Particle Filtering with Progressive Gaussian Approximations to the Optimal Importance Density

Pete Bunch and Simon Godsill

Cambridge University Engineering Department
Signal Processing & Communications Lab

17th December, 2012

The Plan

- A quick review of particle filtering and importance densities.
- The problem with the usual choices of importance density.
- The new progressive proposal method.
- How it relates to existing algorithms.
- Some simulation results.

Particle Filtering

$$
\begin{aligned}
x_n \sim f(x_n|x_{n-1}) &= \mathcal{N}\left(x_n \,|\phi(x_{n-1}), Q\right) \\
y_n \sim \quad g(y_n|x_n) &= \mathcal{N}\left(y_n \,|\psi(x_n), R\right) \\
x_1 \sim \quad p(x_1) &= \mathcal{N}\left(x_1 \,|m_1, P_1\right),
\end{aligned}
$$

Particle Filtering

$$
\begin{aligned}
x_n \sim f(x_n|x_{n-1}) &= \mathcal{N}\left(x_n \,|\phi(x_{n-1}), Q\,\right) \\
y_n \sim \quad g(y_n|x_n) &= \mathcal{N}\left(y_n \,|\psi(x_n), R\,\right) \\
x_1 \sim \quad p(x_1) &= \mathcal{N}\left(x_1 \,|m_1, P_1\,\right),
\end{aligned}
$$

- Particle filter approximates:

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

- Select an $(n-1)$ particle and sample a new state, $x_n^{(i)} \sim q(x_n|x_{n-1}^{(i)})$.
- Update weight $w_n^{(i)} = \frac{f(x_n^{(i)}|x_{n-1}^{(i)})g(y_n|x_n^{(i)})}{q(x_n^{(i)}|x_{n-1}^{(i)})}$.

Importance Densities

Importance Densities

Bootstrap filter

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

Importance Densities

Bootstrap filter

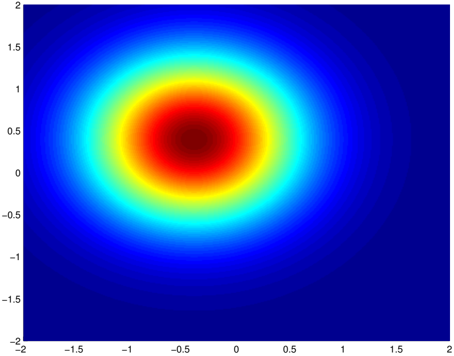
$$q(x_n|x_{n-1}) = f(x_n|x_{n-1}).$$

Optimal Importance Density

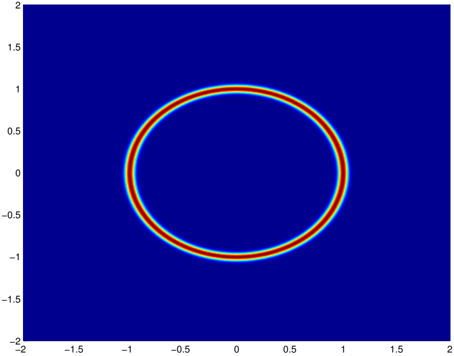
$$q(x_n|x_{n-1}) = \frac{f(x_n|x_{n-1})g(y_n|x_n)}{\int f(x_n|x_{n-1})g(y_n|x_n)dx_n}.$$

Approximate by linearisation or sigma-point approximation.

