Kalman Filter using MATLAB
2: Mathematical preliminaries

by
Midhun T. Augustine

Overview

1. Basic probability and statistics
   - Probability space
   - Probability density function
   - Expectation and variance
   - Gaussian distribution
2. Conditional probability
   - Conditional expectation and variance

Basic probability and statistics

Probability space

- **Random experiment:** is an experiment or process for which outcomes cannot be predicted exactly.
- **Probability space:** is denoted as $(\mathbb{S}, \mathbb{E}, Pr)$ which contains:
  1. Sample space $\mathbb{S}$: is the set of all possible outcomes.
  2. Event space $\mathbb{E}$: contains the set of all events (i.e., the set of all subsets of $\mathbb{S}$).
  3. Probability function Pr: assigns each event in $\mathbb{E}$ a real number between 0 and 1, known as the probability measure of the event, i.e., $Pr : \mathbb{E} \rightarrow [0, 1]$.
- **Random variable:** A random variable x is a real-valued function defined on the sample space $\mathbb{S}$, i.e., $x : \mathbb{S} \rightarrow \mathbb{R}$, and a random vector $\mathbf{x} = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix}^T$ contains random variables as its elements, i.e., $\mathbf{x} : \mathbb{S} \rightarrow \mathbb{R}^n$.

Probability density function

- **Probability density function** (PDF) $f(x)$: is a function that assigns each random variable, a number between 0 and 1, i.e. $f : \mathbb{R} \to [0, 1]$. Similarly, for a random vector $\mathbf{x}$, we have $f : \mathbb{R}^n \to [0, 1]$.
- For continuous random variables, the PDF is used to specify the probability of the random variable to take a value within an interval:

$$Pr(a \leq x \leq b) = \int_a^b f(x)dx \tag{1}$$

- For discrete random variables, the PDF, which is known by the name probability mass function (PMF), is used to specify the probability of the random variable to take a particular value:

$$Pr(x = a) = f(a) \tag{2}$$

Expectation

- For a random variable x with PDF $f(x)$ the expectation is a parameter that indicates the average value, and is defined as

$$E(x) = \int x f(x) dx = \sum x_i f(x_i) \tag{3}$$

- For a random vector $\mathbf{x} = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix}^T$ the expectation becomes

$$\mathbf{E}(\mathbf{x}) = \begin{bmatrix} E(x_1) & E(x_2) & \dots & E(x_n) \end{bmatrix}^T \tag{4}$$

Variance

- For a random variable x the variance is defined as

$$V(x) = E([x - E(x)][x - E(x)]) = E([x - E(x)]^2) \tag{5}$$

which indicates the measure of deviation from the mean.

- For a random vector $\mathbf{x}$, the variance is defined as

$$\mathbf{V}(\mathbf{x}) = \mathbf{E}\big([\mathbf{x} - \mathbf{E}(\mathbf{x})][\mathbf{x} - \mathbf{E}(\mathbf{x})]^T\big) \tag{6}$$

and $\mathbf{V}(\mathbf{x}) \in \mathbb{R}^{n \times n}$ which is known by the names: variance matrix, covariance matrix, etc.

Covariance

- For the random variables x and y the covariance is defined as

$$V(x, y) = E([x - E(x)][y - E(y)]). \tag{7}$$

- For the random vectors $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{y} \in \mathbb{R}^p$ the covariance matrix becomes

$$\begin{aligned} \mathbf{V}(\mathbf{x}, \mathbf{y}) &= \mathbf{E}\big([\mathbf{x} - \mathbf{E}(\mathbf{x})][\mathbf{y} - \mathbf{E}(\mathbf{y})]^T\big) \\ &= \begin{bmatrix} V(x_1, y_1) & V(x_1, y_2) & \dots & V(x_1, y_p) \\ V(x_2, y_1) & V(x_2, y_2) & \dots & V(x_2, y_p) \\ \vdots & \vdots & & \vdots \\ V(x_n, y_1) & V(x_n, y_2) & \dots & V(x_n, y_p) \end{bmatrix} \end{aligned} \tag{8}$$

  and $\mathbf{V}(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^{n \times p}$, $\mathbf{V}(\mathbf{x}, \mathbf{x}) = \mathbf{V}(\mathbf{x}) \in \mathbb{R}^{n \times n}$.

- $\mathbf{V}(\mathbf{x}, \mathbf{y}) = \mathbf{0}$, if the random vectors $\mathbf{x}$ and $\mathbf{y}$ are independent.

