TAM 598 Lecture 20: State Space Models – Kalman Filters

Announcements:

- HW 5 covers lectures 17-20; due on Fri Apr 18 last time:

initial conditions $\underset{\sim}{x}_0 = \underline{\mu}_0 + \underline{z}_0$ $\quad \underline{z}_0 \sim N(\underline{0}, \underline{\underline{V}}_0)$

transitions $\underline{x}_{t+1} = \underline{\underline{A}}\,\underline{x}_t + \underline{\underline{B}}\,\underline{u}_t + \underline{z}_t$ $\quad \underset{\sim}{z}_t \sim N(\underline{0}, \underline{\underline{Q}})$

emissions $\underline{y}_t = \underline{\underline{C}}\,\underline{x}_t + \underline{w}_t$ $\quad \underset{\sim}{w}_t \sim N(\underline{0}, \underline{\underline{R}})$

Kalman Filter - recursively computes the posterior distribution $p(\underline{x}_t \mid \underline{y}_{1:t}, \underline{u}_{1:t})$ using two steps:

$$\begin{cases} \text{(1) prediction step} - \text{obtain prior } p(\underset{\sim}{x}_t \mid \underline{y}_{1:t-1}, \underline{u}_{1:t}) \\ \text{(2) update step} - \text{obtain posterior } p(\underline{x}_t \mid \underline{y}_{1:t}, \underline{u}_{1:t}) \end{cases}$$

prediction step goal $p(\underline{x}_t \mid \underline{y}_{1:t-1}, \underline{u}_{1:t})$

say we have $p(\underline{x}_{t-1} \mid \underline{y}_{1:t-1}, \underline{u}_{1:t-1}) = N(\underline{x}_{t-1} \mid \underline{\mu}_{t-1}, \underline{\underline{V}}_{t-1})$

we use our update rules and marginalize over $\underline{x}_{t-1}$:

$$p(\underline{x}_t \mid \underline{y}_{1:t-1}, \underline{u}_{1:t}) =$$

$$= \int p(\underline{x}_t \mid \underline{x}_{t-1}, \underline{u}_t)\, p(\underline{x}_{t-1} \mid \underline{y}_{1:t-1}, \underline{u}_{1:t-1})\, d\underline{x}_{t-1}$$

$$= \int N(\underline{x}_t \mid \underline{\underline{A}}\,\underline{x}_{t-1} + \underline{\underline{B}}\,\underline{u}_t, \underline{\underline{Q}})\, N(\underline{x}_{t-1} \mid \underline{\mu}_{t-1}, \underline{\underline{V}}_{t-1})\, d\underline{x}_{t-1}$$

comes from prop's of Gaussians

$$= N\Big(\underline{x}_t \,\Big|\, \underbrace{\underline{\underline{A}}\,\underline{\mu}_{t-1} + \underline{\underline{B}}\,\underline{u}_t}_{\underline{\mu}_t^P}, \underbrace{\underline{\underline{Q}} + \underline{\underline{A}}\,\underline{\underline{V}}_{t-1}\underline{\underline{A}}^T}_{\underline{V}_t^P}\Big)$$

update step

goal: $p(\underline{x}_t \mid \underline{y}_{1:t}, \underline{u}_{1:t})$

we now condition our prior on a new measurement $\underline{y}_t$ to obtain the posterior using Bayes' Rule:

$$p(\underline{x}_t \mid \underline{y}_{1:t}, \underline{u}_{1:t}) \propto p(\underline{y}_t \mid \underline{x}_t)\, p(\underline{x}_t \mid \underline{y}_{1:t-1}, \underline{u}_{1:t})$$

$$= N(\underline{y}_t \mid \underline{\underline{C}}\,\underline{x}_t, \underline{\underline{R}})\, N(\underline{x}_t \mid \underline{\mu}_t^P, \underline{\underline{V}}_t^P)$$

$$= N(\underline{x}_t \mid \underline{\mu}_t, \underline{\underline{V}}_t)$$

where

$$\left.\begin{aligned} \underline{\mu}_t &= \underline{\mu}_t^P + \underline{\underline{K}}_t(\underline{y}_t - \underline{\underline{C}}\,\underline{\mu}_t^P) \\ \underline{\underline{V}}_t &= \underline{\underline{V}}^P - \underline{\underline{K}}_t \underline{\underline{C}}\, \underline{\underline{V}}_t^P \end{aligned}\right\}$$

and Kalman gain matrix

$$\underline{\underline{K}}_t = \underline{\underline{V}}_t^P \underline{\underline{C}}^T (\underline{\underline{C}}\,\underline{\underline{V}}_t^P \underline{\underline{C}}^T + \underline{\underline{R}})^{-1}$$

