Particle Filters

a.k.a. “Where are the matrices?”

Recap from last session

- Let’s Step back from Kalman for a sec.

Measurement Model

Process Model

$$\boxed{p(\mathbf{x}_0)}$$

Initial condition

$$\boxed{p\left(\mathbf{x}_k \middle| \mathbf{z}_k^*\right)} = \kappa_k \boxed{p(\mathbf{z}_k | \mathbf{x}_k)} p\left(\mathbf{x}_k \middle| \mathbf{z}_{k-1}^*\right)$$

Update

$$\boxed{p\left(\mathbf{x}_{k+1} \middle| \mathbf{z}_k^*\right)} = \int \boxed{p(\mathbf{x}_{k+1} | \mathbf{x}_k)} p\left(\mathbf{x}_k \middle| \mathbf{z}_k^*\right) d\mathbf{x}_k$$

Prediction

Measurement Model

Process Model

$$p(\mathbf{x}_0)$$

Initial condition

$$\boxed{p\left(\mathbf{x}_k|\mathbf{z}_k^*\right)} = \kappa_k \boxed{p\left(\mathbf{z}_k|\mathbf{x}_k\right)} p\left(\mathbf{x}_k|\mathbf{z}_{k-1}^*\right)$$

Update

$$\boxed{p\left(\mathbf{x}_{k+1}|\mathbf{z}_k^*\right)} = \int \boxed{p\left(\mathbf{x}_{k+1}|\mathbf{x}_k\right)} p\left(\mathbf{x}_k|\mathbf{z}_k^*\right) d\mathbf{x}_k$$

Prediction

- Fundamental equations of a Recursive Bayesian Filter (p.d.f.s everywhere, no algebra)
- Generally no analytical form of solutions to these eq.
- “Simplifying assumptions” turn them into “useful” recursive solutions, e.g. Gaussian + linearity (on both process & measurement models) ⇒ Kalman eqs

Recursive Kalman filter

$$E\left(\mathbf{x}_k | \mathbf{z}_k^*\right) = \hat{\mathbf{x}}_k$$

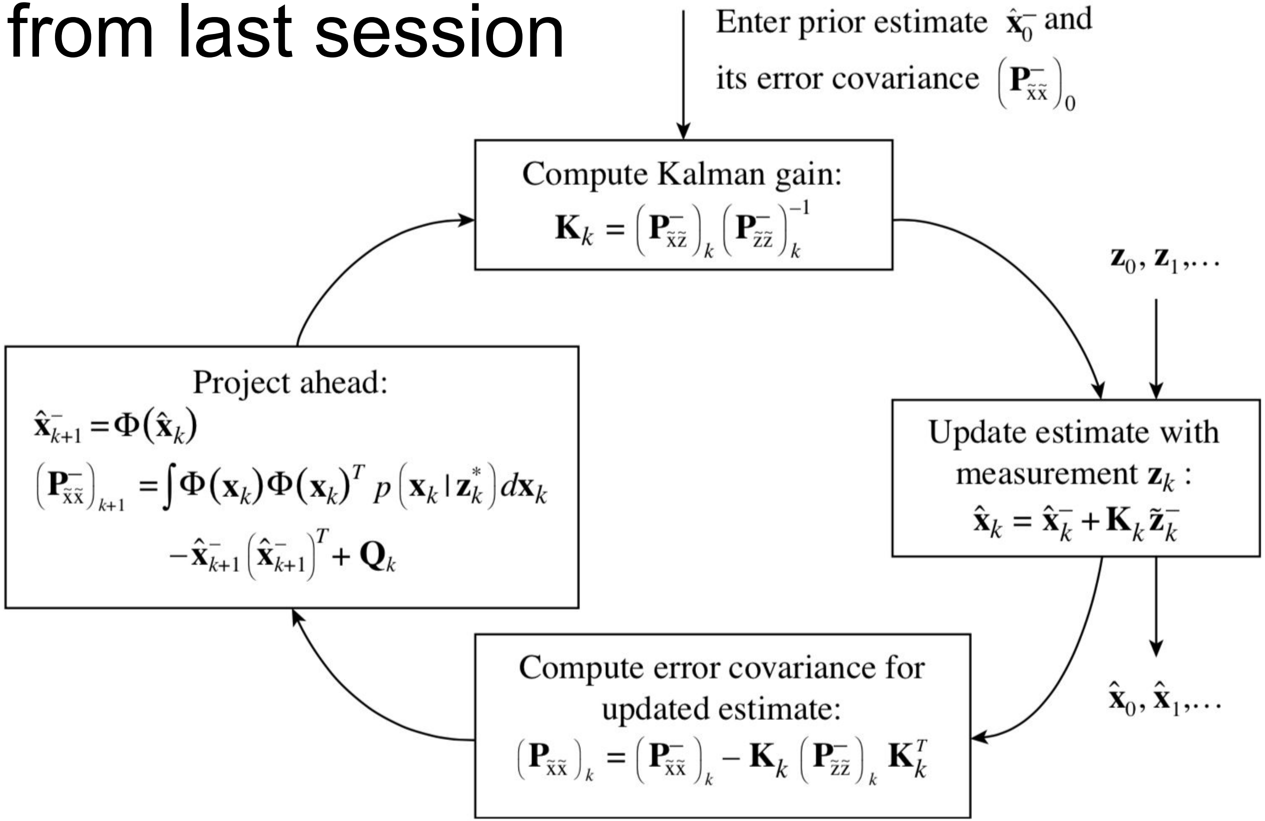
$$Cov\left(\mathbf{x}_k | \mathbf{z}_k^*\right) = \mathbf{P}_k$$

$$p\left(\mathbf{x}_k | \mathbf{z}_k^*\right) = \frac{p\left(\mathbf{z}_k, \mathbf{x}_k | \mathbf{z}_{k-1}^*\right)}{p\left(\mathbf{z}_k | \mathbf{z}_{k-1}^*\right)}$$

$$\mathcal{N}\left(\hat{\mathbf{x}}, \mathbf{P}_{\tilde{\mathbf{x}}\tilde{\mathbf{x}}}\right) = \frac{\mathcal{N}\left(\begin{bmatrix} \hat{\mathbf{x}}^- \\ \hat{\mathbf{z}}^- \end{bmatrix}, \begin{bmatrix} \mathbf{P}_{\tilde{\mathbf{x}}\tilde{\mathbf{x}}}^- & \mathbf{P}_{\tilde{\mathbf{x}}\tilde{\mathbf{z}}}^- \\ \mathbf{P}_{\tilde{\mathbf{z}}\tilde{\mathbf{x}}}^- & \mathbf{P}_{\tilde{\mathbf{z}}\tilde{\mathbf{z}}}^- \end{bmatrix}\right)}{\mathcal{N}\left(\hat{\mathbf{z}}^-, \mathbf{P}_{\tilde{\mathbf{z}}\tilde{\mathbf{z}}}^-\right)}$$

