Kalman Filter for State Estimation

Arun Kumar Singh

Institute of Technology, University of Tartu

arun.singh@ut.ee, aks1812@gmail.com

November 23, 2021

Overview

State Estimation

- State Estimation is a framework which allows us to use multiple sources of information to best answer "where you are" or in more general "what your state is?"
- Not all information sources carry equal weight while computing the answer.

Random Variable: Intuitive Understanding

- Let's consider a variable x that stores how much distance you move when you take one step.
- In real-world, value of x might be very difficult to know exactly. Every time you move, the distance moved might be different.
- We call variables like x, a random variable. Each time you query x, you get a different answer.

A Sampling Experiment

Suppose, you take one-step and measure the distance you moved (x). Now, repeat this experiment say 1000 times and record the different x.

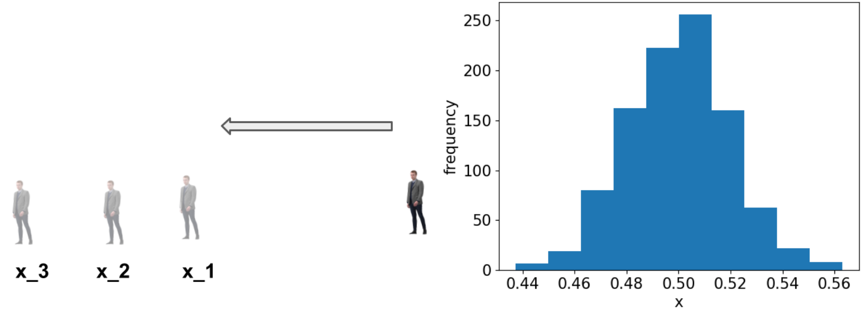


Figure: An Experiment where you have to guess the distance moved in a step.

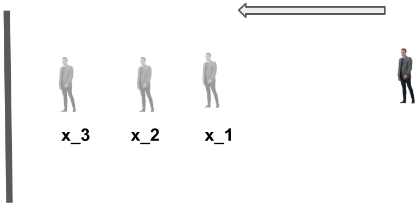
We can see that the most likely value of x is around 0.5 which occurs almost 250 times.

The notion of Probability

- The notion of frequency can be given a more technical term called probability $P(x)$.
- We can roughly say frequency and $P(x)$ are synonymous.
- Is there a way to give more accurate answer of how much distance we moved in one step?

Bringing an additional information

Suppose, you can measure the distance of your current position with respect to a wall. Will that improve your answer?



Formalizing our Example

x	Distance moved
z	Distance from the wall

We want to compute $P(x|z)$

$P(x|z)$	the probability that you moved x units given that you are at a distance z from the wall.

Bayes' Rule over two variables

x	Distance moved
z	Distance from the wall

We want to compute $P(x|z)$

$$\boxed{P(x|z)} = \eta P(z|x)P(x) \tag{1}$$

$P(x\|z)$	The probability that you moved x units given that you are at a distance z from the wall.
$P(z\|x)$	The probability that you are at a distance z from the wall given that you moved x units.
$P(x)$	Probability that distance moved is x units
η	Some normalizing constant (not so important, don't worry about it).

Bringing in additional variables

$$P(x|z,u) = \eta P(z|x,u)P(x|u) \tag{2}$$

u	Number of steps taken
$P(x\|z,u)$	Probability that distance moved is x units given that you are at a distance z from the wall and number of number of steps taken is u
$P(z\|x,u)$	Probability that you are at a distance z from the wall given that distance moved is x units and number of steps taken is u.
$P(x\|u)$	Probability that distance moved is x units given that number of steps taken is u.

Bayes' rule to Bayes' Filter

- Recursively apply Bayes' rule for k iterations by repeatedly taking some steps and noting down the distance to the wall (sensor measurements).
- Gives rise to so called Bayes' Filter. Simple and (can be) computationally efficient. Works quite well most of the times.

