The Extended Kalman Filter or EKF

D.A. Forsyth,

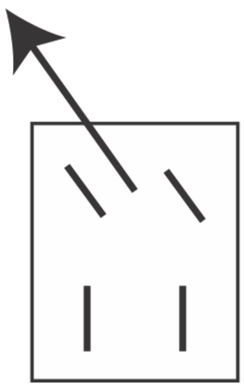
University of Illinois at Urbana-Champaign

Kalman filter is wonderful, but…

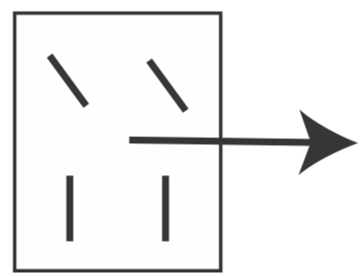
- Linear measurement isn’t always helpful
  - Example:
    - our localization procedures
- Linear model of movement isn’t always helpful
  - Example:
    - simple car model (next)

- What to do about non-linearities?

Example: Nasty dynamical model



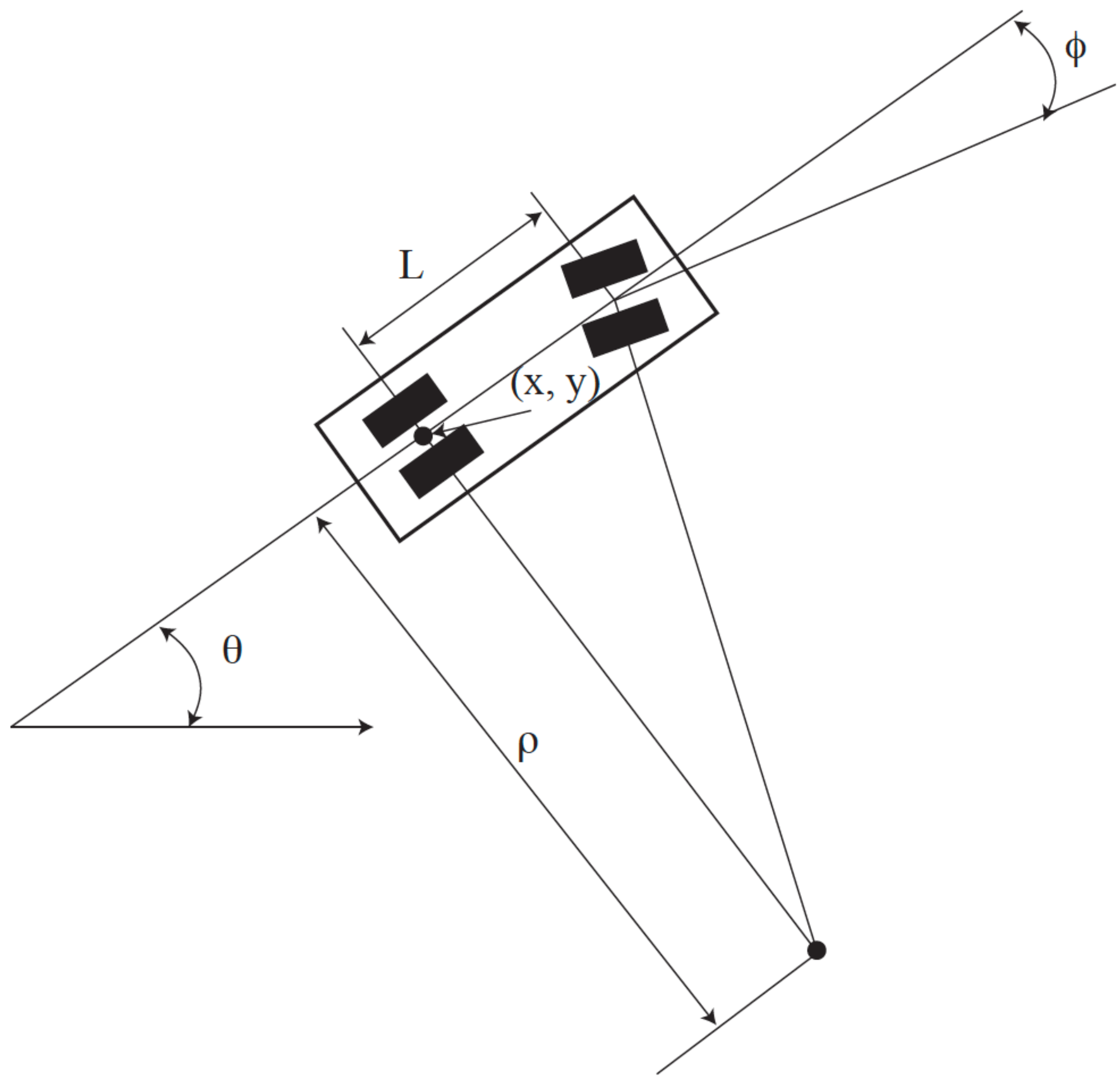
OK



Not OK

Formally: car is non-holonomic

Notation for dynamics



Representing state

- Must have:
  - x, y - position of vehicle
  - phi – angle of steering relative to midline

