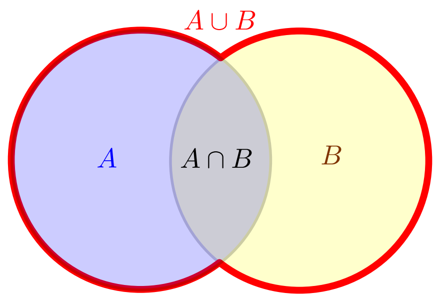
Lecture 19: Recursive Bayesian Estimation

Bayes' Theorem

Let us begin by reviewing Bayes' Theorem, which you may have encountered in a previous course on probability or statistics. Let A and B be two related *events* with probabilities $P(A)$ and $P(B)$. In the context of state estimation, you can think of A as being the state history of the system (i.e., $\boldsymbol{X}_{1:k}$) and B as the noisy past measurement history of that state (i.e., $\boldsymbol{Y}_{1:k}$)—the connection will be made clear later on. Draw a Venn diagram of these events with an intersecting region $A \cap B$.



By comparing the area of B and $A \cap B$ we may intuitively write the conditional probability of "A given B" as a ratio of areas between the probability of "A and B" and probability of just "B":

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \,, \tag{1}$$

and similarly,

$$P(B|A) = \frac{P(A \cap B)}{P(A)} \,. \tag{2}$$

Rearrange (2) for $P(A \cap B)$ and substitute into (1) to give *Bayes' Theorem*:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \,. \tag{3}$$

Bayes' Theorem can be extended to continuous random variables X and Y by letting event $A = \{X = x\}$ and event $B = \{Y = y\}$. Then, (3) becomes

$$P(X = x|Y = y) = \frac{P(Y = y|X = x)P(X = x)}{P(Y = y)} \,. \tag{4}$$

In terms of probability density functions (p.d.f.s), (1) and (2) are:

$$p_{X|Y=y}(x) = \frac{p_{X,Y}(x,y)}{p_Y(y)} \qquad \text{and} \qquad p_{Y|X=x}(x) = \frac{p_{X,Y}(x,y)}{p_X(x)} \tag{5}$$

so that Bayes' Theorem (3) becomes

$$p_{X|Y=y}(x) = \frac{p_{Y|X=x}(y) \cdot p_X(x)}{p_Y(y)} \,. \tag{6}$$

When it is clear from context we often drop the subscripts for each p.d.f. and Bayes' Theorem for two related random vectors is written

$$p(\boldsymbol{x}|\boldsymbol{y}) = \frac{p(\boldsymbol{y}|\boldsymbol{x}) \cdot p(\boldsymbol{x})}{p(\boldsymbol{y})} \,, \tag{7}$$

and for three random vectors:

$$p(\boldsymbol{x}|\boldsymbol{y},\boldsymbol{z}) = \frac{p(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z})}{p(\boldsymbol{y},\boldsymbol{z})} \tag{8}$$

$$= \frac{p(\boldsymbol{y},\boldsymbol{x},\boldsymbol{z})}{p(\boldsymbol{y},\boldsymbol{z})} \tag{9}$$

$$= \frac{p(\boldsymbol{y}|\boldsymbol{x},\boldsymbol{z})p(\boldsymbol{x},\boldsymbol{z})}{p(\boldsymbol{y}|\boldsymbol{z})p(\boldsymbol{z})} \tag{10}$$

$$= \frac{p(\boldsymbol{y}|\boldsymbol{x},\boldsymbol{z})p(\boldsymbol{x}|\boldsymbol{z})p(\boldsymbol{z})}{p(\boldsymbol{y}|\boldsymbol{z})p(\boldsymbol{z})} \tag{11}$$

$$= \frac{p(\boldsymbol{y}|\boldsymbol{x},\boldsymbol{z})p(\boldsymbol{x}|\boldsymbol{z})}{p(\boldsymbol{y}|\boldsymbol{z})} . \tag{12}$$

Bayesian Filtering

We will now apply Bayes' Theorem to the state-estimation problem. Let $\boldsymbol{X}_{1:k} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k\}$ be the set the set of all states up to timestep t_k. Similarly, let $\boldsymbol{Y}_{1:k} = \{\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_k\}$ be the set of all measurements up to timestep t_k. If there is a known physical relationship between $\boldsymbol{X}_{1:k}$ and $\boldsymbol{Y}_{1:k}$, then we may use Bayes' Theorem (7) to update the state given new measurements:

$$p(\boldsymbol{X}_{1:k}|\boldsymbol{Y}_{1:k}) = \frac{p(\boldsymbol{Y}_{1:k}|\boldsymbol{X}_{1:k})p(\boldsymbol{X}_{1:k})}{p(\boldsymbol{Y}_{1:k})} . \tag{13}$$

