Budapest University of Technology and Economics
Faculty of Transportation Engineering and Vehicle Engineering
Department of Control for Transportation and Vehicle Systems

Automotive environment sensors
BMEKOKAM708

Lecture 02

Olivér Törő

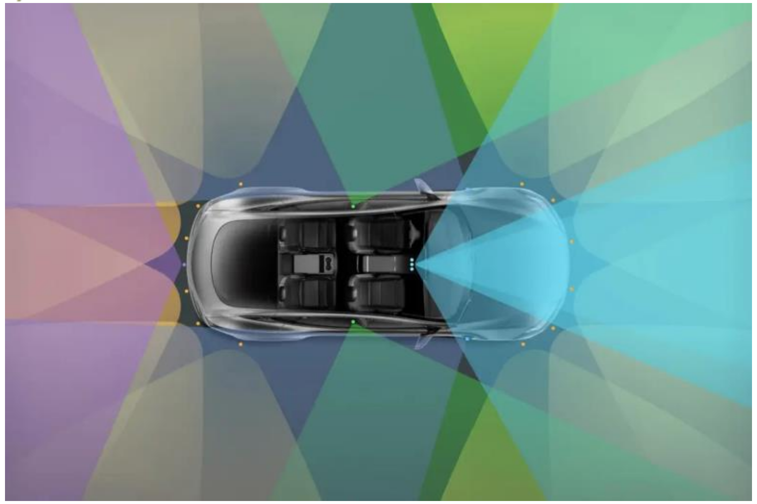
2023

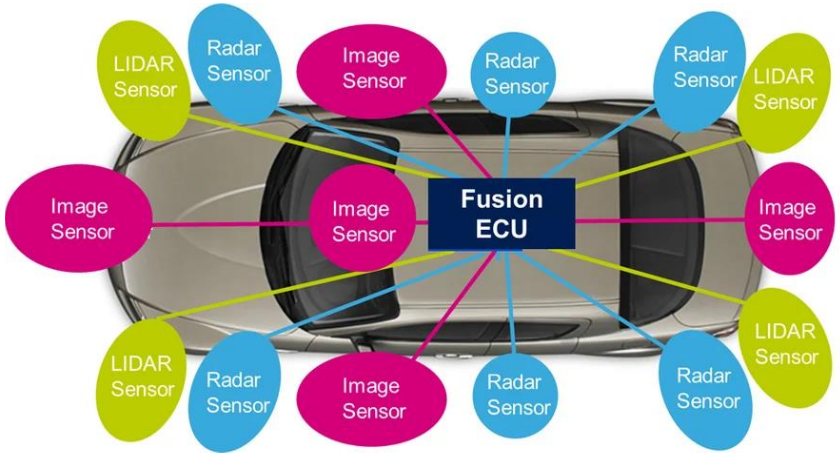
Sensor fusion / data fusion

Enhance data authenticity or availability

- Improve detection, confidence
- Extend spatial and temporal coverage
- Combine field of view
  - US sensors, radars
- Use cheaper devices
- Increase reliability
- Decrease uncertainty

- Fuse different types of data
  - Velocity + acceleration
- Create new type of data
  - Position from velocity and acceleration





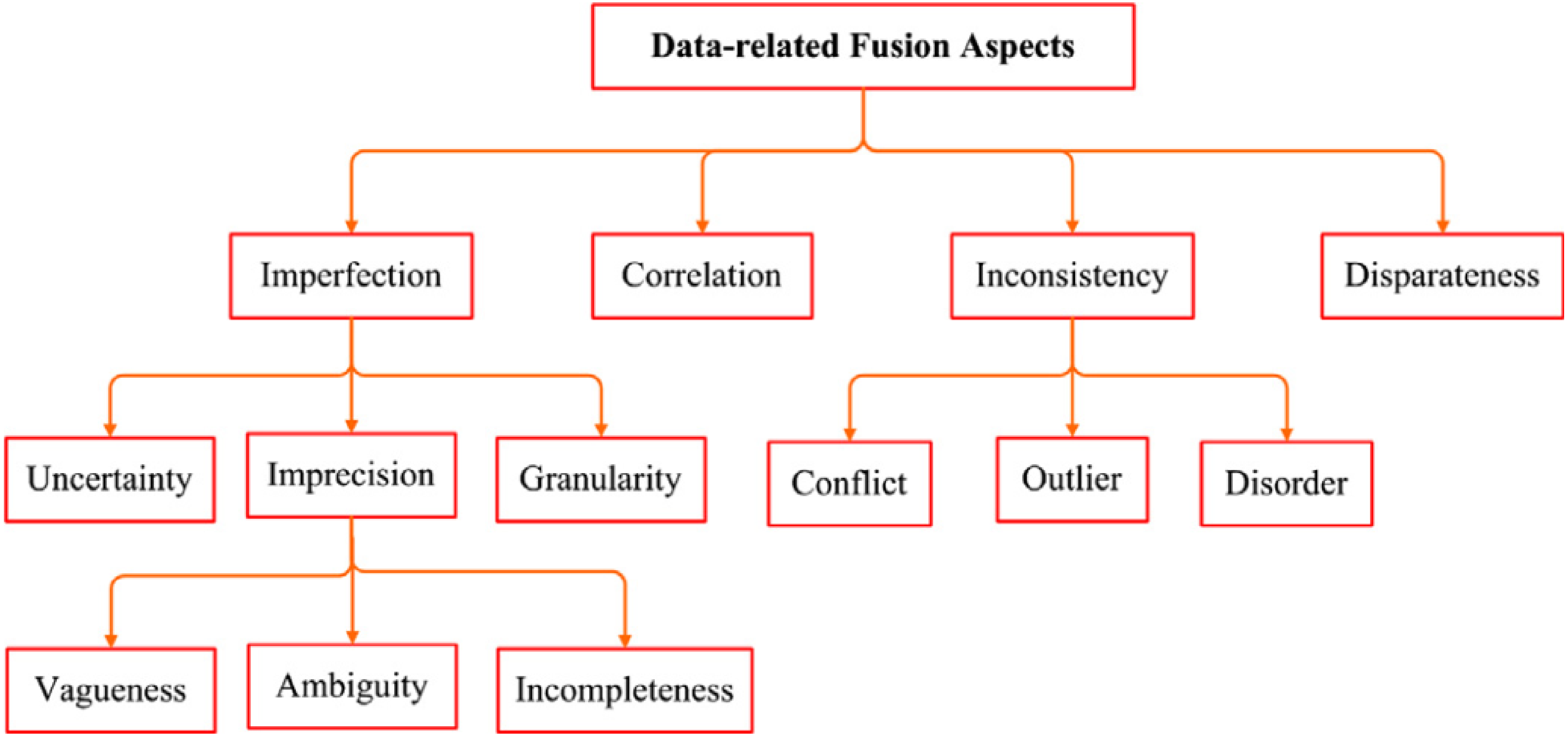
Sensor fusion / data fusion

Difficulties

- Data association
  - In multi-object setups
  - Measurement-to-track, Track-to-track
- Data alignment
  - In multi-sensor setups
  - calibration, coordinate system
- Sensor imperfection
- Sensor diversity
- Noisy environment
- Problems with data

