NA 568 - Winter 2026

Particle Filtering

Maani Ghaffari

January 16, 2026

UNIVERSITY OF MICHIGAN

Nonlinear Dynamic Systems Excited by Noise

- Nonlinear process model:

$$x_k = f(u_k, x_{k-1}, w_k)$$

- Nonlinear measurement model:

$$z_k = h(x_k, v_k)$$


Q. How to map belief (a probability distribution) through a nonlinear function?

Key ideas:

- Linearization via Taylor expansion
  $\longrightarrow$ Extended Kalman Filter (EKF)
- Unscented Transform (deterministic sampling)
  $\longrightarrow$ Unscented Kalman Filter (UKF)
- **Monte Carlo methods (random sampling)**
  $\longrightarrow$ **Sequential Monte Carlo methods (Particle Filters)**

Sequential Monte Carlo methods

Sequential Monte Carlo (SMC) methods are a set of simulation-based methods for computing posterior distributions.

- Observations arrive sequentially in time and we wish to perform online inference;
- The posterior distribution is updated as data become available (recursive Bayesian estimation/learning);
- SMC methods are used when dealing with non-Gaussian, high-dimensionality, and nonlinearity where often obtaining an analytical solution is not possible.

Remark

SMC methods can be used for inferring both filtering and smoothing posterior distributions.

我们有n个独立同分布随机样本，每个样本的长度为k+1

Suppose we can simulate n independent and identically distributed (i.i.d.) random samples (particles), $\{x_{0:k}^i\}_{i=1}^n$ according to $p(x_{0:k}|z_{1:k})$. An empirical estimate of this distribution is given by

联合分布为上述p

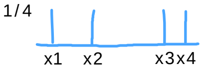
$$p_n(x_{0:k} - x_{0:k}^{1:n}|z_{1:k}) = \frac{1}{n}\sum_{i=1}^{n}\delta(x_{0:k} - x_{0:k}^i)$$

delta函数表示，在每一个值为样本xi的点，其概率分布为1/n（整个概率密度积分之后仍旧为1，虽然密度为无限大，但是测度也为0）。全部求和得到一个完整的分布于各个样本的离散均匀分布。

Then, the following integral can be computed:

$$\text{E(f)} = \mathrm{I}_n(f) = \int f(x_{0:k})p_n(x_{0:k} - x_{0:k}^{1:n}|z_{1:k}) = \frac{1}{n}\sum_{i=1}^{n} f(x_{0:k}^i)$$

经过积分，我们会得到，正好在每一个xi位置，
pn=1/n * delta，对delta积分之后对应得到所有位置的f(xi)求和。



假设是离散情况（好理解），实际上是连续情况

Notation: Dirac delta

The Dirac delta function, for $x \in \mathbb{R}$, is defined by the properties

$$\delta(x) = \begin{cases} \infty & \text{if } x = 0 \\ 0 & \text{if } x \neq 0 \end{cases} \quad \text{and} \quad \int_{-\infty}^{\infty} \delta(x) \mathrm{d}x = 1$$

For any smooth function f and $a \in \mathbb{R}$, we have:

$$\int_{-\infty}^{\infty} \delta(x - a) f(x) = f(a)$$

which is a Lebesgue integral with respect to the measure δ (thought as a point mass). This can be generalized to $\mathbb{R}^n$ or any set with the similar idea to define δ measure or unit mass concentrated at a point.

Perfect Monte Carlo Sampling

Suppose we could draw i.i.d. samples $x_{0:k}^{1:n} \sim p(x_{0:k} \mid z_{1:k})$. Then we can approximate expectations by a sample average:

$$I(f) \;=\; \mathbb{E}[f(x_{0:k})] \;\approx\; \hat{I}_n(f) \;=\; \frac{1}{n}\sum_{i=1}^{n} f\left(x_{0:k}^{i}\right).$$

- *Unbiased and concentrates:* $\hat{I}_n(f) \to I(f)$ as $n \to \infty$ (law of large numbers).
- *Error rate:* standard deviation scales like

  $$\operatorname{Std}\left[\hat{I}_n(f)\right] \;=\; O\left(\frac{1}{\sqrt{n}}\right),$$

  and this rate is *independent of the dimension* of $x_{0:k}$.
- *But:* we usually *cannot* sample efficiently from $p(x_{0:k} \mid z_{1:k})$.

