~~Kalman Filter~~ ~~Lecture~~ Kalman ~~Kalman~~ Filter Preliminaries

Joshua Mangelson

Bayes Filter Review

Definitions

state at time t $\quad x_t \quad$ n-vector

control at time t $\quad u_t \quad$ m-vector

measurement at time t $\quad z_t \quad$ k-vector

Kalman Filter

Optimal Filter when update/meas update state is linear ~~and noise~~ ~~is addt independent~~ w/ additive ~~Gaussian~~ normally distributed ~~iid~~ noise

zero mean

Bayes Filter

Goal is to track: $Bel(x_t) = p(x_t \mid u_{1:t}, z_{1:t})$

Given: $Bel(x_0)$ Initial belief

$u_1, \ldots, u_t$ all control inputs

$z_1, \ldots, z_t$ all measurements

Recursive Update + Correction Steps.

work Backwards from Goal

$$Bel(x_t) = p(x_t \mid u_{1:t}, z_{1:t})$$

$$= \frac{p(x_t, z_t \mid u_{1:t}, z_{1:t-1})}{p(z_t \mid u_{1:t}, z_{1:t-1})}$$

conditioning Joint Distribution of x_t, z_t on z_t Given prior measurements + controls

If z_t only dependent on x_t, then we need $p(x_t \mid u_{1:t}, z_{1:t-1})$ to form joint

Let

$$\overline{Bel}(x_t) = p(x_t \mid u_{1:t}, z_{1:t-1})$$

$$= \int p(x_t, x_{t-1} \mid u_{1:t}, z_{1:t-1})\, dx_{t-1}$$

~~Conditioning~~ Marginalizing x_{t-1} ~~Joint dis~~ from joint distribution of x_{t-1} and x_t given prior measurements + controls

markov assumption

$$= \int p(x_t \mid u_t; x_{t-1})\, Bel(x_{t-1})\, dx_{t-1}$$

Kalman Filter Assumptions / Formulation

State Update (Process Model)

$$\underset{n\times 1}{x_t} = \underset{n\times n}{A_t}\ \underset{n\times 1}{x_{t-1}} + \underset{n\times m}{B_t}\ \underset{m\times 1}{u_t} + \underset{n\times 1}{\epsilon_t}$$

A_t: How state evolves from $t-1$ to t

B_t: How control changes state from $t-1$ to t

ϵ_t: Process noise
Assume independent from state + control and normaly distributed according to

$$\epsilon_t \sim N(\underset{\text{zero mean}}{0}, \underset{\text{cov.}}{R_t})$$

Measurement Update (Meas. Model)

$$\underset{k\times 1}{z_t} = \underset{k\times n}{C_t}\ \underset{n\times 1}{x_t} + \underset{n\times 1}{\delta_t}$$

C_t: projection of state to an observation

δ_t: Measurement Noise
Assume independent from state and normaly distributed

$$\delta_t \sim N(\underset{\text{zero mean}}{0}, Q_t)$$

~~Bel~~ Initial Belief

~~Gaussian~~

Remember?

~~$Bel(x_0) \sim N(x_0; \mu_0$~~ ~~$Bel(x_0; \mu_0, \Sigma_0)$~~

$$Bel(x_0) \sim N(x_0; \mu_0, \Sigma_0)$$

- Linear Functions of Gaussians
  state will remain Gaussian

$\alpha\alpha$	$\alpha\beta$
$\beta\alpha$	$\beta\beta$

Gaussian Distributions

Marginalization

$$p(\alpha) = \int p(\alpha, \beta)\, d\beta$$

$$\mu = \mu_\alpha$$

$$\Sigma = \Sigma_{\alpha\alpha}$$

Conditioning

$$p(\alpha \mid \beta) = \frac{p(\alpha, \beta)}{p(\beta)}$$

$$\mu' = \mu_\alpha + \Sigma_{\alpha\beta}\Sigma_{\beta\beta}^{-1}(\beta - \mu_\beta)$$

$$\Sigma' = \Sigma_{\alpha\alpha} - \Sigma_{\alpha\beta}\Sigma_{\beta\beta}^{-1}\Sigma_{\beta\alpha}$$