- Relates a priori states and cov estimates to a posteriori estimates
- Converts a “*p.d.f.*” problem into an *algebra* problem

Recap from last session



Recap from last session

$$\boxed{E\left(\mathbf{x}_k|\mathbf{z}_k^*\right) = \hat{\mathbf{x}}_k}$$

$$Cov\left(\mathbf{x}_k|\mathbf{z}_k^*\right) = \mathbf{P}_k$$

$$\mathrm{E}[\mathbf{g}(\mathbf{x}) \mid \mathbf{z}_k^*] = \int \mathbf{g}(\mathbf{x})\, p(\mathbf{x} \mid \mathbf{z}_k^*)\, \mathrm{d}\mathbf{x}$$

$$\mathbf{g} : \mathbb{R}^n \rightarrow \mathbb{R}^m$$

- In Bayesian Filtering, the main inference problem can often be reduced into computing this Expectation over the posterior p.d.f.

- Let's convert this "*p.d.f.*" problem into an *integral estimation* problem this time

Recap from last session

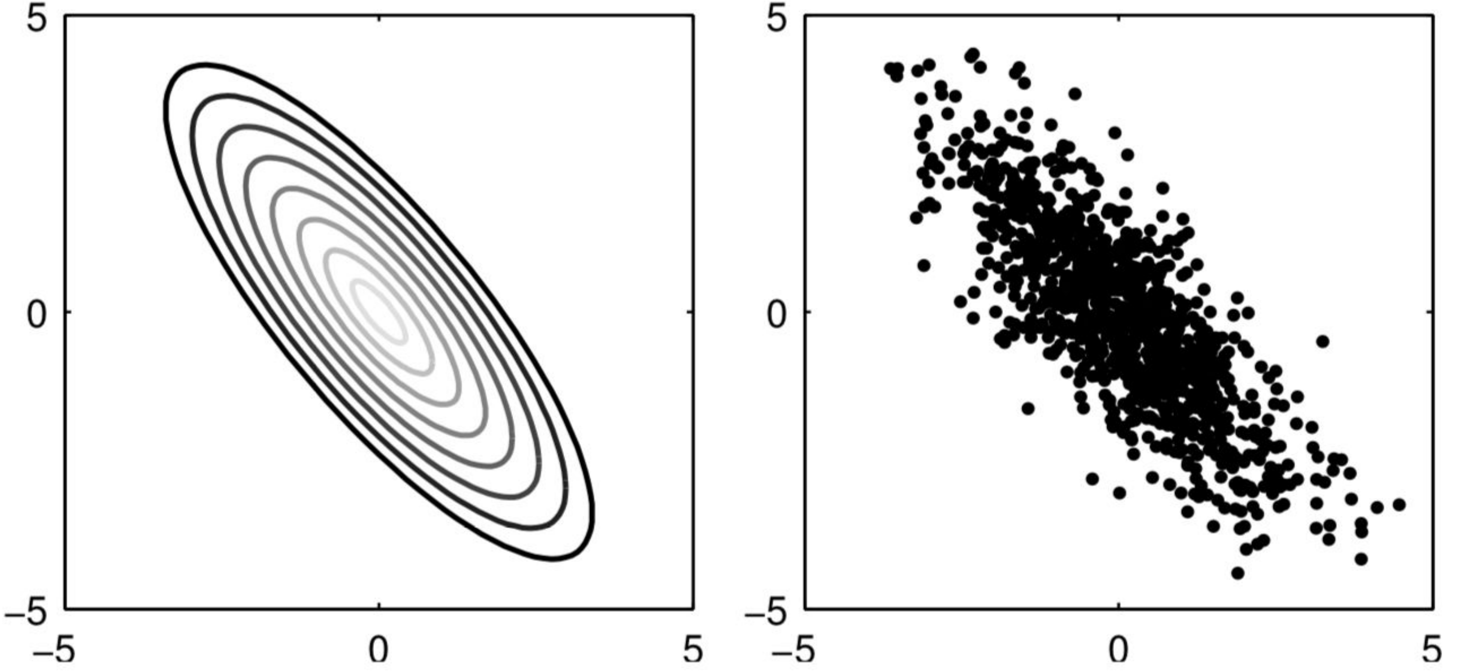
$$\boxed{E\left(\mathbf{x}_k|\mathbf{z}_k^*\right) = \hat{\mathbf{x}}_k}$$

$$Cov\left(\mathbf{x}_k|\mathbf{z}_k^*\right) = \mathbf{P}_k$$

$$\mathrm{E}[\mathbf{g}(\mathbf{x}) \mid \mathbf{z}_k^* \;] = \int \mathbf{g}(\mathbf{x})\; p(\mathbf{x} \mid \mathbf{z}_k^* \;)\; \mathrm{d}\mathbf{x}$$

