5 Bayesian estimation

Traditional estimation only uses measurements; the most common technique being maximum likelihood estimation. Bayesian estimation also includes prior information regarding the belief of the state.

The BLU estimator used in a Kalman filter is a Bayesian estimator but only if the beliefs are Gaussian. When the beliefs are not Gaussian, Bayes' theorem needs to be applied to correctly determine the posterior PDF.

5.1 Bayes' theorem

Bayes' theorem provides a more general method for fusing measurements with prior knowledge to obtain a posterior PDF. In terms of probabilities, Bayes' theorem is:

$$P(X|Z) = \frac{P(Z|X)P(X)}{P(Z)}. \tag{5.1}$$

In terms of continuous random variables, Bayes' theorem is:

$$f_{X|Z}(x|z) = \frac{L_{X|Z}(x|z)f_X(x)}{f_Z(z)}. \tag{5.2}$$

This specifies the posterior PDF for $X = x$ given $Z = z$, where Z is the measurement random variable and z is a specific measurement.

There are four different quantities involved:

$f_{X|Z}(x|z)$ is the *posterior PDF* of X,

$L_{X|Z}(x|z)$ is the *measurement likelihood function*,

$f_X(x)$ is the *prior PDF* of X, and

$f_Z(z)$ is the *evidence PDF*.

5.1.1 *Posterior normalisation*

The evidence, $f_Z(z)$, is not known. However, since it is independent of x, it can be replaced by a normalising factor, $1/\eta$, so the posterior PDF integrates to one,

$$f_{X|Z}(x|z) = \eta\, f_X(x) L_{X|Z}(x|z). \tag{5.3}$$

The normalising factor, η, is calculated using

$$\eta = \frac{1}{\int_{-\infty}^{\infty} f_X(x) L_{X|Z}(x|z) \mathrm{d}x}. \tag{5.4}$$

5.1.2 *Bayes' theorem example*

Figure 5.1 shows an example of Bayes' theorem.

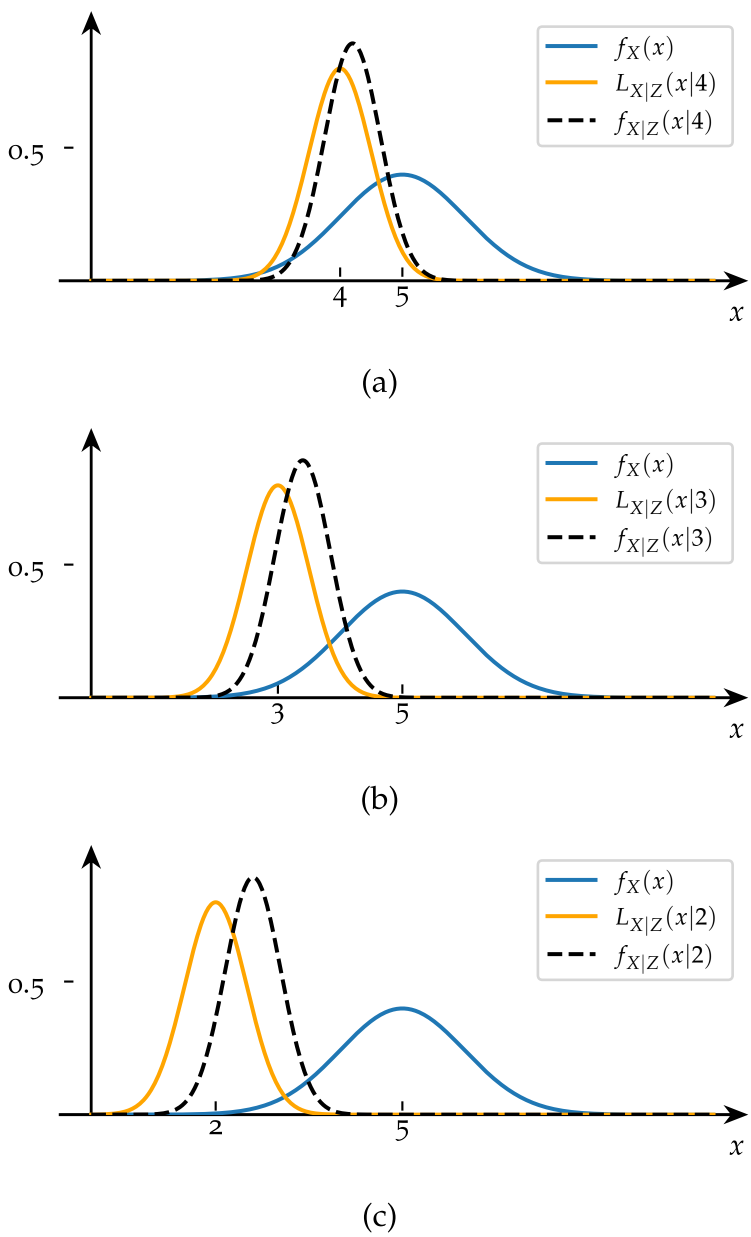

Figure 5.1: Bayes' theorem example 1: (a) $z = 4$, (b) $z = 3$, (c) $z = 2$. In each case the posterior distribution, $f_{X|Z}(x|z)$, has the smallest variance. Note, the measured values only affect the mean of the posterior and not the variance.

