NA 568 - Winter 2026

Bayes Filters

Maani Ghaffari

January 9, 2026

UNIVERSITY OF MICHIGAN

- Goal: estimate the (hidden) state x_t using controls $u_{1:t}$ and measurements $z_{1:t}$.

$u_{1:t} = \{u_1, u_2, \dots, u_t\}$

- Output: belief (posterior)

置信=后验=基于观测和输入得到状态

$$\mathrm{bel}(x_t) \;:=\; p(x_t \mid z_{1:t}, u_{1:t}).$$

- Core loop (every time step):
  - **Predict** using motion model: $p(x_t \mid x_{t-1}, u_t)$
  - **Update** using sensor model: $p(z_t \mid x_t)$
  - **Normalize** so probabilities sum / integrate to 1
- Today: Bayes filter recursion + intuition (discrete example).
- Next: Kalman filter = Bayes filter with linear-Gaussian models.

Remark: `belief = probability distribution over state`.

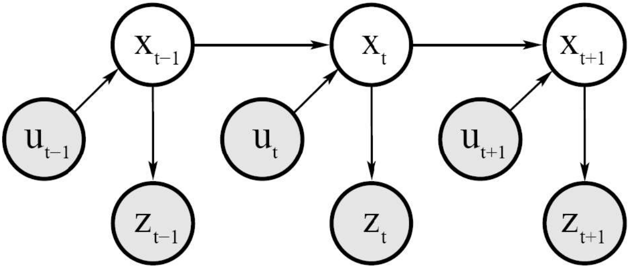
- Given: 三要素：观测与动作、观测模型、动作模型
  - Stream of observations $z_{1:t}$ and action data $u_{1:t}$
  - Sensor/measurement model $p(z_t|x_t)$
  - Action/motion/transition model $p(x_t|x_{t-1},u_t)$
- Wanted: 目标：系统状态、后验概率
  - The state x_t of dynamical system
  - The posterior of state is called belief $bel(x_t) = p(x_t|z_{1:t}, u_{1:t})$

从xt递推到，z和u直接推出x

Assumption (Markov Property)

*The Markov property states that "**the future is independent of the past if the present is known.**" A stochastic process that has this property is called a **Markov process**.*

Dynamic Bayesian Network for Controls, States, and Sensations



Markov Assumptions

- Filtering is possible because of two conditional independencies:
- State evolution (1st-order Markov):

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

- Measurement model (memoryless sensor):

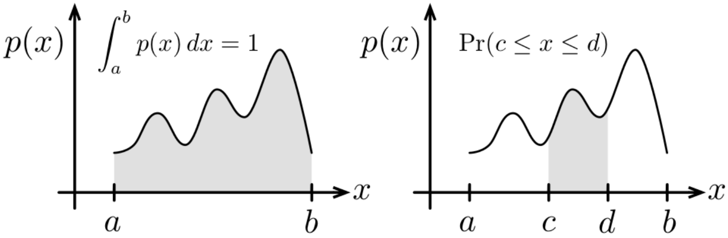
$$p(z_t \mid x_{0:t}, z_{1:t-1}, u_{1:t}) \ = \ p(z_t \mid x_t)$$

- Dependency "diagram":

$$x_{t-1} \ \to \ x_t \ \to \ x_{t+1}, \qquad u_t \ \to \ x_t, \qquad x_t \ \to \ z_t$$

- Takeaway: *the belief at time t can be updated recursively from the belief at time $t-1$.*

Probability Density Functions



Courtesy: T. Barfoot

Joint and Conditional Distribution

Let X and Y be two random variables.

- The joint distribution of X and Y is:
  $p(x,y) = p(X = x \text{ and } Y = y)$;

- The conditional probability of X given Y is:
  $p(x|y) = \frac{p(x,y)}{p(y)} \qquad p(y) > 0.$

- If X and Y are independent then $p(x,y) = p(x)p(y)$

- Given $p(x,y)$, the marginal distribution of X can be computed by summing (integration) over Y.

$$p(x) = \sum_{y \in \mathcal{Y}} p(x,y)$$

- The law of total probability is its variant which uses the conditional probability definition

$$p(x) = \sum_{y \in \mathcal{Y}} p(x|y)p(y)$$

and for continuous random variables, it is

$$p(x) = \int_{y \in \mathcal{Y}} p(x,y)dy = \int_{y \in \mathcal{Y}} p(x|y)p(y)dy$$

- $$p(x,y) = p(x|y)p(y) = p(y|x)p(x)$$

- $$p(x|y) = \frac{p(y|x)p(x)}{p(y)} = \frac{p(y|x)p(x)}{\sum_{x \in \mathcal{X}} p(y|x)p(x)}$$

