Lecture 10: Extended Kalman Filters

CS 344R/393R: Robotics

Benjamin Kuipers

Up To Higher Dimensions

- Our previous Kalman Filter discussion was of a simple one-dimensional model.
- Now we go up to higher dimensions:
  - State vector: $\mathbf{x} \in \Re^n$
  - Sense vector: $\mathbf{z} \in \Re^m$
  - Motor vector: $\mathbf{u} \in \Re^l$
- First, a little statistics.

Expectations

- Let x be a random variable.
- The expected value $E[x]$ is the mean:

$$E[x] = \int x\ p(x)\ dx \approx \bar{x} = \frac{1}{N}\sum_{1}^{N} x_i$$

  - The probability-weighted mean of all possible values. The sample mean approaches it.
- Expected value of a vector $\mathbf{x}$ is by component.

$$E[\mathbf{x}] = \bar{\mathbf{x}} = [\bar{x}_1, \cdots \bar{x}_n]^T$$

Variance and Covariance

- The variance is $E[\ (x\text{-}E[x])^2\ ]$

$$\sigma^2 = E[(x-\bar{x})^2] = \frac{1}{N}\sum_{1}^{N}(x_i - \bar{x})^2$$

- Covariance matrix is $E[\ (\mathbf{x}\text{-}E[\mathbf{x}])(\mathbf{x}\text{-}E[\mathbf{x}])^T\ ]$

$$C_{ij} = \frac{1}{N}\sum_{k=1}^{N}(x_{ik} - \bar{x}_i)(x_{jk} - \bar{x}_j)$$

  - Divide by N–1 to make the sample variance an *unbiased estimator* for the population variance.

Covariance Matrix

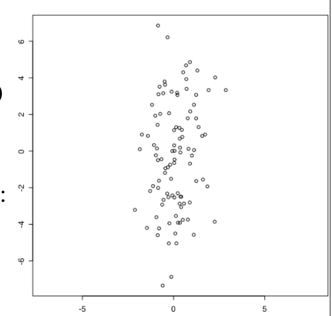
- Along the diagonal, C_{ii} are variances.
- Off-diagonal C_{ij} are essentially correlations.

$$\begin{bmatrix} C_{1,1} = \sigma_1^2 & C_{1,2} & & C_{1,N} \\ C_{2,1} & C_{2,2} = \sigma_2^2 & & \\ & & \ddots & \vdots \\ C_{N,1} & & \cdots & C_{N,N} = \sigma_N^2 \end{bmatrix}$$

Independent Variation

- x and y are Gaussian random variables (N=100)
- Generated with σ_x=1 σ_y=3
- Covariance matrix:

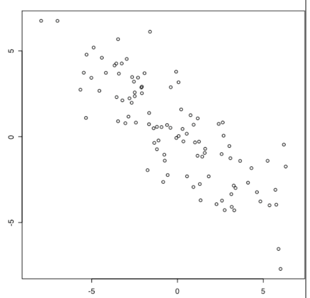
$$C_{xy} = \begin{bmatrix} 0.90 & 0.44 \\ 0.44 & 8.82 \end{bmatrix}$$



Dependent Variation

- c and d are random variables.
- Generated with $c=x+y \quad d=x-y$
- Covariance matrix:

$$C_{cd} = \begin{bmatrix} 10.62 & -7.93 \\ -7.93 & 8.84 \end{bmatrix}$$



Discrete Kalman Filter

- Estimate the state $\mathbf{x} \in \Re^n$ of a linear stochastic difference equation

$$\mathbf{x}_k = \mathbf{A}\mathbf{x}_{k-1} + \mathbf{B}\mathbf{u}_{k-1} + \mathbf{w}_{k-1}$$

  - process noise $\mathbf{w}$ is drawn from $N(0,\mathbf{Q})$, with covariance matrix $\mathbf{Q}$.
- with a measurement $\mathbf{z} \in \Re^m$

$$\mathbf{z}_k = \mathbf{H}\mathbf{x}_k + \mathbf{v}_k$$

  - measurement noise $\mathbf{v}$ is drawn from $N(0,\mathbf{R})$, with covariance matrix $\mathbf{R}$.
- $\mathbf{A}, \mathbf{Q}$ are $n \times n$. $\mathbf{B}$ is $n \times l$. $\mathbf{R}$ is $m \times m$. $\mathbf{H}$ is $m \times n$.

Estimates and Errors

- $\hat{\mathbf{x}}_k \in \Re^n$ is the estimated state at time-step k.
- $\hat{\mathbf{x}}_k^- \in \Re^n$ after prediction, before observation.
- Errors: $\mathbf{e}_k^- = \mathbf{x}_k - \hat{\mathbf{x}}_k^-$

$$\mathbf{e}_k = \mathbf{x}_k - \hat{\mathbf{x}}_k$$

- Error covariance matrices:

$$\mathbf{P}_k^- = E[\mathbf{e}_k^- \mathbf{e}_k^{-T}]$$

$$\mathbf{P}_k = E[\mathbf{e}_k \mathbf{e}_k^T]$$

- Kalman Filter's task is to update $\hat{\mathbf{x}}_k \quad \mathbf{P}_k$

Time Update (Predictor)