- BUT
  - only 2D velocity
    - speed - along the midline
    - rate of turn - rate at which steering angle changes

- Can’t go sideways < == > 2D velocity, 3D state

Dynamics

- Write
  - s – speed along the midline
  - r - rate of change of steering angle
- Get:

$$\frac{d}{dt}\begin{bmatrix} x \\ y \\ \theta \end{bmatrix} = \begin{bmatrix} s\cos\theta \\ s\sin\theta \\ \frac{s}{L}\tan r \end{bmatrix}$$

Derivation at: https://msl.cs.uiuc.edu/planning/node658.html

The extended Kalman filter

- If state update, measurement aren't linear
  - linearize and approximate (EKF)
    - (and hope!)
  - idea is straightforward, mechanics require some care

state

$$\mathbf{x}_i = f(\mathbf{x}_{i-1}, \mathbf{n})$$

Noise - normal, mean 0, Cov known

$$\mathbf{y}_i = g(\mathbf{x}_i, \mathbf{n})$$

observation

The steps, KF:

Have:

Mean and covariance of posterior
after i-1'th measurement

Construct:

Mean and covariance of predictive
distribution just before i'th measurement

Measurement arrives:

Now construct:

Mean and covariance of posterior
distribution just before i'th measurement

posterior mean is weighted combo
of prior mean and measurement

posterior covar is weighted combo
of prior covar, measurement
matrix and measurement covar

Linearization and noise

- Two ways in which noise could affect x_i
  - x_{i-1} is noisy
  - AND there is n to account for

Possibly noisy input

$$\mathbf{x}_i = f(\mathbf{x}_{i-1}, \mathbf{n})$$

Noise supplied to function

State update

- Now consider some nonlinear function h(x)
  - first case (input is noisy)

$$h(\mathbf{x}) \text{ where } \mathbf{x} \sim N(\bar{\mathbf{x}}, \Sigma_x)$$

$$h(\bar{\mathbf{x}} + \zeta) \text{ where } \zeta \sim N(0, \Sigma_x)$$

$$J_{h,x} = \begin{bmatrix} \frac{\partial h_1}{\partial x_1} & \cdots & \cdots \\ \cdots & \frac{\partial h_i}{\partial x_j} & \cdots \\ \cdots & \cdots & \cdots \end{bmatrix}$$

Jacobian === derivative

Approximate

$$h(\bar{\mathbf{x}} + \zeta) \approx h(\bar{\mathbf{x}}) + J_{h,x}\zeta$$

Yields

$$h(\mathbf{x}) \sim N(h(\bar{\mathbf{x}}), J_{h,x}\Sigma_x J_{h,x}^T)$$

State update

- Now consider some nonlinear function h(x)
    - second case (noise supplied to function)

$$h(\mathbf{x}, \mathbf{n}) \text{ where } \mathbf{n} \sim N(0, \sigma_n)$$

Approximate

$$h(\mathbf{x}, \mathbf{n}) \approx h(\mathbf{x}, \mathbf{0}) + J_{h,n}\mathbf{n}$$

$$J_{h,n} = \begin{bmatrix} \frac{\partial h_1}{\partial n_1} & \cdots & \cdots \\ \cdots & \frac{\partial h_i}{\partial n_j} & \cdots \\ \cdots & \cdots & \cdots \end{bmatrix}$$

Yields

$$h(\mathbf{x}, \mathbf{n}) \sim N(h(\mathbf{x}, \mathbf{0}), J_{h,n}\Sigma_n J_{h,n}^T)$$

Jacobian === derivative

State update

- Linearize:

$$\mathbf{x}_i = f_i(\mathbf{x}_{i-1}, \mathbf{n})$$

$$\mathcal{F}_x = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \cdots \\ \cdots & \frac{\partial f_i}{\partial x_j} & \cdots \end{bmatrix}$$

$$\mathcal{F}_n = \begin{bmatrix} \frac{\partial f_1}{\partial n_1} & \cdots & \cdots \\ \cdots & \frac{\partial f_i}{\partial n_j} & \cdots \end{bmatrix}$$