Problems with data

- Probabilistic uncertainty
  - Noise: additive, multiplicative, etc
- Ambiguity
  - Interval measurement
- Vagueness
  - Natural language (“small”, “large”)
- Incomplete
  - Partial information (only upper limit)


Data-related Fusion Aspects
Imperfection
Correlation
Inconsistency
Disparateness
Uncertainty
Imprecision
Granularity
Conflict
Outlier
Disorder
Vagueness
Ambiguity
Incompleteness

Probabilistic data fusion

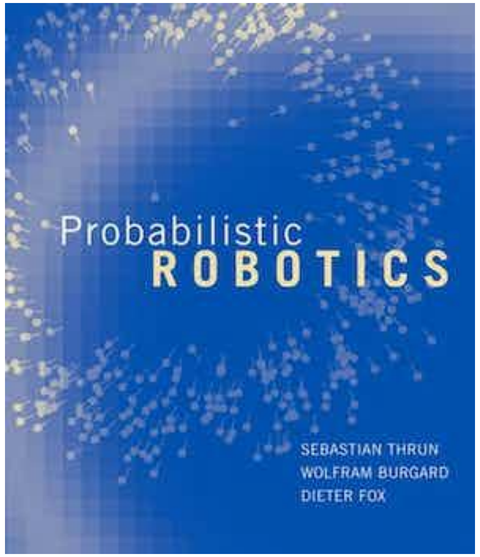
- Consider the uncertainty
- Probabilistic model
- Bayesian inference

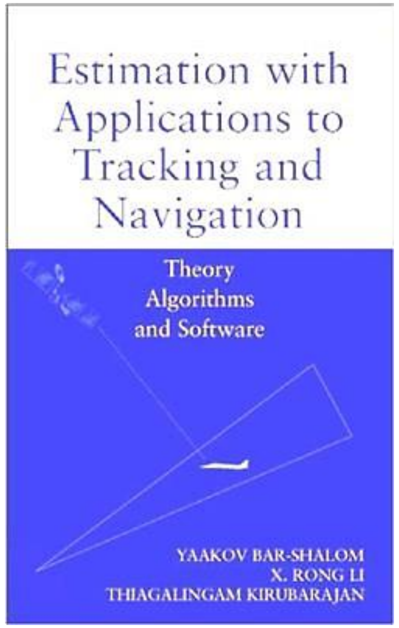
- **Model based**

- Stochastic dynamic process
- Noisy measurements

A robot that carries a notion of its own uncertainty and that acts accordingly is superior to one that does not.

Lemma 2.	Uncertainty is everywhere.
Theorem 2.	Uncertainty cannot be conquered.
Corollary 2.	Embrace it!





Uncertainty

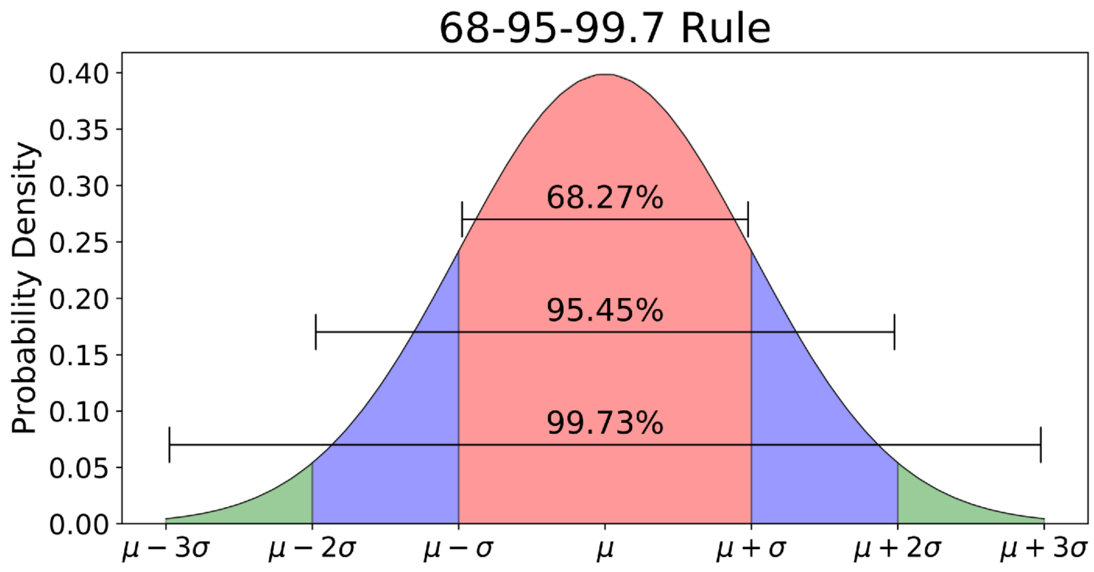
- Notion of uncertainty
- Propagation of uncertainty

- Uncertainty of prediction
- Uncertainty of estimation

- Uncertainty after fusion should be lower

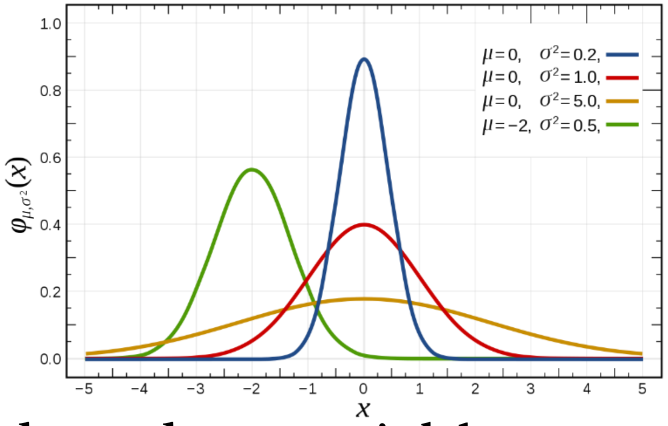
- Probabilistic model: random variable
  - Mean, variance

$$f(x;\mu,\sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$



Normal distribution

Budapest University of Technology and Economics
Faculty of Transportation Engineering and Vehicle Engineering
Department of Control for Transportation and Vehicle Systems

