Gaussian Approximations of the Optimal Proposal in Particle Filtering

Particle Filtering in Dynamical Systems

Sai Sampath Kedari

sampath@umich.edu

Contents

1 Gaussian Approximation of the Predictive Measurement Distribution

In the optimal proposal particle filter, updating the particle weights requires the predictive likelihood

$$\mathbb{P}(Y_k \mid x_{k-1}^{(i)}),$$

while sampling requires the conditional distribution

$$\mathbb{P}(X_k \mid x_{k-1}^{(i)}, Y_k).$$

Throughout this section, the particle index i is fixed and $x_{k-1}^{(i)}$ is treated as deterministic and known.

Assume Gaussian process noise and Gaussian measurement noise, yielding the exact conditional models

$$X_k \mid x_{k-1}^{(i)} \sim \mathcal{N}\big(\Phi(x_{k-1}^{(i)}), \Sigma\big),$$
$$Y_k \mid X_k \sim \mathcal{N}\big(h(X_k), \Gamma\big).$$

Our objective is to characterize the predictive measurement distribution $\mathbb{P}(Y_k \mid x_{k-1}^{(i)})$.

1.1 Predictive Measurement Distribution

By marginalization,

$$\mathbb{P}(Y_k \mid x_{k-1}^{(i)}) = \int \mathbb{P}(Y_k, X_k \mid x_{k-1}^{(i)}) \, dX_k.$$

Using the conditional independence implied by the state-space model,

$$\mathbb{P}(Y_k, X_k \mid x_{k-1}^{(i)}) = \mathbb{P}(Y_k \mid X_k) \, \mathbb{P}(X_k \mid x_{k-1}^{(i)}),$$

so that

$$\mathbb{P}(Y_k \mid x_{k-1}^{(i)}) = \int \mathbb{P}(Y_k \mid X_k) \, \mathbb{P}(X_k \mid x_{k-1}^{(i)}) \, dX_k.$$

Since $h(\cdot)$ is nonlinear, this distribution is generally non-Gaussian. We therefore introduce a Gaussian approximation by matching the first two moments:

$$Y_k \mid x_{k-1}^{(i)} \approx \mathcal{N}(\mu_k, S_k).$$

1.1.1 Predictive Measurement Mean

By definition,

$$\mu_k \equiv \mathbb{E}[Y_k \mid x_{k-1}^{(i)}] = \int y_k \, \mathbb{P}(y_k \mid x_{k-1}^{(i)}) \, dy_k.$$

Substituting the marginalization expression and interchanging integrals,

$$\mu_k = \int \left[\int y_k \, \mathbb{P}(y_k \mid x_k) \, dy_k \right] \mathbb{P}(x_k \mid x_{k-1}^{(i)}) \, dx_k.$$

Since $Y_k \mid x_k \sim \mathcal{N}(h(x_k), \Gamma)$,

$$\int y_k \, \mathbb{P}(y_k \mid x_k) \, dy_k = \mathbb{E}[Y_k \mid x_k] = h(x_k).$$

Thus,

$$\mu_k = \int h(x_k) \, \mathbb{P}(x_k \mid x_{k-1}^{(i)}) \, dx_k.$$

With

$$X_k \mid x_{k-1}^{(i)} \sim \mathcal{N}(m_k^-, C_k^-), \qquad m_k^- = \Phi(x_{k-1}^{(i)}), \quad C_k^- = \Sigma,$$

we obtain

$$\boxed{\mu_k = \int h(x_k) \, \mathcal{N}(x_k; m_k^-, C_k^-) \, dx_k.}$$

1.1.2   Predictive Measurement Covariance

Define

$$S_k \;\equiv\; \mathrm{Cov}(Y_k \mid x_{k-1}^{(i)}) = \mathbb{E}\Big[(Y_k - \mu_k)(Y_k - \mu_k)^T \mid x_{k-1}^{(i)}\Big].$$

Using $Y_k = h(X_k) + \eta_k$,

$$Y_k - \mu_k = \big(h(X_k) - \mu_k\big) + \eta_k.$$

Expanding and taking conditional expectation given X_k, with $\mathbb{E}[\eta_k \mid X_k] = 0$ and $\mathbb{E}[\eta_k \eta_k^T \mid X_k] = \Gamma$, yields

$$\mathbb{E}\big[(Y_k - \mu_k)(Y_k - \mu_k)^T \mid X_k\big] = \big(h(X_k) - \mu_k\big)\big(h(X_k) - \mu_k\big)^T + \Gamma.$$

