y f_Y(y)dy \\
&= E[X]E[Y]
\end{aligned}
$$

Weaker notion: Correlation of Two Random Variables

How to express that two random variables X and Y are **not** independent?

Observe that $E[XY] = E[X]E[Y]$ in case of independent variables

So, try $E[XY] - E[X]E[Y]$ as a measure

Yet: this depends on absolute sizes of $E[X]$ and $E[Y]$ $\implies$ normalise

Correlation coefficient

$$\rho = \frac{E[XY] - E[X]E[Y]}{\sigma_x \sigma_y}$$

Numerator of Correlation is: Covariance

Covariance

$$\mathrm{cov}(X,Y) := E[(X-E[X])\,(Y-E[Y])]$$

$$\begin{aligned}
\mathrm{cov}(X,Y) &= E[(X-E[X])\,(Y-E[Y])] \\
&= \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} (X-E[X])\,(Y-E[Y])\, f_2(x,y)dy \\
&= \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} (xy - yE[X] - xE[Y] + E[X]E[Y])\, f_2(x,y)dy \\
&= E[XY] - E[X] \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} y f_2(x,y)dy \\
&\quad - E[Y] \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} x f_2(x,y)dy + E[X]E[Y] \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} f_2(x,y)dy \\
&= E[XY] - E[X] \int_{-\infty}^{\infty} y f(y)dy - E[Y] \int_{-\infty}^{\infty} x f(x)dx + E[X]E[Y] \\
&= E[XY] - E[X]E[Y]
\end{aligned}$$

Illustration for ρ / Empirical Corr. Coef.

$$
\begin{aligned}
\rho \underset{defin.}{:=} & \frac{E[XY] - E[X]E[Y]}{\sigma_x \sigma_y} \\
\underset{algebra}{=} & \frac{E[(X - E[X])\,(Y - E[Y])]}{\sigma_x \sigma_y} \\
\underset{empir.}{=} & \underbrace{\frac{\sum_{i=1,\ldots,n} (x_i - \bar{x})\,(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2}}}_{m} \cdot \frac{1}{\sqrt{\sum_i (y_i - \bar{y})^2}}
\end{aligned}
$$

- ρ is a measure of the (signed) strength of the linear relationship between two quantitative variables x and y.
- $\rho \pm 1 \implies$ strong (linear correlation), $\rho \approx 0 \implies$ little or no correlation.
- last equation is also called Pearson's correlation coefficient $r_{x,y}$.
- main use case for m : $f(x) = m \cdot x + (\bar{y} - m\bar{x})$ is the best linear fit to the empirical data (x_i, y_i) in least mean square (LMS) sense.

Covariance Matrix

Given a vector of random variables $\mathbf{x} = (X_1, \ldots, X_n)^T$, we have

Covariance matrix

$$\begin{aligned} \mathrm{cov}(\mathbf{x}) &= E\left[(\mathbf{x} - E[\mathbf{x}])\,(\mathbf{x} - E[\mathbf{x}])^T\right] \\ &= \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} (\mathbf{x} - E[\mathbf{x}])\,(\mathbf{x} - E[\mathbf{x}])^T\, p(\mathbf{x}) dx_1 \ldots dx_n \end{aligned}$$

Properties of the covariance matrix

The matrix is symmetric and positive definite

The diagonal entries are the variances of the random variables, i.e. $\mathrm{cov}(\mathbf{x})_{i,i} = \sigma^2_{X_i}$

The off-diagonal entries are the covariances ($= 0$ if the RVs are independent)

Covariance Example: Motivation

Assume that a laser scanner measures polar coordinates (d, α), such that the measurements of d and α are normally distributed, i.e. $d \sim \mathcal{N}(\mu_d, \sigma_d^2)$, $\alpha \sim \mathcal{N}(\mu_\alpha, \sigma_\alpha^2)$

The measurements have to be mapped to Cartesian (x, y) coordinates via $F\left((d, \alpha)^T\right) = (d\cos(\alpha), d\sin(\alpha))^T$