Intuitive understanding of Bayes' Filter: Possible locations after one step motion

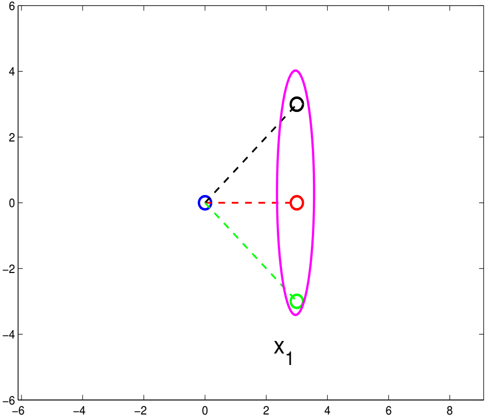


Figure: Motion update iteration 1

Narrowing down possibilities with measurement

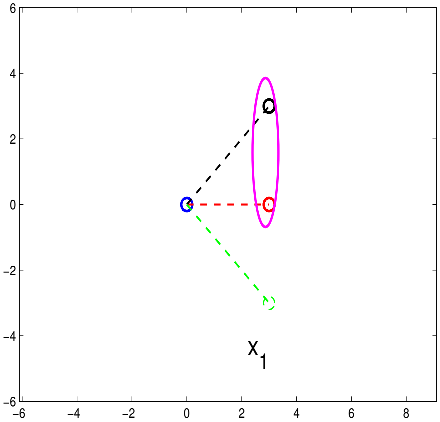


Figure: Measurement update iteration 1

Possible positions after second control step

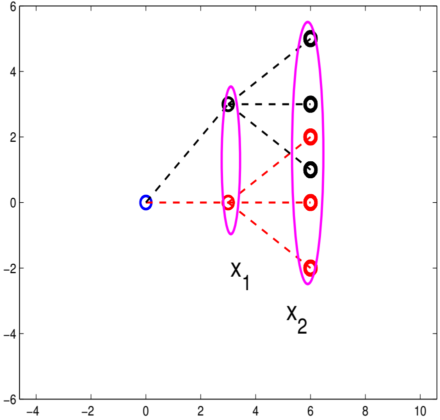


Figure: Motion update iteration 2

Narrowing down possibilities with measurement

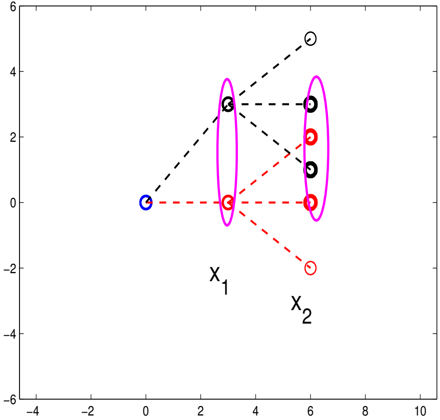


Figure: Measurement update iteration 2

Explosion of possible positions without measurement update

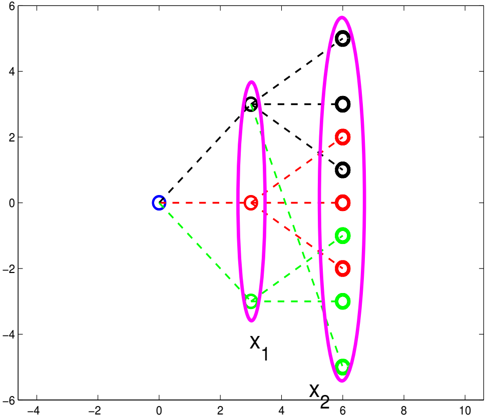


Figure: Iteration 2 without measurement update

Mathematical formulation for Bayes' Filter: Notations

$\mathbf{u}_k$	Control input at k^{th} iteration.
$\mathbf{z}_k$	Measurement obtained at k^{th} iteration.
$\mathbf{x}_k$	Position/State after the k^{th} iteration .

