Probabilistic Fundamentals in Robotics

Basic Concepts in Probability

Basilio Bona
DAUIN – Politecnico di Torino

Course Outline

- Motivations
- Basic mathematical framework
- Probabilistic models of mobile robots
- Mobile robot localization problem
- Robotic mapping
- Probabilistic planning and control

Reference textbook [TBF2006]
http://www.probabilistic-robotics.org/
Basic mathematical framework

- Basic concepts in probability
- Recursive state estimation
  - Robot environment
  - Bayes filters
- Gaussian filters
  - Kalman filter
  - Extended Kalman Filter
  - Unscented Kalman filter
  - Information filter
- Nonparametric filters
  - Histogram filter
  - Particle filter

Basic concepts in probability

- In binary logic, a proposition about the state of the world is only True or False; no third hypothesis is considered
- Bayesian probability is a measure of the degree of belief of a proposition, or an objective degree of rational belief, given the evidence

\[
\begin{align*}
Pr(\text{True}) &= 1 \\
Pr(\text{False}) &= 0 \\
0 &\leq Pr(A) \leq 1 \\
Pr(A \cup B) &= Pr(A) + Pr(B) - Pr(A \cap B)
\end{align*}
\]
Other axioms

\[ \begin{align*}
\Pr(A \cup (\neg A)) &= \Pr(A) + \Pr(\neg A) - \Pr(A \cap (\neg A)) \\
\Pr(\text{True}) &= \Pr(A) + \Pr(\neg A) - \Pr(\text{False}) \\
1 &= \Pr(A) + \Pr(\neg A) - 0 \\
\Pr(\neg A) &= 1 - \Pr(A)
\end{align*} \]

Random variables

- A random variable \( X \) (or stochastic variable) is a variable whose value is a function of the outcome of a statistical experiment.
- \( X \) can take a countable number of values in \( \{x_1, x_2, \ldots, x_n\} \)
- \( P(X = x_i) \), or \( P(x_i) \) is the probability that the random variable \( X \) takes the value \( x_i \)
- \( P(\cdot) \) is called “probability mass function”
Continuous random variables

- $X$ takes values in $\mathbb{R}$
- $p(X = x)$ or $p(x)$ is a probability density function

$$\Pr(x \in [a, b]) = \int_{a}^{b} p(x) \, dx$$

Univariate Gaussian distribution

$$p(x) = (2\pi\sigma^2)^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} \frac{(x - \mu)^2}{\sigma^2} \right\}$$
Normal distribution

\[ N(x; \mu, \sigma^2) \]

Normal distribution

Area = \( \Phi(0.7) \)

Mean = 0

Variance = 1

0.399

\[ \Phi(0.7) \]

Standard normal PDF

\[ \Phi(y) \]

Standard normal CDF
Multi-variate Gaussian distribution

\[ p(\mathbf{x}) = \det(2\pi \Sigma)^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \Sigma^{-1}(\mathbf{x} - \boldsymbol{\mu}) \right\} \]

Joint and conditional probabilities

- \( P((X = x) \cap (Y = y)) = P(x, y) \) joint probability
- If \( X \) and \( Y \) are independent random variables
  \[ P(x, y) = P(x)P(y) \]
- \( P(x|y) \) is the probability of \( x \) given \( y \) (conditional probability)
  \[ P(x, y) = P(x|y)P(y) = P(y|x)P(x) \]
- If \( X \) and \( Y \) are independent
  \[ P(x|y)P(y) = P(x, y) = P(x)P(y) \Rightarrow P(x|y) = P(x) \]
Marginal and total Probability

Discrete

\[ P(x) = \sum_y P(x, y) = \sum_y P(x|y)P(y) \]
\[ \sum_x P(x) = 1 \]

Continuous

\[ p(x) = \int_y p(x, y)dy = \int_y p(x|y)p(y)dy \]
\[ \int_x p(x) = 1 \]

Posterior probability and Bayes rule

\[ P(x, y) = P(x|y)P(y) = P(y|x)P(x) \]
\[ P(x|y) = \frac{P(y|x)P(x)}{P(y)} \]
Bayes rule conditioned by another variable

\[
P(x|y) = \frac{P(y|x)P(x)}{P(y)}
\]

\[
P(x|y, z) = \frac{P(y|x, z)P(x|z)}{P(y|z)}
\]

as long as \(P(y|z) > 0\)

Normalization

\[
P(x|y) = \frac{P(y|x)P(x)}{P(y)}
\]

\[
P(x|y) = \frac{P(y|x)P(x)}{\sum_x P(y|x)P(x)} = \eta P(y|x)P(x)
\]
Marginal probability

\[ P(x) = \int P(x, z) dz \]
\[ P(x) = \int P(x|z) P(z) dz \]
\[ P(x|y) = \int P(x|y, z) P(z|y) dz \]