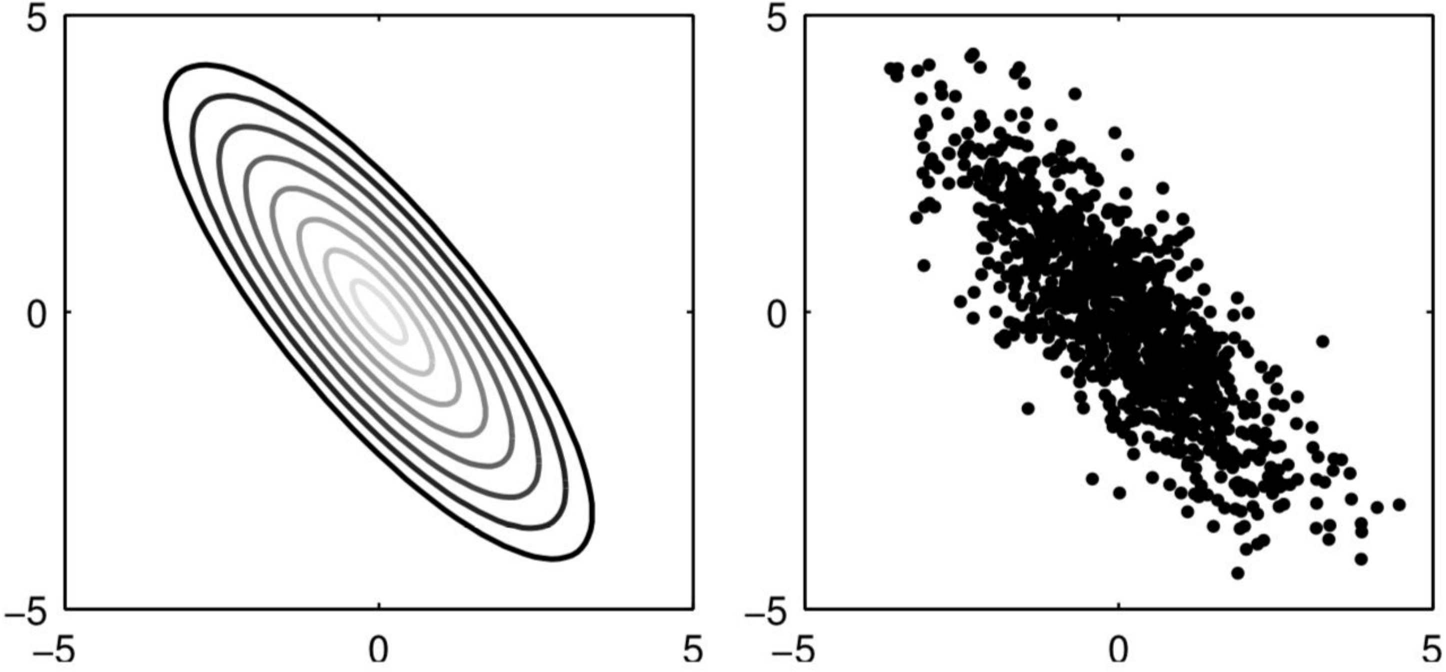
$$\mathbf{g} \;:\; \mathbb{R}^n \;\rightarrow\; \mathbb{R}^m$$

- In Bayesian Filtering, the main inference problem can often be reduced into computing this Expectation over the posterior p.d.f.
- Let's convert this "*p.d.f.*" problem into an *integral estimation* problem this time
- This is the purpose of **Particle Filters**



(a) (b)

$$\mathrm{E}[\mathbf{g}(\mathbf{x}) \mid \mathbf{z}_k^*] \approx \frac{1}{N} \sum_{i=1}^{N} \mathbf{g}(\mathbf{x}^{(i)})$$



(a) (b)

$$\mathrm{E}[\mathbf{g}(\mathbf{x}) \mid \mathbf{y}_{1:T}] \approx \frac{1}{N} \sum_{i=1}^{N} \mathbf{g}(\mathbf{x}^{(i)}).$$

Recap from last session

$$\boxed{E\left(\mathbf{x}_k|\mathbf{z}_k^*\right) = \hat{\mathbf{x}}_k}$$

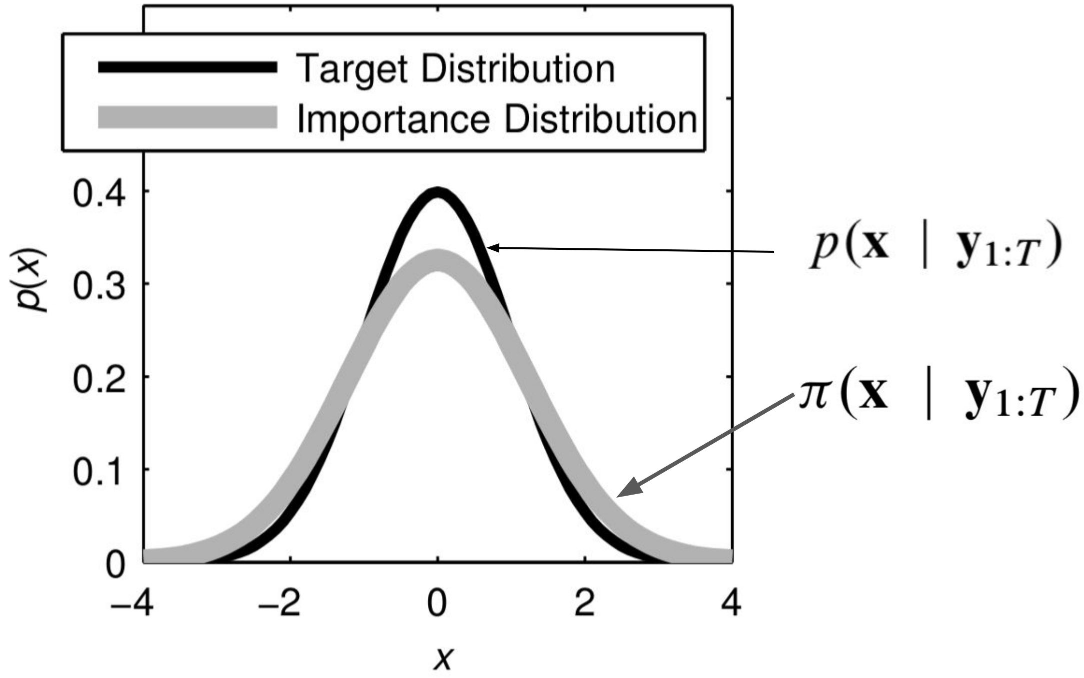
$$Cov\left(\mathbf{x}_k|\mathbf{z}_k^*\right) = \mathbf{P}_k$$

$$\mathrm{E}[\mathbf{g}(\mathbf{x}) \mid \mathbf{y}_{1:T}] = \int \mathbf{g}(\mathbf{x})\, p(\mathbf{x} \mid \mathbf{y}_{1:T})\, \mathrm{d}\mathbf{x}$$

$$\mathbf{g} : \mathbb{R}^n \to \mathbb{R}^m$$

- Practically, often impossible to obtain samples directly from $p(\mathbf{x} \mid \mathbf{y}_{1:T})$
- We can use an *approximate* distribution, $\pi(\mathbf{x} \mid \mathbf{y}_{1:T})$ from which we can easily draw samples
- This is the purpose of **Importance Sampling**

