430.457

Introduction to Intelligent Systems

Prof. Songhwai Oh

ECE, SNU

Vector Case

KALMAN FILTERING

Linear Gaussian Model

- Dynamic model

$$x_{t+1} = Ax_t + Gw_t$$

- Measurement model

$$y_t = Cx_t + v_t$$

- Initial state distribution: $x_0 \sim \mathcal{N}(\mu_0, \Sigma_0)$ (assume $\mu_0 = 0$ for now)
- Noises: $w_t \sim \mathcal{N}(0, Q)$, $v_t \sim \mathcal{N}(0, R)$. w_t, v_t, and x_0 are independent.

- Unconditional distribution of x_t (i.e., y_t are not observed)

$$\begin{aligned}
\mathbb{E}(x_t) &= 0 \\
\mathbf{var}(x_{t+1}) &= \Sigma_{t+1} = \mathbb{E}(x_{t+1}x_{t+1}^T) \\
&= \mathbb{E}\left((Ax_t + Gw_t)(Ax_t + Gw_t)^T\right) \\
&= A\mathbb{E}(x_t x_t^T)A^T + G\mathbb{E}(w_t w_t^T)G^T \\
&= A\Sigma_t A^T + GQG^T
\end{aligned}$$

Filtering

- Dynamic update

$$P(x_t|y_0,\ldots,y_t) \to P(x_{t+1}|y_0,\ldots,y_t) = \int P(x_{t+1}|x_t)P(x_t|y_0,\ldots,y_t)dx_t$$

- Measurement update

$$P(x_{t+1}|y_0,\ldots,y_t) \to P(x_{t+1}|y_0,\ldots,y_t,y_{t+1}) \propto P(y_{t+1}|x_{t+1})P(x_{t+1}|y_0,\ldots,y_t)$$

  - $P(x_{t+1}, y_{t+1}|y_0,\ldots,y_t) \sim \mathcal{N}(\mu,\Sigma)$, where

$$\mu = \begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix}, \qquad \Sigma = \begin{pmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{yx} & \Sigma_{yy} \end{pmatrix}$$

Conditional PDF of Multivariate Gaussian

Theorem 10.2 (Conditional PDF of Multivariate Gaussian) *If* $\mathbf{x}$ *and* $\mathbf{y}$ *are jointly Gaussian, where* $\mathbf{x}$ *is* $k \times 1$ *and* $\mathbf{y}$ *is* $l \times 1$, *with mean vector* $[E(\mathbf{x})^T \; E(\mathbf{y})^T]^T$ *and partitioned covariance matrix*

$$\mathbf{C} = \begin{bmatrix} \mathbf{C}_{xx} & \mathbf{C}_{xy} \\ \mathbf{C}_{yx} & \mathbf{C}_{yy} \end{bmatrix} = \begin{bmatrix} k \times k & k \times l \\ l \times k & l \times l \end{bmatrix} \tag{10.23}$$

so that

$$p(\mathbf{x}, \mathbf{y}) = \frac{1}{(2\pi)^{\frac{k+l}{2}} \det^{\frac{1}{2}}(\mathbf{C})} \exp\left[-\frac{1}{2}\left(\begin{bmatrix} \mathbf{x} - E(\mathbf{x}) \\ \mathbf{y} - E(\mathbf{y}) \end{bmatrix}\right)^T \mathbf{C}^{-1} \left(\begin{bmatrix} \mathbf{x} - E(\mathbf{x}) \\ \mathbf{y} - E(\mathbf{y}) \end{bmatrix}\right)\right],$$

then the conditional PDF $p(\mathbf{y}|\mathbf{x})$ *is also Gaussian and*

$$E(\mathbf{y}|\mathbf{x}) = E(\mathbf{y}) + \mathbf{C}_{yx}\mathbf{C}_{xx}^{-1}(\mathbf{x} - E(\mathbf{x})) \tag{10.24}$$

$$\mathbf{C}_{y|x} = \mathbf{C}_{yy} - \mathbf{C}_{yx}\mathbf{C}_{xx}^{-1}\mathbf{C}_{xy}. \tag{10.25}$$

Source: Steven M. Kay, "Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory", Prentice Hall, 1993.