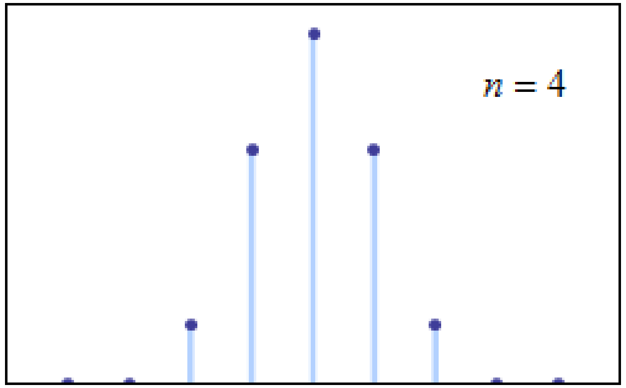


- Is the limit of the
  - Binomial distribution: $B(k;n,p) \rightarrow N(k;np,np(1-p))$
  - Poisson distribution: $P(k;\lambda) \rightarrow N(k;\lambda,\lambda)$
  - Chi-squared distribution: $\chi^2(k) \rightarrow N(k,2k)$
- Generally, the sum of independent, identically distributed random variables tends toward a normal distribution
- For a given mean and variance this is the maximum entropy distribution
  - It is the least informative distribution
  - It minimizes the information that we assume to be there
  - Physical systems generally move towards equilibrium, that is maximum entropy state
- It has nice mathematical properties

Central limit theorem


```
%% Central limit theorem
% Dice roll
n = 1e4;

R = sum(round(6*rand(n)));
histogram(R)
```



Tossing a coin n times and getting k heads

- https://phet.colorado.edu/sims/html/plinko-probability/latest/plinko-probability_hu.html

Create Gaussian variable

- Usually we have a random number generator
  - We can generate a random number in the interval 0…1
  - The standard deviation is $\frac{1}{\sqrt{12}}$
  - The mean is 0.5

Algorithm

1. Add 12 random numbers $(\mu = 6, \sigma = 1)$
2. Subtract 6 $(\mu = 0, \sigma = 1)$
3. Multiply by the desired STD
4. Add the desired mean

```
x = sum(rand(12,1e4));

x = x - 6;
x = x * 3;
x = x + 8;

histogram(x,'normalization',
'pdf')
hold on
t=(-3*sigma:0.1:3*sigma)+mu;
plot(t,normpdf(t,8,3))
```

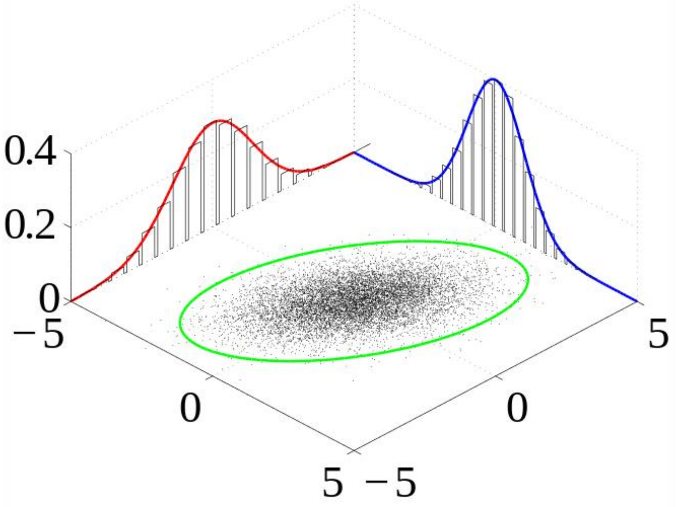
Gaussian vs White noise

- Gaussian noise and white noise are not synonyms
  - Gaussian refers the distribution of the amplitude
  - White means that the values are not correlated in time. The intensity is the same at all frequencies and the PDF can be any
- A random process can be white and Gaussian
  - This is a desired property
  - Tractable analytic models
  - Can be a good approximation of real-world situations
- Additive White Gaussian Noise (AWGN)

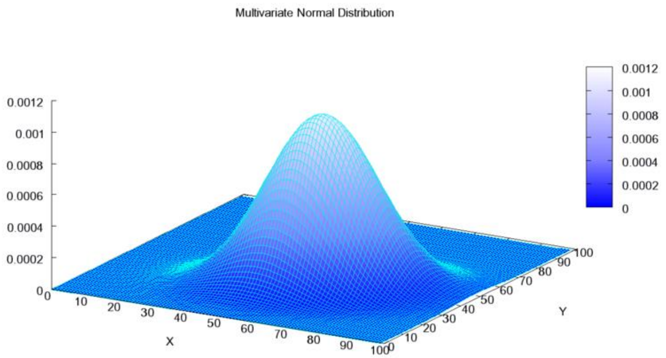
Multivariate normal distribution

Budapest University of Technology and Economics
Faculty of Transportation Engineering and Vehicle Engineering
Department of Control for Transportation and Vehicle Systems

- Joint and multivariate distributions are synonyms



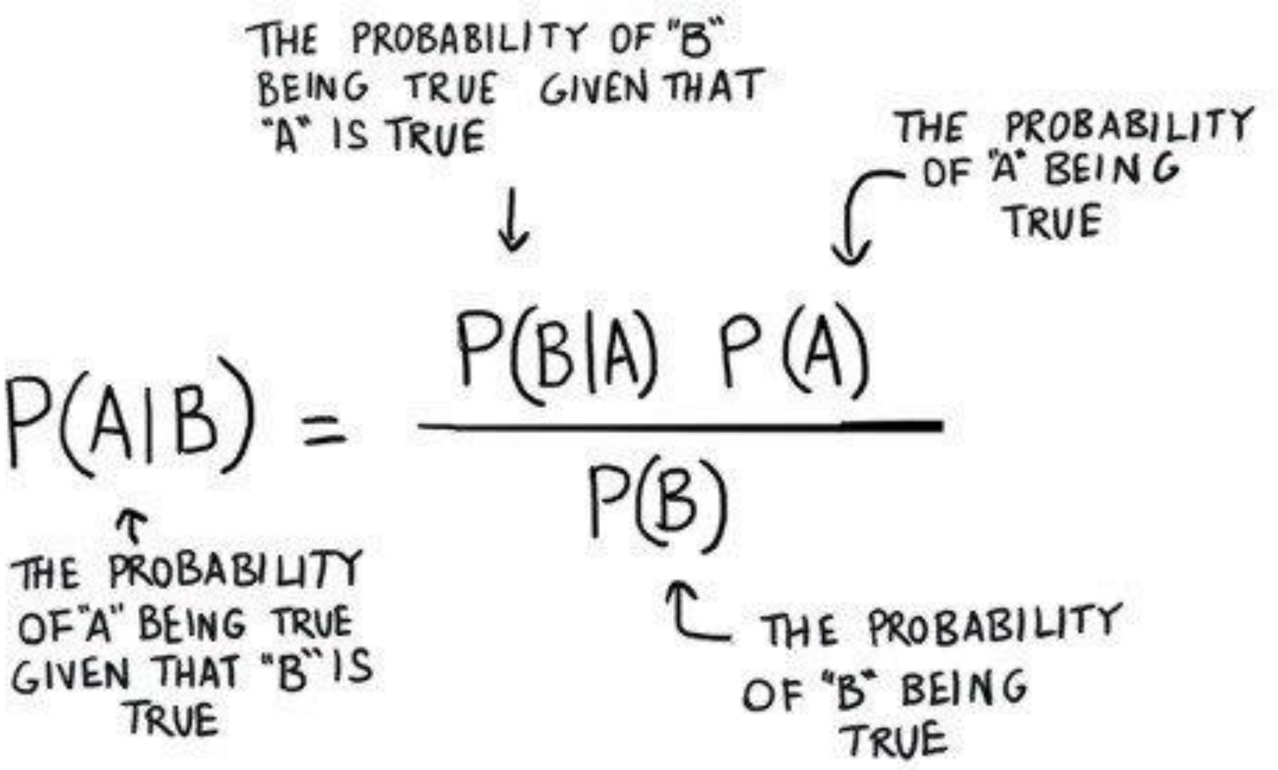
$$f(\mathbf{x}) = f(x_1, x_2, \dots, x_k) = \frac{1}{\sqrt{(2\pi)^k \det(\Sigma)}} \exp\left(-\frac{1}{2}(\mathbf{x}-\mu)^T \Sigma^{-1} (\mathbf{x}-\mu)\right)$$