The above equation is, however, not very useful in its current form. The quantities involved in this expression: $\boldsymbol{X}_{1:k} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k\}$ and $\boldsymbol{Y}_{1:k} = \{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k\}$ are large (and growing!) histories of all the states and measurements. Defining the probability density functions over these objects and applying Bayes' rule is not tractable. Thus, we wish to express this as a simpler recursive relation (i.e., as a function of $p(\boldsymbol{X}_{1:k-1}|\boldsymbol{Y}_{1:k-1})$), and ultimately arrive at an expression for $p(\boldsymbol{x}_k|\boldsymbol{Y}_{1:k})$ rather than $p(\boldsymbol{X}_{1:k}|\boldsymbol{Y}_{1:k})$ (i.e., the probability of the *current* state $\boldsymbol{x}_k$ rather than the state *history*).

Recursive Bayesian Filtering

Begin by splitting the measurements $\boldsymbol{Y}_{1:k}$ as $\boldsymbol{Y}_{1:k} = (\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_k) = (\boldsymbol{y}_k, \boldsymbol{Y}_{1:k-1})$, and similarly with the state, $\boldsymbol{X}_{1:k} = (\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k) = (\boldsymbol{x}_k, \boldsymbol{X}_{1:k-1})$. Expand the first term in the numerator on the RHS of (13) using the conditional probability law to obtain:

$$p(\boldsymbol{Y}_{1:k}|\boldsymbol{X}_{1:k}) = p(\boldsymbol{y}_k, \boldsymbol{Y}_{1:k-1}|\boldsymbol{X}_{1:k}) = p(\boldsymbol{y}_k|\boldsymbol{Y}_{1:k-1}, \boldsymbol{X}_{1:k})p(\boldsymbol{Y}_{1:k-1}|\boldsymbol{X}_{1:k}) . \tag{14}$$

Causaulity assumption. From the causality principle, measurements at time $k-1$ do not depend on the object states at time $\geq k$. Thus, the last term above can be simplified as $p(\boldsymbol{Y}_{1:k-1}|\boldsymbol{X}_{1:k}) = p(\boldsymbol{Y}_{1:k-1}|\boldsymbol{X}_{1:k-1})$ and

$$p(\boldsymbol{Y}_{1:k}|\boldsymbol{X}_{1:k}) = p(\boldsymbol{y}_k|\boldsymbol{Y}_{1:k-1}, \boldsymbol{X}_{1:k})p(\boldsymbol{Y}_{1:k-1}|\boldsymbol{X}_{1:k-1}) . \tag{15}$$

Next, expand the second term in the RHS of (13) as $p(\boldsymbol{X}_{1:k}) = p(\boldsymbol{x}_k, \boldsymbol{X}_{1:k-1})$ and the denomiantor as $p(\boldsymbol{Y}_{1:k}) = p(\boldsymbol{y}_k, \boldsymbol{Y}_{1:k-1})$ and use (15):

$$\begin{aligned}
p(\boldsymbol{X}_{1:k}|\boldsymbol{Y}_{1:k}) &= \frac{p(\boldsymbol{y}_k|\boldsymbol{Y}_{1:k-1}, \boldsymbol{X}_{1:k})p(\boldsymbol{Y}_{1:k-1}|\boldsymbol{X}_{1:k-1})p(\boldsymbol{x}_k, \boldsymbol{X}_{1:k-1})}{p(\boldsymbol{y}_k, \boldsymbol{Y}_{1:k-1})} \\
&= \frac{p(\boldsymbol{y}_k|\boldsymbol{Y}_{1:k-1}, \boldsymbol{X}_{1:k})p(\boldsymbol{Y}_{1:k-1}|\boldsymbol{X}_{1:k-1})p(\boldsymbol{x}_k|\boldsymbol{X}_{1:k-1})p(\boldsymbol{X}_{1:k-1})}{p(\boldsymbol{y}_k|\boldsymbol{Y}_{1:k-1})p(\boldsymbol{Y}_{1:k-1})} \\
&= \frac{p(\boldsymbol{y}_k|\boldsymbol{Y}_{1:k-1}, \boldsymbol{X}_{1:k})p(\boldsymbol{x}_k|\boldsymbol{X}_{1:k-1})}{p(\boldsymbol{y}_k|\boldsymbol{Y}_{1:k-1})} \left[\frac{p(\boldsymbol{Y}_{1:k-1}|\boldsymbol{X}_{1:k-1})p(\boldsymbol{X}_{1:k-1})}{p(\boldsymbol{Y}_{1:k-1})}\right] \\
&= \frac{p(\boldsymbol{y}_k|\boldsymbol{Y}_{1:k-1}, \boldsymbol{X}_{1:k})p(\boldsymbol{x}_k|\boldsymbol{X}_{1:k-1})}{p(\boldsymbol{y}_k|\boldsymbol{Y}_{1:k-1})} \underbrace{[p(\boldsymbol{X}_{1:k-1}|\boldsymbol{Y}_{1:k-1})]}_{\text{prior p.d.f. of object state}} \ .
\end{aligned}$$

