Principles of Robot Autonomy I

Introduction to localization and filtering theory
Module 3

Knowledge

Localization
Map Building

- Environmental model
- Local map

Information extraction

- Raw data

Sensing

Decision making
Motion planning

- Position
- Global map

Trajectory execution

- Trajectory
- Actuator commands

Actuation

Real world environment

Mission goals

See-think-act
Today’s lecture

• Aim
  • Learn basic concepts about Bayesian filtering

• Readings
Localization

• Two main approaches:
  1. Behavioral approach: design a set of behaviors that together result in the desired robot motion (no need for a map)
  2. Map-based approaches: robot *explicitly* attempts to localize by collecting sensor data, then updating belief about its position with respect to a map

• We will focus on map-based approaches; two main aspects:
  • Map representation: how to represent the environment?
  • Belief representation: how to model the belief regarding the position within the map?
Probabilistic map-based localization

- **Key idea**: represent belief as a probability distribution
  1. Encodes sense of position
  2. Maintains notion of robot’s uncertainty

- **Belief representation**:
  1. Single-hypothesis vs. multiple hypothesis
  2. Continuous vs. discretized

- Today we will overview basic concepts in Bayesian filtering
Basic concepts in probability

• **Key idea:** quantities such as sensor measurements, states of a robot, and its environment are modeled as random variables (RVs)

• **Discrete RV:** the space of all the values that a random variable $X$ can take on is *discrete*; characterized by probability mass function (pmf)

$$p(X = x) \quad \text{or} \quad p(x), \quad \sum_x p(X = x) = 1$$

• **Continuous RV:** the space of all the values that a random variable $X$ can take on is *continuous*; characterized by probability density function (pdf)

$$P(a \leq X \leq b) = \int_a^b p(x) \, dx, \quad \int_{-\infty}^{\infty} p(x) \, dx = 1$$
Joint distribution, independence, and conditioning

- Joint distribution of two random variables $X$ and $Y$ is denoted as
  \[ p(x, y) := p(X = x \text{ and } Y = y) \]

- If $X$ and $Y$ are independent
  \[ p(x, y) = p(x)p(y) \]

- Suppose we know that $Y = y$ (with $p(y) > 0$); conditioned on this fact, the probability that the $X$’s value is $x$ is given by
  \[ p(x \mid y) := \frac{p(x, y)}{p(y)} \]
  Note: if $X$ and $Y$ are independent
  \[ p(x \mid y) := p(x)! \]
Theorem of total probability

• For discrete RVs:
  \[ p(x) = \sum_{y} p(x, y) = \sum_{y} p(x \mid y)p(y) \]

• For continuous RVs:
  \[ p(x) = \int p(x, y) dy = \int p(x \mid y)p(y) dy \]

• Note: if \( p(y) = 0 \), define the product \( p(x \mid y)p(y) = 0 \)
Bayes’ rule

• Key relation between $p(x \mid y)$ and its “inverse,” $p(y \mid x)$

• For discrete RVs:

$$p(x \mid y) = \frac{p(y \mid x)p(x)}{p(y)} = \frac{p(y \mid x)p(x)}{\sum_{x'} p(y \mid x')p(x')}$$

• For continuous RVs:

$$p(x \mid y) = \frac{p(y \mid x)p(x)}{p(y)} = \frac{p(y \mid x)p(x)}{\int p(y \mid x')p(x') \, dx'}$$
Bayes’ rule and probabilistic inference

• Assume $x$ is a quantity we would like to infer from $y$
• Bayes rule allows us to do so through the inverse probability, which specifies the probability of data $y$ assuming that $x$ was the cause

\[
p(x \mid y) = \frac{p(y \mid x)p(x)}{\int p(y \mid x')p(x') \, dx'}.
\]

• Notational simplification

\[
p(x \mid y) = \eta p(y \mid x)p(x)
\]
More on Bayes’ rule and independence

• Extension of Bayes rule: conditioning Bayes rule on $Z=z$ gives

$$p(x \mid y, z) = \frac{p(y \mid x, z)p(x \mid z)}{p(y \mid z)}$$

• Extension of independence: *conditional independence*

$$p(x, y \mid z) = p(x \mid z)p(y \mid z), \quad \text{equivalent to} \quad \begin{cases} p(x \mid z) = p(x \mid z, y) \\ p(y \mid z) = p(y \mid z, x) \end{cases}$$

• Note: in general

$$p(x, y \mid z) = p(x \mid z)p(y \mid z) \implies p(x, y) = p(x)p(y)$$

$$p(x, y) = p(x)p(y) \implies p(x, y \mid z) = p(x \mid z)p(y \mid z)$$
Expectation of a RV