Main Equation

- The end game is to compute the distribution $P(\mathbf{x}_k|\mathbf{z}_{1:k}, \mathbf{u}_{1:k})$
- From Bayes' rule, we have the following

$$\underbrace{P(\mathbf{x}_k|\mathbf{z}_{1:k}, \mathbf{u}_{1:k})}_{bel(\mathbf{x}_k)} = \eta P(\mathbf{z}_k|\mathbf{x}_k, \mathbf{z}_{1:k-1}, \mathbf{u}_{1:k}) \underbrace{P(\mathbf{x}_k|\mathbf{z}_{1:k-1}, \mathbf{u}_{1:k})}_{\overline{bel(\mathbf{x}_k)}} \tag{3}$$

- $bel(\mathbf{x}_k)$ is called the posterior belief. $\overline{bel(\mathbf{x}_k)}$ is called the prior belief.
- From Markovian assumption, we have

$$P(\mathbf{z}_k|\mathbf{x}_k, \mathbf{z}_{1:k-1}, \mathbf{u}_{1:k}) = P(\mathbf{z}_k|\mathbf{x}_k) \tag{4}$$

- Thus:

$$bel(\mathbf{x}_k) = P(\mathbf{z}_k|\mathbf{x}_k)\overline{bel(\mathbf{x}_k)} \tag{5}$$

$$\underbrace{P(\mathbf{x}_k|\mathbf{z}_{1:k},\mathbf{u}_{1:k})}_{bel(\mathbf{x}_k)} = \eta P(\mathbf{z}_k|\mathbf{x}_k,\mathbf{z}_{1:k-1},\mathbf{u}_{1:k})\underbrace{P(\mathbf{x}_k|\mathbf{z}_{1:k-1},\mathbf{u}_{1:k})}_{\overline{bel(\mathbf{x}_k)}}$$

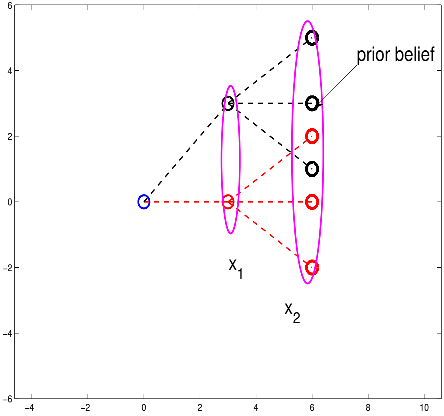


Figure: Prior belief without getting measurement at iteration 2

$$\underbrace{P(\mathbf{x}_k|\mathbf{z}_{1:k},\mathbf{u}_{1:k})}_{bel(\mathbf{x}_k)} = \eta P(\mathbf{z}_k|\mathbf{x}_k,\mathbf{z}_{1:k-1},\mathbf{u}_{1:k})\underbrace{P(\mathbf{x}_k|\mathbf{z}_{1:k-1},\mathbf{u}_{1:k})}_{\overline{bel(\mathbf{x}_k)}}$$

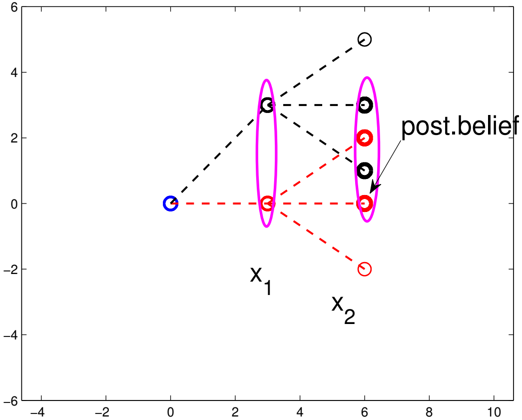


Figure: Post belief after getting measurement at iteration 2

Simplifying prior belief equations

$$\underbrace{P(\mathbf{x}_k|\mathbf{z}_{1:k-1},\mathbf{u}_{1:k})}_{bel(\mathbf{x}_k)} = \int P(\mathbf{x}_k|\mathbf{x}_{k-1},\mathbf{z}_{1:k-1},\mathbf{u}_k)\underbrace{P(\mathbf{x}_{k-1}|\mathbf{z}_{1:k-1},\mathbf{u}_{1:k-1})}_{bel(\mathbf{x}_{k-1})} d\mathbf{x}_{k-1}$$