Importance Sampling



Importance Sampling

$$\int \mathbf{g}(\mathbf{x})\, p(\mathbf{x} \mid \mathbf{y}_{1:T})\, \mathrm{d}\mathbf{x} = \int \left[\mathbf{g}(\mathbf{x})\, \frac{p(\mathbf{x} \mid \mathbf{y}_{1:T})}{\pi(\mathbf{x} \mid \mathbf{y}_{1:T})} \right] \pi(\mathbf{x} \mid \mathbf{y}_{1:T})\, \mathrm{d}\mathbf{x}$$

Importance Sampling

$$\int \mathbf{g}(\mathbf{x})\, p(\mathbf{x} \mid \mathbf{y}_{1:T})\, \mathrm{d}\mathbf{x} = \int \left[\mathbf{g}(\mathbf{x}) \frac{p(\mathbf{x} \mid \mathbf{y}_{1:T})}{\pi(\mathbf{x} \mid \mathbf{y}_{1:T})} \right] \pi(\mathbf{x} \mid \mathbf{y}_{1:T})\, \mathrm{d}\mathbf{x}$$

N samples

$$\mathbf{x}^{(i)} \sim \pi(\mathbf{x} \mid \mathbf{y}_{1:T}), \qquad i = 1, \ldots, N,$$

$$\begin{aligned} \mathrm{E}[\mathbf{g}(\mathbf{x}) \mid \mathbf{y}_{1:T}] &\approx \frac{1}{N} \sum_{i=1}^{N} \frac{p(\mathbf{x}^{(i)} \mid \mathbf{y}_{1:T})}{\pi(\mathbf{x}^{(i)} \mid \mathbf{y}_{1:T})}\, \mathbf{g}(\mathbf{x}^{(i)}) \\ &= \sum_{i=1}^{N} \tilde{w}^{(i)}\, \mathbf{g}(\mathbf{x}^{(i)}), \end{aligned}$$

Importance Sampling

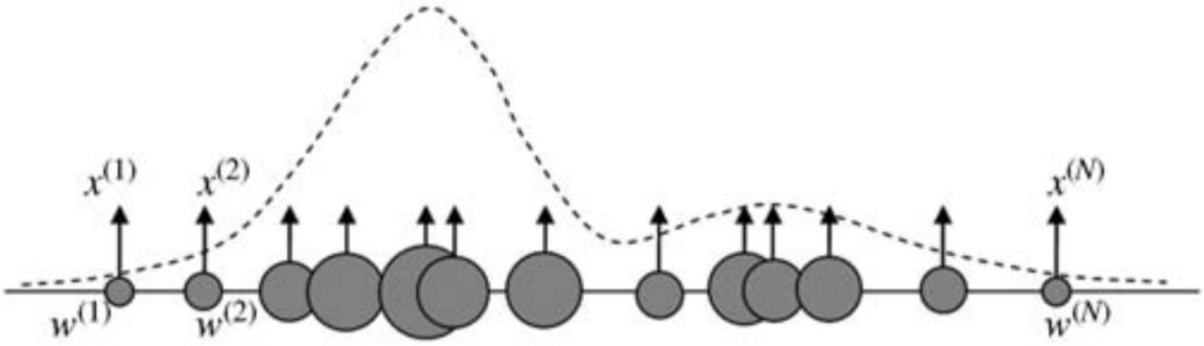
$$\int \mathbf{g}(\mathbf{x})\, p(\mathbf{x} \mid \mathbf{y}_{1:T})\, \mathrm{d}\mathbf{x} = \int \left[\mathbf{g}(\mathbf{x}) \frac{p(\mathbf{x} \mid \mathbf{y}_{1:T})}{\pi(\mathbf{x} \mid \mathbf{y}_{1:T})} \right] \pi(\mathbf{x} \mid \mathbf{y}_{1:T})\, \mathrm{d}\mathbf{x}$$

N samples

$$\mathbf{x}^{(i)} \sim \pi(\mathbf{x} \mid \mathbf{y}_{1:T}), \qquad i = 1, \ldots, N,$$

$$\begin{aligned} \mathrm{E}[\mathbf{g}(\mathbf{x}) \mid \mathbf{y}_{1:T}] &\approx \frac{1}{N} \sum_{i=1}^{N} \frac{p(\mathbf{x}^{(i)} \mid \mathbf{y}_{1:T})}{\pi(\mathbf{x}^{(i)} \mid \mathbf{y}_{1:T})} \mathbf{g}(\mathbf{x}^{(i)}) \\ &= \sum_{i=1}^{N} \tilde{w}^{(i)} \mathbf{g}(\mathbf{x}^{(i)}), \end{aligned}$$

where: $\tilde{w}^{(i)} = \frac{1}{N} \frac{p(\mathbf{x}^{(i)} \mid \mathbf{y}_{1:T})}{\pi(\mathbf{x}^{(i)} \mid \mathbf{y}_{1:T})}$



$x^{(1)}$
$x^{(2)}$
$x^{(N)}$
$w^{(1)}$
$w^{(2)}$
$w^{(N)}$

Importance Sampling

$$\tilde{w}^{(i)} = \frac{1}{N} \frac{p(\mathbf{x}^{(i)} \mid \mathbf{y}_{1:T})}{\pi(\mathbf{x}^{(i)} \mid \mathbf{y}_{1:T})}$$

- Direct Importance Sampling requires us to be able to evaluate $p(\mathbf{x}^{(i)} \mid \mathbf{y}_{1:T})$
- … which we assume we can't!

