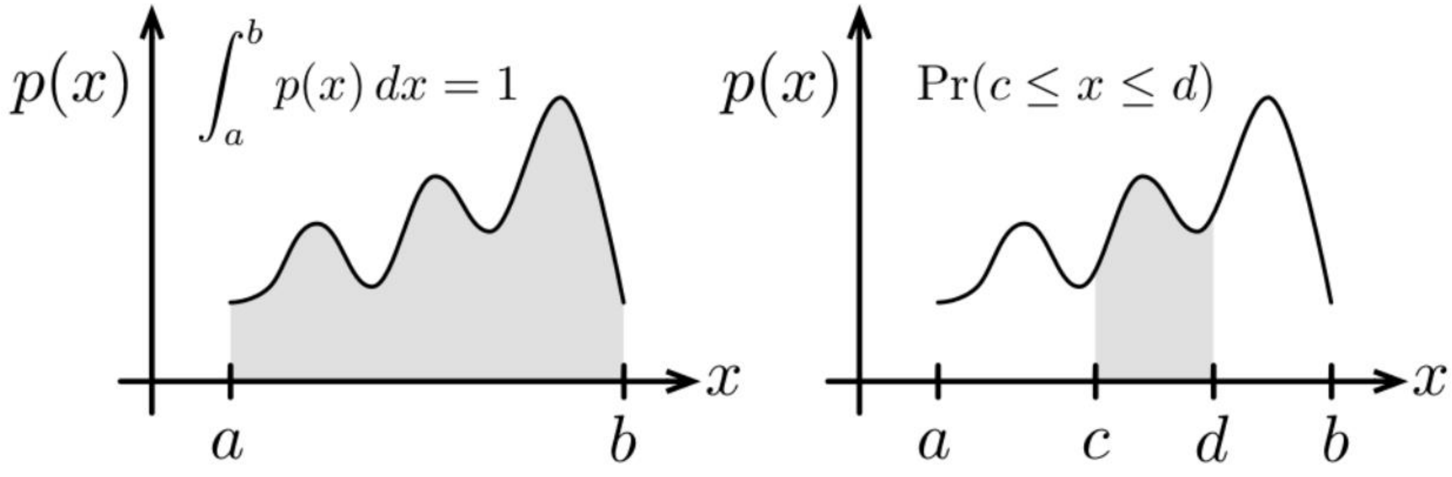
Bayesian filtering

Ali AlBeladi

January 14, 2025

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:

$$p(x, y) = p(X = x \; and \; Y = y);$$

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

- The conditional probability of X given Y is:

$$p(x|y) = \frac{p(x, y)}{p(y)} \qquad p(y) > 0.$$

Marginalization

- Given $p(x,y)$, the marginal distribution of X can be computed by summing (integration) over Y.
- The law of total probability is its variant which uses the conditional probability definition

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

and for continuous random variables is

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

Bayes' Rule

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

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

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

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

Bayes' Rule (Example)

An autonomous car is approaching a traffic light which can be either green, yellow, or red. The car is programmed to be conservative and thus it will stop if it detects a yellow or red light; otherwise it will continue driving. Previous tests have demonstrated that due to sensor imperfections, the car will drive through (without stopping) 10% of yellow lights, 95% of green lights, and 1% of red lights. The traffic light is on a continuous cycle (30 seconds green, 5 seconds yellow, 25 seconds red). You are riding in the car and are busy working on your Mobile Robotics project (i.e., not watching the road, light, etc.). You feel the car stop as it approaches the traffic light described above. What is the probability that the traffic light was yellow when the vehicle sensed it?

Bayes' Rule (Example)

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:* $P(Y|S)$

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

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

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

Bayes' Rule with Prior Knowledge

- 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.

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

- Given Z, if X and Y are conditionally independent then

$$p(x,y|z) \ = \ p(x|z)p(y|z)$$

- Height and vocabulary are not independent; but they are conditionally independent if age is given

Univariate Normal Distribution

The univariate (one-dimensional) *Gaussian (or normal) distribution* with mean μ and variance σ^2 has the following Probability Density Function (PDF).

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp(-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2})$$

We often write $x \sim \mathcal{N}(\mu, \sigma^2)$ or $\mathcal{N}(x; \mu, \sigma^2)$ to imply that x follows a Gaussian distribution with mean $\mu = \mathbb{E}[x]$ and variance $\sigma^2 = \mathbb{V}[x]$.

