线性动态系统(Linear Dynamic System)

一、背景

对于 Linear Dynamic System (Kalman Filter)，其状态变量和观测变量均服从高斯分布，因此也叫做线性高斯模型(Linear Guassian Model)

Dynamic Model (state space Model)
- HMM → 状态离散
- Linear Dynamic System (Kalman Filter) → Linear Guassian Model
- Particle Filter → None-linear, non-Guassian

Kalman Filter (Linear Guassian Model)

Linear：体现在两个方面：

$$z_t = A\cdot z_{t-1} + B + \varepsilon$$
$$x_t = C\cdot z_t + D + \delta$$

Guassian：体现在噪声 ε 和 δ 上，

$$\varepsilon \sim N(0, Q)$$
$$\delta \sim N(0, R)$$

在HMM中，$\lambda = (\pi, A, B)$，其中转移矩阵 $A = [a_{ij}]$，$a_{ij} = P(i_{t+1} = q_j \mid i_t = q_i)$，它描述了前一时刻状态和后一时刻状态的关系；$B = [b_j(k)]$，$b_j(k) = P(o_t = v_k \mid i_t = q_j)$，它描述了 i_t 和 o_t 的关系。

对于 Kalman Filter，其式子：

$$z_t = A\cdot z_{t-1} + B + \varepsilon$$
$$x_t = C\cdot z_t + D + \delta$$

也可以写为：

$$P(z_t \mid z_{t-1}) \sim N(Az_{t-1} + B, Q)$$
$$P(x_t \mid z_t) \sim N(Cz_t + D, R)$$

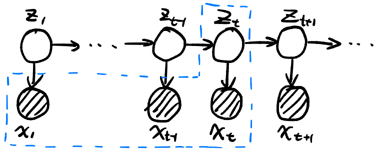
它们同样描述状态和状态之间的关系，x_t 和 z_t 的关系，只不过不同点是它们是连续的，所以会满足一定的线性关系，且是服从高斯分布的。

系统初始状态 $z_1 \sim N(\mu_1, \Sigma_1)$

综上，Kalman Filter 的参数 $\theta = (A, B, C, D, Q, R, \mu_1, \Sigma_1)$

二. Filtering 问题

Kalman Filter 模型如下：



$$\begin{cases} P(z_t|z_{t-1}) \sim N(Az_{t-1}+B, Q) \\ P(x_t|z_t) \sim N(Cz_t+D, R) \\ P(z_1) \sim N(\mu_1, \Sigma_1) \end{cases}$$

$$\begin{cases} z_t = Az_{t-1}+B+\varepsilon, & \varepsilon \sim N(0,Q) \\ x_t = Cz_t+D+\delta, & \delta \sim N(0,R) \end{cases}$$

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

Filtering (marginal posterior)

$$\boxed{P(z_t|x_1,\cdots,x_t)} = \frac{P(x_1,\cdots,x_t,z_t)}{\underbrace{P(x_1,\cdots,x_t)}_{\text{观测变量的联合概率，可求出！}}} \propto P(x_1,x_2,\cdots,x_t,z_t)$$

$$= P(x_t|x_1,x_2,\cdots,x_{t-1},z_t)\cdot P(x_1,x_2,\cdots,x_{t-1},z_t)$$

$$= P(x_t|z_t)\cdot P(x_1,x_2,\cdots,x_{t-1},z_t)$$

$$= P(x_t|z_t)\cdot \underbrace{P(z_t|x_1,x_2,\cdots,x_{t-1})}_{\text{prediction}}\cdot \underbrace{P(x_1,x_2,\cdots,x_{t-1})}_{\text{可求出}}$$

$$\propto P(x_t|z_t)\cdot \underbrace{P(z_t|x_1,x_2,\cdots,x_{t-1})}$$

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

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

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

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

很明显这是个online的过程

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

由于 $t=2$ 时获取到观测变量再求 $P(z_2|x_1,x_2)$，是对上一时间步预测的一个修正，因此update也叫做correction

⋮

Filtering 问题求解.

$$\text{已知}\begin{cases} p(x) \sim N(x|\mu, \Lambda^{-1}) \\ p(y|x) \sim N(y|Ax+b, L^{-1}) \end{cases}$$

$$\text{则}\begin{cases} p(y) \sim N(y|A\mu+b, L^{-1}+A\Lambda^{-1}A^T) \\ p(x|y) \sim N \end{cases}$$

Filtering:

step 1: prediction

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

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

step 2: Update

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

$$N(\mu_t, \Sigma_t) \qquad N(x_t|C \cdot z_t+D, R) \cdot N(\mu_t^*, \Sigma_t^*)$$

prediction:

由红框知，$p(x)$ 对应的是 $P(z_{t-1}|x_1,\cdots,x_{t-1})$，$p(y|x)$ 对应的 $P(z_t|z_{t-1})$，$p(y)$ 对应的是 $P(z_t|x_1,\cdots,x_{t-1})$

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

update:

由红框知，将 $P(z_t|x_1,\cdots,x_{t-1})$ 看作 $p(x)$，将 $P(x_t|z_t)$ 看作 $p(y|x)$，将 $x_1,\cdots,x_{t-1}$ 为观测变量，可看作常数省略，则 $P(z_t|x_1,\cdots x_t) \propto P(z_t|x_t)$，可将其看作 $p(x|y)$

套用公式可得 μ_t 和 Σ_t