$\Rightarrow$ Next: **importance sampling** + **sequential factorization** to make it *recursive*.

设 $Y_i = f(x^i)$。因为 i.i.d.：

- $\mathbb{E}[Y_i] = I(f)$
- $\mathrm{Var}(Y_i) = \sigma_f^2$（只要方差有限）

那么

$$\mathrm{Var}\left(\hat{I}_n(f)\right) = \mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^{n} Y_i\right) = \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{Var}(Y_i) = \frac{\sigma_f^2}{n}.$$

所以

$$\mathrm{Std}[\hat{I}_n(f)] = \frac{\sigma_f}{\sqrt{n}} = O\left(\frac{1}{\sqrt{n}}\right).$$

直观：误差随样本量按平方根下降，想把误差减半需要 4 倍样本。

重要性采样

- We introduce an *importance sampling distribution* (also called *proposal distribution*), $\pi(x_{0:k}|z_{1:k})$;

- we also assume the support of $\pi(x_{0:k}|z_{1:k})$ includes the support of $p(x_{0:k}|z_{1:k})$; we get:

$$
\begin{aligned}
\mathrm{I}(f) &= \frac{\int f(x_{0:k}) w(x_{0:k}) \pi(x_{0:k}|z_{1:k}) \mathrm{d}x_{0:k}}{\int w(x_{0:k}) \pi(x_{0:k}|z_{1:k}) \mathrm{d}x_{0:k}} \\
&= \frac{\int f(x_{0:k}) p(x_{0:k}|z_{1:k}) \mathrm{d}x_{0:k}}{\int p(x_{0:k}|z_{1:k}) \mathrm{d}x_{0:k}} = \int f(x_{0:k}) p(x_{0:k}|z_{1:k}) \mathrm{d}x_{0:k}
\end{aligned}
$$

where $w(x_{0:k})$ is known importance weight:

$$
w(x_{0:k}) = \frac{p(x_{0:k}|z_{1:k})}{\pi(x_{0:k}|z_{1:k})}
$$

假设 "$\pi(x_{0:k} \mid z_{1:k})$ 的支持集包含 $p(x_{0:k} \mid z_{1:k})$ 的支持集"，意思是：

支持集概念

> 只要 $p(x) > 0$ 的地方，一定也有 $\pi(x) > 0$。

这样 $w(x) = \frac{p(x)}{\pi(x)}$ 才不会在 $p(x) > 0$ 的地方出现除以 0（否则权重会无穷/不定义）。

2) 关键恒等式：把 p 写成 $w\pi$

定义

$$w(x_{0:k}) = \frac{p(x_{0:k} \mid z_{1:k})}{\pi(x_{0:k} \mid z_{1:k})}.$$

那么在 $\pi > 0$ 的地方就有

$$w(x_{0:k})\,\pi(x_{0:k} \mid z_{1:k}) = p(x_{0:k} \mid z_{1:k}).$$

3) 为什么要写成"分子/分母"的形式？

你真正想要的是

$$I(f) = \int f(x)\,p(x)\,dx$$

（我这里把 $x_{0:k}$ 简写成 x）。

如果 p 已经是**归一化的密度**，那 $\int p(x)dx = 1$，所以也可以写成

$$I(f) = \frac{\int f(x)p(x)\,dx}{\int p(x)\,dx}.$$

接下来用上面那条 $p = w\pi$ 代入：

$$I(f) = \frac{\int f(x)\,w(x)\,\pi(x)\,dx}{\int w(x)\,\pi(x)\,dx}.$$

一句话总结

重要性采样就是把目标分布 p 的积分，改写成在更容易采样的提议分布 π 下的加权积分；

Drawing n i.i.d. particles according to $\pi(x_{0:k}|z_{1:k})$:

这里粒子以pi为概率生成。然后用标准化的权重（什么值有更高的概率），来重新分配概率从而计算期望。

$$\hat{\mathrm{I}}_n(f) = \frac{\frac{1}{n}\sum_{i=1}^{n} f(x_{0:k}^i)w(x_{0:k}^i)}{\frac{1}{n}\sum_{i=1}^{n} w(x_{0:k}^i)} = \sum_{i=1}^{n} f(x_{0:k}^i)\tilde{w}_k^i$$

where the *normalized importance weights* are given by

$$\tilde{w}_k^i = \frac{w(x_{0:k}^i)}{\sum_{i=1}^{n} w(x_{0:k}^i)}$$

权重/权重和=归一化权重

but this is not adequate for recursive estimation!