Multivariate Normal Distribution

The multivariate Gaussian (normal) distribution of an n-dimensional random vector $x \sim \mathcal{N}(\mu, \Sigma)$, with mean $\mu = \mathbb{E}[x]$ and covariance $\Sigma = \mathrm{Cov}[x] = \mathbb{E}[(x-\mu)(x-\mu)^{\mathsf{T}}]$ is

$$p(x) = (2\pi)^{-\frac{n}{2}} |\Sigma|^{-\frac{1}{2}} \exp(-\frac{1}{2}(x-\mu)^{\mathsf{T}}\Sigma^{-1}(x-\mu))$$

Let $x = \mathrm{vec}(x_1, x_2)$ and $x \sim \mathcal{N}(\mu, \Sigma)$ where

$$\mu = \begin{bmatrix} 0.0 \\ 0.5 \end{bmatrix}, \Sigma = \begin{bmatrix} 0.8 & 0.3 \\ 0.3 & 1.0 \end{bmatrix}$$

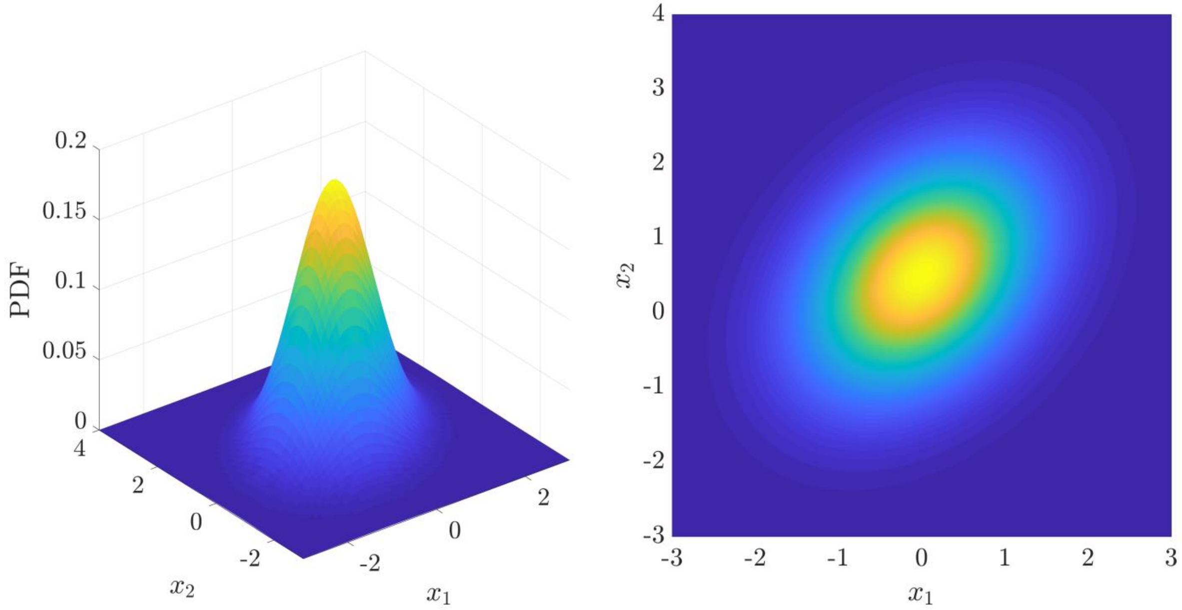


Figure: Left, two-dimensional PDF; right, top view of the first plot.

Marginalization and Conditioning of Normal Distribution

Let x and y be jointly Gaussian random vectors

$$\begin{bmatrix} x \\ y \end{bmatrix} \sim \mathcal{N}(\begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix}, \begin{bmatrix} A & C \\ C^\mathsf{T} & B \end{bmatrix})$$

then the marginal distribution of x is

$$x \sim \mathcal{N}(\mu_x, A)$$

and the conditional distribution of x given y is

$$x|y \sim \mathcal{N}(\mu_x + CB^{-1}(y - \mu_y), A - CB^{-1}C^\mathsf{T})$$

Marginalization and Conditioning of Normal Distribution

Let $x = \mathrm{vec}(x_1, x_2)$ and $x \sim \mathcal{N}(\mu, \Sigma)$ where

$$\mu = \begin{bmatrix} 0.0 \\ 0.5 \end{bmatrix}, \Sigma = \begin{bmatrix} 0.8 & 0.3 \\ 0.3 & 1.0 \end{bmatrix}$$