- $$p(\text{hypothesis}|\text{data}) = \frac{p(\text{data}|\text{hypothesis})p(\text{hypothesis})}{p(\text{data})}$$

- $$\text{Posterior} = \frac{\text{Likelihood} \times \text{Prior}}{\text{Evidence (Marginal Likelihood)}}$$

Causal vs. Diagnostic Reasoning

- $p(\text{hypothesis}|\text{data})$ is **diagnostic**.
- $p(\text{data}|\text{hypothesis})$ is **causal**.
- Often causal knowledge is easier to obtain.
- Bayes rule allows us to use causal knowledge for diagnostic reasoning:

$$p(\text{hypothesis}|\text{data}) = \frac{p(\text{data}|\text{hypothesis})p(\text{hypothesis})}{p(\text{data})}$$

Example

An autonomous car approaches a traffic light that can be green, yellow, or red. It is programmed to stop if it detects yellow or red; otherwise it continues. Due to sensor imperfections, it drives through 10% of yellow lights, 95% of green lights, and 1% of red lights. The light cycles continuously: 30 seconds green, 5 seconds yellow, 25 seconds red. You are not watching the road and feel the car stop as it approaches the light. What is the probability the light was yellow when the vehicle sensed it?

Answer

Let S represent the event that the vehicle stopped, G the event that the light was green, Y that it was yellow, R that is was red.

- *Given:* $P(S|Y) = 0.90, P(S|G) = 0.05, P(S|R) = 0.99, P(Y) = 5/60, P(R) = 25/60, P(G) = 30/60$
- *Find:* $P(Y|S)$

$$P(Y|S) = \frac{P(S|Y)P(Y)}{P(S)}$$

$$P(Y|S) = \frac{P(S|Y)P(Y)}{P(S|Y)P(Y) + P(S|R)P(R) + P(S|G)P(G)}$$

$$P(Y|S) = \frac{0.90(5/60)}{0.90(5/60) + 0.99(25/60) + 0.05(30/60)} = 14.63\%$$

From Bayes' Rule to Bayes Filters

- Bayes' rule (one-shot inference):

$$p(x \mid z) \propto p(z \mid x)\, p(x)$$

- Robotics is sequential: we receive controls and measurements over time

$$u_{1:t},\, z_{1:t}$$

- Bayes filter = Bayes' rule applied repeatedly:
  - **Predict:** propagate belief forward using the motion model
  - **Update:** incorporate the new measurement via likelihood
  - **Normalize:** keep a valid distribution

Key idea: *belief is a distribution over state, updated recursively*.

State Estimation

- Estimate the state x of a system given observations z and controls u
- **Goal:**

$$p(x \mid z, u)$$

Courtesy: C. Stachniss

Recursive Bayes Filter 1

$$bel(x_t) = \underline{p(x_t \mid z_{1:t}, u_{1:t})}$$

Definition of the belief

Courtesy: C. Stachniss

Recursive Bayes Filter 2

$$bel(x_t) = p(x_t \mid z_{1:t}, u_{1:t})$$

$$= \underline{\eta \, p(z_t \mid x_t, z_{1:t-1}, u_{1:t}) \, p(x_t \mid z_{1:t-1}, u_{1:t})}$$

Bayes' rule

$$\eta = \frac{1}{p(z_t \mid z_{1:t-1}, u_{1:t})}$$

Courtesy: C. Stachniss

Recursive Bayes Filter 3

$$bel(x_t) = p(x_t \mid z_{1:t}, u_{1:t})$$

$$= \eta \, p(z_t \mid x_t, z_{1:t-1}, u_{1:t}) \, p(x_t \mid z_{1:t-1}, u_{1:t})$$

$$= \eta \, \underline{p(z_t \mid x_t)} \, p(x_t \mid z_{1:t-1}, u_{1:t})$$

Markov assumption

Courtesy: C. Stachniss

Recursive Bayes Filter 4

$$
\begin{aligned}
bel(x_t) &= p(x_t \mid z_{1:t}, u_{1:t}) \\
&= \eta\, p(z_t \mid x_t, z_{1:t-1}, u_{1:t})\, p(x_t \mid z_{1:t-1}, u_{1:t}) \\
&= \eta\, p(z_t \mid x_t)\, p(x_t \mid z_{1:t-1}, u_{1:t}) \\
&= \eta\, p(z_t \mid x_t) \int \underline{p(x_t \mid x_{t-1}, z_{1:t-1}, u_{1:t})} \\
&\qquad \underline{p(x_{t-1} \mid z_{1:t-1}, u_{1:t})}\, dx_{t-1}
\end{aligned}
$$