Minimum Mean Square Error (MMSE) Estimator

- Estimator error vector: $\mathbf{x}_{e_k} = \mathbf{x}_k - \hat{\mathbf{x}}_k$
- Mean Square Error

$$\begin{aligned} MSE &= E(x_{e_{k1}}^2 + ... + x_{e_{kn}}^2) = E(\mathbf{x}_{e_k}^T \mathbf{x}_{e_k}) \\ &= E(Trace(\mathbf{x}_{e_k} \mathbf{x}_{e_k}^T)) = Trace(\mathbf{E}(\mathbf{x}_{e_k} \mathbf{x}_{e_k}^T)) \\ &= Trace(\mathbf{V}(\mathbf{x}_{e_k})) \end{aligned} \tag{9}$$

- MMSE Estimator: is the estimator achieving minimal MSE

$$\begin{aligned} \hat{\mathbf{x}}_{k_{MMSE}} &= \underset{\hat{\mathbf{x}}_k}{argmin}\ MSE \\ &= \underset{\hat{\mathbf{x}}_k}{argmin}\ Trace(\mathbf{V}(\mathbf{x}_{e_k})) \end{aligned} \tag{10}$$

Gaussian distribution

- **Gaussian distribution (Normal distribution):** is characterised by the probability density function:

$$f(x) = \frac{1}{\sqrt{2\pi V(x)}} e^{-\frac{1}{2}\frac{[x-E(x)]^2}{V(x)}} \tag{11}$$

- In Gaussian distribution, for $x = E(x)$ we obtain $f(x) = \frac{1}{\sqrt{2\pi V(x)}}$, and for $x = E(x) \pm \sqrt{V(x)}$ we obtain $f(x) = \frac{1}{\sqrt{2\pi V(x)}} e^{-\frac{1}{2}} = \frac{1}{\sqrt{2\pi e V(x)}}$.

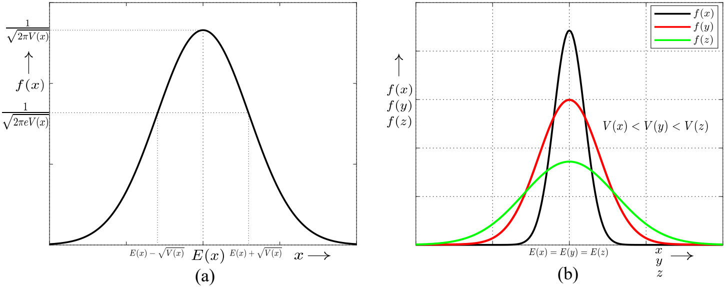

Figure 1: Gaussian distribution graphical representation

Gaussian distribution

- Let x be a random variable with Gaussian PDF $f(x)$, then

$$
\begin{aligned}
&\int_{E(x)-\sqrt{V(x)}}^{E(x)+\sqrt{V(x)}} f(x)dx = 0.68 \\
&\int_{E(x)-2\sqrt{V(x)}}^{E(x)+2\sqrt{V(x)}} f(x)dx = 0.95 \\
&\int_{-\infty}^{\infty} f(x)dx = 1
\end{aligned}
\tag{12}
$$

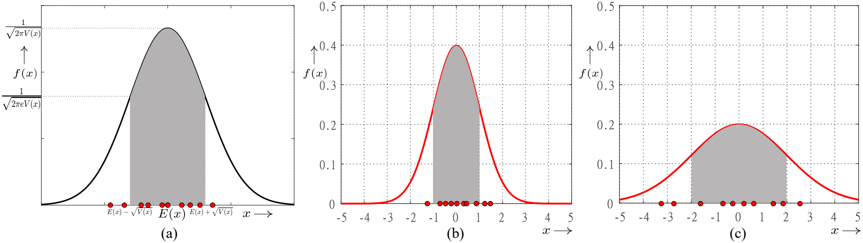

Figure 2: Gaussian distribution: (a) with variance V(x) (b) with variance 1 (c) with variance 4

Multivariate Gaussian distribution

- **Multivariate (Joint) Gaussian distribution**: for the random vector $\mathbf{x}$ is defined as

$$f(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^n det(\mathbf{V}(\mathbf{x}))}} e^{-\frac{1}{2}[\mathbf{x}-\mathbf{E}(\mathbf{x})]^T \mathbf{V}(\mathbf{x})^{-1}[\mathbf{x}-\mathbf{E}(\mathbf{x})]} \tag{13}$$

