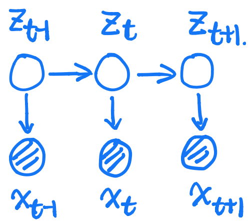
15. 卡曼滤波 Kalman Filter

P1 Dynamic Model (State Space Model).

$\begin{cases}\text{HMM (State is discrete). 状态是离散的} \\ \text{Linear Dynamic Model (Kalman Filter) Linear Gaussian Model. 状态是连续的} \\ \text{Particle Filter (Non-Linear, Non-Gaussian).}\end{cases}$

图 + 时序 $\longrightarrow$ Dynamic Model (动态模型)



Learning: $\lambda = (\pi, A, B)$

Inference:
- decoding: $P(z_1, z_2, \cdots, z_t \mid x_1, x_2, \cdots, x_t)$.
- Prob of evidence: $P(X \mid \theta) = P(x_1, x_2, \cdots, x_T \mid \theta)$.
- filtering: $P(z_t \mid x_1, \cdots, x_t)$ online.
- smoothing: $P(z_t \mid x_1, \cdots, x_T)$ offline.
- prediction: $\begin{cases} P(z_{t+1}, z_{t+2} \mid x_1, \cdots, x_t) \\ P(x_{t+1}, x_{t+2} \mid x_1, \cdots, x_t) \end{cases}$

Linear:
$$z_t = A \cdot z_{t-1} + B + \varepsilon \quad \text{Gaussian Distribution.}$$
$$x_t = C \cdot z_t + D + \delta \quad \varepsilon \sim N(0, Q) \quad \delta \sim N(0, R).$$

HMM: $\lambda = (\pi, A, B)$

状态转移矩阵:

$A = [a_{ij}] \quad a_{ij} = P(i_{t+1} = q_j \mid i_t = q_i)$.

发射矩阵:

$B = [b_j(k)] \quad b_j(k) = P(o_t = v_k \mid i_t = q_j)$

Kalman Filter

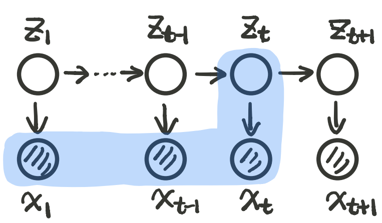
$z_t \mid z_{t-1} \sim N(A \cdot z_{t-1} + B, Q)$. 状态转移概率

$x_t \mid z_t \sim N(C \cdot z_t + D, R)$ 发射概率

$z_1 \sim N(\mu_1, \Sigma_1)$ 初始状态

$\theta = (A, B, C, D, Q, R, \mu_1, \Sigma_1)$ 模型参数

P2 Kalman Filter



$\begin{cases}\text{learning} \\ \text{inference. } P(Z \mid X) \quad \text{filtering}\end{cases}$

(marginal posterior).

$P(z_t \mid z_{t-1}) = N(A \cdot z_{t-1} + B, Q)$

$P(x_t \mid z_t) = N(C \cdot z_t + D, R)$

$$P(z_t \mid x_1, \cdots, x_t) = \frac{P(x_1, \cdots, x_t, z_t)}{P(x_1, \cdots, x_t)} \text{ const.}$$

$$\propto P(x_1, \cdots, x_t, z_t) = \underbrace{P(x_t \mid x_1, \cdots, x_{t-1}, z_t)}_{P(x_t \mid z_t)} \cdot P(x_1, \cdots, x_{t-1}, z_t)$$

$$= P(x_t \mid z_t) \cdot P(x_1, \cdots, x_{t-1}, z_t)$$

$$= P(x_t \mid z_t) \cdot \underbrace{P(z_t \mid x_1, \cdots, x_{t-1})}_{\text{prediction问题}} \cdot \overset{\text{const}}{P(x_1, \cdots, x_{t-1})}$$

$$\propto P(x_t \mid z_t) \cdot P(z_t \mid x_1, \cdots, x_{t-1})$$

$$= \int_{z_{t-1}} P(z_t, z_{t-1} \mid x_1, \cdots, x_{t-1}) \, dz_{t-1}.$$

$$= \int_{z_{t-1}} \underbrace{P(z_t \mid z_{t-1}, x_1, \cdots, x_{t-1})}_{P(z_t \mid z_{t-1})} \cdot P(z_{t-1} \mid x_1, \cdots, x_{t-1}) \, dz_{t-1}$$

步骤: $t=1$ $\begin{cases} P(z_1|x_1) & \text{update (correction)} \\ P(z_2|x_1) & \text{prediction} \end{cases}$

(online)

$t=2$ $\begin{cases} P(z_2|x_1,x_2) & \text{update (correction)} \\ P(z_3|x_1,x_2) & \text{prediction} \end{cases}$

$\vdots$

$t=T$ $\begin{cases} P(z_T|x_1,\cdots,x_T) \\ P(z_{T+1}|x_1,\cdots,x_T) \end{cases}$

P3 Filtering:

① prediction.

$$P(z_t|x_1,\cdots,x_{t-1}) = \int_{z_{t-1}} P(z_t|z_{t-1})\cdot \underbrace{P(z_{t-1}|x_1,\cdots,x_{t-1})}_{P(z_{t-1})}\, dz_{t-1}.$$

$$N(\mu_t^*, \Sigma_t^*) = \int N(z_t|Az_{t-1}+B, Q)\cdot N(\mu_{t-1},\Sigma_{t-1})$$

$$\boxed{\begin{aligned} &p(x) = N(x|\mu, \Lambda^{-1}) && (1)\\ &p(y|x) = N(y|Ax+b, L^{-1}) && (2)\\ &P(y) = N(y|A\mu+b, L^{-1}+A\Lambda^{-1}A^T) && (3)\end{aligned}}$$

令 $x=z_{t-1}$, $y=z_t$. 代入 (1)(2)(3)

$$\begin{cases} \mu_t^* = A\mu_{t-1}+B \\ \Sigma_t^* = Q + A\Sigma_{t-1}A^T \end{cases}$$

② update.

$$\underbrace{P(z_t|x_1,\cdots,x_t)}_{p(x|y)} \propto \underbrace{P(x_t|z_t)}_{p(y|x)}\cdot \underbrace{P(z_t|x_1,\cdots,x_{t-1})}_{p(x)}$$

$$N(\mu_t,\Sigma_t) \propto N(x_t|Cz_t+D, R)\cdot N(\mu_t^*,\Sigma_t^*)$$

$$\Rightarrow \begin{cases} \mu_t = ? \\ \Sigma_t = ? \end{cases}$$