- Expectation for discrete RVs: \( E[X] = \sum_x x p(x) \)
- Expectation for continuous RVs: \( E[X] = \int x p(x) \, dx \)
- Expectation is a linear operator: \( E[aX + b] = a \, E[X] + b \)
- Expectation of a vector of RVs is simply the vector of expectations
- Covariance

\[
\]
Model for robot-environment interaction

• Two fundamental types of robot-environment interactions: the robot can influence the state of its environment through control actions, and gather information about the state through measurements

• State $x_t$: collection at time $t$ of all aspects of the robot and its environment that can impact the future
  • Robot pose (e.g., robot location and orientation)
  • Robot velocity
  • Locations and features of surrounding objects in the environment, etc.

• Useful notation: $x_{t_1:t_2} := x_{t_1}, x_{t_1+1}, x_{t_1+2}, \ldots, x_{t_2}$

• A state $x_t$ is called complete if no variables prior to $x_t$ can influence the evolution of future states -> Markov property
Measurement and control data

• **Measurement data** $z_t$: information about state of the environment at time $t$; useful notation

$$z_{t_1:t_2} := z_{t_1}, z_{t_1+1}, z_{t_1+2}, \ldots, z_{t_2}$$

• **Control data** $u_t$: information about the change of state at time $t$; useful notation

$$u_{t_1:t_2} := u_{t_1}, u_{t_1+1}, u_{t_1+2}, \ldots, u_{t_2}$$

• Key difference: measurement data tends to increase robot’s knowledge, while control actions tend to induce a loss of knowledge
State equation

• General probabilistic generative model

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

• Key assumption: state is complete (i.e., the Markov property holds)

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

Convention: first take control action and then take measurement

State transition probability

• In other words, we assume *conditional independence*, with respect to conditioning on \( x_{t-1} \)
Measurement equation and overall stochastic model

• Assuming $x_t$ is complete

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

Measurement probability

• Overall dynamic Bayes network model (also referred to as hidden Markov model)
Belief distribution

- **Belief distribution**: reflects internal knowledge about the state
- A belief distribution assigns a probability to each possible hypothesis with regard to the true state
- Formally, belief distributions are posterior probabilities over state variables conditioned on the available data

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

- Similarly, the *prediction* distribution is defined as

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

- Calculating \( \text{bel}(x_t) \) from \( \overline{\text{bel}}(x_t) \) is called correction or measurement update
Bayes filter algorithm

- **Bayes’ filter algorithm**: most general algorithm for calculating beliefs
- **Key assumption**: state is complete

• **Recursive algorithm**
  - Step 1 (prediction): compute \( \overline{\text{bel}}(x_t) \)
  - Step 2 (measurement update): compute \( \text{bel}(x_t) \)

• Algorithm initialized with \( \text{bel}(x_0) \) (e.g., uniform or points mass)

**Update rule**

\[
\begin{align*}
\text{Data}: & \quad \text{bel}(x_{t-1}), u_t, z_t \\
\text{Result}: & \quad \text{bel}(x_t) \\
\text{foreach } x_t \text{ do} & \\
& \begin{align*}
& \text{bel}(x_t) = \int p(x_t | u_t, x_{t-1}) \text{bel}(x_{t-1}) dx_{t-1} \\
& \text{bel}(x_t) = \eta p(z_t | x_t) \overline{\text{bel}}(x_t); \\
& \text{end}
\end{align*}
\]

Return \( \text{bel}(x_t) \)
Derivation: measurement update

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

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

\[ := \eta^{-1} \]

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

\[ = \text{bel}(x_t) \]

Bayes rule

Markov property
Derivation: correction update

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

Total probability

Markov

For general output feedback policies, \( u_t \) does not provide additional information on \( x_{t-1} \)
Discrete Bayes’ filter

• Discrete Bayes’ filter algorithm: applies to problems with finite state spaces

• Belief $\text{bel}(x_t)$ represented as pmf $\{p_{k,t}\}$

Data: $\{p_{k,t-1}\}, u_t, z_t$
Result: $\{p_{k,t}\}$
foreach $k$ do
  $\bar{p}_{k,t} = \sum_i p(X_t = x_k \mid u_t, X_{t-1} = x_i) p_{i,t-1}$;
  $p_{k,t} = \eta p(z_t \mid X_t = x_k) \bar{p}_{k,t}$;
end
Return $\{p_{k,t}\}$
Next time