State Update ~~2018~~ 2019 16833 ~~L9.2~~ L8.2

$p(x_t, x_{t-1} \mid u_{1:t}, z_{1:t-1})$ followed by marginalization

$$\begin{bmatrix} x_t \\ x_{t-1} \end{bmatrix} = \begin{bmatrix} A_t \\ I \end{bmatrix} x_{t-1} + \begin{bmatrix} B_t \\ 0 \end{bmatrix} u_t + \begin{bmatrix} I \\ 0 \end{bmatrix} \varepsilon_t$$

$$\xi = F_t \, x_{t-1} + G_t \, u_t + H_t \, \varepsilon_t$$

$$\mu_\xi = E[\xi] = F_t \, \mu_{t-1} + G_t \, u_t + 0 = \begin{bmatrix} A_t \mu_{t-1} + B_t u_t \\ \mu_{t-1} \end{bmatrix}$$

$$\Sigma_\xi = E[(\xi - \mu_\xi)(\xi - \mu_\xi)^T] = E[(F_t (x_{t-1} - \mu_{t-1}) + H_t \varepsilon_t) \cdot (\ldots)^T]$$

$$= F_t \Sigma_{t-1} F_t^T + F_t \, E[\cancel{(x_{t-1} - \mu_{t-1}) \varepsilon_t^T}]^{\to 0} H_t^T$$

$$+ H_t \, E[\cancel{\varepsilon_t (x_{t-1} - \mu_{t-1})^T}]^{\to 0} F_t^T + H_t R_t H_t^T$$

$$= \begin{bmatrix} A_t \\ I \end{bmatrix} \Sigma_{t-1} [A_t^T \; I] + \begin{bmatrix} I \\ 0 \end{bmatrix} R_t [I \; 0]$$

$$= \begin{bmatrix} A_t \Sigma_{t-1} A_t^T + R_t & A_t \Sigma_{t-1} \\ \Sigma_{t-1} A_t^T & \Sigma_{t-1} \end{bmatrix}$$

$$\overline{Bel}(x_t) = p(x_t \mid u_{1:t}, z_{1:t-1})$$

$$= \int p(x_t, x_{t-1} \mid u_{1:t}, z_{1:t-1}) \, dx_{t-1} \sim N(x_t ; \bar{\mu}_t, \bar{\Sigma}_t)$$

marginalization:

$$\left.\begin{aligned} \bar{\mu}_t &= A_t \mu_{t-1} + B_t u_t \\ \bar{\Sigma}_t &= A_t \Sigma_{t-1} A_t^T + R_t \end{aligned}\right\} \text{Kalman prediction}$$

Measurement Update

$P(x_t, z_t \mid z_{1:t-1}, u_{1:t})$ followed by conditioning

$$\underbrace{\begin{bmatrix} x_t \\ z_t \end{bmatrix}}_{\zeta_t} = \underbrace{\begin{bmatrix} I \\ C \end{bmatrix}}_{D_t} x_t + \underbrace{\begin{bmatrix} 0 \\ I \end{bmatrix}}_{E_t} \delta_t$$

$$\mu_\zeta = E[\zeta_t] = D_t \bar{\mu}_t = \begin{bmatrix} \bar{\mu} \\ C\bar{\mu} \end{bmatrix}$$

$$\Sigma_\zeta = E\left[(\zeta_t - \mu_\zeta)(\zeta_t - \mu_\zeta)^T\right] = D_t \bar{\Sigma}_t D_t^T + E_t Q_t E_t^T$$

$$= \begin{bmatrix} \bar{\Sigma}_t & \bar{\Sigma}_t C^T \\ C\bar{\Sigma}_t & C\bar{\Sigma}_t C^T + Q_t \end{bmatrix}$$