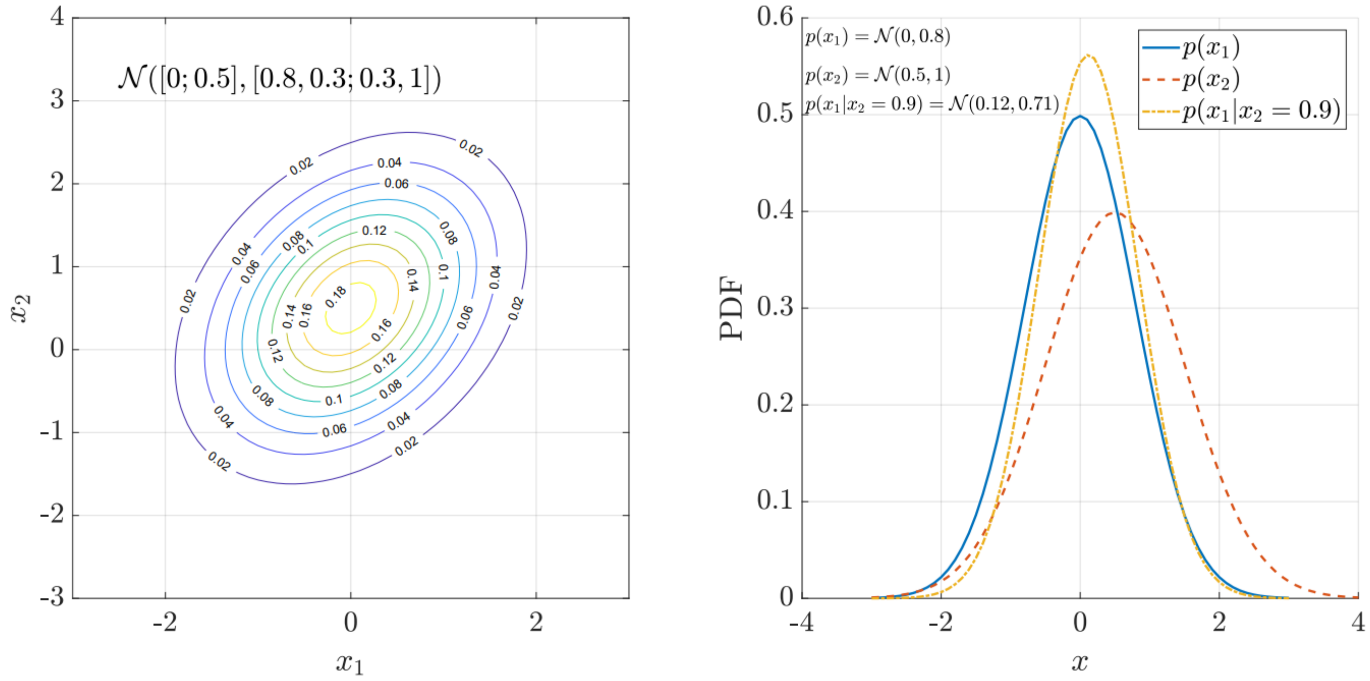


Figure: Left, the contour plot of the PDF; right, the marginals and the conditional distribution of $p(x_1|x_2 = 0.9)$.

Affine Transformation of a Multivariate Gaussian

Suppose $x \sim \mathcal{N}(\mu, \Sigma)$ and $y = Ax + b$.

Then $y \sim \mathcal{N}(A\mu + b, A\Sigma A^{\mathsf{T}})$.

$$\mathbb{E}[y] = \mathbb{E}[Ax + b] = A\mathbb{E}[x] + b = A\mu + b$$

$$\begin{aligned}
&\mathrm{Cov}[y] = \mathbb{E}[(y - \mathbb{E}[y])(y - \mathbb{E}[y])^{\mathsf{T}}] \\
&= \mathbb{E}[(Ax - A\mu)(Ax - A\mu)^{\mathsf{T}}] = A\mathbb{E}[(x - \mu)(x - \mu)^{\mathsf{T}}]A^{\mathsf{T}} \\
&= A\Sigma A^{\mathsf{T}}
\end{aligned}$$

Bayes Filters: Framework

Given:

- 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:

- The state x_t of dynamical system
- The posterior of state is called belief $bel(x_t) = p(x_t|z_{1:t}, u_{1:t})$

Algorithm 1 `Bayes-filter`

Require: Belief $bel(x_{t-1}) = p(x_{t-1}|z_{1:t-1}, u_{1:t-1})$, action u_t, measurement z_t;

1: **for** all state variables **do**
2: $\overline{bel}(x_t) = \int p(x_t|x_{t-1}, u_t) bel(x_{t-1}) \mathrm{d}x_{t-1}$ // Predict using action/control input u_t
3: $bel(x_t) = \eta p(z_t|x_t)\overline{bel}(x_t)$ // Update using perceptual data z_t
4: **return** $bel(x_t)$

Bayes Filters: Implementation Examples

Linear:

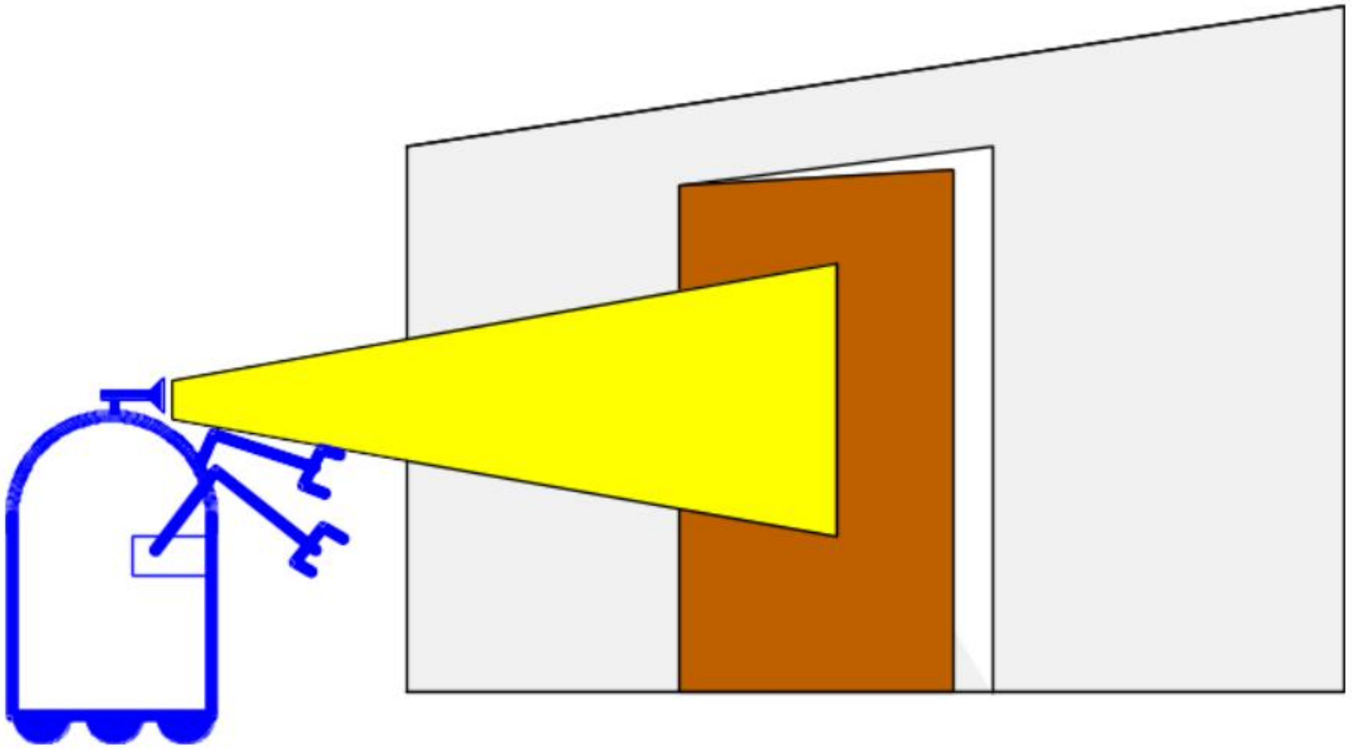
- Kalman Filter: unimodal linear filter
- Information Filter: unimodal linear filter

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

Simple Example of State Estimation

Suppose a robot obtains measurement z, e.g., using its camera;



What is $p(\text{open}|z)$?