这还不足以循环估计。

我们通常令“过去的轨迹部分”只依赖过去观测：

$$\pi(x_{0:k-1} \mid z_{1:k}) \approx \pi(x_{0:k-1} \mid z_{1:k-1})$$

（也可理解为：我们在时间 k 更新时，不回头再改 $x_{0:k-1}$，只在末尾扩展一个 x_k。）

$$p(A, B \mid C) = p(A \mid C)\, p(B \mid A, C)$$

Let us factor the importance sampling distribution as follows: 见上一张ppt

这个东西看着很反直觉，事实也是，概率公式应该写为

$$\pi(x_{0:k}|z_{1:k}) = \pi(x_{0:k-1}|z_{1:k-1})\pi(x_k|x_{0:k-1}, z_{1:k})$$

$\pi(x_{0:k-1}|z_{1:k})$

$$= \pi(x_0)\prod_{j=1}^{k}\pi(x_j|x_{0:j-1}, z_{1:j})$$

自回归模型，联合概率为一步步过渡概率相乘。

Note: $p(x_2,x_1 \mid z) = p(x_1 \mid z)p(x_2 \mid x_1,z)$

then:

$$\tilde{w}_k^i \propto \tilde{w}_{k-1}^i \frac{p(z_k|x_k^i)p(x_k^i|x_{k-1}^i)}{\pi(x_k^i|x_{0:k-1}^i, z_{1:k})} \quad \left(w(x_{0:k}) = \frac{p(x_{0:k}|z_{1:k})}{\pi(x_{0:k}|z_{1:k})}\right)$$

Recall: $p(x_{0:k}^i|z_{1:k}) = p(x_{0:k-1}^i|z_{1:k-1})\dfrac{p(z_k|x_k^i)p(x_k^i|x_{k-1}^i)}{p(z_k|z_{1:k-1})}$

为什么wk递推公式成立？

0. 先把模型假设写清楚（这是推导能成立的根）

我们在标准状态空间模型里，假设：

1. **初始分布**

$$p(x_0)$$

2. **一阶马尔可夫转移**

$$p(x_k \mid x_{0:k-1}) = p(x_k \mid x_{k-1})$$

3. **观测条件独立（给定当前状态）**

$$p(z_k \mid x_{0:k}, z_{1:k-1}) = p(z_k \mid x_k)$$

1. 证明 “Recall” 那一行（后验递推）完全等号成立

目标是证明：

$$p(x_{0:k} \mid z_{1:k}) = p(x_{0:k-1} \mid z_{1:k-1}) \frac{p(z_k \mid x_k)\, p(x_k \mid x_{k-1})}{p(z_k \mid z_{1:k-1})}.$$

Step 1：从 Bayes 定理（对新增观测 z_k 做一次更新）

把 $z_{1:k}$ 写成 $(z_{1:k-1}, z_k)$。Bayes 给出：

z1:k-1为共有条件

$$p(x_{0:k} \mid z_{1:k}) = \frac{p(z_k \mid x_{0:k}, z_{1:k-1})\, p(x_{0:k} \mid z_{1:k-1})}{p(z_k \mid z_{1:k-1})}.$$

这一步只是 Bayes 定义，没有任何近似。

$p(x_{0:k} \mid z_k, z_{1:k-1})$

Step 2：用观测条件独立，把 $p(z_k \mid x_{0:k}, z_{1:k-1})$ 简化

由假设 3：

$$p(z_k \mid x_{0:k}, z_{1:k-1}) = p(z_k \mid x_k).$$

代回 Step 1：

$$p(x_{0:k} \mid z_{1:k}) = \frac{p(z_k \mid x_k)\, p(x_{0:k} \mid z_{1:k-1})}{p(z_k \mid z_{1:k-1})}.$$

Step 3：把 $p(x_{0:k} \mid z_{1:k-1})$ 用链式法则展开

链式法则（概率乘法公式）永远成立：

$$p(x_{0:k} \mid z_{1:k-1}) = p(x_{0:k-1} \mid z_{1:k-1})\, p(x_k \mid x_{0:k-1}, z_{1:k-1}).$$