- Bayes' rule to the rescue:

Importance Sampling

$$\tilde{w}^{(i)} = \frac{1}{N} \frac{p(\mathbf{x}^{(i)} \mid \mathbf{y}_{1:T})}{\pi(\mathbf{x}^{(i)} \mid \mathbf{y}_{1:T})}$$

- Direct Importance Sampling requires us to be able to evaluate $p(\mathbf{x}^{(i)} \mid \mathbf{y}_{1:T})$
- … which we assume we can't!

- Bayes' rule to the rescue:

Often, we can easily evaluate those terms

$$p(\mathbf{x}^{(i)} \mid \mathbf{y}_{1:T}) = \frac{p(\mathbf{y}_{1:T} \mid \mathbf{x}^{(i)})\, p(\mathbf{x}^{(i)})}{\int p(\mathbf{y}_{1:T} \mid \mathbf{x})\, p(\mathbf{x})\, \mathrm{d}\mathbf{x}}$$

As for this one, much less so...

Importance Sampling

$$\tilde{w}^{(i)} = \frac{1}{N} \frac{p(\mathbf{x}^{(i)} \mid \mathbf{y}_{1:T})}{\pi(\mathbf{x}^{(i)} \mid \mathbf{y}_{1:T})}$$

- Direct Importance Sampling requires us to be able to evaluate $p(\mathbf{x}^{(i)} \mid \mathbf{y}_{1:T})$
- … which we assume we can't!

- Bayes' rule to the rescue:

Often, we can easily evaluate those terms

$$p(\mathbf{x}^{(i)} \mid \mathbf{y}_{1:T}) = \frac{\boxed{p(\mathbf{y}_{1:T} \mid \mathbf{x}^{(i)})}\,\boxed{p(\mathbf{x}^{(i)})}}{\boxed{\int p(\mathbf{y}_{1:T} \mid \mathbf{x})\, p(\mathbf{x})\, \mathrm{d}\mathbf{x}}}$$

Let's just do Importance Sampling again, this time on this integral normalization constant!

$$
\begin{aligned}
\mathrm{E}[\mathbf{g}(\mathbf{x}) \mid \mathbf{y}_{1:T}] &= \int \mathbf{g}(\mathbf{x})\, p(\mathbf{x} \mid \mathbf{y}_{1:T})\, \mathrm{d}\mathbf{x} \\
&= \frac{\int \mathbf{g}(\mathbf{x})\, p(\mathbf{y}_{1:T} \mid \mathbf{x})\, p(\mathbf{x})\, \mathrm{d}\mathbf{x}}{\int p(\mathbf{y}_{1:T} \mid \mathbf{x})\, p(\mathbf{x})\, \mathrm{d}\mathbf{x}}
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{E}[\mathbf{g}(\mathbf{x}) \mid \mathbf{y}_{1:T}] &= \int \mathbf{g}(\mathbf{x})\, p(\mathbf{x} \mid \mathbf{y}_{1:T})\, \mathrm{d}\mathbf{x} \\
&= \frac{\int \mathbf{g}(\mathbf{x})\, p(\mathbf{y}_{1:T} \mid \mathbf{x})\, p(\mathbf{x})\, \mathrm{d}\mathbf{x}}{\int p(\mathbf{y}_{1:T} \mid \mathbf{x})\, p(\mathbf{x})\, \mathrm{d}\mathbf{x}} \\
&= \frac{\int \left[\frac{p(\mathbf{y}_{1:T} \mid \mathbf{x})\, p(\mathbf{x})}{\pi(\mathbf{x} \mid \mathbf{y}_{1:T})}\, \mathbf{g}(\mathbf{x}) \right] \pi(\mathbf{x} \mid \mathbf{y}_{1:T})\, \mathrm{d}\mathbf{x}}{\int \left[\frac{p(\mathbf{y}_{1:T} \mid \mathbf{x})\, p(\mathbf{x})}{\pi(\mathbf{x} \mid \mathbf{y}_{1:T})} \right] \pi(\mathbf{x} \mid \mathbf{y}_{1:T})\, \mathrm{d}\mathbf{x}}
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{E}[\mathbf{g}(\mathbf{x}) \mid \mathbf{y}_{1:T}] &= \int \mathbf{g}(\mathbf{x})\, p(\mathbf{x} \mid \mathbf{y}_{1:T})\, \mathrm{d}\mathbf{x} \\
&= \frac{\int \mathbf{g}(\mathbf{x})\, p(\mathbf{y}_{1:T} \mid \mathbf{x})\, p(\mathbf{x})\, \mathrm{d}\mathbf{x}}{\int p(\mathbf{y}_{1:T} \mid \mathbf{x})\, p(\mathbf{x})\, \mathrm{d}\mathbf{x}} \\
&= \frac{\int \left[\frac{p(\mathbf{y}_{1:T} | \mathbf{x})\, p(\mathbf{x})}{\pi(\mathbf{x} | \mathbf{y}_{1:T})}\, \mathbf{g}(\mathbf{x}) \right] \pi(\mathbf{x} \mid \mathbf{y}_{1:T})\, \mathrm{d}\mathbf{x}}{\int \left[\frac{p(\mathbf{y}_{1:T} | \mathbf{x})\, p(\mathbf{x})}{\pi(\mathbf{x} | \mathbf{y}_{1:T})} \right] \pi(\mathbf{x} \mid \mathbf{y}_{1:T})\, \mathrm{d}\mathbf{x}} \\
&\approx \frac{\frac{1}{N} \sum_{i=1}^{N} \frac{p(\mathbf{y}_{1:T} | \mathbf{x}^{(i)})\, p(\mathbf{x}^{(i)})}{\pi(\mathbf{x}^{(i)} | \mathbf{y}_{1:T})}\, \mathbf{g}(\mathbf{x}^{(i)})}{\frac{1}{N} \sum_{j=1}^{N} \frac{p(\mathbf{y}_{1:T} | \mathbf{x}^{(j)})\, p(\mathbf{x}^{(j)})}{\pi(\mathbf{x}^{(j)} | \mathbf{y}_{1:T})}} \\
&= \sum_{i=1}^{N} \underbrace{\left[\frac{\frac{p(\mathbf{y}_{1:T} | \mathbf{x}^{(i)})\, p(\mathbf{x}^{(i)})}{\pi(\mathbf{x}^{(i)} | \mathbf{y}_{1:T})}}{\sum_{j=1}^{N} \frac{p(\mathbf{y}_{1:T} | \mathbf{x}^{(j)})\, p(\mathbf{x}^{(j)})}{\pi(\mathbf{x}^{(j)} | \mathbf{y}_{1:T})}} \right]}_{w^{(i)}} \mathbf{g}(\mathbf{x}^{(i)}).
\end{aligned}
$$