Law of total probability

Courtesy: C. Stachniss

Recursive Bayes Filter 5

$$
\begin{aligned}
bel(x_t) &= p(x_t \mid z_{1:t}, u_{1:t}) \\
&= \eta\, p(z_t \mid x_t, z_{1:t-1}, u_{1:t})\, p(x_t \mid z_{1:t-1}, u_{1:t}) \\
&= \eta\, p(z_t \mid x_t)\, p(x_t \mid z_{1:t-1}, u_{1:t}) \\
&= \eta\, p(z_t \mid x_t) \int p(x_t \mid x_{t-1}, z_{1:t-1}, u_{1:t}) \\
&\qquad\qquad p(x_{t-1} \mid z_{1:t-1}, u_{1:t})\, dx_{t-1} \\
&= \eta\, p(z_t \mid x_t) \int \underline{p(x_t \mid x_{t-1}, u_t)}\, p(x_{t-1} \mid z_{1:t-1}, u_{1:t})\, dx_{t-1}
\end{aligned}
$$

Markov assumption

Courtesy: C. Stachniss

Recursive Bayes Filter 6

$$
\begin{aligned}
bel(x_t) &= p(x_t \mid z_{1:t}, u_{1:t}) \\
&= \eta \, p(z_t \mid x_t, z_{1:t-1}, u_{1:t}) \, p(x_t \mid z_{1:t-1}, u_{1:t}) \\
&= \eta \, p(z_t \mid x_t) \, p(x_t \mid z_{1:t-1}, u_{1:t}) \\
&= \eta \, p(z_t \mid x_t) \int p(x_t \mid x_{t-1}, z_{1:t-1}, u_{1:t}) \\
&\qquad\qquad p(x_{t-1} \mid z_{1:t-1}, u_{1:t}) \, dx_{t-1} \\
&= \eta \, p(z_t \mid x_t) \int p(x_t \mid x_{t-1}, u_t) \, p(x_{t-1} \mid z_{1:t-1}, u_{1:t}) \, dx_{t-1} \\
&= \eta \, p(z_t \mid x_t) \int p(x_t \mid x_{t-1}, u_t) \, p(x_{t-1} \mid z_{1:t-1}, u_{1:\underline{t-1}}) \, dx_{t-1}
\end{aligned}
$$

Markov assumption

Courtesy: C. Stachniss

Recursive Bayes Filter 7

$$
\begin{aligned}
bel(x_t) &= p(x_t \mid z_{1:t}, u_{1:t}) \\
&= \eta\, p(z_t \mid x_t, z_{1:t-1}, u_{1:t})\, p(x_t \mid z_{1:t-1}, u_{1:t}) \\
&= \eta\, p(z_t \mid x_t)\, p(x_t \mid z_{1:t-1}, u_{1:t}) \\
&= \eta\, p(z_t \mid x_t) \int p(x_t \mid x_{t-1}, z_{1:t-1}, u_{1:t}) \\
&\qquad\qquad p(x_{t-1} \mid z_{1:t-1}, u_{1:t})\, dx_{t-1} \\
&= \eta\, p(z_t \mid x_t) \int p(x_t \mid x_{t-1}, u_t)\, p(x_{t-1} \mid z_{1:t-1}, u_{1:t})\, dx_{t-1} \\
&= \eta\, p(z_t \mid x_t) \int p(x_t \mid x_{t-1}, u_t)\, p(x_{t-1} \mid z_{1:t-1}, u_{1:t-1})\, dx_{t-1} \\
&= \underbrace{\eta\, p(z_t \mid x_t)}_{\text{即下一页的correction项}} \underbrace{\int p(x_t \mid x_{t-1}, u_t)\, bel(x_{t-1})\, dx_{t-1}}_{\text{即下一页的prediction步骤}}
\end{aligned}
$$

预测步只关注这一项，因为这一项的先验，使我们已知的所有条件，这就是prediction，未知当前观测

Recursive term

Courtesy: C. Stachniss

预测：成为xt的概率（未知zt） = 积分xt-1（已知xt-1和ut变成xt的概率 * 成为xt-1的概率）
校正：成为xt的概率（已知zt）=xt下得到zt观测的概率 * 成为xt的概率（未知zt）

Prediction and Correction Step

- Bayes filter can be written as a two step process
- **Prediction step** 不引入当前时刻的观测zt，只使用z1:t-1,u1:t得到的预测

$$\overline{bel}(x_t) = \int p(x_t \mid u_t, x_{t-1})\ bel(x_{t-1})\, dx_{t-1}$$

- **Correction step** 引入当前时刻的观测模型，以校正预测

$$bel(x_t) = \eta\, p(z_t \mid x_t)\ \overline{bel}(x_t)$$

Courtesy: C. Stachniss

Motion and Observation Model

- Prediction step

$$\overline{bel}(x_t) = \int \underbrace{p(x_t \mid u_t, x_{t-1})}_{\text{motion model}}\, bel(x_{t-1})\, dx_{t-1}$$