Bayes' theorem
Bayesian inference

Bayes’ theorem

An Essay towards solving a Problem in the Doctrine of Chances (1763)



Bayes' theorem – examples

Applications

- COVID19 test
- Drug test
- Diagnosis
  - Humans and machines
- Genetics
  - Inheritance
  - Carrier of recessive gene

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B)}$$

Bayes' theorem – examples

Applications

- **COVID19 test**
- Drug test
- Diagnosis
  - Humans and machines
- Genetics
  - Inheritance
  - Carrier of recessive gene

Has COVID given positive test result

Positive test result given has COVID

Has COVID

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B)}$$

Positive test result

Bayes' theorem – examples

Applications

- **COVID19 test**
- Drug test
- Diagnosis
  - Humans and machines
- Genetics
  - Inheritance
  - Carrier of recessive gene

Result

Has COVID given positive test result

Measurement

Positive test result given has COVID

Initial belief

Has COVID

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B)}$$

Positive test result

True positive + False positive

Bayes' theorem – examples

Applications

- COVID19 test
- **Drug test**
- Diagnosis
  - Humans and machines
- Genetics
  - Inheritance
  - Carrier of recessive gene

Result

Uses drug given positive test result

Measurement

Positive test result given uses drug

Initial belief

Prevalence (and other factors)

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B)}$$

Positive test result

True positive + False positive

Bayes' theorem – examples

Applications

- COVID19 test
- Drug test
- **Diagnosis**
  - Humans and machines
- Genetics
  - Inheritance
  - Carrier of recessive gene

Result

Has cancer given a symptom

Measurement

Has symptom given has cancer

Initial belief

Prevalence (and other factors)

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B)}$$

Positive test result

True positive + False positive

Bayes’ theorem – examples

Applications

- COVID19 test
- Drug test
- Diagnosis
  - Humans and machines
- Genetics
  - Inheritance
  - Carrier of recessive gene

Posterior

What we are looking for

Likelihood

Information from measurement

Prior

Our initial belief

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B)}$$

Normalizing factor

Probability of the outcome

Bayes' theorem – examples

Likelihood

Posterior

Information from measurement

Prior

What we are looking for

Our initial belief

$$P(A|B) = \frac{P(B|A)\ P(A)}{P(B)}$$

Normalizing factor

Probability of the outcome

Probability of the outcome

True positive + False positive

- $P(B) = P(B|A)P(A) + P(B|\neg A)P(\neg A)$

- Law of total probability

Assigning probabilities for events

- Probability is a value in the interval $[0,1]$
- Event → probability $A \rightarrow P(A)$
- Set of all possible events: Ω $P(\Omega) = 1$

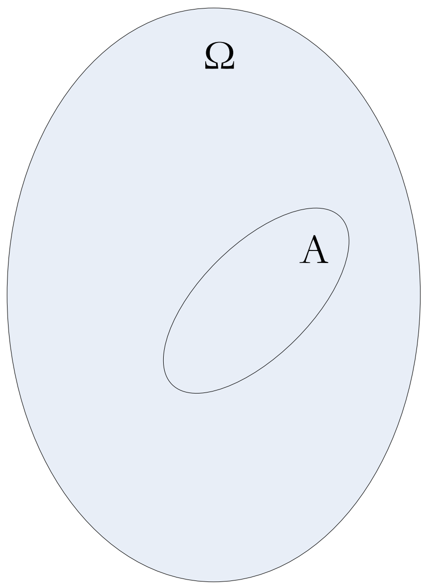
How to assign a value to $P(A)$?

- Classical interpretation of Laplace:

$$\frac{\#\, favorable\ events}{\#\, all\ possible\ events}$$ (Principle of indifference)

- Frequentist interpretation

$$P(A) = \lim_{N\to\infty} \frac{\#\, trials\ A\ occured}{\#\, total\ trials}$$ (relative frequency)



$$P(A) = \frac{area(A)}{area(\Omega)}$$

Kolmogorov axioms

- **Sample space (Ω):** set of all possible events
- **Events ($\boldsymbol{A}, \boldsymbol{B}, \ldots$):** subsets of Ω, can be elementary or complex
- Probability of an event: $P(E)$

1. $P(E) \geq 0$ for every event
2. $P(\Omega) = 1$
3. $P(\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} P(E_i)$

How to choose a specific function P is not part of the axioms

Dice Roll

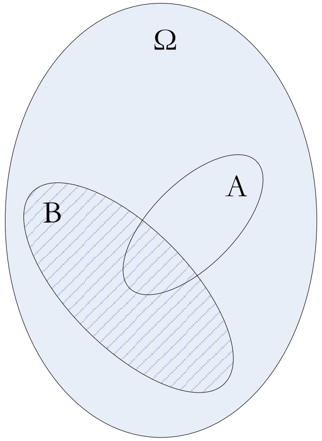
- **Sample space (Ω):** {1,2,3,4,5,6} or {even, odd}
- **Elementary events (ω):** {1,2,3,4,5,6}
- **Set of considered events ($\boldsymbol{F}$):** eg.: {Ø,1,2,3,4,5,6, even, >3, etc.}
- **Events ($\boldsymbol{A}, \boldsymbol{B}, \ldots$):** {2, even, greater than 3 and odd, 4&5, etc}
- **Probability measure $P: F \rightarrow [0,1]$:** "favorable cases/possible cases" (Laplace)
- An event has probability: e.g. $P(A), P(\neg A), P(A \cap B)$ etc.
- The triplet (Ω, F, P) defines a probability space