Importance Sampling

1 Draw N samples from the importance distribution:

$$\mathbf{x}^{(i)} \sim \pi(\mathbf{x} \mid \mathbf{y}_{1:T}), \qquad i = 1, \ldots, N. \tag{7.9}$$

Importance Sampling

1 Draw N samples from the importance distribution:

$$\mathbf{x}^{(i)} \sim \pi(\mathbf{x} \mid \mathbf{y}_{1:T}), \qquad i = 1, \ldots, N. \tag{7.9}$$

2 Compute the unnormalized weights by

$$w^{*(i)} = \frac{p(\mathbf{y}_{1:T} \mid \mathbf{x}^{(i)})\, p(\mathbf{x}^{(i)})}{\pi(\mathbf{x}^{(i)} \mid \mathbf{y}_{1:T})}, \tag{7.10}$$

and the normalized weights by

$$w^{(i)} = \frac{w^{*(i)}}{\sum_{j=1}^{N} w^{*(j)}}. \tag{7.11}$$

Importance Sampling

1 Draw N samples from the importance distribution:

$$\mathbf{x}^{(i)} \sim \pi(\mathbf{x} \mid \mathbf{y}_{1:T}), \qquad i = 1, \ldots, N. \tag{7.9}$$

2 Compute the unnormalized weights by

$$w^{*(i)} = \frac{p(\mathbf{y}_{1:T} \mid \mathbf{x}^{(i)})\, p(\mathbf{x}^{(i)})}{\pi(\mathbf{x}^{(i)} \mid \mathbf{y}_{1:T})}, \tag{7.10}$$

and the normalized weights by

$$w^{(i)} = \frac{w^{*(i)}}{\sum_{j=1}^{N} w^{*(j)}}. \tag{7.11}$$

3 The approximation to the posterior expectation of $\mathbf{g}(\mathbf{x})$ is then given as

$$\mathrm{E}[\mathbf{g}(\mathbf{x}) \mid \mathbf{y}_{1:T}] \approx \sum_{i=1}^{N} w^{(i)}\, \mathbf{g}(\mathbf{x}^{(i)}). \tag{7.12}$$

Importance Sampling

- Equivalent interpretation:

$$\mathrm{E}[\mathbf{g}(\mathbf{x}) \mid \mathbf{y}_{1:T}] = \int \mathbf{g}(\mathbf{x})\, p(\mathbf{x} \mid \mathbf{y}_{1:T})\, \mathrm{d}\mathbf{x}$$

$$\mathrm{E}[\mathbf{g}(\mathbf{x}) \mid \mathbf{y}_{1:T}] \approx \sum_{i=1}^{N} w^{(i)}\, \mathbf{g}(\mathbf{x}^{(i)})$$

$$\Longleftrightarrow \quad p(\mathbf{x} \mid \mathbf{y}_{1:T}) \approx \sum_{i=1}^{N} w^{(i)}\, \delta(\mathbf{x} - \mathbf{x}^{(i)}).$$

Sequential Importance Sampling

- Assuming a generic state-space model,

$$\mathbf{x}_k \sim p(\mathbf{x}_k \mid \mathbf{x}_{k-1}),$$
$$\mathbf{y}_k \sim p(\mathbf{y}_k \mid \mathbf{x}_k),$$

- And using the Markovian property of the model,

$$\begin{aligned} p(\mathbf{x}_{0:k} \mid \mathbf{y}_{1:k}) &\propto p(\mathbf{y}_k \mid \mathbf{x}_{0:k}, \mathbf{y}_{1:k-1})\, p(\mathbf{x}_{0:k} \mid \mathbf{y}_{1:k-1}) \\ &= p(\mathbf{y}_k \mid \mathbf{x}_k)\, p(\mathbf{x}_k \mid \mathbf{x}_{0:k-1}, \mathbf{y}_{1:k-1})\, p(\mathbf{x}_{0:k-1} \mid \mathbf{y}_{1:k-1}) \\ &= p(\mathbf{y}_k \mid \mathbf{x}_k)\, p(\mathbf{x}_k \mid \mathbf{x}_{k-1})\, p(\mathbf{x}_{0:k-1} \mid \mathbf{y}_{1:k-1}). \end{aligned}$$

Sequential Importance Sampling

- We can then compute the importance weights

$$w_k^{(i)} \propto \frac{p(\mathbf{y}_k \mid \mathbf{x}_k^{(i)})\, p(\mathbf{x}_k^{(i)} \mid \mathbf{x}_{k-1}^{(i)})\, p(\mathbf{x}_{0:k-1}^{(i)} \mid \mathbf{y}_{1:k-1})}{\pi(\mathbf{x}_{0:k}^{(i)} \mid \mathbf{y}_{1:k})}$$

- Assume the importance distribution follows a recursive form

$$\pi(\mathbf{x}_{0:k} \mid \mathbf{y}_{1:k}) = \pi(\mathbf{x}_k \mid \mathbf{x}_{0:k-1}, \mathbf{y}_{1:k})\, \pi(\mathbf{x}_{0:k-1} \mid \mathbf{y}_{1:k-1})$$

Sequential Importance Sampling

- We can then compute the importance weights

$$w_k^{(i)} \propto \frac{p(\mathbf{y}_k \mid \mathbf{x}_k^{(i)})\, p(\mathbf{x}_k^{(i)} \mid \mathbf{x}_{k-1}^{(i)})\, p(\mathbf{x}_{0:k-1}^{(i)} \mid \mathbf{y}_{1:k-1})}{\pi(\mathbf{x}_{0:k}^{(i)} \mid \mathbf{y}_{1:k})}$$