An Example

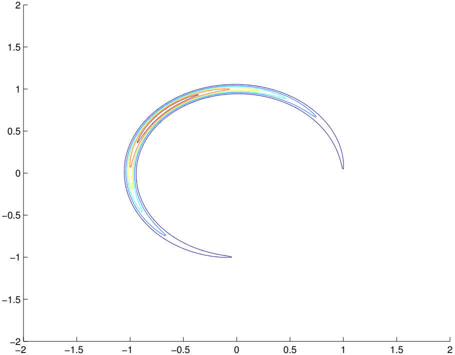


Prior

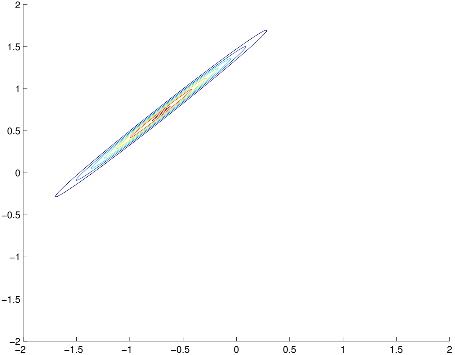


Likelihood

An Example - Linearisation

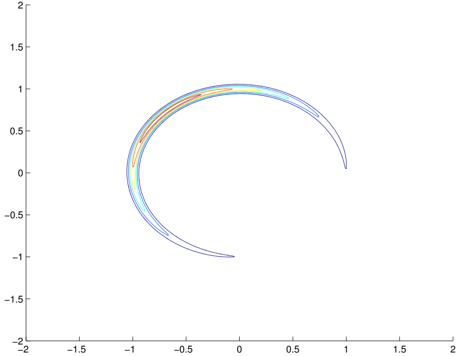


True Posterior

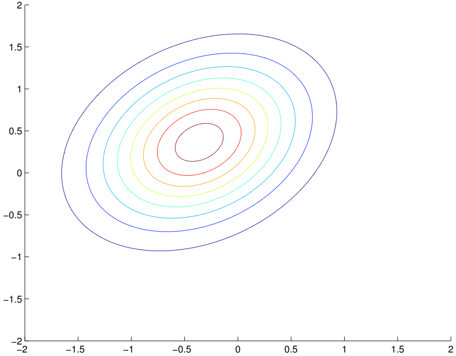


Approximated by Linearisation

An Example - Unscented Transform



True Posterior



Approximated by Unscented Transform

Progressive Principle

Introduce the observation gradually using a stretch of "pseudo-time", $\lambda \in [0, 1]$.

Progressive Principle

Introduce the observation gradually using a stretch of "pseudo-time", $\lambda \in [0, 1]$.

Smooth sequence of target distributions:

$$\tilde{\pi}_{n,\lambda}(x_{1:n-1}, x_{n,\lambda}) \propto g(y_n|x_{n,\lambda})^{\lambda} f(x_{n,\lambda}|x_{n-1}) p(x_{1:n-1}|y_{1:n-1})$$

Progressive Principle

Introduce the observation gradually using a stretch of "pseudo-time", $\lambda \in [0, 1]$.

Smooth sequence of target distributions:

$$\tilde{\pi}_{n,\lambda}(x_{1:n-1}, x_{n,\lambda}) \propto g(y_n|x_{n,\lambda})^{\lambda} f(x_{n,\lambda}|x_{n-1}) p(x_{1:n-1}|y_{1:n-1})$$

Smooth sequence of optimal importance densities:

$$\pi_{n,\lambda}(x_{n,\lambda}|x_{n-1}^{(j)}) \propto g(y_n|x_{n,\lambda})^{\lambda} f(x_{n,\lambda}|x_{n-1}^{(j)})$$

Partially Linear Gaussian Models

$$f(x_n|x_{n-1}) = \mathcal{N}\left(x_n \,|\phi(x_{n-1}), Q\right)$$
$$g(y_n|x_n) = \mathcal{N}\left(y_n \,|Hx_n, R\right)$$

Partially Linear Gaussian Models

$$f(x_n|x_{n-1}) = \mathcal{N}(x_n\,|\phi(x_{n-1}), Q)$$
$$g(y_n|x_n) = \mathcal{N}(y_n\,|Hx_n, R)$$

Optimal importance density is analytically tractable.

$$\pi_\lambda(x_\lambda|x_{n-1}) = \mathcal{N}(x_\lambda\,|m_\lambda, P_\lambda),$$

$$P_\lambda = \left[Q^{-1} + \lambda H^T R^{-1} H\right]^{-1}$$
$$m_\lambda = P_\lambda \left[Q^{-1}\phi(x_{n-1}) + \lambda H^T R^{-1} y_n\right].$$