Filtering

- Dynamic update

$$P(x_t|y_0,\ldots,y_t) \to P(x_{t+1}|y_0,\ldots,y_t) = \int P(x_{t+1}|x_t)P(x_t|y_0,\ldots,y_t)dx_t$$

- Measurement update

$$P(x_{t+1}|y_0,\ldots,y_t) \to P(x_{t+1}|y_0,\ldots,y_t,y_{t+1}) \propto P(y_{t+1}|x_{t+1})P(x_{t+1}|y_0,\ldots,y_t)$$

  - $P(x_{t+1}, y_{t+1}|y_0,\ldots,y_t) \sim \mathcal{N}(\mu, \Sigma)$, where

$$\mu = \begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix}, \qquad \Sigma = \begin{pmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{yx} & \Sigma_{yy} \end{pmatrix}$$

  - $P(x_{t+1}|y_0,\ldots,y_t,y_{t+1}) \sim \mathcal{N}(\mu_{x|y}, \Sigma_{x|y})$, where

$$\begin{aligned} \mu_{x|y} &= \mu_x + \Sigma_{xy}\Sigma yy^{-1}(y-\mu_y) \\ \Sigma_{x|y} &= \Sigma_{xx} - \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx} \end{aligned}$$

Dynamic Update

$$
\begin{aligned}
\hat{x}_{t|t} &:= \mathbb{E}(x_t | y_0, \ldots, y_t) \\
P_{t|t} &= \mathbb{E}(\tilde{x}_{t|t} \tilde{x}_{t|t}^T | y_0, \ldots, y_t) \\
\tilde{x}_{t|t} &= x_t - \hat{x}_{t|t}
\end{aligned}
\quad \Longrightarrow \quad
\begin{aligned}
\hat{x}_{t+1|t} &:= \mathbb{E}(x_{t+1} | y_0, \ldots, y_t) \\
P_{t+1|t} &= \mathbb{E}\left(\tilde{x}_{t+1|t} \tilde{x}_{t+1|t}^T | y_0, \ldots, y_t\right) \\
\tilde{x}_{t+1|t} &= x_{t+1} - \hat{x}_{t+1|t}
\end{aligned}
$$

$$
\begin{aligned}
\hat{x}_{t+1|t} &= \mathbb{E}(x_{t+1} | y_0, \ldots, y_t) \\
&= \mathbb{E}(Ax_t + Gw_t | y_0, \ldots, y_t) \\
&= A\hat{x}_{t|t}
\end{aligned}
$$

$$x_{t+1} = Ax_t + Gw_t$$

$$
\begin{aligned}
P_{t+1|t} &= \mathbb{E}\left[\tilde{x}_{t+1|t} \tilde{x}_{t+1|t}^T | y_0, \ldots, y_t\right] \\
&= \mathbb{E}\left[(Ax_t + Gw_t - A\hat{x}_{t|t})(Ax_t + Gw_t - A\hat{x}_{t|t})^T | y_0, \ldots, y_t\right] \\
&= \mathbb{E}\left[(A(x_t - \hat{x}_{t|t}) + Gw_t)(A(x_t - \hat{x}_{t|t}) + Gw_t)^T | y_0, \ldots, y_t\right] \\
&= AP_{t|t}A^T + GQG^T
\end{aligned}
$$

Measurement Update (1)

$$\begin{aligned} \hat{x}_{t+1|t} &:= \mathbb{E}(x_{t+1}|y_0,\ldots,y_t) \\ P_{t+1|t} &= \mathbb{E}\left(\tilde{x}_{t+1|t}\tilde{x}_{t+1|t}^T|y_0,\ldots,y_t\right) \\ \tilde{x}_{t+1|t} &= x_{t+1} - \hat{x}_{t+1|t} \end{aligned} \quad \Longrightarrow \quad \begin{aligned} \hat{x}_{t+1|t+1} &:= \mathbb{E}(x_{t+1}|y_0,\ldots,y_{t+1}) \\ P_{t+1|t+1} &= \mathbb{E}(\tilde{x}_{t+1|t+1}\tilde{x}_{t+1|t+1}^T|y_0,\ldots,y_{t+1}) \\ \tilde{x}_{t+1|t+1} &= x_{t+1} - \hat{x}_{t+1|t+1} \end{aligned}$$