- Correction step

$$bel(x_t) = \eta\, \underbrace{p(z_t \mid x_t)}_{\text{sensor or observation model}}\, \overline{bel}(x_t)$$

Courtesy: C. Stachniss

Bayes Filter

Algorithm `Bayes-filter` (Predict → Update → Normalize)

Require: $bel(x_{t-1}) = p(x_{t-1} \mid z_{1:t-1}, u_{1:t-1})$, control u_t, measurement z_t

Prediction: *propagate belief through motion model (weighted integral / sum)*

for all x_t **do**

$\quad \overline{bel}(x_t) \;=\; \int p(x_t \mid x_{t-1}, u_t)\, bel(x_{t-1})\, \mathrm{d}x_{t-1}$ ▷ discrete case: replace $\int$ with $\sum$

Update: *reweight by likelihood and normalize*

for all x_t **do**

$\quad bel(x_t) \;=\; \eta\, p(z_t \mid x_t)\, \overline{bel}(x_t)$ 对于每一种可能的xt计算概率

return $bel(x_t)$

where $\eta^{-1} = \int p(z_t \mid x_t)\, \overline{bel}(x_t)\, \mathrm{d}x_t$ (or $\sum$ in discrete).

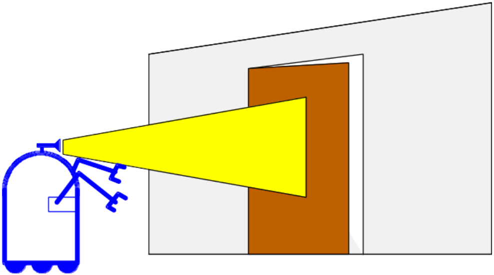
Ingredients:

- Bayes' rule
- Conditional independence (Markov assumptions)
- Law of total probability (marginalization)

Simple Example of State Estimation

- Suppose a robot obtains measurement z, e.g., using its camera;
- What is $p(\mathrm{open}|z)$?

  在得到观测的情况下，开门的概率。



Example

Sensor model (likelihood):

- $p(z = \text{sense\_open}|\text{open}) = 0.6$ 开门的时候检测开概率为0.6
- $p(z = \text{sense\_open}|\neg\text{open}) = 0.3$ 关门的时候检测开的概率为0.3

Prior knowledge (non-informative in this case):

- $p(\text{open}) = p(\neg\text{open}) = 0.5$ 开门关门的先验为0.5

Update/Correction:

$$p(\text{open}|z) = \frac{p(z|\text{open})p(\text{open})}{p(z|\text{open})p(\text{open}) + p(z|\neg\text{open})p(\neg\text{open})}$$

$$p(\text{open}|z = \text{sense\_open}) = \frac{0.6 \times 0.5}{0.6 \times 0.5 + 0.3 \times 0.5} = 0.6667$$

Remark

z *raises the probability that the door is open.*

如果有第二个观测怎么办？如何融合新的信息？

- Suppose our robot obtains another observation z_2.
- How can we integrate this new information?
- More generally, how can we estimate $p(x|z_1, \ldots, z_n)$?

- Given three random variables X, Y, and Z, Bayes' rule relates the prior probability distribution, $p(x|z)$, and the likelihood function, $p(y|x,z)$, as follows.

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

有两个观测，那么始终以其中一个观测z为条件得到全概率。

- Given Z, if X and Y are **conditionally independent** then $p(x,y|z) = p(x|z)p(y|z)$

Example

Height and vocabulary are not independent; but they are conditionally independent if you add age.

https://en.wikipedia.org/wiki/Conditional_independence#Examples

Bayes Filters: Implementation Examples

Linear:

- Kalman Filter: unimodal linear filter
- Information Filter: unimodal linear filter

Nonlinear:

- Extended Kalman Filter: unimodal nonlinear filter with Gaussian noise assumption
- Extended Information Filter: unimodal nonlinear filter with Gaussian noise assumption
- Particle Filter: multimodal nonlinear filter

Readings

- Probabilistic Robotics: Ch. 1 and 2, Understand Example 2.4.2
- State Estimation for Robotics: Ch. 2
- Lecture Notes for Mobile Robotics: Ch. 1

Next Time

- Kalman Filtering
- Readings:
  - Probabilistic Robotics: Ch. 3
  - State Estimation for Robotics: Ch. 3
  - Lecture Notes for Mobile Robotics: Ch. 2 and 6