Partially Linear Gaussian Models

Any Gaussian random variable may be expressed as,

$$
\begin{aligned}
x &= m + P^{\frac{1}{2}} z \\
z &\sim \mathcal{N}\left(z \,|\, 0, I\right).
\end{aligned}
$$

Partially Linear Gaussian Models

Any Gaussian random variable may be expressed as,

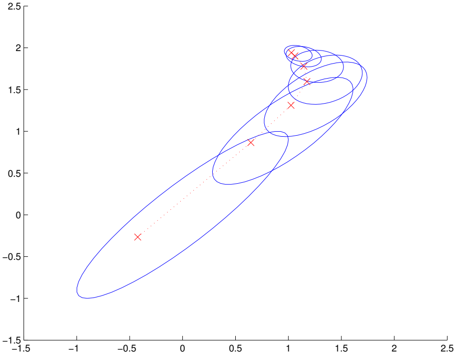
$$x = m + P^{\frac{1}{2}} z$$
$$z \sim \mathcal{N}\left(z \,|\, 0, I\right).$$

Progress particle from λ_0 to λ_1 with a linear transformation.

$$x_{\lambda_1} = m_{\lambda_1} + P_{\lambda_1}^{\frac{1}{2}} P_{\lambda_0}^{-\frac{1}{2}} (x_{\lambda_0} - m_{\lambda_0}).$$

If $x_{\lambda_0} \sim \pi_{\lambda_0}(x_{\lambda_0} | x_{n-1})$, then $x_{\lambda_1} \sim \pi_{\lambda_1}(x_{\lambda_1} | x_{n-1})$

Partially Linear Gaussian Models



Nonlinear Gaussian Models

$$f(x_n|x_{n-1}) = \mathcal{N}(x_n|\phi(x_{n-1}), Q)$$
$$g(y_n|x_n) = \mathcal{N}(y_n|\psi(x_n), R)$$

Nonlinear Gaussian Models

$$f(x_n|x_{n-1}) = \mathcal{N}\left(x_n \,|\phi(x_{n-1}), Q\right)$$
$$g(y_n|x_n) = \mathcal{N}\left(y_n \,|\psi(x_n), R\right)$$

Approximate the optimal importance density with a Gaussian at each point in pseudo-time.

$$\pi_\lambda(x_\lambda|x_{n-1}) \approx \mathcal{N}\left(x_\lambda \,|m_\lambda, P_\lambda\right)$$

Nonlinear Gaussian Models

$$f(x_n|x_{n-1}) = \mathcal{N}\left(x_n \,|\phi(x_{n-1}), Q\right)$$
$$g(y_n|x_n) = \mathcal{N}\left(y_n \,|\psi(x_n), R\right)$$

Approximate the optimal importance density with a Gaussian at each point in pseudo-time.

$$\pi_\lambda(x_\lambda|x_{n-1}) \approx \mathcal{N}\left(x_\lambda \,|m_\lambda, P_\lambda\right)$$

To update from λ_0 to λ_1,

$$\begin{aligned}\pi_{\lambda_1}(x|x_{n-1}) &\propto \pi_{\lambda_0}(x|x_{n-1})g(y_n|x)^{\lambda_1-\lambda_0} \\ &\propto \mathcal{N}\left(x\,|m_{\lambda_0}, P_{\lambda_0}\right)\mathcal{N}\left(y_n\,|\psi(x), R\right)^{\lambda_1-\lambda_0} \\ &\propto \mathcal{N}\left(x\,|m_{\lambda_0}, P_{\lambda_0}\right)\mathcal{N}\left(y_n \,\middle|\, \psi(x), \frac{R}{\lambda_1-\lambda_0}\right).\end{aligned}$$

Nonlinear Gaussian Models

Approximate with linearisation.

$$\hat{H}_{x_{\lambda_0}} = \left.\frac{\partial \psi}{\partial x}\right|_{x_{\lambda_0}}$$