Conditional probability

Budapest University of Technology and Economics
Faculty of Transportation Engineering and Vehicle Engineering
Department of Control for Transportation and Vehicle Systems

- Conditional probability (definition)

$$P(A|B) := \frac{P(A \cap B)}{P(B)} \qquad P(A \cap B) = P(A|B)P(B) = P(B|A)P(A)$$

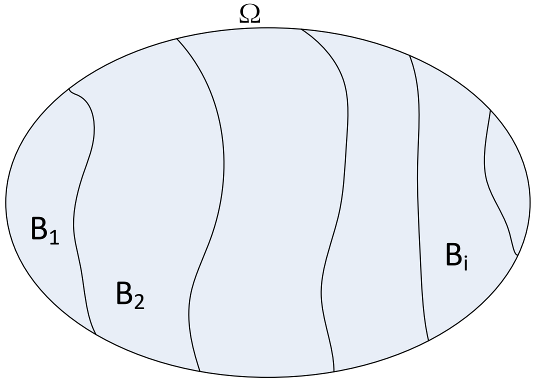


- Independent events

$$P(A|B) = P(A) \ \text{ and } \ P(B|A) = P(B)$$

$$P(A \cap B) = P(A)P(B)$$

- Collectively exhaustive events

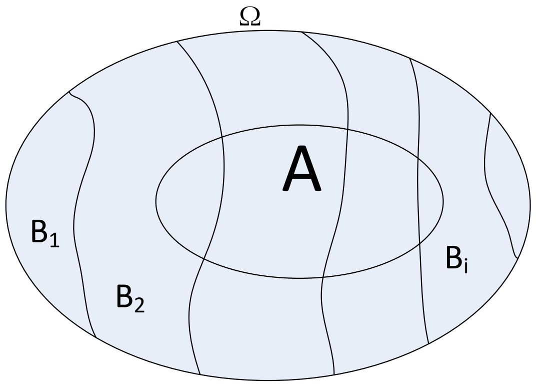


$$\bigcup_{i=1}^{N} B_i = \Omega \qquad B_i \cap B_j = \emptyset$$

Bayes' theorem

- Law of total probabilities

$$P(A) = \sum_{i=1}^{N} P(A \cap B_i) = \sum_{i=1}^{N} P(A|B_i)P(B_i)$$



- Bayes-theorem

$$\boxed{\boldsymbol{P(B_k|A) = \frac{P(A|B_k)P(B_k)}{P(A)} = \frac{P(A|B_k)P(B_k)}{\sum_{i=1}^{N} P(A|B_i)P(B_i)}}}$$

Usual terminology

Posterior: $P(B_k|A)$

Prior: $P(B_k)$

Likelihood: $P(A|B_k)$

Evidence, marginal likelihood: $P(A)$

Bayesian inference

Budapest University of Technology and Economics
Faculty of Transportation Engineering and Vehicle Engineering
Department of Control for Transportation and Vehicle Systems

Application of the Bayes' theorem for hypothesis testing

- We have a prior probability, that hypothesis H is true: $P(H)$
- We observe an event E, which is the evidence or observation and associate the probability: $P(E)$
- The likelihood that E happens given H is true is: $P(E|H)$
- The posterior probability that H is true given E is

$$\boxed{P(H|E) = \frac{P(E|H)P(H)}{P(E)} = \frac{P(E|H)P(H)}{P(E|H)P(H) + P(E|\neg H)P(\neg H)}}$$

Hypothesis test – loaded coin

- Someone is tossing a coin in the next room and tells us the results
- We have two hypotheses
  - The coin is loaded and produces $<\text{heads}>$ with 70% (L)
  - The coin is fair and does $50\% - 50\%$ $(\neg L)$
- We give probability $\mathrm{P}_0(L)$ that the coin is loaded (at the beginning)
- Based on what we hear, how shall we change our belief?
- The probabilities of the outcomes conditioned on the hypotheses are:

$$P(<\text{heads}>|L) = 0.7 \qquad P(<\text{tails}>|L) = 0.3$$
$$P(<\text{heads}>|\neg L) = 0.5 \qquad P(<\text{tails}>|\neg L) = 0.5$$

Hypothesis test – loaded coin

- Say the first toss gives $<\text{heads}>$ which results in:

$$P_1(L) = P_0(L| <\text{heads}>)$$

$$P_1(L) = \frac{P_0(<\text{heads}>|L)P_0(L)}{P_0(<\text{heads}>|L)P_0(L) + P_0(<\text{heads}>|\neg L)P_0(\neg L)}$$

$$P_1(L) = \frac{0.7P_0(L)}{0.7P_0(L) + 0.5(1 - P_0(L))}$$

- If we would have $<\text{tails}>$ instead:

$$P_1(L) = \frac{P_0(<\text{tails}>|L)P_0(L)}{P_0(<\text{tails}>|L)P_0(L) + P_0(<\text{tails}>|\neg L)P_0(\neg L)}$$

$$P_1(L) = \frac{0.3P_0(L)}{0.3P_0(L) + 0.5(1 - P_0(L))}$$

Hypothesis test – loaded coin

With a concrete prior belief: $P_0(L) = 0.2$

- 1. outcome: < heads >:

$$P_1(L) = \frac{0.7 \times 0.2}{0.7 \times 0.2 + 0.5 \times (1 - 0.2)} = 0.26$$

- 1. outcome: < tails >:

$$P_1(L) = \frac{0.3 \times 0.2}{0.3 \times 0.2 + 0.5 \times (1 - 0.2)} = 0.13$$

Hypothesis test – loaded coin

If we get two $<\text{heads}>$ in a row:

$$P_2(L) = P_1(L| < \text{heads} >)$$

$$P_2(L) = \frac{0.7 \times 0.26}{0.7 \times 0.26 + 0.5 \times (1 - 0.26)} = 0.33$$

- The second evidence also increases our belief
- This is a **recursive process** where we use the last result as prior
- We can have more than one concurrent hypotheses about a parameter (or a variable)
- In fact we can have continuously many hypotheses (from a parameter space or a state space)

