Particle filters for deterministic problems

The stochastic model set-up:
(specified to obs every time and linear obs fctn)

$$x_{n+1} = f(x_n) + v_n \quad , \quad v_n \sim N(0,Q) \ , \text{iid}$$

$$y_{n+1} = H x_{n+1} + \eta_n \quad , \quad \eta_n \sim N(0,R) \ . \text{iid}$$

PFs:

target
$$p(x_{0:n} | y_{1:n}) \propto p(x_{0:n-1} | y_{1:n-1}) \underbrace{p(x_n | x_{n-1})\, p(y_n | x_n)}_{p(y_n | x_{n-1})\, p(x_n | x_{n-1}, y_n)}$$

proposal
$$q(x_{0:n} | y_{1:n}) \propto q_0(x_0) \prod_{j=1}^{n} q_j(x_j | x_{j-1}, y_j)$$

Idea: Consider limit $Q \to 0$ to get stochastic $\to$ deterministic

OPF:
$$x_n^j = f(x_{n-1}^j) + K(y_n - Hf(x_{n-1}^j))$$

$$w^j \propto \exp\left[-\tfrac{1}{2} (y_n - Hf(x_{n-1}^j))^T (HQH^T + R)^{-1} (y_n - Hf(x_{n-1}^j))\right]$$

$$K = QH^T(HQH^T + R)^{-1}$$

For $Q \to 0$: $\quad K = 0$

$$x_n^j = f(x_{n-1}^j) \quad \rightarrow \text{This means run the model!}$$

$$w^j \propto \exp\Big(-\tfrac{1}{2}\big(y_n - H\underbrace{f(x_{n-1}^j)}_{x_n^j}\big) R^{-1} \big(y_n - H\underbrace{f(x_{n-1}^j)}_{x_n^j}\big)\Big)$$

$\Rightarrow$ OPF becomes the standard PF!

Particle filters for deterministic problems

The stochastic model set-up:
(specified to obs every time and linear obs fctn)

$$x_{n+1} = f(x_n) + v_n \quad , \quad v_n \sim \mathcal{N}(0, Q) \ , \text{iid}$$

$$y_{n+1} = H x_{n+1} + \eta_n \quad , \quad \eta_n \sim \mathcal{N}(0, R) \ . \ \text{iid}$$

PFs: $p(x_{0:n} | y_{1:n}) \propto p(x_{0:n-1} | y_{1:n-1}) \underbrace{p(x_n | x_{n-1})\, p(y_n | x_n)}_{p(y_n | x_{n-1})\, p(x_n | x_{n-1}, y_n)}$

target

proposal $q(x_{0:n} | y_{1:n}) \propto q_0(x_0) \prod_{j=1}^{n} q_j(x_j | x_{j-1}, y_j)$

Idea: Consider limit $Q \to 0$ to get stochastic $\to$ deterministic

OPF:

$$x_n^j = f(x_{n-1}^j) + K(y_n - H f(x_{n-1}^j))$$

$$w^j \propto \exp\left[-\tfrac{1}{2} (y_n - H f(x_{n-1}^j))^T (HQH^T + R)^{-1} (y_n - H f(x_{n-1}^j))\right]$$

$$K = QH^T (HQH^T + R)^{-1}$$

For $Q \to 0$: $K = 0$

$$x_n^j = f(x_{n-1}^j) \quad \rightsquigarrow \text{This means run the model!}$$

$$w^j \propto \exp\Big(-\tfrac{1}{2} (y_n - H \underbrace{f(x_{n-1}^j)}_{x_n^j}) R^{-1} (y_n - H \underbrace{f(x_{n-1}^j)}_{x_n^j})\Big)$$

$\Rightarrow$ OPF becomes the standard PF!

... filter can be viewed as a sequential importance sampler for the posterior distribution $p(x_n|y_{1:n})$

Recall: $p(x_n|y_{1:n}) \propto p(y_n|x_n) \underbrace{p(x_n|y_{1:n-1})}$

Proposal of standard PF for deterministic problems

$x_{n-1} \sim p(x_{n-1}|y_{1:n-1})$

$x_n = f(x_{n-1})$

$x_n \sim p(x_n|y_{1:n-1})$

SPF/OPF: $q(x_n) \propto p(x_n|y_{1:n-1})$

$w \propto p(y_n|x_n)$

↳ This is exactly what we have just obtained by looking at $Q \to 0$ limit.