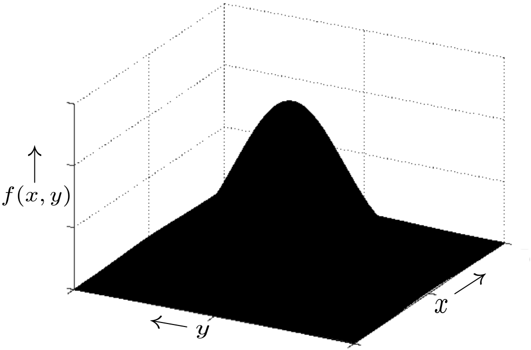

Figure 3: Joint Gaussian distribution: graphical representation

Conditional probability

Conditional probability

- **Conditional probability**: Let x and y be two random variables, the conditional PDF $f(x|y)$ denotes the probability of x given y.
- $f(x|y=a)$ gives the probability of x given that y takes the value a, where $a \in \mathbb{R}$ is any scalar.
- **Conditional expectation:** For the random variables x, y the conditional expectation is

$$E(x|y) = \sum x_i f(x_i|y) \tag{14}$$

- Similarly, for the random vectors $\mathbf{x}$ and $\mathbf{y}$, the conditional expectation is denoted by $\mathbf{E}(\mathbf{x}|\mathbf{y})$ and

$$\mathbf{E}(\mathbf{x}|\mathbf{y}) = \begin{bmatrix} E(x_1|\mathbf{y}) & E(x_2|\mathbf{y}) & \dots & E(x_n|\mathbf{y}) \end{bmatrix}^T \tag{15}$$

Conditional probability

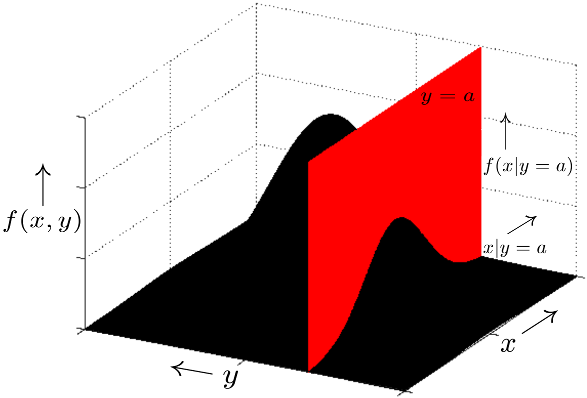

Figure 4: Conditional probability: graphical representation

- For jointly Gaussian random variables x and y, the conditional distribution $f(x|y = a)$ is Gaussian (see Fig. 4).

Conditional Expectation and Variance: Derivation

- If the random variables x and y are independent, then $E(x|y) = E(x)$.
- Use the transformation $\mathbf{z} = \mathbf{x} - \mathbf{L}\mathbf{y}$ with $\mathbf{L} = \mathbf{V}(\mathbf{x},\mathbf{y})\mathbf{V}(\mathbf{y})^{-1}$ results in

$$\begin{aligned} \mathbf{V}(\mathbf{z},\mathbf{y}) &= \mathbf{V}(\mathbf{x}-\mathbf{L}\mathbf{y},\mathbf{y}) = \mathbf{V}(\mathbf{x},\mathbf{y}) - \mathbf{L}\mathbf{V}(\mathbf{y}) \\ &= \mathbf{V}(\mathbf{x},\mathbf{y}) - \mathbf{V}(\mathbf{x},\mathbf{y})\mathbf{V}(\mathbf{y})^{-1}\mathbf{V}(\mathbf{y}) = \mathbf{0} \end{aligned} \tag{16}$$

  which implies $\mathbf{z}$ and $\mathbf{y}$ are independent.
- This gives the expression for conditional expectation as

$$\begin{aligned} \mathbf{E}(\mathbf{x}|\mathbf{y}) &= \mathbf{E}(\mathbf{z}+\mathbf{L}\mathbf{y}|\mathbf{y}) = \mathbf{E}(\mathbf{z}|\mathbf{y}) + \mathbf{E}(\mathbf{L}\mathbf{y}|\mathbf{y}) \\ &= \mathbf{E}(\mathbf{z}) + \mathbf{L}\mathbf{y} = \mathbf{E}(\mathbf{x}) - \mathbf{L}\mathbf{E}(\mathbf{y}) + \mathbf{L}\mathbf{y} \qquad \because\ \mathbf{E}(\mathbf{z}|\mathbf{y}) = \mathbf{E}(\mathbf{z}) \\ &= \mathbf{E}(\mathbf{x}) + \mathbf{L}[\mathbf{y} - \mathbf{E}(\mathbf{y})] = \mathbf{E}(\mathbf{x}) + \mathbf{V}(\mathbf{x},\mathbf{y})\mathbf{V}(\mathbf{y})^{-1}[\mathbf{y} - \mathbf{E}(\mathbf{y})]. \end{aligned} \tag{17}$$

- Similarly, an expression for the conditional variance can be obtained as

$$\mathbf{V}(\mathbf{x}|\mathbf{y}) = \mathbf{V}(\mathbf{x}) - \mathbf{L}\mathbf{V}(\mathbf{y},\mathbf{x}) = \mathbf{V}(\mathbf{x}) - \mathbf{V}(\mathbf{x},\mathbf{y})\mathbf{V}(\mathbf{y})^{-1}\mathbf{V}(\mathbf{y},\mathbf{x}). \tag{18}$$

Thank you