Recursive Bayesian estimation

- If we get more and more measurements we can make a recursive estimation

$$P_2(H|E) = \frac{P_2(E|H)P_2(H)}{P_2(E)} \longleftarrow P_1(H|E) = \frac{P_1(E|H)P_1(H)}{P_1(E)}$$

- If there is some kind of "motion" (which isn't in the coin toss example) we should consider it and change out prior belief accordingly (green arrow)

Bayes' theorem for continuous problems

- Estimate x given z at timestep 1

$$p(x_1|z_1) = \frac{p(z_1|x_1)\,p(x_1)}{p(z_1)}$$

- At timestep 2 we have 2 measurements and 2 states

$$p(x_{1:2}|z_{1:2}) = \frac{p(z_{1:2}|x_{1:2})\,p(x_{1:2})}{p(z_{1:2})}$$

- And so on… $x_{1:k}$ and $z_{1:k}$ gets bigger at every timestep

$$p(x_{1:k}|z_{1:k}) = \frac{p(z_{1:k}|x_{1:k})\,p(x_{1:k})}{p(z_{1:k})}$$

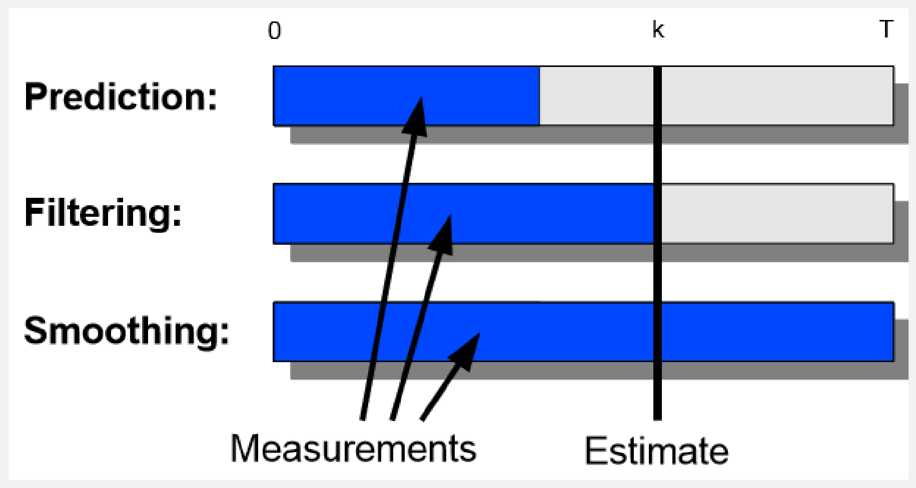
This is the full (or batch) Bayesian estimation problem

- It is generally intractable (computational complexity: $O(k^3)$)
- Usually we can simplify the problem

Bayesian estimation

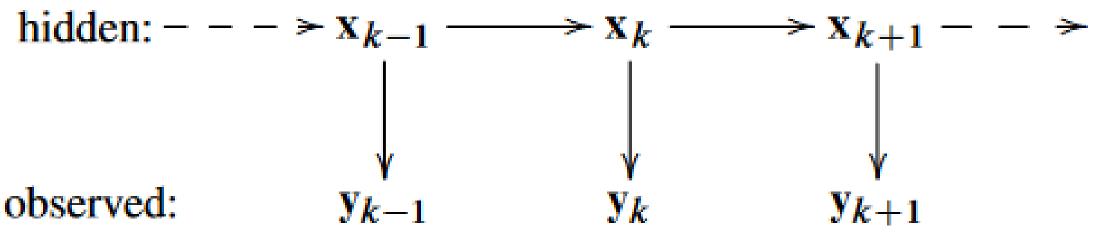
Using the measurements up to timestep *k*:

- **Filtering** we need the actual state x_k: $p(x_k|z_{1:k})$
- **Prediction** we need the next states x_{k+n}: $p(x_{k+n}|z_{1:k})$
- **Smoothing** we re-estimate the states x_{k-n}: $p(x_{k-n}|z_{1:k})$



Recursive estimation

- Filtering is generally achieved by a recursive estimation process
  - $p(x_k|z_{1:k}) = \dfrac{p(z_{1:k}|x_k)p(x_k)}{p(z_{1:k})}$

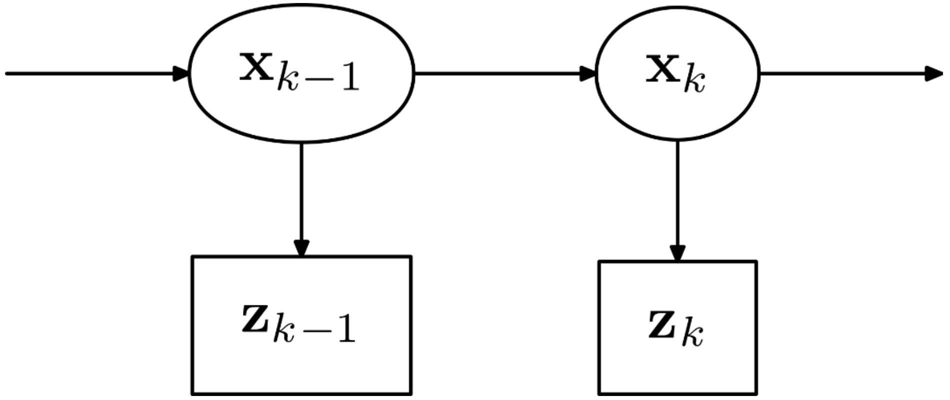


  - $p(x_k|z_{1:k}) = \dfrac{p(z_k|x_k)\,p(x_k|z_{1:k-1})}{p(z_k|z_{1:k-1})}$

Hidden Markov model (HMM)

Budapest University of Technology and Economics
Faculty of Transportation Engineering and Vehicle Engineering
Department of Control for Transportation and Vehicle Systems

- In the context of state estimation (robotics) the value to be estimated is the state (or state vector in general) of an object or an ensemble of objects
- The state in unknown to us (hidden) and possibly evolves in time: the system has dynamics
- We can observe the system and obtain a limited amount of information, for example
  - Partial observation of the state
  - Noisy measurements