How does the original covariance matrix change after the mapping? (1. linear $\Longrightarrow$ 2. general)

Linear Case: Map Expectations

In the linear case, $\mathbf{u}$ is a linear map of $\mathbf{x}$

$$\mathbf{u} = F(\mathbf{x}) = A\mathbf{x} + \mathbf{b}$$

$$\begin{aligned} \implies \mu_{\mathbf{u}} &= E[\mathbf{u}] = E[A\mathbf{x} + \mathbf{b}] \\ &= \int\int\int \dots \int (A\mathbf{x} + \mathbf{b})\, p(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n \end{aligned}$$

Taking component j, we have

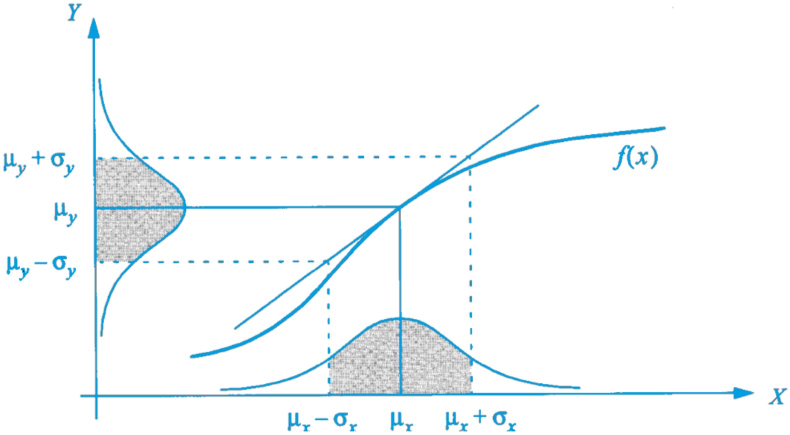
$$\begin{aligned} \mu_{uj} &= \int\int\int \dots \int \left(\sum_i a_{ij} x_i \right) p(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n + \\ &\int\int\int \dots \int b_j p(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n \\ &= \sum_i a_{ij} \int\int\int \dots \int x_i p(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n + b_j \\ &= \sum_i a_{ij} E[X_i] + b_j \end{aligned}$$

Linear Case: Map Covariances

$$
\begin{aligned}
\operatorname{cov}(\mathbf{u}) &= E\left[(A\mathbf{x} + \mathbf{b} - (AE[\mathbf{x}] + \mathbf{b}))\,(A\mathbf{x} + \mathbf{b} - (AE[\mathbf{x}] + \mathbf{b}))^T\right] \\
&= E\left[(A\mathbf{x} - AE[\mathbf{x}])(A\mathbf{x} - AE[\mathbf{x}])^T\right] \\
&= E\left[A(\mathbf{x} - E[\mathbf{x}])(\mathbf{x} - E[\mathbf{x}])^T A^T\right] \\
&= A \operatorname{cov}(\mathbf{x}) A^T
\end{aligned}
$$

Nonlinear Case: Approximate use Taylor

Use a Taylor expansion $f(x+h) = f(x_0) + hf'(x_0) + \epsilon$



General Nonlinear Case: Taylor expansion

$$\mathbf{y} = F(\mathbf{x}) = \begin{pmatrix} f_1(\mathbf{x}) \\ \vdots \\ f_n(\mathbf{x}) \end{pmatrix}$$

Via Taylor, we get

$$F(\mathbf{x}+\mathbf{h}) = F(\mathbf{x}) + \mathrm{Jac}(F)\mathbf{h} + O(\|\mathbf{h}\|)$$

where the **Jacobian** of F (def: see later MRC lecture) is given as

$$\mathrm{Jac}(F) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & & \frac{\partial f_2}{\partial x_n} \\ \vdots & & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \cdots & \frac{\partial f_n}{\partial x_n} \end{pmatrix}$$