$$Bel(x_t) = p(x_t \mid z_t, z_{1:t-1}, u_{1:t}) \sim N(x_t; \mu_t, \Sigma_t)$$

conditioning

$$\mu_t = \underbrace{\bar{\mu}_t}_{\mu_\alpha} + \underbrace{\bar{\Sigma}_t C_t^T}_{\Sigma_{\alpha\beta}} \underbrace{(C_t \bar{\Sigma}_t C_t^T + Q_t)^{-1}}_{\Sigma_{\beta\beta}^{-1}} \underbrace{(z_t - C_t \bar{\mu}_t)}_{(\beta - \mu_\beta)}$$

$$\Sigma_t = \underbrace{\bar{\Sigma}_t}_{\Sigma_{\alpha\alpha}} - \underbrace{\underbrace{\bar{\Sigma}_t C_t^T}_{\Sigma_{\alpha\beta}} \underbrace{(C_t \bar{\Sigma}_t C_t^T + Q_t)^{-1}}_{\Sigma_{\beta\beta}^{-1}}}_{K_t \text{ Kalman gain}} \underbrace{C_t \bar{\Sigma}_t}_{\Sigma_{\beta\alpha}}$$

$$\left.\begin{aligned} \mu_t &= \bar{\mu}_t + K_t (z_t - C_t \bar{\mu}_t) \\ \Sigma_t &= \bar{\Sigma}_t - K_t C_t \bar{\Sigma}_t \\ &= (I - K_t C_t) \bar{\Sigma}_t \end{aligned}\right\} \text{Kalman update}$$

Kalman filter recap

use slides?

Kalman prediction:

$$\bar{\mu}_t = A_t \mu_{t-1} + B_t u_t$$ A_t often I, u_t eg. wheel ticks

$$\bar{\Sigma}_t = A_t \Sigma_{t-1} A_t^T + R_t$$ uncertainty grows

Kalman update:

$$\mu_t = \bar{\mu}_t + K_t (\underbrace{z_t - C_t \bar{\mu}_t}_{\text{measurement - predicted measurement}})$$ K converts this into update on state estimate

$$\Sigma_t = (I - K_t C_t) \bar{\Sigma}_t$$ uncertainty shrinks

with Kalman gain $K_t = \bar{\Sigma}_t C_t^T (C_t \bar{\Sigma}_t C_t^T + Q_t)^{-1}$

depends on "ratio" between state and meas. uncert.

(A) $\bar{\Sigma}_t \to \infty$ or $Q_t = 0$

$$K_t = \bar{\Sigma}_t C_t^T (C_t \bar{\Sigma}_t C_t^T + \underbrace{\cancel{Q_t}}_{0 \text{ or other term dominates}})^{-1} = \bar{\Sigma}_t \underbrace{C_t^T C_t^{-T}} \bar{\Sigma}_t^{-1} C_t^{-1} = C_t^{-1}$$

$$\mu_t = \underbrace{\bar{\mu}_t + C_t^{-1}(z_t - C_t \bar{\mu}_t)} = C_t^{-1} z_t$$ measurement dominates (assumes invertible)

(B) $\bar{\Sigma}_t = 0$ or $Q_t \to \infty$

then $K_t = 0$, $\mu_t = \bar{\mu}_t$ measurement ignored

Complexity

Polynomial in measurement dimensionality k and state dimensionality n

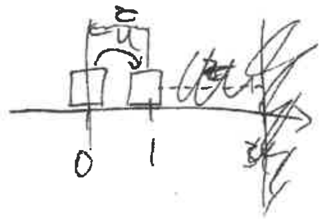
$$O(k^{2.376} + n^2)$$

$$(\underbrace{C_t \bar{\Sigma}_t C_t^T + Q_t}_{k \times k})^{-1} \to k^{2.376}$$

$$\Sigma_t = (I - K_t C_t) \bar{\Sigma}_t$$ $\Rightarrow \sim O(n^2)$ complexity

dense (K_t), typically sparse $O(n)$ (C_t) in general: $O(n^3)$

1D Example



initial: $x_0 = 0m$

$\sigma_0 = 0.3m$

motion command: $u = 1.2m$ (but robot ends up at 1m)