Posterior covariance of x_{i-1}

$$\mathbf{x}_i \sim N(f_i(\bar{\mathbf{x}}^+_{i-1}, \mathbf{0}), \mathcal{F}_x \Sigma^+_{i-1} \mathcal{F}_x^T + \mathcal{F}_n \Sigma_{n,i} \mathcal{F}_n^T)$$

Noise covariance

The steps, EKF:

Have:

$$\overline{X}_{i-1}^{+} \qquad \Sigma_{i-1}^{+}$$

Construct:

$$\bar{X}_i^- = f_i(\bar{\mathbf{x}}_{i-1}^{+}, \mathbf{0}) \qquad \Sigma_i^- = \mathcal{F}_x \Sigma_{i-1}^{+} \mathcal{F}_x^T + \mathcal{F}_n \Sigma_{n,i} \mathcal{F}_n^T$$

Measurement arrives:

Now construct:

Where:

The steps, EKF:

Have:

$$\overline{X}^+_{i-1} \qquad \Sigma^+_{i-1}$$

Construct:

$$\bar{X}_i^- = f_i(\bar{\mathbf{x}}^+_{i-1}, \mathbf{0}) \qquad \Sigma_i^- = \mathcal{F}_x \Sigma^+_{i-1} \mathcal{F}_x^T + \mathcal{F}_n \Sigma_{n,i} \mathcal{F}_n^T$$

Measurement arrives: $\mathbf{y}_i = g_i(\mathbf{x}_i, \mathbf{n})$

Now construct:

Where:

Measurement

- Linearize:

$$\mathbf{y}_i = g_i(\mathbf{x}_i, \mathbf{n})$$

$$\mathcal{G}_x = \begin{bmatrix} \frac{\partial g}{\partial x_1} & \cdots & \cdots \\ \cdots & \frac{\partial g}{\partial x_1} & \cdots \end{bmatrix}$$

$$\mathcal{G}_n = \begin{bmatrix} \frac{\partial g}{\partial n_1} & \cdots & \cdots \\ \cdots & \frac{\partial g}{\partial n_1} & \cdots \end{bmatrix}$$

$$\mathbf{y}_i \approx \mathcal{N}(g_i(\bar{X}_i^-, \mathbf{0}), \mathcal{G}_x \Sigma_i^- \mathcal{G}_x^T + \mathcal{G}_n \Sigma_{m,i} \mathcal{G}_n^T)$$

Recall: The steps, KF:

Have:

$$\overline{X}_{i-1}^{+} \qquad \Sigma_{i-1}^{+}$$

Construct:

$$\overline{X}_{i}^{-} = \mathcal{D}_i \overline{X}_{i-1}^{+} \qquad \Sigma_i^{-} = \Sigma_{d_i} + \mathcal{D}_i \Sigma_{i-1}^{-} \mathcal{D}_i^T$$

Measur

$(\mathcal{M}_i \mathbf{x}_i; \Sigma_{mi})$

Difference between predicted and observed measurement

Now construct:

$$\bar{X}_i^{+} = \bar{X}_i^{-} + \mathcal{K}_i \left[\mathbf{y}_i - \mathcal{M}_i \bar{X}_i^{-} \right] \qquad \Sigma_i^{+} = \left[\mathcal{I} - \mathcal{K}_i \mathcal{M}_i \right] \Sigma_i^{-}$$

Where:

$$\mathcal{K}_i = \Sigma_i^{-} \mathcal{M}_i^T \left[\mathcal{M}_i \Sigma_i^{-} \mathcal{M}_i^T + \Sigma_{m_i} \right]^{-1}$$

The steps, EKF:

Have:

$$\bar{X}_{i-1}^{+} \qquad \Sigma_{i-1}^{+}$$

Construct:

$$\bar{X}_i^{-} = f_i(\bar{\mathbf{x}}_{i-1}^{+}, \mathbf{0}) \qquad \Sigma_i^{-} = \mathcal{F}_x \Sigma_{i-1}^{+} \mathcal{F}_x^T + \mathcal{F}_n \Sigma_{n,i} \mathcal{F}_n^T$$

Difference between predicted and observed measurement

Measurement

$$\mathbf{y}_i = g_i(\mathbf{x}_i, \mathbf{n})$$

Now construct:

$$\bar{X}_i^{+} = \bar{X}_i^{-} + \mathcal{K}_i \left[\mathbf{y}_i - g_i(\bar{X}_i^{-}, \mathbf{0})\right]$$

Where:

Recall: The steps, KF:

Have:

$$\overline{X}_{i-1}^{+} \qquad \Sigma_{i-1}^{+}$$

Construct:

$$\overline{X}_{i}^{-} = \mathcal{D}_i \overline{X}_{i-1}^{+} \qquad \Sigma_i^{-} = \Sigma_{d_i} + \mathcal{D}_i \Sigma_{i-1}^{-} \mathcal{D}_i^T$$

Measurement arrives: $\mathbf{y}_i \sim N(\mathcal{M}$

Linear measurement model

Now construct:

$$\bar{X}_i^{+} = \bar{X}_i^{-} + \mathcal{K}_i \left[\mathbf{y}_i - \mathcal{M}_i \bar{X}_i^{-} \right] \qquad \Sigma_i^{+} = \left[\mathcal{I} - \mathcal{K}_i \mathcal{M}_i \right] \Sigma_i^{-}$$

Where:

$$\mathcal{K}_i = \Sigma_i^{-} \mathcal{M}_i^T \left[\mathcal{M}_i \Sigma_i^{-} \mathcal{M}_i^T + \Sigma_{m_i} \right]^{-1}$$

The steps, EKF:

Have:

$$\overline{X}^{+}_{i-1} \qquad \Sigma^{+}_{i-1}$$

Construct:

$$\bar{X}_i^{-} = f_i(\bar{\mathbf{x}}^{+}_{i-1}, \mathbf{0}) \qquad \Sigma_i^{-} = \mathcal{F}_x \Sigma^{+}_{i-1} \mathcal{F}_x^T + \mathcal{F}_n \Sigma_{n,i} \mathcal{F}_n^T$$

Measurement arrives: $\mathbf{y}_i = g_i(\mathbf{x}_i, \mathbf{n})$

Now construct:

$$\bar{X}_i^{+} = \bar{X}_i^{-} + \mathcal{K}_i \left[\mathbf{y}_i - g_i(\bar{X}_i^{-}, \mathbf{0})\right] \qquad \Sigma_i^{+} = \left[Id - \mathcal{K}_i \mathcal{G}_x\right] \Sigma_i^{-}$$

Where:

Recall: The steps, KF:

Have:

$$\overline{X}_{i-1}^{+} \qquad \Sigma_{i-1}^{+}$$

Construct:

$$\overline{X}_{i}^{-} = \mathcal{D}_i \overline{X}_{i-1}^{+} \qquad \Sigma_{i}^{-} = \Sigma_{d_i} + \mathcal{D}_i \Sigma_{i-1}^{-} \mathcal{D}_i^T$$

Measurement arrives:

Inverse of the covariance of y_i

Linear measurement model

Now construct:

$$\bar{X}_i^{+} = \bar{X}_i^{-} + \mathcal{K}_i \left[\mathbf{y}_i - \mathcal{M}_i \bar{X}_i^{-} \right] \qquad \Sigma_i^{+} = \left[\mathcal{I} - \mathcal{K}_i \mathcal{M}_i \right] \Sigma_i^{-}$$

Where:

$$\mathcal{K}_i = \Sigma_i^{-} \mathcal{M}_i^T \left[\mathcal{M}_i \Sigma_i^{-} \mathcal{M}_i^T + \Sigma_{m_i} \right]^{-1}$$

The steps, EKF:

Have:

$$\overline{X}_{i-1}^{+} \qquad \Sigma_{i-1}^{+}$$

Construct:

$$\bar{X}_i^- = f_i(\bar{\mathbf{x}}_{i-1}^+, \mathbf{0}) \qquad \Sigma_i^- = \mathcal{F}_x \Sigma_{i-1}^+ \mathcal{F}_x^T + \mathcal{F}_n \Sigma_{n,i} \mathcal{F}_n^T$$

Measurement arrives:

$$\mathbf{y}_i = g_i(\mathbf{x}_i, \mathbf{n})$$

Now construct:

$$\bar{X}_i^+ = \bar{X}_i^- + \mathcal{K}_i \left[\mathbf{y}_i - g_i(\bar{X}_i^-, \mathbf{0})\right] \qquad \Sigma_i^+ = \left[Id - \mathcal{K}_i \mathcal{G}_x\right] \Sigma_i^-$$

Where:

$$\mathcal{K}_i = \Sigma_i^- \mathcal{G}_x^T \left[\mathcal{G}_x \Sigma_i^- \mathcal{G}_x^T + \mathcal{G}_n \Sigma_{m,i} \mathcal{G}_n^T\right]^{-1}$$

Outcome and issues

- Can now filter position/orientation wrt map
    - linearize dynamics following recipe above
    - linearize measurements ditto
- There could be problems
    - EKF's are fine if the linearization is reliable
        - can be awful if not