- Assume the importance distribution follows a recursive form

$$\pi(\mathbf{x}_{0:k} \mid \mathbf{y}_{1:k}) = \pi(\mathbf{x}_k \mid \mathbf{x}_{0:k-1}, \mathbf{y}_{1:k})\, \pi(\mathbf{x}_{0:k-1} \mid \mathbf{y}_{1:k-1})$$

$$\Rightarrow \boxed{w_k^{(i)} \propto \frac{p(\mathbf{y}_k \mid \mathbf{x}_k^{(i)})\, p(\mathbf{x}_k^{(i)} \mid \mathbf{x}_{k-1}^{(i)})}{\pi(\mathbf{x}_k^{(i)} \mid \mathbf{x}_{0:k-1}^{(i)}, \mathbf{y}_{1:k})} \frac{p(\mathbf{x}_{0:k-1}^{(i)} \mid \mathbf{y}_{1:k-1})}{\pi(\mathbf{x}_{0:k-1}^{(i)} \mid \mathbf{y}_{1:k-1})}}$$

Sequential Importance Sampling

$$w_k^{(i)} \propto \frac{p(\mathbf{y}_k \mid \mathbf{x}_k^{(i)})\, p(\mathbf{x}_k^{(i)} \mid \mathbf{x}_{k-1}^{(i)})}{\pi(\mathbf{x}_k^{(i)} \mid \mathbf{x}_{0:k-1}^{(i)}, \mathbf{y}_{1:k})} \boxed{\frac{p(\mathbf{x}_{0:k-1}^{(i)} \mid \mathbf{y}_{1:k-1})}{\pi(\mathbf{x}_{0:k-1}^{(i)} \mid \mathbf{y}_{1:k-1})}}$$

$$\boxed{w_{k-1}^{(i)} \propto \frac{p(\mathbf{x}_{0:k-1}^{(i)} \mid \mathbf{y}_{1:k-1})}{\pi(\mathbf{x}_{0:k-1}^{(i)} \mid \mathbf{y}_{1:k-1})}}$$

Sequential Importance Sampling

$$w_k^{(i)} \propto \frac{p(\mathbf{y}_k \mid \mathbf{x}_k^{(i)})\, p(\mathbf{x}_k^{(i)} \mid \mathbf{x}_{k-1}^{(i)})}{\pi(\mathbf{x}_k^{(i)} \mid \mathbf{x}_{0:k-1}^{(i)}, \mathbf{y}_{1:k})}\, w_{k-1}^{(i)}$$

Sequential Importance Sampling (SIS)

- SIS algo

- *Draw N samples $\mathbf{x}_0^{(i)}$ from the prior*

$$\mathbf{x}_0^{(i)} \sim p(\mathbf{x}_0), \qquad i = 1, \ldots, N,$$

and set $w_0^{(i)} = 1/N$, for all $i = 1, \ldots, N$.

- *For each $k = 1, \ldots, T$ do the following.*

  1 Draw samples $\mathbf{x}_k^{(i)}$ from the importance distributions

$$\mathbf{x}_k^{(i)} \sim \pi(\mathbf{x}_k \mid \mathbf{x}_{0:k-1}^{(i)}, \mathbf{y}_{1:k}), \qquad i = 1, \ldots, N.$$

  2 Calculate new weights according to

$$w_k^{(i)} \propto w_{k-1}^{(i)} \frac{p(\mathbf{y}_k \mid \mathbf{x}_k^{(i)})\, p(\mathbf{x}_k^{(i)} \mid \mathbf{x}_{k-1}^{(i)})}{\pi(\mathbf{x}_k^{(i)} \mid \mathbf{x}_{0:k-1}^{(i)}, \mathbf{y}_{1:k})}$$

  and normalize them to sum to unity.

Sequential Importance Sampling (SIS)

- Two Big Questions remain:

  1. what is / how do we choose the importance distribution?

  2. in practice, does SIS work?

Question 1: *what is / how do we choose the importance distribution?*

- Should be in such a form we can easily draw samples from it
- *And* we can evaluate the probability densities at the sample points
- Litterature is heavy on this, e.g. cf. Doucet et al. (2011)
- Optimal importance distribution is

$$\pi(\mathbf{x}_k \mid \mathbf{x}_{0:k-1}, \mathbf{y}_{1:k}) = p(\mathbf{x}_k \mid \mathbf{x}_{k-1}, \mathbf{y}_k)$$

Question 1: *what is / how do we choose the importance distribution?*

- Importance Distribution can sometimes be obtained by local linearization or a mixture of EKFs / UKFs are used

- The *bootstrap filter* (1993) used the dynamic model as the importance distrib.
  - Implementation super easy but requires N to be very large for fair accuracy

$$\mathbf{x}_k^{(i)} \sim p(\mathbf{x}_k \mid \mathbf{x}_{k-1}^{(i)})$$

Question 2: *does it work in practice?*

- Weights tend to concentrate into a single particle… Degeneracy problem
- ***Resampling*** is meant to deal with that problem
- We draw N new samples from the discrete distribution defined by the weights and replace the old set

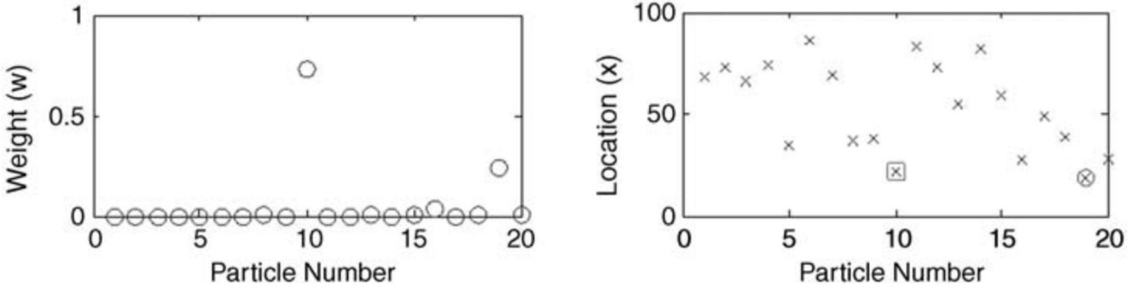
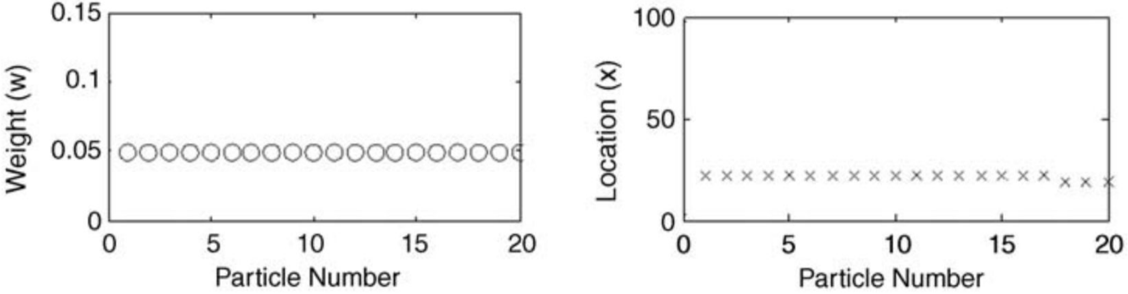


