线性动态系统

HMM 模型适用于隐变量是离散的值的时候，对于连续隐变量的 HMM，常用线性动态系统描述线性高斯模型的态变量，使用粒子滤波来表述非高斯非线性的态变量。

LDS 又叫卡尔曼滤波，其中，线性体现在上一时刻和这一时刻的隐变量以及隐变量和观测之间：

$$
\begin{align}
z_t&=A\cdot z_{t-1}+B+\varepsilon\\
x_t&=C\cdot z_t+D+\delta\\
\varepsilon&\sim\mathcal{N}(0,Q)\\
\delta&\sim\mathcal{N}(0,R)
\end{align}
$$

类比 HMM 中的几个参数：

$$
\begin{align}
p(z_t|z_{t-1})&\sim\mathcal{N}(A\cdot z_{t-1}+B,Q)\\
p(x_t|z_t)&\sim\mathcal{N}(C\cdot z_t+D,R)\\
z_1&\sim\mathcal{N}(\mu_1,\Sigma_1)
\end{align}
$$

在含时的概率图中，除了对参数估计的学习问题外，在推断任务中，包括译码，证据概率，滤波，平滑，预测问题，LDS 更关心滤波这个问题：$p(z_t|x_1,x_2,\cdots,x_t)$。类似 HMM 中的前向算法，我们需要找到一个递推关系。

$$
p(z_t|x_{1:t})=p(x_{1:t},z_t)/p(x_{1:t})=Cp(x_{1:t},z_t) \tag{8}
$$

对于 $p(x_{1:t},z_t)$：

$$
\begin{align}
p(x_{1:t},z_t)&=p(x_t|x_{1:t-1},z_t)p(x_{1:t-1},z_t)=p(x_t|z_t)p(x_{1:t-1},z_t)\nonumber\\
&=p(x_t|z_t)p(z_t|x_{1:t-1})p(x_{1:t-1})=Cp(x_t|z_t)p(z_t|x_{1:t-1}) \tag{9}
\end{align}
$$

我们看到，右边除了只和观测相关的常数项，还有一项是预测任务需要的概率。对这个值：

$$
\begin{align}
p(z_t|x_{1:t-1})&=\int_{z_{t-1}}p(z_t,z_{t-1}|x_{1:t-1})dz_{t-1}\nonumber\\
&=\int_{z_{t-1}}p(z_t|z_{t-1},x_{1:t-1})p(z_{t-1}|x_{1:t-1})dz_{t-1}\nonumber\\
&=\int_{z_{t-1}}p(z_t|z_{t-1})p(z_{t-1}|x_{1:t-1})dz_{t-1} \tag{10}
\end{align}
$$

我们看到，这又化成了一个滤波问题。于是我们得到了一个递推公式：

1. $t=1$，$p(z_1|x_1)$，称为 update 过程，然后计算 $p(z_2|x_1)$，通过上面的积分进行，称为 prediction 过程。
2. $t=2$，$p(z_2|x_2,x_1)$ 和 $p(z_3|x_1,x_2)$

我们看到，这个过程是一个 Online 的过程，对于我们的线性高斯假设，这个计算过程都可以得到解析解。

1. Prediction：

$$p(z_t|x_{1:t-1})=\int_{z_{t-1}}p(z_t|z_{t-1})p(z_{t-1}|x_{1:t-1})dz_{t-1}=\int_{z_{t-1}}\mathcal{N}(Az_{t-1}+B,Q)\mathcal{N}(\mu_{t-1},\Sigma_{t-1})dz_{t-1} \tag{11}$$

其中第二个高斯分布是上一步的 Update 过程，所以根据线性高斯模型，直接可以写出这个积分：

$$p(z_t|x_{1:t-1})=\mathcal{N}(A\mu_{t-1}+B,Q+A\Sigma_{t-1}A^T) \tag{12}$$

2. Update:

$$p(z_t|x_{1:t})\propto p(x_t|z_t)p(z_t|x_{1:t-1}) \tag{14}$$

同样利用线性高斯模型，也可以直接写出这个高斯分布。