Markov assumptions

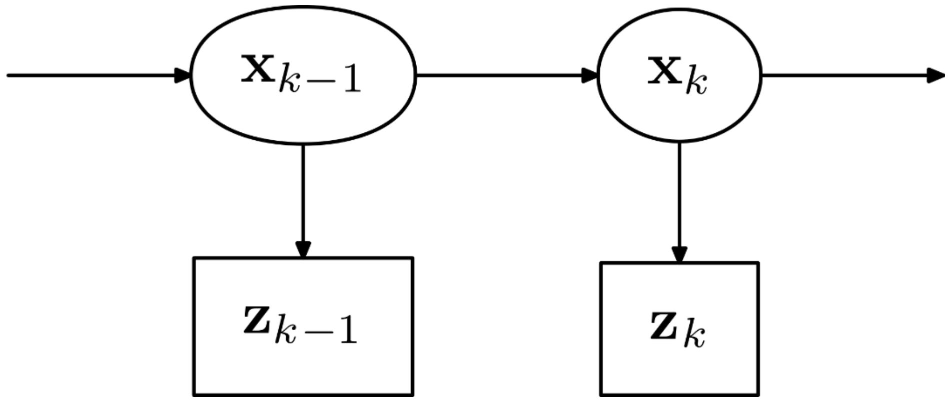
Budapest University of Technology and Economics
Faculty of Transportation Engineering and Vehicle Engineering
Department of Control for Transportation and Vehicle Systems

- The current state depends only on the previous state

$$p(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{x}_{k-2}, \dots, \mathbf{x}_0) = p(\mathbf{x}_k|\mathbf{x}_{k-1})$$

- The measurement depends only on the current state

$$p(\mathbf{z}_k|\mathbf{x}_k, \mathbf{x}_{k-1}, \dots, \mathbf{x}_0) = p(\mathbf{z}_k|\mathbf{x}_k)$$



Modelling uncertainties

- Additive noise acting on the motion and sensor model

$$\mathrm{x}_{k+1|k} = f_k(\mathrm{x}_k) + w_k$$

$$\mathrm{z}_k = h_k(\mathrm{x}_k) + v_k$$

random deterministic random

- How do we create probabilities from these random variables?
- Since x and z are usually continuous variables, the probabilities of taking specific values are zero.
- However, x and z residing in some region S and T have nonzero probabilities

$$P(\mathrm{x}_{k+1|k} \in S|\mathrm{x}_k) \qquad P(\mathrm{z}_k \in T|\mathrm{x}_k)$$

Modelling uncertainties

- The probability mass is given by integrating the probability density over a region

$$P\left(\mathrm{x}_{k+1|k} \in S \middle| \mathrm{x}_k\right) = \int_S p(\mathrm{x}|\mathrm{x}_k)\mathrm{dx} \qquad P(\mathrm{z}_k \in T|\mathrm{x}_k) = \int_T p(\mathrm{z}|\mathrm{x}_k)\mathrm{dz}$$

  - $p(\mathrm{x}|\mathrm{x}_k)$ is the probability density function associated to the uncertain motion model
  - $p(\mathrm{z}|\mathrm{x}_k)$ is the probability density function associated to the uncertain sensor model
- If the additive noise is zero mean Gaussian

$$\boxed{p(\mathrm{x}|\mathrm{x}_k) = \mathcal{N}(\mathrm{x}; f_k(\mathrm{x}_k), \sigma_w^2)}$$

- Similarly for the sensor model

$$\boxed{p(\mathrm{z}|\mathrm{x}_k) = \mathcal{N}(\mathrm{z}; h_k(\mathrm{x}_k), \sigma_v^2)}$$

Recursive Bayesian estimation (in discrete time)

- Estimate the state vector at timestep k using measurements up to k:

$$\boxed{p(\mathrm{x}_k|\mathrm{z}_{1:k}) = \frac{p(\mathrm{z}_k|\mathrm{x}_k)p(\mathrm{x}_k|\mathrm{z}_{1:k-1})}{p(\mathrm{z}_k|\mathrm{z}_{1:k-1})}}$$

$$P(B_k|A) = \frac{P(A|B_k)P(B_k)}{\sum_{i=1}^{N} P(A|B_i)P(B_i)}$$

This was the Bayes-theorem

- The denominator is constant and can be expressed as

$$p(\mathrm{z}_k|\mathrm{z}_{k-1}) = \int p(\mathrm{z}_k|\mathrm{x}_\mathrm{k})\, p(\mathrm{x}_k|\mathrm{z}_{k-1})\mathrm{dx}_k$$

- The prior, with the help of a model of the system is obtained from the pervious posterior through the time-prediction integral (Chapman-Kolmogorov integral):

$$p(\mathrm{x}_k|\mathrm{z}_{1:k-1}) = \int p(\mathrm{x}_k|\mathrm{x}_{\mathrm{k}-1})\, p(\mathrm{x}_{k-1}|\mathrm{z}_{1:\mathrm{k}-1})\mathrm{dx}_{k-1}$$

motion model previous posterior

Terminology in estimation

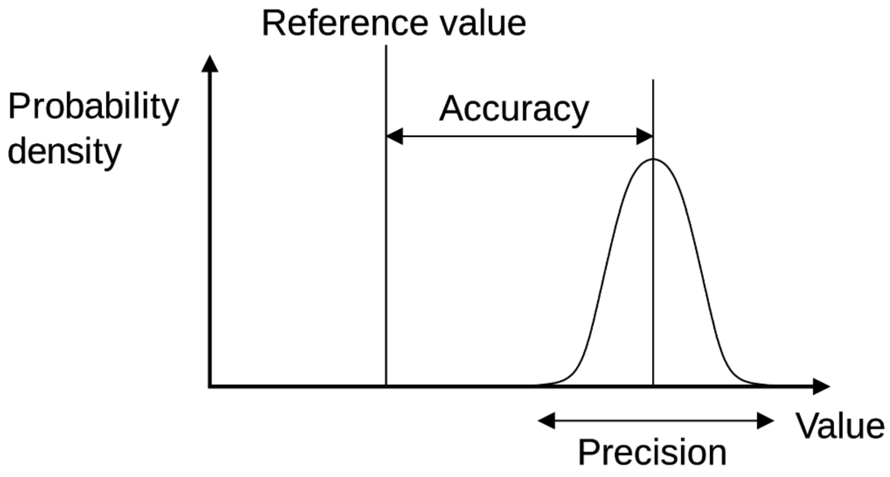
- Statistic: a function of the data
- Estimator: a function of the data that intends to describe some property of the underlying distribution
  - A statistic is not good or bad(or biased or unbiased). It is just a function
  - An estimator can be good (unbiased, minimum variance etc.). E.g.: the sample mean is an unbiased estimator of the expected value