$$= \int P(\mathbf{x}_k|\mathbf{x}_{k-1},\mathbf{u}_k) bel(\mathbf{x}_{k-1}) d\mathbf{x}_{k-1} \tag{6}$$

Two ingredients to derive the above equation

- From law of total probability, we have

$$P(a|b) = \int P(a|b,c)P(c|b)dc \tag{7}$$

- The Markov assumption

$$P(\mathbf{x}_k|\mathbf{x}_{k-1},\mathbf{z}_{1:k-1},\mathbf{u}_k) = P(\mathbf{x}_k|\mathbf{x}_{k-1},\mathbf{u}_k) \tag{8}$$

Algorithm for Bayesian Filter

for k = 1:n do

- Step 1: Compute $\overline{bel(\mathbf{x}_k)} = \int P(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{u}_k) bel(\mathbf{x}_{k-1})$
- Step 2 : Compute $bel(\mathbf{x}_k) = \eta P(\mathbf{z}_k|\mathbf{x}_k)\overline{bel(\mathbf{x}_k)}$

Kalman Filter: Basic Assumptions

- If we assume that every probability density function in Bayesian Filter is Gaussian, we get the Kalman Filter. Essentially, we assume:

$$bel(\mathbf{x}_k) = N(\mathbf{x}_k; \boldsymbol{\mu}_k, \Sigma_k)$$
$$= det(2\pi\Sigma_k)^{\frac{-1}{2}} \exp(-\frac{1}{2}(\mathbf{x}_k - \boldsymbol{\mu}_k)\Sigma_k^{-1}(\mathbf{x}_k - \boldsymbol{\mu}_k)^T \tag{9}$$

$$\overline{bel(\mathbf{x}_k)} = N(\mathbf{x}_k; \overline{\boldsymbol{\mu}}_k, \overline{\Sigma}_k) \tag{10}$$

$$\overline{bel(\mathbf{x}_k)} = N(\mathbf{x}_k; \overline{\boldsymbol{\mu}}_k, \overline{\Sigma}_k)$$
$$= det(2\pi\overline{\Sigma}_k)^{\frac{-1}{2}} \exp(-\frac{1}{2}(\mathbf{x}_k - \overline{\boldsymbol{\mu}}_k)\Sigma_k^{-1}(\mathbf{x}_k - \overline{\boldsymbol{\mu}}_k)^T \tag{11}$$

- Kalman Filter boils down to obtaining the relationship between $\boldsymbol{\mu}_k$ and $\overline{\boldsymbol{\mu}}_k$ and between Σ_k and $\overline{\Sigma}_k$ (Recall Bayes Filter)

Key Ingredients: Motion Model

$\mathbf{A}$	State matrix of appropriate dimension
$\mathbf{B}$	Control matrix of appropriate dimension
ϵ_k	a random vector, $\epsilon_k = N(0, \mathbf{R}_k)$

Recall

$$x(k) = x(k-1) + (v_x(k) + \epsilon_x)\Delta t \tag{12}$$

$$y(k) = y(k-1) + (v_y(k) + \epsilon_y)\Delta t \tag{13}$$

$$\epsilon_x \sim N(0, \sigma_x^2), \epsilon_x \sim N(0, \sigma_y^2) \tag{14}$$

$$\begin{bmatrix} x(k) \\ y(k) \end{bmatrix} = \mathbf{A}\mathbf{x}_{k-1} + \mathbf{B}\mathbf{u}_k + \boldsymbol{\epsilon}_k \tag{15}$$