代回上式：

$$p(x_{0:k} \mid z_{1:k}) = \frac{p(z_k \mid x_k)\, p(x_{0:k-1} \mid z_{1:k-1})\, p(x_k \mid x_{0:k-1}, z_{1:k-1})}{p(z_k \mid z_{1:k-1})}.$$

Step 4：用一阶马尔可夫，把 $p(x_k \mid x_{0:k-1}, z_{1:k-1})$ 简化

由假设 2（状态马尔可夫性，且观测不影响动力学转移）：

$$p(x_k \mid x_{0:k-1}, z_{1:k-1}) = p(x_k \mid x_{k-1}).$$

代回：

$$p(x_{0:k} \mid z_{1:k}) = p(x_{0:k-1} \mid z_{1:k-1}) \frac{p(z_k \mid x_k)\, p(x_k \mid x_{k-1})}{p(z_k \mid z_{1:k-1})}.$$

继续证明权重
递推公式成立

Step 1：把 proposal π 按照图里的方式分解（链式法则）

图中给：

$$\pi(x_{0:k} \mid z_{1:k}) = \pi(x_{0:k-1} \mid z_{1:k-1})\, \pi(x_k \mid x_{0:k-1}, z_{1:k}).$$

对粒子 i：

$$\pi(x^i_{0:k} \mid z_{1:k}) = \pi(x^i_{0:k-1} \mid z_{1:k-1})\, \pi(x^i_k \mid x^i_{0:k-1}, z_{1:k}).$$

Step 2：把目标后验 $p(x_{0:k} \mid z_{1:k})$ 用上一节的 “Recall” 等式替换

上一节已经证明：

$$p(x_{0:k} \mid z_{1:k}) = p(x_{0:k-1} \mid z_{1:k-1}) \frac{p(z_k \mid x_k)\, p(x_k \mid x_{k-1})}{p(z_k \mid z_{1:k-1})}.$$

对粒子 i：

$$p(x^i_{0:k} \mid z_{1:k}) = p(x^i_{0:k-1} \mid z_{1:k-1}) \frac{p(z_k \mid x^i_k)\, p(x^i_k \mid x^i_{k-1})}{p(z_k \mid z_{1:k-1})}.$$

Step 3：把 Step 1 和 Step 2 同时代回权重定义 $\tilde{w}_k^i = \frac{p}{\pi}$

$$\tilde{w}_k^i = \frac{p(x_{0:k}^i \mid z_{1:k})}{\pi(x_{0:k}^i \mid z_{1:k})}.$$

代入分子（Step 2）与分母（Step 1）：

为什么二者不同？

分子为p递推

对pi有假设pi(x0:k-1|z1:k)=pi(x0:k-1|z1:k-1),从而直接简化表达式。但是p将zk看做一个正式的条件变量。

$$\tilde{w}_k^i = \frac{p(x_{0:k-1}^i \mid z_{1:k-1}) \frac{p(z_k|x_k^i)\, p(x_k^i|x_{k-1}^i)}{p(z_k|z_{1:k-1})}}{\pi(x_{0:k-1}^i \mid z_{1:k-1})\, \pi(x_k^i \mid x_{0:k-1}^i, z_{1:k})}.$$

分母为pi递推

根本原因是：p 是目标后验（由 Bayes 强制），π 是提议分布（由算法设计），为了顺序生成，常做"不回头"的结构化选择。

把乘法分开整理：

$$\tilde{w}_k^i = \underbrace{\frac{p(x_{0:k-1}^i \mid z_{1:k-1})}{\pi(x_{0:k-1}^i \mid z_{1:k-1})}}_{\tilde{w}_{k-1}^i} \cdot \frac{p(z_k \mid x_k^i)\, p(x_k^i \mid x_{k-1}^i)}{\pi(x_k^i \mid x_{0:k-1}^i, z_{1:k})} \cdot \frac{1}{p(z_k \mid z_{1:k-1})}.$$

所以得到**严格等号形式**：

$$\boxed{\tilde{w}_k^i = \tilde{w}_{k-1}^i \cdot \frac{p(z_k \mid x_k^i)\, p(x_k^i \mid x_{k-1}^i)}{\pi(x_k^i \mid x_{0:k-1}^i, z_{1:k})} \cdot \frac{1}{p(z_k \mid z_{1:k-1})}}.$$

于是wk的递推式形式得证！

Remark

In filtering, we often set $\pi(x_k|x_{0:k-1}, z_{1:k}) = \pi(x_k|x_{k-1}, z_k)$ *so that the importance sampling distribution only depends on* x_{k-1} *and* z_k.

