Choice of importance distribution

Sensor fusion & nonlinear filtering

Lars Hammarstrand

- A carefully selected importance distribution, $q(\mathbf{x}_k | \mathbf{x}_{k-1}, \mathbf{y}_k)$, can slow down the degeneracy and improve performance.

Intuition: if most particles are placed in ”high probability regions” there is less need to get rid of useless particles.

Optimal importance density

- The optimal importance density is

$$q(\mathbf{x}_k | \mathbf{x}_{k-1}, \mathbf{y}_k) = p(\mathbf{x}_k | \mathbf{x}_{k-1}, \mathbf{y}_k).$$

- Unfortunately, in most nonlinear settings, $p(\mathbf{x}_k | \mathbf{x}_{k-1}, \mathbf{y}_k)$, is difficult to both draw samples from and to evaluate.

- We can approximate $p(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{y}_k)$ using, e.g., linearization.
- The most common choice is still, $q(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{y}_k) = p(\mathbf{x}_k|\mathbf{x}_{k-1})$, and the bootstrap algorithm.

The bootstrap PF

At each time k:

- Draw $\mathbf{x}_k^{(i)} \sim p(\mathbf{x}_k|\mathbf{x}_{k-1}^{(i)})$, for $i = 1, \ldots, N$.
- Calculate $w_k^{(i)} \propto w_{k-1}^{(i)} p(\mathbf{y}_k|\mathbf{x}_k^{(i)})$ and normalize to 1.
- Resample.

- Note: if we resample at every time step, we get $w_k^{(i)} \propto p(\mathbf{y}_k|\mathbf{x}_k^{(i)})$ since $w_{k-1}^{(i)} = 1/N \ \forall i$ after resampling.
- Note 2: the Auxiliary PF (APF) is variation of the SIR algorithm that makes use of $\mathbf{y}_k$.

- Particle filters (PFs) can handle highly nonlinear and non-Gaussian systems.
- Particle filters are asymptotically exact as you increase N.
- The complexity is roughly $O(N)$ but the gain in performance flattens out as you increase N.
- Unfortunately, PFs suffer from the curse of dimensionality and are intractable in higher dimensions.

- The output from a PF is an approximation

$$p(\mathbf{x}_k|\mathbf{y}_{1:k}) \approx \sum_{i=1}^{N} w_k^{(i)} \delta(\mathbf{x}_k - \mathbf{x}_k^{(i)})$$

$\Leftrightarrow \Pr\{x_k = x' | y_{1:k}\} = \begin{cases} \omega_k^{(i)} & \text{if } x' = x_k^{(i)} \\ 0 & \text{otherwise} \end{cases}$

which implies that

$$\mathbb{E}\left[\mathbf{g}(\mathbf{x}_k)|\mathbf{y}_{1:k}\right] \approx \sum_{i=1}^{N} w_k^{(i)} \mathbf{g}(\mathbf{x}_k^{(i)}).$$