SMOOTHING - now we want the posterior $p(\underline{x}_t \mid \underline{y}_{1:T}, \underline{u}_{1:T})$ of the distribution of the state at t, given all measurements and control inputs up to some future time T.

use info from future to improve est of state at time t.

- start w/ filtering distribution $p(\underline{x}_T \mid \underline{y}_{1:T}, \underline{u}_{1:T})$, which is gaussian with mean $\underline{\mu}_T$, covariance $\underline{\underline{V}}_T$

- for the previous time step $p(\underline{x}_{T-1} \mid \underline{y}_{1:T}, \underline{u}_{1:T})$, consider the joint distribution of $\underline{x}_T, \underline{x}_{T-1}$

$$p(\underline{x}_T, \underline{x}_{T-1} \mid \underline{y}_{1:T}, \underline{u}_{1:T})$$

$$= \underbrace{p(\underline{x}_{T-1} \mid \underline{x}_T, \underline{y}_{1:T}, \underline{u}_{1:T})}_{\text{consider on next page}} \underbrace{p(\underline{x}_T \mid \underline{y}_{1:T}, \underline{u}_{1:T})}_{\text{known}}$$

$$p(\underline{x}_{T-1} \mid \underline{x}_T, \underline{y}_{1:T}, \underline{u}_{1:T})$$

$$= p(\underline{x}_{T-1} \mid \underline{x}_T, \underline{y}_{1:T-1}, \underline{u}_T)$$

to trace $\underline{x}_{T-1}$ back from $\underline{x}_T$, only $\underline{u}_T$ matters.

only obs up to y_{T-1} too

$$= \frac{① \; p(\underline{x}_T \mid \underline{x}_{T-1}, \underline{y}_{1:T-1}, \underline{u}_T) \; ② \; p(\underline{x}_{T-1} \mid \underline{y}_{1:T-1}, \underline{u}_T)}{③ \; p(\underline{x}_T \mid \underline{y}_{1:T-1}, \underline{u}_T)}$$

③ prediction step of Kalman step $\quad N(\underline{x}_T \mid \underline{\mu}_T^P, \underline{\underline{V}}_T^P)$

② filtering dist at time $T-1 \quad N(\underline{x}_{T-1} \mid \underline{\mu}_{T-1}, \underline{\underline{V}}_{T-1})$

① use Markov property

$$p(\underline{x}_T \mid \underline{x}_{T-1}, \underline{u}_T) = N(\underline{x}_T \mid \underline{\underline{A}}\,\underline{x}_{T-1} + \underline{\underline{B}}\,\underline{u}_T, \underline{\underline{Q}})$$

from here:

① our joint distribution $p(\underline{x}_T, \underline{x}_{T-1} \mid \underline{y}_{1:T}, \underline{u}_{1:T})$ is gaussian

② integrate out $\underline{x}_T$ to get our smoothed distribution

$$\underbrace{p(\underline{x}_{T-1} \mid \underline{y}_{1:T}, \underline{u}_{1:T})}_{\text{SMOOTHED}} = \int p(\underline{x}_T, \underline{x}_{T-1} \mid y_{1:T}\ u_{1:T})\, d\underline{x}_T$$

③ Result is gaussian with

mean $\underline{\mu}_{T-1|T} = \underline{\mu}_{T-1} + \underline{\underline{J}}_T (\underline{\mu}_T - \underline{\mu}_T^P)$

covariance $\underline{\underline{V}}_{T-1|T} = \underline{\underline{V}}_{T-1} + \underline{\underline{J}}_T (\underline{\underline{V}}_T - \underline{\underline{V}}_T^P) \underline{\underline{J}}_T^T$

where $\underline{\underline{J}}_T = \underline{\underline{V}}_{T-1} \underline{\underline{A}}^T \underline{\underline{V}}_T^{-1}$ is the smoothed gain matrix

then go back, one step at a time, to obtain the distribution for all time steps