Distribution of y_{t+1} given $y_0,\ldots,y_t$:

$$\begin{aligned} \hat{y}_{t+1|t} &:= \mathbb{E}(y_{t+1}|y_0,\ldots,y_t) \\ &= \mathbb{E}(Cx_{t+1} + v_{t+1}|y_0,\ldots,y_t) = C\hat{x}_{t+1|t} \end{aligned}$$

$$\begin{aligned} &\mathbb{E}\left[(y_{t+1} - \hat{y}_{t+1|t})(y_{t+1} - \hat{y}_{t+1|t})^T|y_0,\ldots,y_t\right] \\ &= \mathbb{E}\left[(Cx_{t+1} + v_{t+1} - C\hat{x}_{t+1|t})(Cx_{t+1} + v_{t+1} - C\hat{x}_{t+1|t})^T|y_0,\ldots,y_t\right] \\ &= \mathbb{E}\left[(C(x_{t+1} - \hat{x}_{t+1|t}) + v_{t+1})(C(x_{t+1} - \hat{x}_{t+1|t}) + v_{t+1})^T|y_0,\ldots,y_t\right] \\ &= CP_{t+1|t}C^T + R \end{aligned}$$

$$\begin{aligned} &\mathbb{E}\left[(y_{t+1} - \hat{y}_{t+1|t})(x_{t+1} - \hat{x}_{t+1|t})|y_0,\ldots,y_t\right] \\ &= \mathbb{E}\left[(Cx_{t+1} + v_{t+1} - C\hat{x}_{t+1|t})(x_{t+1} - \hat{x}_{t+1|t})|y_0,\ldots,y_t\right] \\ &= CP_{t+1|t} \end{aligned}$$

Measurement Update (2)

x_{t+1} and y_{t+1} have the conditional joint multivariate Gaussian distribution

$$P(x_{t+1}, y_{t+1} | y_0, \ldots, y_t) \sim \mathcal{N}(\mu, \Sigma),$$

where

$$\mu = \begin{pmatrix} \hat{x}_{t+1|t} \\ C\hat{x}_{t+1|t} \end{pmatrix}, \qquad \Sigma = \begin{pmatrix} P_{t+1|t} & P_{t+1|t}C^T \\ CP_{t+1|t} & CP_{t+1|t}C^T + R \end{pmatrix}$$

- $P(x_{t+1}, y_{t+1} | y_0, \ldots, y_t) \sim \mathcal{N}(\mu, \Sigma)$, where

$$\mu = \begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix}, \qquad \Sigma = \begin{pmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{yx} & \Sigma_{yy} \end{pmatrix}$$

- $P(x_{t+1} | y_0, \ldots, y_t, y_{t+1}) \sim \mathcal{N}(\mu_{x|y}, \Sigma_{x|y})$, where

$$\begin{aligned} \mu_{x|y} &= \mu_x + \Sigma_{xy} \Sigma yy^{-1} (y - \mu_y) \\ \Sigma_{x|y} &= \Sigma_{xx} - \Sigma_{xy} \Sigma_{yy}^{-1} \Sigma_{yx} \end{aligned}$$

$P(x_{t+1} | y_0, \ldots, y_t, y_{t+1}) \sim \mathcal{N}(\hat{x}_{t+1|t+1}, P_{t+1|t+1})$, where

$$\begin{aligned} \hat{x}_{t+1|t+1} &= \hat{x}_{t+1|t} + P_{t+1|t}C^T (CP_{t+1|t}C^T + R)^{-1} (y_{t+1} - C\hat{x}_{t+1|t}) \\ P_{t+1|t+1} &= P_{t+1|t} - P_{t+1|t}C^T (CP_{t+1|t}C^T + R)^{-1} CP_{t+1|t} \end{aligned}$$

Kalman Filter

- Dynamic update

$$\begin{aligned}\hat{x}_{t+1|t} &= A\hat{x}_{t|t} \\ P_{t+1|t} &= AP_{t|t}A^T + GQG^T\end{aligned}$$

- Measurement update

$$\begin{aligned}\hat{x}_{t+1|t+1} &= \hat{x}_{t+1|t} + K_{t+1}(y_{t+1} - C\hat{x}_{t+1|t}) \\ P_{t+1|t+1} &= P_{t+1|t} - K_{t+1}CP_{t+1|t}\end{aligned}$$

- Kalman gain

$$K_{t+1} = P_{t+1|t}C^T(CP_{t+1|t}C^T + R)^{-1}$$

- Dynamic model: $x_{t+1} = Ax_t + Gw_t$
- Measurement model: $y_t = Cx_t + v_t$
- Initial state distribution: $x_0 \sim \mathcal{N}(\mu_0, \Sigma_0)$
- Noises: $w_t \sim \mathcal{N}(0, Q)$, $v_t \sim \mathcal{N}(0, R)$. w_t, v_t, and x_t are independent.