$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} \Delta t & 0 \\ 0 & \Delta t \end{bmatrix}, \mathbf{u}_k = \begin{bmatrix} v_x(k) \\ v_y(k) \end{bmatrix}, \boldsymbol{\epsilon}_k = (\epsilon_x, \epsilon_y)^T \tag{16}$$

$\mathbf{A}$	State matrix of appropriate dimension
$\mathbf{B}$	Control matrix of appropriate dimension
ϵ_k	a random vector, $\epsilon_k = N(0, \mathbf{R}_k)$

$$\mathbf{R}_k = \mathbf{B} \begin{bmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{bmatrix} \mathbf{B}^T \tag{17}$$

$$P(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{u}_k) = \overbrace{\mathbf{A}\mathbf{x}_{k-1} + \mathbf{B}\mathbf{u}_k}^{\text{deterministic part}} + \overbrace{\epsilon_k}^{\text{Gaussian Noise}}$$
$$= N(\mathbf{x}_k; -\mathbf{A}\mathbf{x}_{k-1} - \mathbf{B}\mathbf{u}_k, \mathbf{R}_k)$$
$$= det(2\pi\mathbf{R}_k)^{\frac{-1}{2}} \exp(-\frac{1}{2}(\mathbf{x}_k - \mathbf{A}\mathbf{x}_{k-1} - \mathbf{B}\mathbf{u}_k)\mathbf{R}_k^{-1}(\mathbf{x}_k - \mathbf{A}\mathbf{x}_{k-1} - \mathbf{B}\mathbf{u}_k)^T \tag{18}$$

The above equation is just a manifestation of the following rule

$$x + N(0, \sigma) = N(x, \sigma) \tag{19}$$

Key Ingredients: Measurement Update

$\mathbf{C}$	Observation matrix
δ_k	a random vector, $\delta_k = N(0, \mathbf{Q}_k)$

$$
\begin{aligned}
P(\mathbf{z}_k|\mathbf{x}_k) &= \mathbf{C}\mathbf{x}_k + \boldsymbol{\delta}_k \\
&= N(\mathbf{z}_k; \mathbf{x}_k, \mathbf{Q}_k) \\
&= det(2\pi\mathbf{Q}_k)^{\frac{-1}{2}} \exp(-\frac{1}{2}(\mathbf{z}_k - \mathbf{C}\mathbf{x}_{k-1})\mathbf{Q}_k^{-1}(\mathbf{z}_k - \mathbf{C}\mathbf{x}_{k-1})^T
\end{aligned}
\tag{20}
$$

Plugging all distributions into Bayesian Filter

$\overline{bel(\mathbf{x}_k)} = \int P(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{u}_k) bel(\mathbf{x}_{k-1})$ reduces to

$$\overline{bel(\mathbf{x}_k)} = \eta \int exp(-\frac{1}{2}(\mathbf{x}_k - \mathbf{A}\mathbf{x}_{k-1} - \mathbf{B}\mathbf{u}_k)\mathbf{R}_k^{-1}(\mathbf{x}_k - \mathbf{A}\mathbf{x}_{k-1} - \mathbf{B}\mathbf{u}_k)^T$$
$$\exp(-\frac{1}{2}(\mathbf{x}_{k-1} - \boldsymbol{\mu}_k)\Sigma_k^{-1}(\mathbf{x}_{k-1} - \boldsymbol{\mu}_k)^T d\mathbf{x}_{k-1} \tag{21}$$