Likelihood assumption. If it is assumed that current measurements do not depend on past measurements, and instead only depend on the current state, then the following term (called the likelihood function) simplifies: $p(\boldsymbol{y}_k|\boldsymbol{Y}_{1:k-1}, \boldsymbol{X}_{1:k}) = p(\boldsymbol{y}_k|\boldsymbol{x}_k)$.

Markovian assumption. Another assumption we make is that the system is Markovian (i.e., that the present state only depends on the last state, and not on previous ones before that), simplifies the prior so that $p(\boldsymbol{x}_k|\boldsymbol{X}_{1:k-1}) = p(\boldsymbol{x}_k|\boldsymbol{x}_{k-1})$. Then the recursive Bayesian solution is re-written as

$$p(\boldsymbol{X}_{1:k}|\boldsymbol{Y}_{1:k}) = \frac{p(\boldsymbol{y}_k|\boldsymbol{x}_k)}{p(\boldsymbol{y}_k|\boldsymbol{Y}_{1:k-1})} p(\boldsymbol{x}_k|\boldsymbol{x}_{k-1})p(\boldsymbol{X}_{1:k-1}|\boldsymbol{Y}_{1:k-1}) \ . \tag{16}$$

Since we are interested in $\boldsymbol{x}_k$ rather than $\boldsymbol{X}_{1:k}$, we integrate over all of the previous states $\boldsymbol{X}_{1:k-1}$ as follows (using (16)):

$$\begin{aligned}
p(\boldsymbol{x}_k|\boldsymbol{Y}_{1:k}) &= \int_{\boldsymbol{x}_{k-1}} \dots \int_{\boldsymbol{x}_0} p(\boldsymbol{X}_{1:k}|\boldsymbol{Y}_{1:k}) d\boldsymbol{x}_{k-1} \dots d\boldsymbol{x}_0 \\
&= \int_{\boldsymbol{x}_{k-1}} \dots \int_{\boldsymbol{x}_0} \frac{p(\boldsymbol{y}_k|\boldsymbol{x}_k)}{p(\boldsymbol{y}_k|\boldsymbol{Y}_{1:k-1})} p(\boldsymbol{x}_k|\boldsymbol{x}_{k-1})p(\boldsymbol{X}_{1:k-1}|\boldsymbol{Y}_{1:k-1}) d\boldsymbol{x}_{k-1} \dots d\boldsymbol{x}_0 \\
&= \frac{p(\boldsymbol{y}_k|\boldsymbol{x}_k)}{p(\boldsymbol{y}_k|\boldsymbol{Y}_{1:k-1})} \int_{\boldsymbol{x}_{k-1}} \dots \int_{\boldsymbol{x}_0} p(\boldsymbol{x}_k|\boldsymbol{x}_{k-1})p(\boldsymbol{X}_{1:k-1}|\boldsymbol{Y}_{1:k-1}) d\boldsymbol{x}_{k-1} \dots d\boldsymbol{x}_0 \\
&= \frac{p(\boldsymbol{y}_k|\boldsymbol{x}_k)}{p(\boldsymbol{y}_k|\boldsymbol{Y}_{1:k-1})} \int_{\boldsymbol{x}_{k-1}} \dots \int_{\boldsymbol{x}_0} p(\boldsymbol{x}_k|\boldsymbol{x}_{k-1})p(\boldsymbol{x}_{k-1}, \boldsymbol{X}_{1:k-2}|\boldsymbol{Y}_{1:k-1}) d\boldsymbol{x}_{k-1} \dots d\boldsymbol{x}_0 \\
&= \frac{p(\boldsymbol{y}_k|\boldsymbol{x}_k)}{p(\boldsymbol{y}_k|\boldsymbol{Y}_{1:k-1})} \int_{\boldsymbol{x}_{k-1}} p(\boldsymbol{x}_k|\boldsymbol{x}_{k-1}) \left[\int_{\boldsymbol{x}_{k-2}} \dots \int_{\boldsymbol{x}_0} p(\boldsymbol{x}_{k-1}, \boldsymbol{x}_{k-2}, \dots, \boldsymbol{x}_0|\boldsymbol{Y}_{1:k-1}) d\boldsymbol{x}_0 \dots d\boldsymbol{x}_{k-2}\right] d\boldsymbol{x}_{k-1} \\
&= \frac{p(\boldsymbol{y}_k|\boldsymbol{x}_k)}{p(\boldsymbol{y}_k|\boldsymbol{Y}_{1:k-1})} \underbrace{\int_{\boldsymbol{x}_{k-1}} p(\boldsymbol{x}_k|\boldsymbol{x}_{k-1}) \left[p(\boldsymbol{x}_{k-1}|\boldsymbol{Y}_{1:k-1})\right] d\boldsymbol{x}_{k-1}}_{\text{Chapman-Kolmogorov Equation } = \ p(\boldsymbol{x}_k|\boldsymbol{Y}_{1:k-1})} \ .
\end{aligned}$$