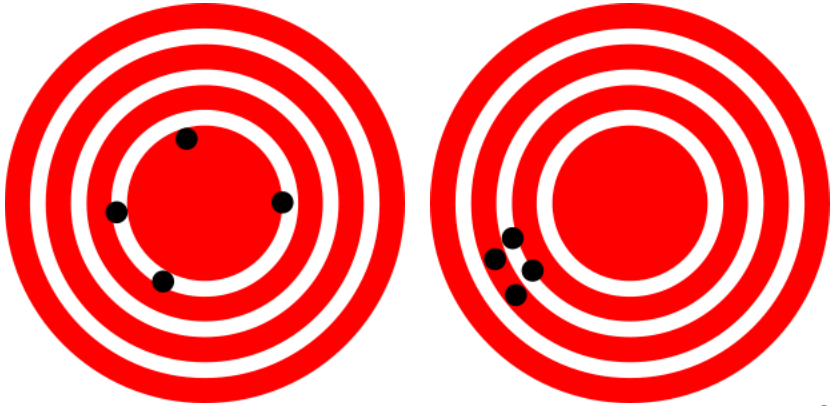
- Filtering: estimate x_t based on measurements $z_{1:t}$
- Prediction: estimate $x_{t+\tau}$ based on measurements $z_{1:t}$
- Smoothing: estimate $x_{t-\tau}$ based on measurements $z_{1:t}$

Accuracy, precision

The quality of a sensor can be described by its precision and accuracy

- Accuracy
  - Measures the systematic error (bias)
  - Related to the mean of the measurement
- Precision
  - Measure the random error (variability)
  - Related to the variance (standard deviation) of the measurement





System model

Approach

- Differential equations
- Continuous time
- Discrete time
- Probabilistic description
- Motion model
- Measurement model

$$\begin{aligned}\dot{\boldsymbol{x}}(t) &= f(t, \boldsymbol{x}, \boldsymbol{w}) \\ \boldsymbol{z}(t) &= h(t, \boldsymbol{x}, \boldsymbol{v})\end{aligned}$$

$$\begin{aligned}\dot{\boldsymbol{x}}(t) &= f(t, \boldsymbol{x}) + \boldsymbol{\Gamma}\boldsymbol{w} \\ \boldsymbol{z}(t) &= h(t, \boldsymbol{x}) + \boldsymbol{v}\end{aligned}$$

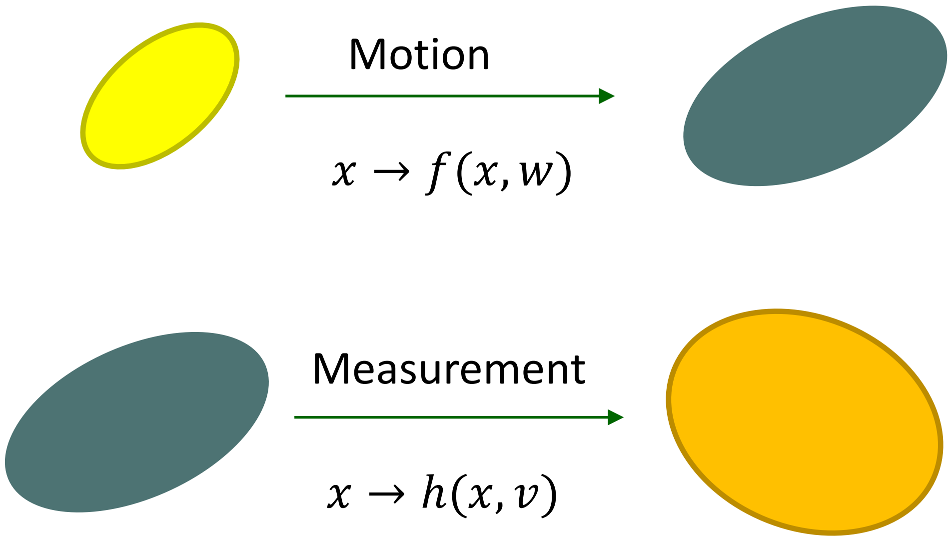
$$\begin{aligned}\mathrm{x}_{k+1} &= f_k(\mathrm{x}_k) + w_k \\ \mathrm{z}_k &= h_k(\mathrm{x}_k) + v_k\end{aligned}$$

random deterministic random

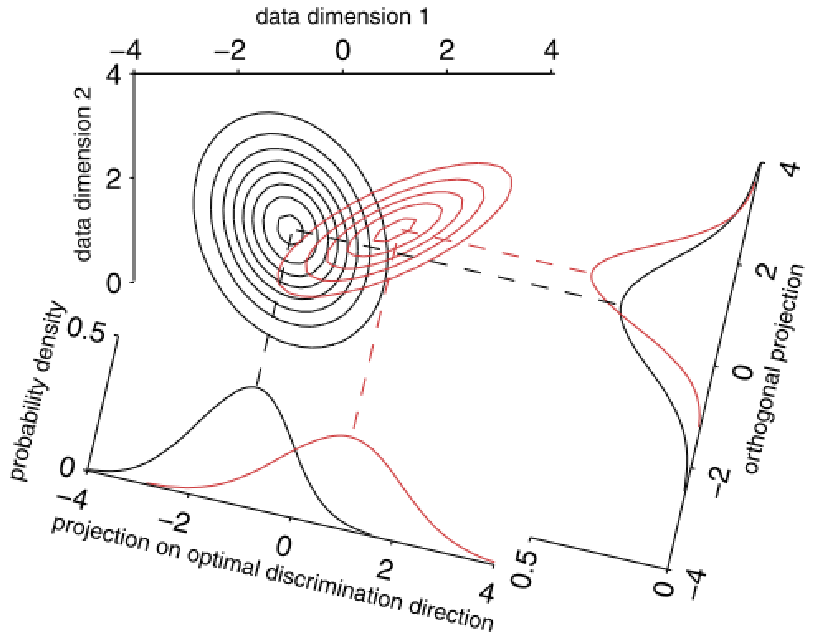
$$\begin{aligned}\mathrm{x}_{k+1} &= F_k \mathrm{x}_k + G_k w_k \\ \mathrm{z}_k &= H_k \mathrm{x}_k + v_k\end{aligned}$$

Propagating uncertainties

- Uncertain initial position
- Uncertain motion
- Uncertain measurement



- The motion model inserts additional noise
  - The uncertainty increases
- The measurement model inserts additional noise and projections
  - Partial observation + uncertainty increase



Propagating uncertainties

Budapest University of Technology and Economics
Faculty of Transportation Engineering and Vehicle Engineering
Department of Control for Transportation and Vehicle Systems

- Adding two random variables: $Z = X \pm Y$
  - $\sigma_Z^2 = \sigma_X^2 + \sigma_Y^2 \pm 2\sigma_{AB}$
  - $\sigma_Z^2 = \sigma_X^2 + \sigma_Y^2$ if X and Y are independent
- Affine transformation of random vector: $Z = AX + b$
  - $X = [\mathrm{x}_1, \mathrm{x}_2, \dots, x_n]$
  - b is a constant vector, A is a constant matrix
  - $\mathrm{E}[Z] \; = \mathrm{E}[AX + b] = A\mathrm{E}[X] + b = A\bar{X} + b$
  - $\mathrm{cov}(Z) = A \,\mathrm{cov}(X)\, A^\top$