Dynamic Model

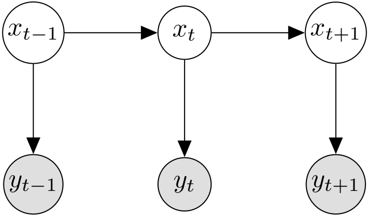
Thomas Bayes

May 14, 2018

Introduction

We are going to introduce the state space models and their corresponding algorithms, **Kalman Filter** and **Particle Filter**.

State Space Model



In state space model, we have state variable sequence $x_1, \ldots, x_t$ and observed variables sequence $y_1, \ldots, y_t$. State variables means the hidden state behind the world. We now have the representation of state space model, now assume we know the all parameters of our model. What if we want to do inference of the model? In such model, we are particularly interested in two tasks.

- Prediction

$$p(y_t | x_{1:t})$$

- Update

$$p(y_t | x_{1:t-1})$$

Inference of Kalman Filter

In Kalman Filter, we assume that

$$p(x_t|x_{t-1}) \sim \mathcal{N}(Ax_{t-1} + B, Q)$$

and

$$p(y_t|x_t) \sim \mathcal{N}(Hx_t, R)$$

So we say that Kalman Filter is a linear model. In inference task, we assume that we already know all the parameters of the model, A, B, Q, H, R, and based on the observation $y_{1:t}$, we infer the probability distribution of the underlying state x_t. Here note that we can't use information from future, y_{t+1} to infer the underlying state x_t.

From Update to Prediction

If

$$p(x) \sim \mathcal{N}(x|\mu, \Sigma) \qquad p(y|x) \sim \mathcal{N}(y|Ax + b, L)$$

we have

$$p(y) = \int_x p(y|x)p(x)dx \sim \mathcal{N}(y|A\mu + b, L + A\Sigma A^\top)$$

So for prediction task, we have

$$\begin{aligned} p(x_t|y_{1:t-1}) \sim \mathcal{N}(\bar{\mu}_t, \bar{\Sigma}_t) &= \int_{x_{t-1}} p(x_t|x_{t-1})p(x_{t-1}|y_{1:t-1})dx_{t-1} \\ &= \int_{x_{t-1}} \mathcal{N}(x_t|Ax_{t-1} + B, Q)\mathcal{N}(x_{t-1}|\hat{\mu}_{t-1}, \hat{\Sigma}_{t-1}) \\ &= \mathcal{N}(x_t|A\hat{\mu}_{t-1} + B, A\hat{\Sigma}_{t-1}A^\top + Q) \end{aligned}$$

So we use the parameters $\hat{\mu}_{t-1}$ and $\hat{\Sigma}_{t-1}$ of previous **update** to infer the **prediction** of state t.

From Prediction to Update I

For this task, we use an alternative approach, by taking advantage of property of Gaussian distribution

$$x_{t-1}|y_{1:t-1} \sim \mathcal{N}(\mathbb{E}[x_{t-1}], \hat{\Sigma}_{t-1})$$

We attempt to write Δx_t and Δy_t in term of Δx_{t-1}

$$x_t = Ax_{t-1} + w_t \quad w_t \sim \mathcal{N}(0, Q)$$

And we get

$$\begin{aligned} x_t|y_{1:t-1} &= A(\Delta x_{t-1} + \mathbb{E}[x_{t-1}]) + w_t \\ &= A\mathbb{E}[x_{t-1}] + B + \underbrace{A\Delta x_{t-1} + w_t}_{\Delta x_t|y_{1:t-1}} \end{aligned}$$

From Prediction to Update II

For y_t,

$$y_t = Hx_{t-1} + v_t \quad v_t \sim \mathcal{N}(0, R)$$

$$\begin{aligned} y_t|y_{1:t-1} &= H(A\mathbb{E}[x_{t-1}] + B + A\Delta x_{t-1} + w_t) + v_t \\ &= HA\mathbb{E}[x_{t-1}] + HB + \underbrace{HA\Delta x_{t-1} + Hw_t + v_t}_{\Delta y_t|y_{1:t-1}} \end{aligned}$$

So we have the expression of $x_t|y_{1:t-1}$ and $y_t|y_{1:t-1}$. And we are aiming to get the expression of $x_t|y_{1:t}$. For Gaussian distribution, we have the property

$$p(u) = \mathcal{N}(\mu_u, \Sigma_{uu}) \qquad p(v) = \mathcal{N}(\mu_v, \Sigma_{vv})$$

and

$$p(u|v) = \mathcal{N}(\mu_u + \Sigma_{uv}\Sigma_{vv}^{-1}(v - \mu_v), \Sigma_{uu} - \Sigma_{uv}\Sigma_{vv}^{-1}\Sigma_{vu})$$

From Prediction to Update III

Imagine $p(u) = p(x_t|y_{1:t-1})$ and $p(v) = p(y_t|y_{1:t-1})$, so $p(u|v) = p(x_t|y_{1:t})$. Now we just need to calculate $\Sigma_{uu}, \Sigma_{uv}, \Sigma_{vv}$.

$$\Sigma_{uu} = A\hat{\Sigma}_{t-1}A^\top + Q = \bar{\Sigma}_t$$

$$\Sigma_{vv} = HA\hat{\Sigma}_{t-1}A^\top H^\top + HQH^\top + R = H\bar{\Sigma}_t H^\top + R$$

$$\Sigma_{uv} = H(A\hat{\Sigma}_{t-1}A^\top + Q) = H\bar{\Sigma}_t$$

We substitute these into the following and will get the result.

$$p(u|v) = \mathcal{N}(\mu_u + \Sigma_{uv}\Sigma_{vv}^{-1}(v - \mu_v), \Sigma_{uu} - \Sigma_{uv}\Sigma_{vv}^{-1}\Sigma_{vu})$$

So we have used $\bar{\Sigma}_t, \hat{\mu}_t$ and parameters to express $p(x_t|y_{1:t})$. So from the update parameter $p(x_{t-1}|y_{1:t-1}) \sim \mathcal{N}(\hat{\mu}_{t-1}, \hat{\Sigma}_{t-1})$ we get the prediction parameter of $p(x_t|y_{1:t-1})$. And from parameter of prediction $\bar{\Sigma}_t, \hat{\mu}_t$, we can get the paramter of update.

From Prediction to Update IV

So the pipeline is like

$$\text{prediction}_t \longrightarrow \text{update}_t \longrightarrow \text{prediction}_{t+1}$$

Note

Here we present a new way of deriving formula by using moment representation of Gaussian distribution. So next time when we encounter derivation concerning Gaussian distribution, we might be using moment representation of Gaussian distribution.

Learning Kalman Filter I