Conditional independence

\[ P(x, y|z) = P(x|z) P(y|z) \]

Equivalent to

\[ P(x|z) = P(x|z, y) \]
and

\[ P(y|z) = P(y|z, x) \]

This is an important rule in probabilistic robotics. It applies whenever a variable \( y \) carries no information about a variable \( x \), if the value \( z \) of another variable is known.
Conditional independence ≠ absolute independence

\[
p(x, y | z) = p(x | z)p(y | z) \not\Rightarrow p(x, y) = p(x)p(y)
\]

and

\[
p(x, y) = p(x)p(y) \not\Rightarrow p(x, y | z) = p(x | z)p(y | z)
\]

absolute independence

---

Expectation of a random variable

- Features of probabilistic distributions are called **statistics**
- **Expectation** of a random variable (RV) \(X\) is defined as

**Discrete**

\[
E[X] = \sum_x x \ p(x)
\]

**Continuous**

\[
E[X] = \int_x x \ p(x) \ dx
\]

**Linearity property**

\[
E[aX + b] = a \ E[X] + b
\]
Covariance

- **Covariance** measures the squared expected deviation from the mean

\[
\text{Cov}[X] = E[(X - E[X])^2] = E[X^2] - E[X]^2
\]

\[
N(x; \mu, \Sigma) \rightarrow E[x] = \mu; \quad \text{Cov}[x] = \Sigma
\]

Entropy

- **Entropy** measures the expected information that the value of \( x \) carries

\[
H_p(x) = E[- \log_2 p(x)]
\]

**Discrete**

\[
H_p(x) = - \sum_x p(x) \log_2 p(x)
\]

**Continuous**

\[
H_p(x) = - \int x p(x) \log_2 p(x) \, dx
\]

In discrete case is the number of bits required to encode \( x \)
using an optimal encoding, assuming that \( p(x) \) is the probability of observing \( x \)
Robot environment interaction

- *World* or *environment* is a dynamical system that has an internal state
- Robot *sensors* can acquire information about the world internal state
- Sensors are *noisy* and often complete information cannot be acquired
- A *belief* measure about the state of the world is stored by the robot
- Robot influences the world through its *actuators* (e.g., they make it move in the environment)
State

- **State** is denoted by $x$; $x$ can be a single variable, but often it is a vector valued variable $\mathbf{x}$.
- **Pose** includes the cartesian coordinates of the robot and its orientation (roll, pitch, yaw angles).
- State includes robot pose and **velocity**, as well as configuration and **velocities** of its generalized coordinates.
- **Features** of the environment (landmarks) and their location are considered part of the state as well.
- Landmarks can be static or dynamic. If they are dynamic, their velocities add to the state vector.
- Other variables (e.g., battery charge state, etc.) can be also considered states.

Complete state

- A **complete state** $x_t$ at time $t$ is defined as the best predictor of its future.
- No other data (coming from past states, measurements, controls) add information that helps to better predict the future.
- Future may be stochastic, but no variables prior to $x_t$ may influence the evolution of future states.
- Evolution of future states is mediated through $x_t$ only.
- Complete state is impossible to specify. A subset of complete state is called **incomplete state**.
- State where $t$ is both continuous and discrete, is called **hybrid** state.
Stochastic process

Given a probability space \((\Omega, \mathcal{F}, P)\), a stochastic process (or random process) with state space \(X\) is a collection of \(X\)-valued random variables indexed by a set \(T\) ("time"). That is, a stochastic process \(F\) is a collection \(\{F_t : t \in T\}\) where each \(F_t\) is an \(X\)-valued random variable.

Markov chains

- a Markov chain is a discrete random process with the Markov property
- A stochastic process has the Markov property if the conditional probability distribution of future states of the process depend only upon the present state; that is, given the present, the future does not depend on the past.
Environment interaction

- **Measurements**: are perceptual interaction between the robot and the environment obtained through specific equipment (called also *perceptions*).
  
  \[ z_{t_1:t_2} = \{ z_{t_1}, z_{t_1+1}, z_{t_1+2}, \ldots, z_{t_2-1}, z_{t_2} \} \]

- **Control actions**: are change in the state of the world obtained through active asserting forces.
  
  \[ u_{t_1:t_2} = \{ u_{t_1}, u_{t_1+1}, u_{t_1+2}, \ldots, u_{t_2-1}, u_{t_2} \} \]

- **Odometer data**: are of perceptual data that convey the information about the robot change of status; as such they are not considered measurements, but control data, since they measure the effect of control actions.