$\sigma_u = 0.4m$

measurement $z = 1m$

$\sigma_z = 0.1m$

K. prediction

$$\bar{\mu}_1 = \mu_0 + u = 0m + 1.2m = 1.2m$$

$$\bar{\sigma}_1^2 = \sigma_0^2 + \sigma_u^2 = (0.3m)^2 + (0.4m)^2 = (0.5m)^2$$

Kalman update

$$\mu_1 = \bar{\mu}_1 + K(z - C\bar{\mu}_1) \qquad C = 1$$

$$K = \bar{\sigma}_1^2 \underset{1}{C^T}\left(\underset{1}{C}\,\bar{\sigma}^2 \underset{1}{C^T} + \sigma_z^2\right)^{-1}$$

$$= (0.5m)^2\left((0.5m)^2 + (0.1m)^2\right)^{-1}$$

$$= 0.25m^2\left(0.26m^2\right)^{-1}$$

$$\approx 0.925$$

$$= 1.2m + 0.925(1m - 1.2m)$$

$$= 1.2m + 0.925(-0.2m)$$

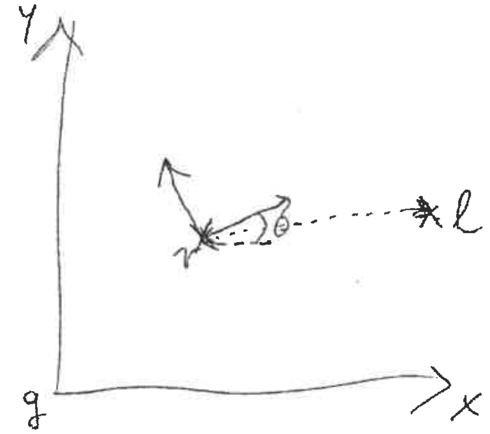
$$\approx 1.2m - 0.185m$$

$$= 1.015m$$

[~~recap of EKF, explain CLT, overview of~~ UKF in detail] ~~2018~~ 2019 16833 ~~L11.1~~ L10.1

Nonlinear measurement example

Robot at $r = (x, y, \theta)$ observing landmark $\ell = (x_\ell, y_\ell)$ by measuring relative translation $z = (x_z, y_z)$



$$h(r,\ell) = \begin{pmatrix} x_z \\ y_z \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \left(\begin{pmatrix} x_\ell \\ y_\ell \end{pmatrix} - \begin{pmatrix} x \\ y \end{pmatrix} \right)$$

$$= \begin{pmatrix} \cos\theta\,(x_\ell - x) + \sin\theta\,(y_\ell - y) \\ -\sin\theta\,(x_\ell - x) + \cos\theta\,(y_\ell - y) \end{pmatrix}$$

Jacobian $A(x)$ $\quad x = \begin{pmatrix} x \\ y \\ \theta \\ x_\ell \\ y_\ell \end{pmatrix}$

$$\begin{array}{c|c} & \begin{array}{ccccc} x & y & \theta & x_\ell & y_\ell \end{array} \\ \begin{array}{c} x_z \\ y_z \end{array} & \begin{pmatrix} -\cos\theta & -\sin\theta & -\sin\theta(x_\ell - x) + \cos\theta(y_\ell - y) & \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta & -\cos\theta(x_\ell - x) - \sin\theta(y_\ell - y) & -\sin\theta & \cos\theta \end{pmatrix} \end{array}$$

$\frac{\partial h}{\partial x}$ vs $\left.\frac{\partial h}{\partial x}\right|_{x_0}$

symbolic

numerical diff.

automatic diff

~~$h(r,\ell) \overset{?}{\approx} A \cdot \begin{pmatrix} x \\ y \\ \theta \\ x_\ell \\ y_\ell \end{pmatrix}$~~ at

$$h(r,\ell) \approx A(x_0)\left(\begin{pmatrix} x \\ y \\ \theta \\ x_\ell \\ y_\ell \end{pmatrix} - x_0 \right) + h(x_0)$$

$$x_0 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \qquad h(x_0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$A(x_0) = \begin{pmatrix} -1 & 0 & \cancel{y_\ell}\ 0 & 1 & 0 \\ 0 & -1 & \cancel{-x_\ell}\ 0 & 0 & 1 \end{pmatrix}$$