↳ Optimality no longer holds (at least nobody proved it)

↳ Recall: building ~~other~~ PF's for deterministic model with proposals other than $p(x_n|y_{1:n-1})$ is very difficult.
MC sampling requires target is known up to a constant.
<u>But</u> $p(x_n|y_{1:n}) \propto p(y_n|x_n) \underbrace{p(x_n|y_{1:n-1})}_{\text{not known up to a constant.}}$

happens to target distribution in this limit?

Consider the first couple of steps:

$$p(x_0, x_1 | y_1) \propto p(x_0)\, \underbrace{p(x_1 | x_0)}\, \underbrace{p(y_1 | x_1)}$$

$$\propto \exp\left(-\tfrac{1}{2}(x_1 - f(x_0))^T Q^{-1} (x_1 - f(x_0))\right)$$

$$\propto \exp\left(-\tfrac{1}{2}(y_1 - H\overset{x_1}{\cancel{f(x_0)}})^T R^{-1} (y_1 - H\cancel{f(x_0)})\right)$$

$\to \delta(x_1 - f(x_0))$ as $Q \to 0$

(qual question?)

$$\propto p(x_0)\, \delta(x_1 - f(x_0))\, p(y_1 | x_p)$$

$$\exp\left(-\tfrac{1}{2}(y_1 - Hx_1)^T R^{-1} (y - Hx_1)\right)$$

$$\int p(x_0, x_1 | y_1)\, dx_1 = p(x_0) \int \delta(x_1 - f(x_0))\, p(y_1 | x_p)\, dx_1$$

$$\underbrace{p(x_0 | y_1) = p(x_0)\, \underbrace{p(y_1 | x_0)}_{\exp\left(-\frac{1}{2}(y_1 - Hf(x_0))^T R^{-1} (y_1 - Hf(x_0))\right)}}$$

→ This suggests sampling $p(x_0 | y_1)$ (~4D-Var) rather than $p(x_1 | y_1)$ (~PF)

→ Connection between PFs and Var!

Idea: (1) Solve 4D-Var problem:

$$F = -\log p(x_0 | y_1)$$

$$\mu = \arg\min F$$

$$P = (2 J^T J)^{-1}$$

(2) Proposal: $q \sim N(\mu, P)$

(or multivariate-t etc.)

or random map

(3) weights: $w \propto \frac{p(x_0)}{q(x_0)}$

→ variational particle smoother

→ implicit sampling

→ IEnKS

(Connections between EnKF and Gauss-Newton)

→ En4D-Var

Optimality? $\operatorname{var}(w) = 0$ if problem is linear $(f(x_0) = M x_0)$

$(h(x_1) = H x_1)$

and Gaussian $(p(x_0) \sim N(\mu_0, P_0)$

$y = H f(x_0) + \eta, \ \eta \sim N(0, R))$

Issues: Consider next steps:

$$p(x_0, x_1, x_2 | y_1, y_2) \rightarrow p(x_0 | y_1, y_2)$$

$$\vdots$$

$$p(x_0 | y_1 \cdots y_n)$$

Good idea in principle, but hard to do in practice if model is chaotic.

→ Difficult to propagate information backward in time for large n.

Also impractical for large n.

[illegible]

we make this sequential?

$$p(x_{n-1} | y_{1:n}) \propto p(y_n | x_n) \underbrace{p(x_{n-1} | y_{1:n-1})}$$

Unknown!

Same problem as with other class of PFs

<u>Idea:</u> As in 4D-Var, approximate $p(x_{n-1} | y_{1:n-1})$ by Gaussian, using results from previous assimilation cycle: $\hat{p}(x_{n-1} | y_{1:n-1})$

Then sample approximate posterior using, e.g. Gaussian proposal;

$$\hat{p}(x_{n-1} | y_{1:n}) \propto p(y_n | x_n) \hat{p}(x_{n-1} | y_{1:n-1})$$

→ This extends 4D-Var

→ Optimal (w/ $\omega = 0$) if problem is linear & Gaussian

→ This can beat EnKF even on L95.

<u>Open questions</u>

- why is Gaussian approximation of $p(x_{n-1} | y_{1:n-1})$ ~~not so bad~~ appropriate? For what class of problems is this appropriate
- What should we do if $p(x_{n-1} | y_{1:n-1})$ is not nearly Gaussian?
- ~~How do~~ we localize this algorithm?
- why is Gaussian proposal "ok" so good that weights have almost no impact on estimates/RMSE/spread?