This gives the posterior of the states $\boldsymbol{x}_k$, conditioned on all the previous measurements. The solution of the above Chapman-Kolmogorov Equation (CKE) gives the predicted state $\boldsymbol{x}_k$ given all previous measurements $\boldsymbol{Y}_{1:k-1}$ (excluding the current kth measurement). When a new measurement $\boldsymbol{y}_k$ is obtained, this prediction is corrected by a likelihood factor and then re-normalized.

The normalization factor can be computed as

$$p(\boldsymbol{y}_k|\boldsymbol{Y}_{1:k-1}) = \int p((\boldsymbol{y}_k, \boldsymbol{x}_k)|\boldsymbol{Y}_{1:k-1}) d\boldsymbol{x}_k \tag{17}$$

$$= \int p(\boldsymbol{y}_k|(\boldsymbol{x}_k, \boldsymbol{Y}_{1:k-1}))p(\boldsymbol{x}_k|\boldsymbol{Y}_{1:k-1}) d\boldsymbol{x}_k \;, \tag{18}$$

where the integrals are generic and refer to the appropriate summations or continuous integrals depending on the state being discrete or continuous. But since $\boldsymbol{y}_k$ is completely determined by $\boldsymbol{x}_k$ it follows that $p(\boldsymbol{y}_k|\boldsymbol{x}_k, \boldsymbol{Y}_{1:k-1}) = p(\boldsymbol{y}_k \mid \boldsymbol{x}_k)$ so the above becomes

$$p(\boldsymbol{y}_k|\boldsymbol{Y}_{1:k-1}) = \int p(\boldsymbol{y}_k|\boldsymbol{x}_k)p(\boldsymbol{x}_k|\boldsymbol{Y}_{1:k-1}) d\boldsymbol{x}_k \;. \tag{19}$$

The term $p(\boldsymbol{y}_k|\boldsymbol{x}_k)$ can be determined from the measurement equation, and $p(\boldsymbol{x}_k|\boldsymbol{Y}_{1:k-1})$ is the CKE. Putting it all together, the Bayesian update law becomes:

$$\begin{aligned}
p(\boldsymbol{x}_k|\boldsymbol{Y}_{1:k}) &= \frac{p(\boldsymbol{y}_k|\boldsymbol{x}_k)}{\underbrace{\int_{\boldsymbol{x}_k} p(\boldsymbol{y}_k|\boldsymbol{x}_k)p(\boldsymbol{x}_k|\boldsymbol{Y}_{1:k-1}) d\boldsymbol{x}_k}_{p(\boldsymbol{y}_k|\boldsymbol{Y}_{1:k-1})}} \underbrace{\int_{\boldsymbol{x}_{k-1}} p(\boldsymbol{x}_k|\boldsymbol{x}_{k-1}) \left[p(\boldsymbol{x}_{k-1}|\boldsymbol{Y}_{1:k-1})\right] d\boldsymbol{x}_{k-1}}_{\text{Chapman-Kolmogorov Equation } = \; p(\boldsymbol{x}_k|\boldsymbol{Y}_{1:k-1})} \\
&= \frac{p(\boldsymbol{y}_k|\boldsymbol{x}_k)}{p(\boldsymbol{y}_k|\boldsymbol{Y}_{1:k-1})} p(\boldsymbol{x}_k|\boldsymbol{Y}_{1:k-1}) \\
\text{(posterior)} &= \frac{\text{(measurement likelihood correction)}}{\text{(normalizing factor)}} \text{(motion prediction)}
\end{aligned}$$