上述证明中没有用到pi的马尔科夫性，但是其实可以满足这一点。

Remark

A **special case** *is when the importance sampling distribution is chosen to be the prior distribution, or in robotics the robot motion model* $p(x_k|x_{k-1}, u_k)$*. Note that* u_k *is deterministic and the motion model does not depend on the measurement* z_k*.*

Then the weights can be computed using:

这里的uk在状态x概率中
是默认存在的。

$$\tilde{w}_k^i \propto \tilde{w}_{k-1}^i p(z_k|x_k^i) \qquad \tilde{w}_k^i \propto \tilde{w}_{k-1}^i \frac{p(z_k|x_k^i)p(x_k^i|x_{k-1}^i)}{\pi(x_k^i|x_{0:k-1}^i, z_{1:k})}$$

课件说“importance sampling distribution chosen to be the prior distribution”，也就是选：

$$\pi(x_k \mid x_{0:k-1}, z_{1:k}) = p(x_k \mid x_{k-1}, u_k).$$

这样分子和分母就可以消掉。

1) 这条式子在说什么

$$\pi(x_k \mid x_{0:k-1}, z_{1:k}) = p(x_k \mid x_{k-1}, u_k)$$

意思是：**proposal 选成先验/运动模型**。也就是你在时刻 k 生成新粒子时：

- 只按运动模型从 x_{k-1}（和控制 u_k）预测出 x_k
- 不用 z_k 来"引导采样"

所以"观测没被用在采样里"这句话是对的。

---

2) 观测在哪里被考虑？——在权重更新里

权重递推的一般式：

$$\tilde{w}_k^i \propto \tilde{w}_{k-1}^i \frac{p(z_k \mid x_k^i)\, p(x_k^i \mid x_{k-1}^i, u_k)}{\pi(x_k^i \mid x_{0:k-1}^i, z_{1:k})}.$$

现在代入 prior proposal（分母 = 转移项）：

$$\tilde{w}_k^i \propto \tilde{w}_{k-1}^i\, p(z_k \mid x_k^i).$$

看到了吗？**观测** z_k 以 **似然** $p(z_k \mid x_k)$ 的形式完整地进来了，而且它正是决定"哪个粒子更可信"的关键。

我的直观理解就是：那些更容易得到当前观测的状态样本，会有更高的权重。

不好的趋势：慢慢wk会变得越来越狭小，只有一个例子有非零的权重

- As time increases, the distribution of the of the weights, $\tilde{w}_k^i$ becomes more and more skewed, in practice, reducing to one particle with non-zero weight after a few iterations;

- the *resampling* idea was introduced to fix this problem.
  使用重采样来解决这个问题

这页在讲：**用"有效样本数" n_{eff} 来衡量粒子权重退化**（degeneracy）——也就是"虽然你有 n 个粒子，但真正有用的可能远少于 n 个"。

- Measure of degeneracy using the effective sample size:

$$n_{\text{eff}} = \frac{1}{\sum_{i=1}^{n} (\tilde{w}_k^i)^2} \qquad 1 < n_{\text{eff}} < n$$

如果只有一个粒子权重为1，那么neff为1，如果全部的例子权重都为1/n，那么neff为n。

- two extreme cases:
  1. all particles have the same weights (uniform): $\forall i \in \{1:n\}$, $\tilde{w}_k^i = \frac{1}{n} \Longrightarrow n_{\text{eff}} = n$;
  2. the entire distribution mass is placed in one particle (singular): $\forall i \in \{1:j-1, j+1:n\}$, $\tilde{w}_k^i = 0$ and $\tilde{w}_k^j = 1 \Longrightarrow n_{\text{eff}} = 1$;

1) 退化是什么?

粒子滤波里你用标准化权重 $\tilde{w}_k^i$（满足 $\sum_i \tilde{w}_k^i = 1$）来近似后验：

$$\hat{p}(dx) \approx \sum_{i=1}^{n} \tilde{w}_k^i \, \delta_{x_k^i}(dx).$$

如果大多数权重都接近 0，只有少数粒子权重大，就叫**权重退化**：很多粒子几乎没贡献，白算了。

Sample Impoverishment

Although resampling step reduces the effect of degeneracy, it introduces a new problem known as *sample impoverishment*.