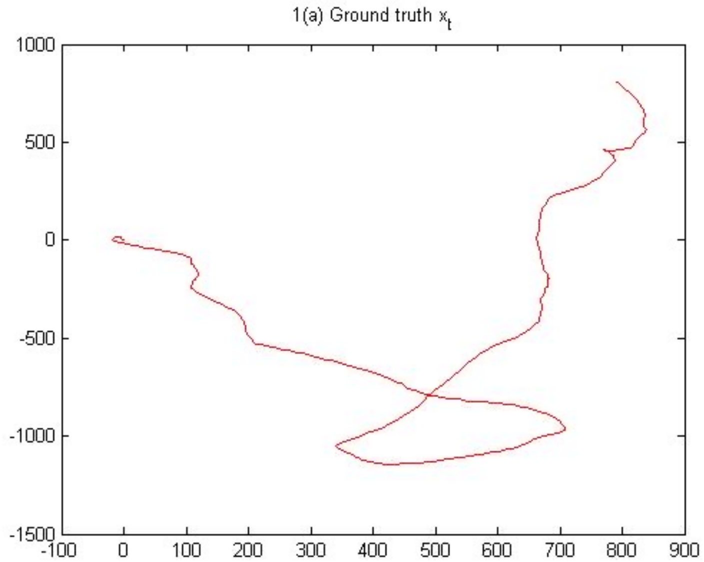
Example (1)

A particle moving in a plane under random forces and damping.

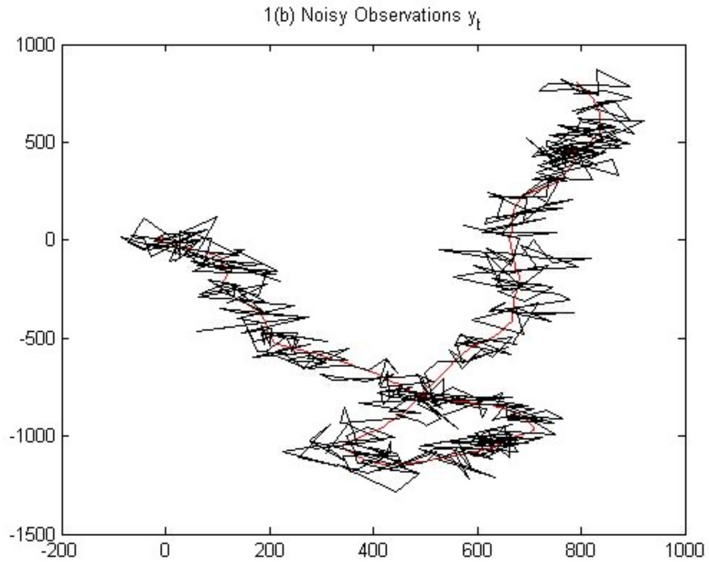
- State of the particle: $x = [x^1, \dot{x}^1, x^2, \dot{x}^2]^T$
- (x^1, x^2): position of the particle; $(\dot{x}^1, \dot{x}^2)$: velocity of the particle
- $x_{t+1}^i = x_t^i + \dot{x}_t^i$
- $\dot{x}_{t+1}^i = 0.98\dot{x}_t^i + w_t^i$
- Dynamic model: $x_{t+1} = Ax_t + Gw_t$
- Measurement model: $y_t = Cx_t + v_t$

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0.98 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0.98 \end{bmatrix} \quad G = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \quad C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

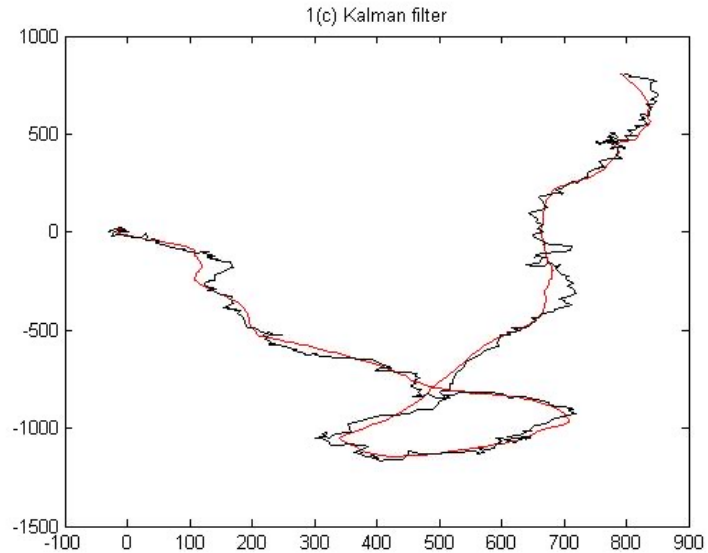
Example (2)



Ground Truth



Noisy Measurements



Kalman filter