Nonlinear Gaussian Models

Approximate with linearisation.

$$\hat{H}_{x_{\lambda_0}} = \left.\frac{\partial \psi}{\partial x}\right|_{x_{\lambda_0}}$$

$$\begin{aligned}
\mu_{\lambda_0} &= \psi(x_{\lambda_0}) + \hat{H}_{x_{\lambda_0}}(m_{\lambda_0} - x_{\lambda_0}) \\
\Sigma_{\lambda_0} &= \hat{H}_{x_{\lambda_0}} P_{\lambda_0} \hat{H}_{x_{\lambda_0}}^T \\
C_{\lambda_0} &= P_{\lambda_0} \hat{H}_{x_{\lambda_0}}^T \\
m_{\lambda_1} &= m_{\lambda_0} + C_{\lambda_0} \left(\Sigma_{\lambda_0} + \frac{R}{\lambda_1 - \lambda_0} \right)^{-1} (y_n - \mu_{\lambda_0}) \\
P_{\lambda_1} &= P_{\lambda_0} - C_{\lambda_0} \left(\Sigma_{\lambda_0} + \frac{R}{\lambda_1 - \lambda_0} \right)^{-1} C_{\lambda_0}^T
\end{aligned}$$

Weight Evolution

- Existing particle at pseudo time λ_0,

$$\left\{x_{1:n-1}^{(i)}, x_{\lambda_0}^{(i)}\right\} \sim \eta_{\lambda_0}(x_{1:n-1}, x_{\lambda_0}).$$

- ... is replaced by new particle at λ_1,

$$\left\{x_{1:n-1}^{(i)}, x_{\lambda_1}^{(i)}\right\} \sim \eta_{\lambda_1}(x_{1:n-1}, x_{\lambda_1}).$$

Weight Evolution

- Existing particle at pseudo time λ_0,

$$\left\{x_{1:n-1}^{(i)}, x_{\lambda_0}^{(i)}\right\} \sim \eta_{\lambda_0}(x_{1:n-1}, x_{\lambda_0}).$$

- ... is replaced by new particle at λ_1,

$$\left\{x_{1:n-1}^{(i)}, x_{\lambda_1}^{(i)}\right\} \sim \eta_{\lambda_1}(x_{1:n-1}, x_{\lambda_1}).$$

- Standard change of variables formula,

$$\eta_{\lambda_1}(x_{1:n-1}, x_{\lambda_1}) = \eta_{\lambda_0}(x_{1:n-1}, x_{\lambda_0}) \times \left|\frac{\partial x_{\lambda_0}}{\partial x_{\lambda_1}}\right|.$$

Weight Evolution

Hence weight update,

$$\begin{aligned} w_{\lambda_1} &= \frac{\tilde{\pi}_{\lambda_1}(x_{1:n-1}, x_{\lambda_1})}{\eta_{\lambda_1}(x_{1:n-1}, x_{\lambda_1})} \\ &= \frac{\tilde{\pi}_{\lambda_0}(x_{1:n-1}, x_{\lambda_0})}{\eta_{\lambda_0}(x_{1:n-1}, x_{\lambda_0})} \times \frac{\tilde{\pi}_{\lambda_1}(x_{1:n-1}, x_{\lambda_1})}{\tilde{\pi}_{\lambda_0}(x_{1:n-1}, x_{\lambda_0})} \times \left| \frac{\partial x_{\lambda_1}}{\partial x_{\lambda_0}} \right| \\ &\propto w_{\lambda_0} \times \frac{g(y_n|x_{\lambda_1})^{\lambda_1} f(x_{\lambda_1}|x_{n-1})}{g(y_n|x_{\lambda_0})^{\lambda_0} f(x_{\lambda_0}|x_{n-1})} \times \sqrt{\frac{|P_{\lambda_1}|}{|P_{\lambda_0}|}}. \end{aligned}$$