- It limits the parallel implementation of the algorithm since all particles must be combined;
- the particles with high weights are selected many times, leading to the loss of diversity.

Sample impoverishment（样本贫化）是粒子滤波在做**重采样（resampling）**之后常见的一个副作用：

> 重采样会把**高权重粒子复制很多份**、把低权重粒子丢掉，导致粒子集合里出现大量**重复粒子**，多样性下降；随后再预测时，粒子可能覆盖不了真实后验的范围，尤其在**过程噪声很小**或**维度高**时更严重。

采样次数越多，不可避免的就是那些权重高的粒子会以更多的次数被抽到。

At each time step k we maintain $\{x_k^i, w_k^i\}_{i=1}^n$.

1. **Sample** (propagate): $x_k^i \sim p(x_k \mid x_{k-1}^i, u_k)$ 先用运动模型采样
2. **Weight** (measurement update): $w_k^i \propto w_{k-1}^i \, p(z_k \mid x_k^i)$ 计算权重
3. **Normalize**: $\tilde{w}_k^i = \dfrac{w_k^i}{\sum_{j=1}^n w_k^j}$ 标准化权重
4. **Resample if needed**: when

$$n_{\text{eff}} = \frac{1}{\sum_{i=1}^n (\tilde{w}_k^i)^2} < n_t \ \text{(e.g., } n/3\text{)}$$

如果有效样本数某个临界，那么重新采样。

PF = *predict with sampling* + *update by likelihood* + *resample to fight degeneracy*.

A Basic Particle Filter Algorithm in Robotics

Algorithm 1 `particle-filter`

Require: particles $\mathcal{X}_{k-1} = \{x_{k-1}^i, \tilde{w}_{k-1}^i\}_{i=1}^n$, action u_k, measurement z_k, resampling threshold n_t (e.g. $n/3$);

$\mathcal{X}_k \leftarrow \varnothing$

for each $x_{k-1}^i \in \mathcal{X}_{k-1}$ **do** 每一个k-1时刻的样本都可以采样得到一个k时刻样本

  draw $x_k^i \sim p(x_k | x_{k-1}^i, u_k)$ ▷ sample from motion model

  $w_k^i \leftarrow \tilde{w}_{k-1}^i p(z_k | x_k^i)$ 更新权重 ▷ update importance weights

归一化

$w_\text{total} \leftarrow \sum_{i=1}^n w_k^i$ ▷ compute total weight to normalize importance weights

$\mathcal{X}_k \leftarrow \mathcal{X}_k \cup \{x_k^i, w_k^i / w_\text{total}\}_{i=1}^n$ ▷ add weighted samples to the new set

$n_\text{eff} \leftarrow 1 / \sum_{i=1}^n (\tilde{w}_k^i)^2$ 计算是否多样性满足要求 ▷ compute effective sample size

if $n_\text{eff} < n_\text{t}$ **then**

  $\mathcal{X}_k \leftarrow$ resample using $\mathcal{X}_k$ ▷ use a resampling algorithm to draw particles with higher weights 如果不满足要求则重采样

return $\mathcal{X}_k$

粒子滤波的目标就是得到每一时刻的状态和状态值样本对应的权重。

1. **Prediction Step**: For each particle i: $x_t^{[i]} = f(x_{t-1}^{[i]}, u_t) + \epsilon_t$ where:

   - $f(\cdot)$ is the motion model
   - u_t is the control input
   - ϵ_t is motion noise

2. **Update Step**: For each particle i: $w_t^{[i]} = p(z_t | x_t^{[i]}) w_{t-1}^{[i]}$ where:

   - $p(z_t | x_t^{[i]})$ is measurement likelihood
   - z_t is the sensor measurement

这更加简单粗暴地使用了贝叶斯滤波对于p(zt|xt)的融入，谁更容易在当前观测的情况下出现，那么谁权重就更高。

3. **Normalization**: $w_t^{[i]} = \frac{w_t^{[i]}}{\sum_{j=1}^{N} w_t^{[j]}}$

4. **Resampling**:

   - Draw N new particles with probability proportional to weights
   - Set all weights to $w_t^{[i]} = \frac{1}{N}$

Belief Representation:

$bel(x_t) \approx \sum_{i=1}^{N} w_t^{[i]} \delta(x_t - x_t^{[i]})$ where $\delta(\cdot)$ is the Dirac delta function

仅仅在例子滤波采集过的样本点有概率密度，可以理解为权重即在这个点的质量

- Resampling eliminates particles with low weights and multiplies particles with high weights;
- the particles with high weights are selected many times, leading to the loss of diversity (i.e., loss of alternative hypotheses).