- Update expected value of $\mathbf{x}$

$$\hat{\mathbf{x}}_k^- = \mathbf{A}\hat{\mathbf{x}}_{k-1} + \mathbf{B}\mathbf{u}_{k-1}$$

- Update error covariance matrix $\mathbf{P}$

$$\mathbf{P}_k^- = \mathbf{A}\mathbf{P}_{k-1}\mathbf{A}^T + \mathbf{Q}$$

- Previous statements were simplified versions of the same idea:

$$\hat{x}(t_3^-) = \hat{x}(t_2) + u[t_3 - t_2]$$

$$\sigma^2(t_3^-) = \sigma^2(t_2) + \sigma_\varepsilon^2[t_3 - t_2]$$

Measurement Update (Corrector)

- Update expected value

$$\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_k^- + \mathbf{K}_k(\mathbf{z}_k - \mathbf{H}\hat{\mathbf{x}}_k^-)$$

  - *innovation* is $\mathbf{z}_k - \mathbf{H}\hat{\mathbf{x}}_k^-$
- Update error covariance matrix

$$\mathbf{P}_k = (\mathbf{I} - \mathbf{K}_k\mathbf{H})\mathbf{P}_k^-$$

- Compare with previous form

$$\hat{x}(t_3) = \hat{x}(t_3^-) + K(t_3)(z_3 - \hat{x}(t_3^-))$$

$$\sigma^2(t_3) = (1 - K(t_3))\sigma^2(t_3^-)$$

The Kalman Gain

- The optimal Kalman gain $\mathbf{K}_k$ is

$$\mathbf{K}_k = \mathbf{P}_k^-\mathbf{H}^T(\mathbf{H}\mathbf{P}_k^-\mathbf{H}^T + \mathbf{R})^{-1}$$

$$= \frac{\mathbf{P}_k^-\mathbf{H}^T}{\mathbf{H}\mathbf{P}_k^-\mathbf{H}^T + \mathbf{R}}$$

- Compare with previous form

$$K(t_3) = \frac{\sigma^2(t_3^-)}{\sigma^2(t_3^-) + \sigma_3^2}$$

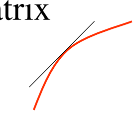
Extended Kalman Filter

- Suppose the state-evolution and measurement equations are non-linear:

$$\mathbf{x}_k = f(\mathbf{x}_{k-1}, \mathbf{u}_{k-1}) + \mathbf{w}_{k-1}$$
$$\mathbf{z}_k = h(\mathbf{x}_k) + \mathbf{v}_k$$

  - process noise $\mathbf{w}$ is drawn from $N(0,\mathbf{Q})$, with covariance matrix $\mathbf{Q}$.
  - measurement noise $\mathbf{v}$ is drawn from $N(0,\mathbf{R})$, with covariance matrix $\mathbf{R}$.

The Jacobian Matrix



- For a scalar function $y=f(x)$,
  $\Delta y = f'(x)\Delta x$
- For a vector function $\mathbf{y}=f(\mathbf{x})$,

$$\Delta\mathbf{y} = \mathbf{J}\Delta\mathbf{x} = \begin{bmatrix}\Delta y_1\\ \vdots \\ \Delta y_n\end{bmatrix} = \begin{bmatrix}\frac{\partial f_1}{\partial x_1}(\mathbf{x}) & \cdots & \frac{\partial f_1}{\partial x_n}(\mathbf{x})\\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1}(\mathbf{x}) & \cdots & \frac{\partial f_n}{\partial x_n}(\mathbf{x})\end{bmatrix} \cdot \begin{bmatrix}\Delta x_1\\ \vdots \\ \Delta x_n\end{bmatrix}$$

Linearize the Non-Linear

- Let $\mathbf{A}$ be the Jacobian of f with respect to $\mathbf{x}$.

$$\mathbf{A}_{ij} = \frac{\partial f_i}{\partial x_j}(\mathbf{x}_{k-1}, \mathbf{u}_{k-1})$$

- Let $\mathbf{H}$ be the Jacobian of h with respect to $\mathbf{x}$.

$$\mathbf{H}_{ij} = \frac{\partial h_i}{\partial x_j}(\mathbf{x}_k)$$

- Then the Kalman Filter equations are almost the same as before!

EKF Update Equations

- Predictor step: $\hat{\mathbf{x}}_k^- = f(\hat{\mathbf{x}}_{k-1}, \mathbf{u}_{k-1})$

$$\mathbf{P}_k^- = \mathbf{A}\mathbf{P}_{k-1}\mathbf{A}^T + \mathbf{Q}$$

- Kalman gain: $\mathbf{K}_k = \mathbf{P}_k^-\mathbf{H}^T(\mathbf{H}\mathbf{P}_k^-\mathbf{H}^T + \mathbf{R})^{-1}$

- Corrector step: $\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_k^- + \mathbf{K}_k(\mathbf{z}_k - h(\hat{\mathbf{x}}_k^-))$

$$\mathbf{P}_k = (\mathbf{I} - \mathbf{K}_k\mathbf{H})\mathbf{P}_k^-$$

Next

- Building a map of landmark locations by combining uncertain observations.