Probabilistic generative laws

- Evolution of state is governed by probabilistic laws.
- If state is complete and Markov, then evolution depends only on present state and control actions

\[
p(x_t|x_{0:t-1}, z_{1:t-1}, u_{1:t}) = p(x_t|x_{t-1}, u_t)
\]

**State transition probability**

- Measurements are generated, according to probabilistic laws, from the present state only

\[
p(z_t|x_{0:t-1}, z_{1:t-1}, u_{1:t}) = p(z_t|x_t)
\]

**Measurement probability**
### Belief distribution

- **What is a belief:** It is a measure of the robot’s internal knowledge about the true state of the environment.

- **Belief is traditionally expressed as conditional probability distributions.**

- **Belief distribution:** Assigns a probability (or a density) to each possible hypothesis about the true state, based upon available data (measurements and controls).

\[
\begin{align*}
\overline{bel}(x_t) &= p(x_t | z_{1:t-1}, u_{1:t}) \quad \text{(State belief (prior))} \\
bel(x_t) &= p(x_t | z_{1:t}, u_{1:t}) \quad \text{(State belief (posterior))}
\end{align*}
\]
Bayes filter

- Basic algorithm

\[
\text{Bayes Filter} \left( bel(x_{t-1}), u_t, z_t \right)
\]

1. forall \( x_t \) do

2.  \[
\overline{bel}(x_t) = \int p(x_t | x_{t-1}, u_t) \, bel(x_{t-1}) \, dx_{t-1}
\]

3.  \[
\text{bel}(x_t) = \eta(z_t | x_t) \, \text{bel}(x_t)
\]

4.  endfor

5.  return \( \text{bel}(x_t) \)

Mathematical formulation of the Bayesian filter (1)

\[
p(x_t | z_{1:t}, u_{1:t}) = \frac{p(z_t | x_t, z_{1:t-1}, u_{1:t}) p(x_t | z_{1:t-1}, u_{1:t})}{p(z_t | z_{1:t-1}, u_{1:t})} = \eta p(z_t | x_t, z_{1:t-1}, u_{1:t}) p(x_t | z_{1:t-1}, u_{1:t})
\]

the state is complete

\[
p(z_t | x_t, z_{1:t}, u_{1:t}) = p(z_t | x_t)
\]

\[
p(x_t | z_{1:t}, u_{1:t}) = \eta(z_t | x_t) \ p(x_t | z_{1:t-1}, u_{1:t})
\]

\[
\text{bel}(x_t) = \eta(z_t | x_t) \overline{\text{bel}}(x_t)
\]
Mathematical formulation of the Bayesian filter (2)

\[
\bar{b}el(x_t) = p(x_t|z_{1:t}, u_{1:t}) = \int p(x_t|x_{t-1}, z_{1:t-1}, u_{1:t}) p(x_{t-1}|z_{1:t-1}, u_{1:t}) \, dx_{t-1}
\]

The state is complete

\[
p(x_t|x_{t-1}, z_{1:t-1}, u_{1:t}) = p(x_t|x_{t-1}, u_{1:t})
\]

\[
p(x_t|x_{t-1}, u_{1:t}) \neq p(x_t|x_{t-1})
\]

Control \(u_t\) can be omitted from \(u_{1:t}\)

\[
p(x_{t-1}|z_{1:t-1}, u_{1:t}) = p(x_{t-1}|z_{1:t-1}, u_{1:t-1})
\]

Mathematical formulation of the Bayesian filter (3)

\[
\bar{b}el(x_t) = \int p(x_t|x_{t-1}, u_t) \, p(x_{t-1}|z_{1:t-1}, u_{1:t-1}) \, dx_{t-1}
\]

\[
bel(x_t) = \eta(z_t|x_t) \, \bar{b}el(x_t)
\]

The filter requires three probability distributions

- \(p(x_0)\): initial belief
- \(p(z_t|x_t)\): measurement probability
- \(p(x_t|u_t, x_{t-1})\): state transition probability
Bayes filter recursion

Causal vs. diagnostic reasoning

A rover obtains a measurement $z$ from a door that can be open ($O$) or closed ($C$)

$P(O|z)$ is **DIAGNOSTIC** knowledge

$P(z|O)$ is **CAUSAL** knowledge

Easier to obtain

$$P(O|z) = \frac{P(z|O)P(O)}{P(z)}$$
Example

\[ P(z|O) = 0.6 \quad P(z|\neg O) = 0.3 \]
\[ P(O) = P(\neg O) = 0.5 \]

\[
P(O|z) = \frac{P(z|O)P(O)}{P(z|O)P(O) + P(z|\neg O)P(\neg O)}
\]

\[
P(O|z) = \frac{0.6 \cdot 0.5}{0.6 \cdot 0.5 + 0.3 \cdot 0.5} = \frac{2}{3} = 0.67 > 0.6
\]

References

- Many textbooks on Probability Theory and Statistics

- Other materials
  - [http://cs.ubc.ca/~arnaud/stat302.html](http://cs.ubc.ca/~arnaud/stat302.html): slides from the course by A. Doucet, University of British Columbia
Thank you. Any question?