

partical filter(粒子滤波)

1 Importance sampling(重要性采样)

$$E(f(x)) = \int_x f(x)p(x)dx \quad (1)$$

其中, x 服从分布 $x \sim p(x)$

进行蒙特卡洛采样:

$$E(f(x)) \approx \frac{1}{N} \sum_{i=1}^N f(x^i) \quad (2)$$

$$x^1, x^2 \dots x^N \sim p(x)$$

对于高维非高斯非线性的概率分布, 我们无法直接获得 $p(x)$ 的概率密度函数, 我们希望用能在其他的分布来代替原来的分布采用。假设我们按照我们设定的分布函数 $q(x)$ 对 x 进行采样, 不是一般性地, (1)式可以改写成:

$$E(f(x)) = \int_x f(x) \frac{p(x)}{q(x)} \cdot q(x) dx \approx \frac{1}{N} \sum_{i=1}^N f(x^i) \frac{p(x^i)}{q(x^i)} \cdot q(x^i) \quad (3)$$

其中, $x^1, x^2 \dots x^N \sim q(x)$.

这里的问题是如何选择合适的 $q(x)$ 分布, 我们在这里记权 $w^i = \frac{p(x^i)}{q(x^i)}$ 。

2 Sequential Importance Sampling(SIS)

2.1 高维度采样

假设:

$$p(x) = \frac{\gamma(x)}{\int f(x)dx} \quad (4)$$

其中 $\int f(x)dx$ 为归一化常数，不会影响样本的相对权重，所以我们只考虑 γ 的分布情况。

对于 N 维的状态，有：

$$x_{1:N} \equiv x_1, x_2 \dots x_N$$

对应的的权为：

$$w_{x_{1:N}}^i = \frac{\gamma(x_{1:N}^i)}{q(x_{1:N}^i)} \quad (5)$$

直接对高维度的状态进行采样对于实际操作来说是十分困难或者说是非常消耗资源的，因为我们可以用边缘概率的思想来递归对高维度状态采样。

对于维度1，可以求 $2 : N$ 维的边缘分布有：

$$\gamma(x_1) = \int_{x_{2:N}} \gamma(x_{1:N}) dx_{2:N} \quad (6)$$

对应的第 i 个样本权重为：

$$w(x_1^i) = \frac{\gamma(x_1^i)}{q(x_1^i)} \quad (7)$$

其中 $x_1^i \sim q(x_1)$ 。

读者需要注意的是本文中上标代表着样本的序号，下表代表的是状态所在的维度。

对于维度2，同理可以得到：

$$\gamma(x_{1:2}) = \int_{x_{3:N}} \gamma(x_{1:N}) dx_{3:N} \quad (6)$$

对应的第 i 个样本权重为：

$$w(x_1^i) = \frac{\gamma(x_{1:2}^i)}{q(x_{1:2}^i)} \quad (7)$$

其中 $x_{1:2}^i \sim q(x_{1:2})$.

对于联合分布的两个状态 $(x_1, x_2) \sim p(x_1, x_2)$ ，如果我们已经对维度1的样本进行采样，根据条件概率有：

$$\begin{aligned} x_1^i &\sim q(x_1) \\ x_2^i &\sim q(x_2 | x_1^i) \end{aligned} \quad (8)$$

推广到 N 维，则有：

$$\begin{aligned} x_{1:N-1}^i &\sim q(x_{1:N-1}) \\ x_N^i &\sim q(x_N | x_{1:N-1}^i) \end{aligned} \quad (8)$$

因此基于(7)式，我们可以推广到 $1 : N$ 维得：

$$\begin{aligned} w(x_{1:N}^i) &= \frac{\gamma(x_{1:N}^i)}{q(x_{1:N}^i)} \\ &= \frac{\gamma(x_{1:N}^i)}{q(x_N | x_{1:N-1}) \cdot q(x_{1:N-1}^i)} \\ &= \frac{\gamma(x_{1:N}^i) \cdot \gamma(x_{1:N-1})}{q(x_N | x_{1:N-1}) \cdot q(x_{1:N-1}^i) \cdot \gamma(x_{1:N-1})} \\ &= w(x_{1:N-1}^i) \frac{\gamma(x_{1:N}^i)}{q(x_N | x_{1:N-1}) \cdot \gamma(x_{1:N-1})} \end{aligned} \quad (9)$$

这样我们得到了 $1 : N - 1$ 到 $1 : N$ 维的递归公式，这样我们可以大大降低高维度的采样难度。

2.2 dynamic model

在时序状态估计中，通常处理不是高维度问题，这里我们可以把上述的高维度问题转化为时间动态问题。对于动态问题，可以获得运动方程和观测方程和相应的条件概率：

状态转移方程	观测方程
$x_t = A_t \cdot x_{t-1} + w_t$	$y_t = H_t \cdot x_t + \eta_t$

状态转移方程	观测方程
$p(x_t x_{t-1})$	$p(y_t x_t)$

上述高维度的 $\gamma(x_{1:N})$ 在动态模型中可以表示为 $p(x_{1:t}, y_{1:t})$

根据(9)式，动态模型的得权重递推关系为：

$$\begin{aligned}
 w(x_{1:t}^i) &= w(x_{1:t-1}^i) \frac{p(x_{1:t}^i, y_{1:t}^i)}{q(x_t^i | x_{1:t-1}^i) \cdot p(x_{1:t-1}^i, y_{1:t-1}^i)} \\
 &= w(x_{1:t-1}^i) \frac{p(y_t^i | x_{1:t}^i, y_{1:t-1}^i) \cdot p(x_{1:t}^i, y_{1:t-1}^i)}{q(x_t^i | x_{1:t-1}^i) \cdot p(x_{1:t-1}^i, y_{1:t-1}^i)}
 \end{aligned} \tag{10}$$