The posterior is proportional to the measurement likelihood and motion prediction

$$p(\boldsymbol{x}_k|\boldsymbol{Y}_{1:k}) \propto p(\boldsymbol{y}_k|\boldsymbol{x}_k)p(\boldsymbol{x}_k|\boldsymbol{Y}_{1:k-1}) \tag{20}$$

and in practice we can use the above expression with explicitly computing the normalizing factor.

Algorithm: Grid-Based Recursive Bayesian Estimation

We can formalize the discrete-state-and-measurement and discrete-time recursive Bayesian filter as follows:

1. Ensure the system satisfies the following assumptions:

   (a) **Discrete state-space.** Assume that the system has a finite state space $\mathcal{Q} = \{\boldsymbol{q}_1, \ldots, \boldsymbol{q}_N\}$ with $\boldsymbol{q}_1 \in \mathbb{R}^n$.

   (b) **State transition model.** Assume that the probability of transitioning to $\boldsymbol{q}_i$ at time k from $\boldsymbol{q}_j$ at time $k-1$ is a known function

   $$\boldsymbol{\pi}(\boldsymbol{q}_i|\boldsymbol{q}_j) := p(\boldsymbol{q}_i|\boldsymbol{q}_j) \tag{21}$$

   (c) **Measurement likelihood model.** Assume that the probability of measuring $\boldsymbol{y}$ given the robot's state of $\boldsymbol{q}_j$ is a known function

   $$L(\boldsymbol{y}_k|\boldsymbol{q}_j) := p(\boldsymbol{y}_k|\boldsymbol{q}_j) \tag{22}$$

   If the sensor has additive Gaussian noise the likelihood function can be computed similar to our discussion on maximum likelihood estimation.

2. **Initialize.** Assume that the p.d.f. of the initial state $p(\boldsymbol{x}_0)$ is known. Writing this prior as a vector

   $$\boldsymbol{p}(\boldsymbol{x}_0|\boldsymbol{y}_0) = \begin{bmatrix} p(\boldsymbol{q}_1) \\ p(\boldsymbol{q}_2) \\ \vdots \\ p(\boldsymbol{q}_N) \end{bmatrix} \tag{23}$$

   such that

   $$\sum_{i=1}^{N} p(\boldsymbol{q}_i) = 1 \tag{24}$$

   where $\boldsymbol{y}_0 = \emptyset$ is the empty set of measurements. In the absence of prior information a uniform probability can be assumed $p(\boldsymbol{q}_i) = 1/N$ for all $i = 1, 2, \ldots, N$.

3. For $k = 1, 2, \ldots,$ perform the following:

   (a) Obtain the measurement $\boldsymbol{y}_k$ from the sensor/data stream.

   (b) Retrieve the posterior from the last step $p(\boldsymbol{x}_{k-1}|\boldsymbol{Y}_{1:k-1})$ that becomes the the prior for this step. Note: in the first iteration where $k = 1$ the posterior $p(\boldsymbol{x}_{k-1}|\boldsymbol{Y}_{1:k-1}) = \boldsymbol{p}(\boldsymbol{x}_0|\boldsymbol{y}_0)$ from Step 2 above.

   (c) **Motion-Update.** The motion update is a prediction of $\boldsymbol{x}_k$ given all measurements up to time $k-1$ (this is the CME)

   $$p(\boldsymbol{x}_k|\boldsymbol{Y}_{1:k-1}) = \begin{bmatrix} \sum_{i=1}^{N} \boldsymbol{\pi}(\boldsymbol{q}_1|\boldsymbol{q}_i) p((\boldsymbol{q}_i)_{k-1}|\boldsymbol{Y}_{1:k-1}) \\ \sum_{i=1}^{N} \boldsymbol{\pi}(\boldsymbol{q}_2|\boldsymbol{q}_i) p((\boldsymbol{q}_i)_{k-1}|\boldsymbol{Y}_{1:k-1}) \\ \vdots \\ \sum_{i=1}^{N} \boldsymbol{\pi}(\boldsymbol{q}_N|\boldsymbol{q}_i) p((\boldsymbol{q}_i)_{k-1}|\boldsymbol{Y}_{1:k-1}) \end{bmatrix} \tag{25}$$