The integral on the right hand side can be shown to be Gaussian with mean $\mathbf{A}\boldsymbol{\mu}_{k-1} + \mathbf{B}\mathbf{u}_k$ and covariance $\mathbf{A}\Sigma_{k-1}\mathbf{A}^T + \mathbf{R}_k$
Thus, we have

$$\overline{\boldsymbol{\mu}_k} = \mathbf{A}\boldsymbol{\mu}_{k-1} + \mathbf{B}\mathbf{u}_k \tag{22a}$$
$$\overline{\Sigma}_k = \mathbf{A}\Sigma_{k-1}\mathbf{A}^T + \mathbf{R}_k \tag{22b}$$

Plugging all distributions into Bayesian Filter

$bel(\mathbf{x}_k) = \eta P(\mathbf{z}_k|\mathbf{x}_k)\overline{bel(\mathbf{x}_k)}$ reduces to

$$bel(\mathbf{x}_k) = \eta \exp(-\frac{1}{2}(\mathbf{z}_k - \mathbf{C}\mathbf{x}_{k-1})\mathbf{Q}_k^{-1}(\mathbf{z}_k - \mathbf{C}\mathbf{x}_{k-1})^T$$
$$\exp((\mathbf{x}_k - \overline{\boldsymbol{\mu}}_k)\overline{\Sigma}_k^{-1}(\mathbf{x}_k - \overline{\boldsymbol{\mu}}_k)^T) \tag{23}$$

The integral on the right hand side can be shown to be Gaussian with mean $\overline{\boldsymbol{\mu}}_k + \mathbf{P}_k(\mathbf{z}_k - \mathbf{C}\overline{\boldsymbol{\mu}}_k)$ and covariance $(I - \mathbf{P}_k\mathbf{C})\overline{\Sigma}_k$.
Thus, we have

$$\boldsymbol{\mu}_k = \overline{\boldsymbol{\mu}}_k + \mathbf{P}_k(\mathbf{z}_k - \mathbf{C}\overline{\boldsymbol{\mu}}_k) \tag{24a}$$
$$\Sigma_k = (I - \mathbf{P}_k\mathbf{C})\overline{\Sigma}_k \tag{24b}$$
$$\mathbf{P}_k = \overline{\Sigma}_k\mathbf{C}^T(\mathbf{C}\overline{\Sigma}_k\mathbf{C}^T + \mathbf{Q}_k)^{-1} \tag{24c}$$

$\mathbf{P}_k$ is called the Kalman Gain.

Final Algorithm : Kalman Filter

for k = 1: n

$$\overline{bel(\mathbf{x}_k)} = \int P(\mathbf{x}_k|\mathbf{x}_{k-1},\mathbf{u}_k)bel(\mathbf{x}_{k-1}) = \begin{cases} \overline{\boldsymbol{\mu}_k} = \mathbf{A}\boldsymbol{\mu}_{k-1} + \mathbf{B}\mathbf{u}_k \\ \overline{\Sigma}_k = \mathbf{A}\Sigma_{k-1}\mathbf{A}^T + \mathbf{R}_k \end{cases} \tag{25}$$

$$bel(\mathbf{x}_k) = \eta P(\mathbf{z}_k|\mathbf{x}_k)\overline{bel(\mathbf{x}_k)} = \begin{cases} \mathbf{P}_k = \overline{\Sigma}_k\mathbf{C}^T(\mathbf{C}\overline{\Sigma}_k\mathbf{C}^T + \mathbf{Q}_k)^{-1} \\ \boldsymbol{\mu}_k = \overline{\boldsymbol{\mu}}_k + \mathbf{P}_k(\mathbf{z}_k - \mathbf{C}\overline{\boldsymbol{\mu}}_k) \\ \Sigma_k = (I - \mathbf{P}_k\mathbf{C})\overline{\Sigma}_k \end{cases} \tag{26}$$

Note from above that if measurement is very noisy, i.e $\mathbf{Q}_k$ is very high, then the Kalman Filter does not prove to be so useful.

Conclusions

- We studied what state estimation is
- Understood Bayes Filter through the lens of Bayes' Rule
- Saw Kalman Filter as a special case of Bayesian Filter