Figure 7.11 After two cycles without resampling using the SIS Particle filter, we see two particle weights dominate over the others.



Question 2: *does it work in practice?*

- So now we do:
    - Draw N samples
    - Compute weights
    - Resample
    - … repeat

- In practice, Resampling is not done at every step

Question 2: does it work in practice?

Algorithm 7.4 (Sequential importance resampling) *The sequential importance resampling (SIR) algorithm, which is also called the particle filter (PF), is the following.*

- *Draw N samples $\mathbf{x}_0^{(i)}$ from the prior*

$$\mathbf{x}_0^{(i)} \sim p(\mathbf{x}_0), \qquad i = 1, \ldots, N, \tag{7.28}$$

  and set $w_0^{(i)} = 1/N$, for all $i = 1, \ldots, N$.
- *For each $k = 1, \ldots, T$ do the following:*

  1 *Draw samples $\mathbf{x}_k^{(i)}$ from the importance distributions*

$$\mathbf{x}_k^{(i)} \sim \pi(\mathbf{x}_k \mid \mathbf{x}_{k-1}^{(i)}, \mathbf{y}_{1:k}), \qquad i = 1, \ldots, N. \tag{7.29}$$

  2 *Calculate new weights according to*

$$w_k^{(i)} \propto w_{k-1}^{(i)} \frac{p(\mathbf{y}_k \mid \mathbf{x}_k^{(i)})\, p(\mathbf{x}_k^{(i)} \mid \mathbf{x}_{k-1}^{(i)})}{\pi(\mathbf{x}_k^{(i)} \mid \mathbf{x}_{k-1}^{(i)}, \mathbf{y}_{1:k})} \tag{7.30}$$

  and normalize them to sum to unity.

  3 *If the effective number of particles* (7.27) *is too low, perform resampling.*

The *Bootstrap Particle Filter*

Algorithm 7.5 (Bootstrap filter) *The bootstrap filter algorithm is as follows.*

1 Draw a new point $\mathbf{x}_k^{(i)}$ *for each point in the sample set* $\{\mathbf{x}_{k-1}^{(i)} \ : \ i = 1, \ldots, N\}$ *from the dynamic model:*

$$\mathbf{x}_k^{(i)} \sim p(\mathbf{x}_k \mid \mathbf{x}_{k-1}^{(i)}), \qquad i = 1, \ldots, N. \tag{7.34}$$

2 Calculate the weights

$$w_k^{(i)} \propto p(\mathbf{y}_k \mid \mathbf{x}_k^{(i)}), \qquad i = 1, \ldots, N, \tag{7.35}$$

and normalize them to sum to unity.

3 Do resampling.

The *Bootstrap Particle Filter*

- Implement one in a Jupyter Notebook

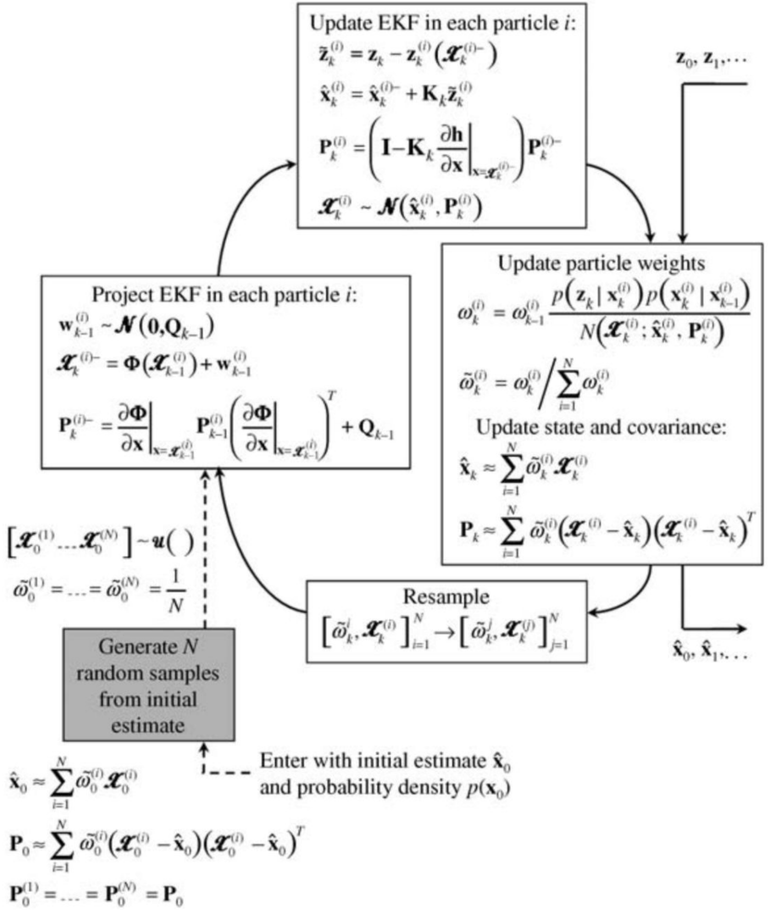


Figure 7.15 Particle filter using Extended Kalman filters for Local Linearization.