The above expression can be computed using two (nested) for loops: one iterating over each element of the above vector and the other computing the required summation as described in the pseudocode below:

Required:

- Discrete state space Q of size $|Q| = N$
- Prior probability vector `p_prior` of size $N \times 1$ (i.e., $p(\boldsymbol{x}_{k-1}|\boldsymbol{Y}_{1:k-1})$)
- State transition model $\boldsymbol{\pi}(\boldsymbol{q}_i|\boldsymbol{q}_j)\ :\ Q^2 \to [0,1]$ that maps two states to a probability

Output:

- Prior probability vector `p_motion` of size $N \times 1$ (i.e., $p(\boldsymbol{x}_k|\boldsymbol{Y}_{1:k-1})$)

---

for $i = \{1, 2, \ldots, N\}$ **do** ▷ for the ith state in the state space (i.e., ith row of (25))
    $r_i \leftarrow 0$ ▷ Initialize ith state probability (temp variable) as zero
    for $j = \{1, 2, \ldots, N\}$ **do** ▷ Iterate over all N states (indexed by j) holding i fixed
        $p_j \leftarrow$ `p_prior`(j) ▷ Prior probability of state $\boldsymbol{q}_j$
        $\tau_{ij} \leftarrow \boldsymbol{\pi}(\boldsymbol{q}_i|\boldsymbol{q}_j)$ ▷ Probability of transitioning to $\boldsymbol{q}_i$ from $\boldsymbol{q}_j$
        $r_i \leftarrow r_i + \tau_{ij} p_j$ ▷ add to summation
    end for
    `p_motion`$(i) \leftarrow r_i$ ▷ Set motion update probability for the ith state
end for

(d) **Measurement Likelihood.** After receiving measurement $\boldsymbol{y}_k$, calculate the likelihood of each state generating that measurement

$$p(\boldsymbol{y}_k|\boldsymbol{x}_k) = \begin{bmatrix} L(\boldsymbol{y}_k|\boldsymbol{q}_1) \\ L(\boldsymbol{y}_k|\boldsymbol{q}_2) \\ \vdots \\ L(\boldsymbol{y}_k|\boldsymbol{q}_N) \end{bmatrix} \tag{26}$$

(e) **Posterior.** Compute the posterior p.d.f. by multiplying (element-wise) the motion update (25) and likelihood (26):

$$p(\boldsymbol{x}_k|\boldsymbol{Y}_{1:k}) \propto p(\boldsymbol{y}_k|\boldsymbol{x}_k) p(\boldsymbol{x}_k|\boldsymbol{Y}_{1:k-1}) = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_N \end{bmatrix} = \begin{bmatrix} L(\boldsymbol{y}_k|\boldsymbol{q}_1) \sum_{i=1}^{N} \boldsymbol{\pi}(\boldsymbol{q}_1|\boldsymbol{q}_i) p((\boldsymbol{q}_i)_{k-1}|\boldsymbol{Y}_{1:k-1}) \\ L(\boldsymbol{y}_k|\boldsymbol{q}_2) \sum_{i=1}^{N} \boldsymbol{\pi}(\boldsymbol{q}_2|\boldsymbol{q}_i) p((\boldsymbol{q}_i)_{k-1}|\boldsymbol{Y}_{1:k-1}) \\ \vdots \\ L(\boldsymbol{y}_k|\boldsymbol{q}_N) \sum_{i=1}^{N} \boldsymbol{\pi}(\boldsymbol{q}_N|\boldsymbol{q}_i) p((\boldsymbol{q}_i)_{k-1}|\boldsymbol{Y}_{1:k-1}) \end{bmatrix} \tag{27}$$

(f) Then normalize

$$p(\boldsymbol{x}_k|\boldsymbol{Y}_{1:k}) = \frac{1}{\sum_{i=1}^{N} b_i} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_N \end{bmatrix} \tag{28}$$

(g) Store the posterior $p(\boldsymbol{x}_k|\boldsymbol{Y}_{1:k})$ to be used as the prior in the next iteration.

References

[1] Sebastian Thrun. Probabilistic robotics. *Communications of the ACM*, 45(3):52–57, 2002.

[2] Dan Simon. *Optimal State Estimation: Kalman, H infinity, and Nonlinear Approaches*. John Wiley & Sons, 2006.