Algorithm 2 `low-variance-resampling`

Require: particles $\mathcal{X}_k = \{x_k^i, \tilde{w}_k^i\}_{i=1}^n$; 计算wi这个离散概率分布的累积分布

$w_c \leftarrow$ compute the vector of cumulative sum of the weights using $\{\tilde{w}_k^i\}_{i=1}^n$

均匀采样r ▷ w_c is the Cumulative Distribution Function (CDF)

$r \leftarrow \texttt{rand}(0, n^{-1})$ ▷ draw a uniform random number between 0 and n^{-1}

$j \leftarrow 1$ ▷ dummy index to climb the CDF and select particles

for all $i \in \{1:n\}$ **do**

  $u \leftarrow r + (i-1)n^{-1}$ 一次性增加1/n ▷ move along the CDF

  while $u > w_c^j$ **do** xi必须取到一个累积分布值达到u的i，不然i继续递增

    $j \leftarrow j + 1$ 那些被跳过的j，不会被采用，因为他们的权重太低了，很可能小于1/n，不需要设置为1/n来减小。

  $x_k^i \leftarrow x_k^j$ ▷ replicate the survived particle

  $\tilde{w}_k^i \leftarrow n^{-1}$ ▷ set the weight to n^{-1} (uniform distribution)

权重设置自动设置为1/n

return $\mathcal{X}_k$

A target is moving in a 2D plane. The ownship position is known and fixed at the origin. We have access to relative noisy range and bearing measurements of the target position at any time step.

$$x_k = f(u_k, x_{k-1}) + w_k = x_{k-1} + w_k$$

$$z_k = h(x_k) + v_k = \left[\begin{array}{c} \sqrt{x_k[[1]^2] + \mathbf{x}_k^{[2]^2}} \\ \mathrm{atan2}(\mathbf{x}_k^{[1]}, \mathbf{x}_k^{[2]}) \end{array} \right] + v_k$$

$$Q_k = 0.1\ I_2,\ R_k = \mathrm{diag}(0.05^2, 0.01^2)$$

Example: PF Target Tracking, Constant Velocity Motion Model

There is no knowledge of the target motion, but this time, we assume a constant velocity random walk motion model and estimate the target velocity along with the position.

$$x_k = f(u_k, x_{k-1}) + w_k = F_k x_{k-1} + w_k$$

$$F_k = \begin{bmatrix} I & \Delta t I \\ 0 & I \end{bmatrix} \qquad \Delta t \text{ : sampling time}$$

$$z_k = h(x_k) + v_k = \begin{bmatrix} \sqrt{{x_k^1}^2 + {x_k^2}^2} \\ \mathrm{atan2}(x_k^1, x_k^2) \end{bmatrix} + v_k$$

$$Q_k = \mathrm{diag}(0.1^2, 0.1^2, 0.01^2, 0.01^2),\ R_k = \mathrm{diag}(0.05^2, 0.01^2)$$

状态（4维）

$$x_k = \begin{bmatrix} p_x \\ p_y \\ v_x \\ v_y \end{bmatrix}$$

运动模型（常速度 + 随机游走）

$$x_k = F x_{k-1} + w_k, \quad w_k \sim \mathcal{N}(0, Q)$$

$$F = \begin{bmatrix} I & \Delta t\, I \\ 0 & I \end{bmatrix} = \begin{bmatrix} 1 & 0 & \Delta t & 0 \\ 0 & 1 & 0 & \Delta t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

观测模型（range-bearing）

$$z_k = h(x_k) + v_k, \quad v_k \sim \mathcal{N}(0, R)$$

$$h(x) = \begin{bmatrix} r \\ \theta \end{bmatrix} = \begin{bmatrix} \sqrt{p_x^2 + p_y^2} \\ \mathrm{atan2}(p_y, p_x) \end{bmatrix}$$

给定图里的

$$Q = \mathrm{diag}(0.1^2, 0.1^2, 0.01^2, 0.01^2), \quad R = \mathrm{diag}(0.05^2, 0.01^2).$$