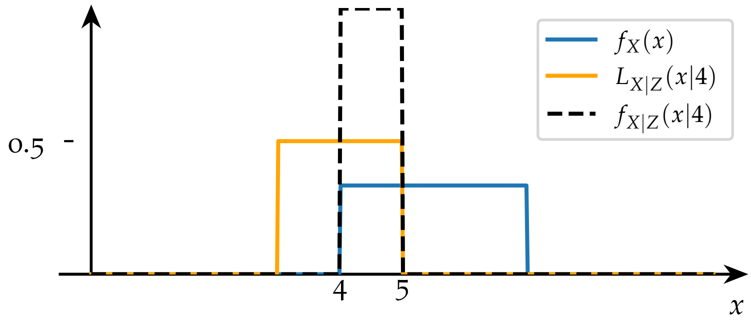

Figure 5.2: Bayes' theorem example for uniform distributions: $z = 4$.

5.2 *Bayesian estimates*

The posterior PDF provides the probability density for each possible value of X. From this information we have to choose a best estimate. There are two common estimators for this: MAP and MMSE.

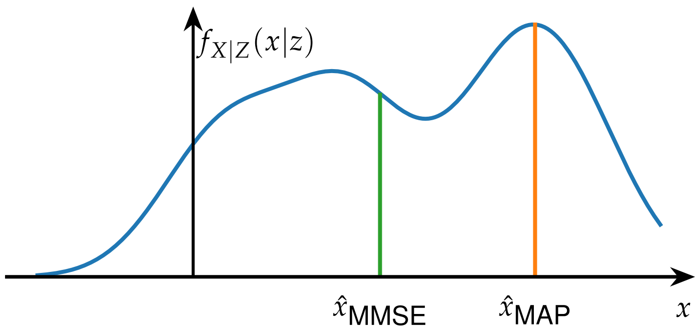

Figure 5.3: MAP and MMSE estimates.

The *maximum a posteriori* (MAP) estimate is the mode of the posterior PDF:

$$\hat{x}_{\text{MAP}} = \arg\max_{x} f_{X|Z}(x|z). \tag{5.5}$$

Note, this is similar to the maximum likelihood (ML) estimate but modified by prior information.

The *minimum mean squared error* (MMSE) estimate is the mean of the posterior PDF:

$$\hat{x}_{\text{MMSE}} = \mathrm{E}\left[X|Z{=}z\right] = \int_{-\infty}^{\infty} x f_{X|Z}(x|z)\mathrm{d}x. \tag{5.6}$$

Note, if the posterior PDF is Gaussian, MAP and MMSE give the same estimate[1].

[1] This is what the Kalman filter tracks.

5.3 *Bayes' theorem for probabilities*

Bayes' theorem[2] is usually stated in probability form as

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B)}, \tag{5.7}$$

since

$$P(A \cap B) = P(A|B)\,P(B) = P(B|A)\,P(A). \tag{5.8}$$

Here $P(A|B)$ is the *conditional probability*[3] for A given B and $P(B|A)$ is the *conditional probability* for B given A.

[2] Bayes' rule is Bayes' theorem in odds form.

[3] A classic example of using conditional probabilities is the Monty Hall game show problem: *Suppose you're on a game show, and you're given the choice of three doors: Behind one door is a car; behind the others a donkey. You pick a door, say No.1, and the host, who knows what's behind the doors, opens another door, say No.3, which has a donkey. He then says to you, "Do you want to pick door No.2?" Is it to your advantage to switch your choice?*

5.4 *Summary*

1. The PDF before a measurement is called the prior PDF.
2. The PDF after a measurement is called the posterior PDF.
3. The posterior PDF is computed from a prior PDF using Bayes' theorem and the likelihood function.
4. The posterior PDF has a smaller variance than the prior PDF.
5. Bayes' theorem for continuous random variables is

$$f_{X|Z}(x|z) = \frac{L_{X|Z}(x|z) f_X(x)}{f_Z(z)}. \tag{5.9}$$

6. Bayes' theorem for continuous random variables simplifies to:

$$f_{X|Z}(x|z) = \eta\, L_{X|Z}(x|z) f_X(x), \tag{5.10}$$

   where η is a normalising factor.

7. If both $f_X(x)$ and $L_{X|Z}(x|z)$ are Gaussian, then the posterior PDF, $f_{X|Z}(x|z)$, is also Gaussian[4],

$$f_{X|Z}(x|z) = \mathcal{N}\left(\frac{\mathrm{Var}\left[Z|X\right] \mathrm{E}\left[X\right] + \mathrm{Var}\left[X\right] z}{\mathrm{Var}\left[Z|X\right] + \mathrm{Var}\left[X\right]}, \frac{\mathrm{Var}\left[Z|X\right] \mathrm{Var}\left[X\right]}{\mathrm{Var}\left[Z|X\right] + \mathrm{Var}\left[X\right]}\right). \tag{5.11}$$

[4] A Gaussian multiplied by a Gaussian produces a Gaussian.