因为 y_t 只与 x_t 有关， x_t 还与 x_{t-1} 有关，其他可省略，式(10)可以写为：

$$\begin{aligned}
 w(x_{1:t}^i) &= w(x_{1:t-1}^i) \frac{p(y_t^i | x_{1:t}^i) \cdot p(x_t^i | x_{t-1}^i) \cdot p(x_{1:t-1}^i, y_{1:t-1}^i)}{q(x_t^i | x_{1:t-1}^i) \cdot p(x_{1:t-1}^i, y_{1:t-1}^i)} \\
 &= w(x_{1:t-1}^i) \frac{p(y_t^i | x_{1:t}^i) \cdot p(x_t^i | x_{t-1}^i)}{q(x_t^i | x_{1:t-1}^i)}
 \end{aligned} \tag{11}$$

其中 $x_t^i \sim q(x_t^i | x_{t-1}^i)$

这里我们把动态模型的递归公式也讲清楚了。但其实前面的一直没有讲到一个大问题，就是 q 的是一个怎么分布的概率函数，我们的采样原则就是按照我们自己定义的概率分布来采样的。下面我们介绍下 q 的是怎么确定。

根据图模型可知， x_t 与 y_t 、 x_{t-1} ，所以最佳的 $q(x_t)$ 分布为：

$$q(x_t) = p(x_t | x_{t-1}, y_t) \tag{12}$$

根据贝叶斯条件概率公式，式(12)可以改写为：

$$\begin{aligned}
p(x_t | x_{t-1}, y_t) &= \frac{p(x_t, x_{t-1}, y_t)}{p(x_{t-1}, y_t)} \\
&= \frac{p(y_t | x_t, x_{t-1}) \cdot p(x_t | x_{t-1}) \cdot p(x_{t-1})}{p(y_t | x_{t-1}) \cdot p(x_{t-1})} \\
&= \frac{p(y_t | x_t, x_{t-1}) \cdot p(x_t | x_{t-1})}{p(y_t | x_{t-1})} \\
&= \frac{p(y_t | x_t) \cdot p(x_t | x_{t-1})}{p(y_t | x_{t-1})}
\end{aligned} \tag{13}$$

将(13)式代入(11)式，得：

$$\begin{aligned}
w(x_{1:t}^i) &= w(x_{1:t-1}^i) \cdot p(y_t^i | x_{1:t}^i) \cdot p(x_t^i | x_{t-1}) \cdot \frac{1}{p(x_t^i | x_{t-1}^i, y_t^i)} \\
&= w(x_{1:t-1}^i) \cdot (y_t^i | x_{1:t}^i) \cdot p(x_t^i | x_{t-1}) \cdot \frac{p(y_t^i | x_{t-1}^i)}{p(y_t^i | x_t^i) \cdot p(x_t^i | x_{t-1}^i)} \\
&= w(x_{1:t-1}^i) \cdot p(y_t^i | x_{t-1}^i)
\end{aligned} \tag{14}$$

要求得 $p(y_t | x_{t-1})$ 的分布，需要对 x_t 进行积分：

$$p(y_t | x_{t-1}) = \int_{x_t} p(y_t | x_t) \cdot p(x_t | x_{t-1}) dx_t \tag{15}$$

要解决式(14)这样的一个问题，我们又需要进行蒙特卡罗采样，这并不是我们想要的。所以要解决 x_t 的采样问题，观察我们上面动态模型，我们从里面选择一个次优解。观测值得概率分布显然不能用于来估计我们的状态，所以状态转移概率分布可以作为我们的 x_t 采样的参考分布。

令：

$$q(x_t) = p(x_t | x_{t-1}) \tag{16}$$

将(16)式代入(11)，最后粒子滤波的权重表达式可以简单表示为：

$$w(x_t^i) = w(x_{t-1}^i) \cdot p(y_t^i | x_t^i) \tag{17}$$

这样的滤波也叫作凝聚滤波(Condensation Filter)。

3 重采样(Resampling)

原理上，粒子滤波不存在重采样这样的操作，但在实际应用中，往往会存在一个问题--粒子退化：随着滤波的进行，少量粒子占据大部分权值，使得实际有效的粒子数目变得很少。为了应对这个问题，所以有了重采样的操作。

一般的，按照下列公式计算有效粒子数目：

$$N_{eff} = \frac{1}{\sum_{i=1}^N (\hat{w}^i)^2} \quad (18)$$

对于给定的判断重采样的阈值 N_{thres} ，当 $N_{eff} < N_{thres}$ 时，进行重采样。

当 $N_{eff} < N_{thres}$ 时，对N个粒子进行重新采样，行如下操作：

```
w_new=w; %新权重初始化
Neff=1/sum(w.*w); %计算有效粒子数目
N=length(w);
Nthres = 75 %粒子数目阈值
if Neff < Nthres
    for i = 1 : N
        r = rand(0,1); %随机产生[0,1]的概率
        sump = 0; %累积概率初始化
        for j = 1 : N
            sump = sump + w(j);
            if sump >= r
                w_new(i)=w(j);
                break;
            end
        end
    end
end
end
```

重采样部分参考<https://blog.csdn.net/u011624019/article/details/80559397> 的内容。但重采样的方法比较多，这里没有仔细研究。

总结

这个文档总结主要是参考了B站up主徐亦达教授的视频《粒子滤波》。之前在网上也找了一些博客来看，不是很直观，最后还是徐老师的视频讲的比较明白，强烈建议去多看几遍。