Relationship to Other Methods

Gradual introduction of likelihood underlies numerous existing particle methods,

- Simon Godsill and Tim Clapp. *Improvement strategies for Monte Carlo particle filters*, pages 139–158. Citeseer, 2001
- Jürgen Gall, Jürgen Potthoff, Christoph Schnörr, Bodo Rosenhahn, and Hans-Peter Seidel. Interacting and annealing particle filters: Mathematics and a recipe for applications. *Journal of Mathematical Imaging and Vision*, 28(1):1–18, 2007
- J. Deutscher, A. Blake, and I. Reid. Articulated body motion capture by annealed particle filtering. In *Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on*, volume 2, pages 126–133 vol.2, 2000
- N. Oudjane and C. Musso. Progressive correction for regularized particle filters. In *Information Fusion, 2000. FUSION 2000. Proceedings of the Third International Conference on*, volume 2, pages THB2/10–THB2/17 vol.2, 2000
- F. Daum and J. Huang. Particle flow for nonlinear filters with log-homotopy. In *Proceedings of SPIE, the International Society for Optical Engineering*, pages 696918–1. Society of Photo-Optical Instrumentation Engineers, 2008
- Sebastian Reich. A dynamical systems framework for intermittent data assimilation. *BIT Numerical Mathematics*, 51:235–249, 2011

Relationship to Other Methods

Distinguishing features of the new progressive proposal method:

- Particles moved deterministically.
- Move one particle at a time — no intermediate interaction steps needed.
- Adaptive step-size control for *each particle*.

Simulations - A Tracking Problem

$$x_n = \begin{bmatrix} p_n^T & v_n^T \end{bmatrix}^T$$

Near-constant velocity model.

Simulations - A Tracking Problem

$$x_n = \begin{bmatrix} p_n^T & v_n^T \end{bmatrix}^T$$

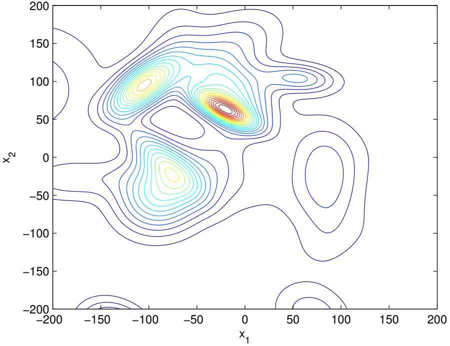
Near-constant velocity model.

$$y_n = \begin{bmatrix} \theta_n & r_n & h_n & s_n \end{bmatrix}^T .$$

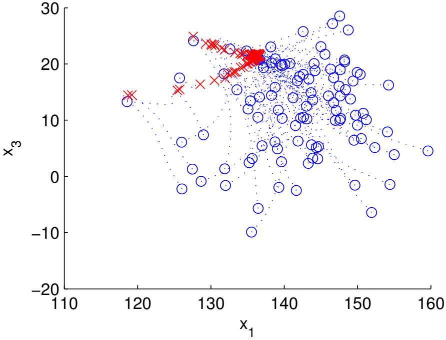
$$\begin{aligned}
\theta_n &= \arctan\left(\frac{p_{n,1}}{p_{n,2}}\right) \\
r_n &= \sqrt{p_{n,1}^2 + p_{n,3}^2 + p_{n,3}^2} \\
h_n &= p_{n,3} - T(p_{n,1}, p_{n,2}) \\
s_n &= \frac{p_n \cdot v_n}{r_n},
\end{aligned}$$

Simulations - Terrain Map

$$T(p_{n,1}, p_{n,2})$$



Simulations - Results



Simulations - Results

Algorithm	N_F	ESS	RMSE
Bootstrap Proposal	6000	1.0	78.6
Unscented Kalman Proposal	460	2.4	70.2
Gaussian Local Maximum Proposal	10	3.1	62.9
Progressive Proposal	180	56.4	22.3

Summary

- The progressive proposal method samples from an effective approximation of the optimal importance density.
- Particles sampled from the transition density, then moved deterministically with a series of approximately optimal steps.
- Lower errors and better sample sizes when filtering with challenging nonlinear models.