Applied Signal Processing
Lecture 12

Tomas McKelvey

Department of Electrical Engineering
Chalmers University of Technology

- Motivating example
- Multivariate Normal Distribution
- Conditional Normal Distribution
- Kalman filter equations

Multivariate Normal Distribution

An n-dimensional real valued *multivariate* random variable Z with a *normal distribution* has the probability density function (PDF)

$$p_Z(z) = \frac{1}{\sqrt{(2\pi)^n \det Q}} \exp\left(-\frac{1}{2}(z-\mu)^T Q^{-1}(z-\mu)\right)$$

$\mathbf{E}\, Z = \mu \in \mathbb{R}^n$ is the mean value

$\mathbf{E}(Z-\mu)(Z-\mu)^T = Q \in \mathbb{R}^{n\times n}$ is the positive definite covariance matrix.

We denote

$$Z \sim N(\mu, Q)$$

Scaled sum of multivariate normal random variables

If the variables $Z_i \sim N(\mu_i, Q_i)$ are statistically independent and normally distributed then

$$\sum_i A_i Z_i \sim N\left(\sum_i A_i \mu_i, \sum_i A_i Q_i A_i^T\right)$$

where A_i are a scaling matrices such that $\sum_i A_i Q_i A_i^T$ is positive definite.

Conditional Distribution

Assume variables X and Y have a joint PDF $p_{X,Y}(x,y)$. The distribution of X given that we know the outcome of $Y = y$ is the conditional distribution

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

Conditional Normal Distribution

Assume we partition a multivariate normally distributed random random variable into two parts $Z = \left[\begin{smallmatrix} X \\ Y \end{smallmatrix}\right]$ and write

$$Z = \begin{bmatrix} X \\ Y \end{bmatrix} \sim N\left(\begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix}, \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{12}^T & Q_{22} \end{bmatrix}\right).$$

Then the conditional distribution of X given that we know $Y = y$ is also a normal distribution

$$p_{X|Y}(x|Y = y) \sim N\left(\mu_x + Q_{12}Q_{22}^{-1}(y - \mu_y), Q_{11} - Q_{12}Q_{22}^{-1}Q_{12}^T\right)$$

Data model

A linear discrete time stochastic process of order n can be written as a multidimensional first order difference equation

$$
\begin{aligned}
x(k+1) &= Ax(k) + w(k) \\
y(k) &= Cx(k) + v(k)
\end{aligned}
$$

where $x(k) \in \mathbb{R}^n$ is the state, $A \in \mathbb{R}^{n \times n}$ is the state-transition matrix, $y(k) \in \mathbb{R}^p$ is the measurement and $C \in \mathbb{R}^{p \times n}$ the measurement matrix.

For the noise we assume The variables $w(k)$ and $v(k)$ are uncorrelated with $x(k)$.

$$
\begin{bmatrix} w(k) \\ v(k) \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix} \right)
$$

For the state we assume

$$
x(k) \sim N\left(\hat{x}_k, P_k\right).
$$

conditioned on the measurements up to $k-1$.

The Joint and Conditional Distributions

The joint distribution of $x(k)$ and $y(k)$ is

$$\begin{bmatrix} x(k) \\ y(k) \end{bmatrix} \sim N\left(\begin{bmatrix} \hat{x}_k \\ C\hat{x}_k \end{bmatrix}, \begin{bmatrix} P_k & P_k C^T \\ CP_k & CP_k C^T + R \end{bmatrix} \right)$$

The conditional distribution of X given $Y = y(k)$ is

$$p_{X|Y}\left(x(k)|Y = y(k)\right) \sim N\left(\hat{x}_k^+, P_k^+\right)$$

where the mean value is

$$\hat{x}_k^+ = \hat{x}_k + P_k C^T (CP_k C^T + R)^{-1} (y(k) - C\hat{x}_k)$$

and covariance

$$P_k^+ = P_k - P_k C^T (CP_k C^T + R)^{-1} CP_k$$

This is known as the *measurment update*

Time update

The distribution for the state variable time update, condistioned on all measurements up to time k, is thus

$$x(k+1) \sim N\left(\hat{x}_{k+1}, P_{k+1}\right)$$

where

$$\hat{x}_{k+1} = A\hat{x}_k^+$$

and

$$P_{k+1} = AP_k^+ A^T + Q.$$

In summary the filtering equations are

$$
\begin{aligned}
\hat{x}_k^+ &= \hat{x}_k + P_k C^T (C P_k C^T + R)^{-1} (y(k) - C\hat{x}_k), \text{ measurement update} \\
P_k^+ &= P_k - P_k C^T (C P_k C^T + R)^{-1} C P_k \\
\hat{x}_{k+1} &= A\hat{x}_k^+, \quad \text{time update (prediction)} \\
P_{k+1} &= A P_k^+ A^T + Q.
\end{aligned}
$$

The recursion equations need to be initialized when started, i.e. values of $\hat{x}_0$ and P_0 must be supplied. The parameters are the mean value and variance of the initial state.

Optimality

If all random variables has a Gaussian distribution then the Kalman filter provides the mean and the covariance for the Gaussian conditional distribution for $x(k)$ given the measurements $y(0), \ldots, y(k)$. (We cannot do better)