3) draw $x_k \sim p(x_k \mid x_{k-1}^i, u_k)$

在本题里就是线性高斯运动模型采样：

$$x_k^i = F x_{k-1}^i + \epsilon_k^i, \quad \epsilon_k^i \sim \mathcal{N}(0, Q)$$

（这里就是你课件 remark 说的 prior/motion proposal。）

4) $w_k^i \leftarrow \tilde{w}_{k-1}^i \, p(z_k \mid x_k^i)$ 这一步最重要，贯彻对于粒子滤波的理解

本题的似然是观测噪声高斯：

$$p(z_k \mid x_k^i) = \mathcal{N}\big(z_k;\ h(x_k^i),\ R\big)$$

更"可计算"的写法是：

- 预测观测：

$$\hat{z}_k^i = h(x_k^i) = \begin{bmatrix} \hat{r}_k^i \\ \hat{\theta}_k^i \end{bmatrix} = \begin{bmatrix} \sqrt{(p_x^i)^2 + (p_y^i)^2} \\ \operatorname{atan2}(p_y^i, p_x^i) \end{bmatrix}$$

- 残差（注意角度要 wrap）：

$$e_k^i = \begin{bmatrix} r_k - \hat{r}_k^i \\ \mathrm{wrap}(\theta_k - \hat{\theta}_k^i) \end{bmatrix}$$

- 似然：

$$p(z_k \mid x_k^i) = \frac{1}{2\pi\sqrt{|R|}} \exp\left(-\frac{1}{2} (e_k^i)^\top R^{-1} e_k^i \right)$$

- 于是：

$$w_k^i = \tilde{w}_{k-1}^i \cdot p(z_k \mid x_k^i)$$

5) $w_{\text{total}} \leftarrow \sum_{i=1}^n w_k^i$

就是把上面所有未归一化权重加起来。

---

6) $\mathcal{X}_k \leftarrow \mathcal{X}_k \cup \{(x_k^i, w_k^i / w_{\text{total}})\}$

得到归一化权重：

$$\tilde{w}_k^i = \frac{w_k^i}{\sum_j w_k^j}$$

并把 $(x_k^i, \tilde{w}_k^i)$ 加入 $\mathcal{X}_k$。

8–9) 如果退化就 resample

若

$$n_{\text{eff}} < n_t$$

则执行重采样：

9) $\mathcal{X}_k \leftarrow$ resample using $\mathcal{X}_k$

按概率 $\tilde{w}_k^i$ 从当前粒子里抽 n 个（有放回），得到新集合：

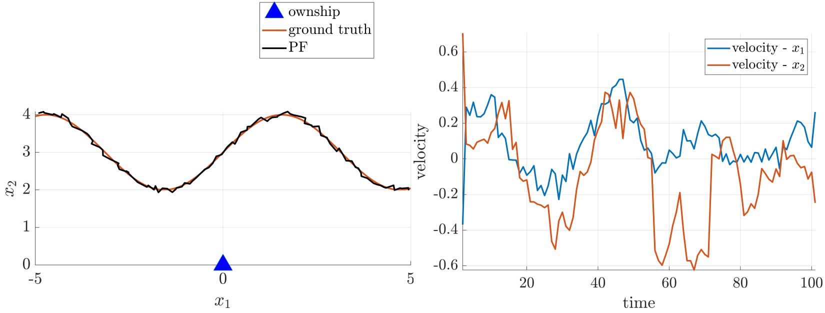
$$\{x_k^{i,\text{new}}\}_{i=1}^n$$

然后**把权重重置为均匀**（这是 resampling 的标准做法）：

$$\tilde{w}_k^{i,\text{new}} = \frac{1}{n}$$

Example: PF Target Tracking, Constant Velocity Motion Model

See `pf_single_target_cv.m` for code.



- SMC methods can solve complex nonlinear, non-Gaussian online estimation problems. For example, dealing with global uncertainty in robot localization and solving the “kidnapped robot” problem.
- The algorithms are applicable to a very large class of models and are often straightforward to implement.
- The price to pay for this simplicity is inefficiency in some application domains.

Readings

- Probabilistic Robotics: Ch. 4
- State Estimation for Robotics: Ch. 4
- Lecture Notes for Mobile Robotics: Ch. 3
- Sequential Monte Carlo Methods in Practice: Ch. 1

Next Time

- Wednesday: Hands-on Lecture on Bayes Filter, KF, and PF.