Taking expectation with respect to $\mathbb{P}(X_k \mid x_{k-1}^{(i)})$ gives

$$S_k = \mathbb{E}\Big[\big(h(X_k) - \mu_k\big)\big(h(X_k) - \mu_k\big)^T \mid x_{k-1}^{(i)}\Big] + \Gamma.$$

Using the Gaussian approximation for $X_k \mid x_{k-1}^{(i)}$,

$$\boxed{S_k = \int \big(h(x_k) - \mu_k\big)\big(h(x_k) - \mu_k\big)^T \, \mathcal{N}(x_k; m_k^-, C_k^-) \, dx_k + \Gamma.}$$

2   Gaussian Approximation of the Conditional State Distribution

We now derive a Gaussian approximation of the conditional distribution

$$\mathbb{P}(X_k \mid x_{k-1}^{(i)}, y_k),$$

which is required for sampling under the optimal proposal.

Throughout this section, the particle state $x_{k-1}^{(i)}$ is fixed and deterministic.

Assume Gaussian process noise and Gaussian measurement noise, yielding the exact conditional models

$$X_k \mid x_{k-1}^{(i)} \sim \mathcal{N}(m_k^-, C_k^-),$$
$$Y_k \mid X_k \sim \mathcal{N}(h(X_k), \Gamma),$$

with

$$m_k^- = \Phi(x_{k-1}^{(i)}), \qquad C_k^- = \Sigma.$$

The Gaussian approximation arises when conditioning on the measurement $Y_k = y_k$.

2.1 Joint Distribution of State and Measurement

Define the stacked random vector

$$Z_k = \begin{bmatrix} X_k \\ Y_k \end{bmatrix}.$$

2.1.1 Joint Mean

From the measurement model,

$$\mathbb{E}[Y_k \mid X_k] = h(X_k).$$

Applying the law of total expectation,

$$\mathbb{E}[Y_k \mid x_{k-1}^{(i)}] = \mathbb{E}_{X_k \mid x_{k-1}^{(i)}}[h(X_k)].$$

Define

$$\mu_k \equiv \mathbb{E}[Y_k \mid x_{k-1}^{(i)}] = \int h(x_k)\,\mathcal{N}(x_k; m_k^-, C_k^-)\,dx_k.$$

The joint mean is therefore

$$\mathbb{E}[Z_k \mid x_{k-1}^{(i)}] = \begin{bmatrix} m_k^- \\ \mu_k \end{bmatrix}.$$

2.1.2 Joint Covariance

The joint covariance admits the block structure

$$\mathrm{Cov}(Z_k \mid x_{k-1}^{(i)}) = \begin{bmatrix} C_k^- & U_k \\ U_k^T & S_k \end{bmatrix},$$

with blocks defined below.

Cross-Covariance Using covariance algebra and independence of X_k and η_k,

$$U_k \equiv \mathrm{Cov}(X_k, Y_k \mid x_{k-1}^{(i)}) = \mathrm{Cov}(X_k, h(X_k) \mid x_{k-1}^{(i)}).$$

Thus,

$$\boxed{U_k = \int (x_k - m_k^-)\big(h(x_k) - \mu_k\big)^T \mathcal{N}(x_k; m_k^-, C_k^-)\,dx_k.}$$

Measurement Covariance   By the law of total variance,

$$\text{Var}(Y_k) = \mathbb{E}[\text{Var}(Y_k \mid X_k)] + \text{Var}(\mathbb{E}[Y_k \mid X_k]).$$

Since $\text{Var}(Y_k \mid X_k) = \Gamma$ and $\mathbb{E}[Y_k \mid X_k] = h(X_k)$, define

$$\boxed{S_k = \int \big(h(x_k) - \mu_k\big)\big(h(x_k) - \mu_k\big)^T \mathcal{N}(x_k; m_k^-, C_k^-)\, dx_k + \Gamma.}$$

2.2   Conditioning on the Measurement

The joint Gaussian distribution of (X_k, Y_k) conditioned on $x_{k-1}^{(i)}$ is

$$\begin{bmatrix} X_k \\ Y_k \end{bmatrix} \mid x_{k-1}^{(i)} \sim \mathcal{N}\left(\begin{bmatrix} m_k^- \\ \mu_k \end{bmatrix}, \begin{bmatrix} C_k^- & U_k \\ U_k^T & S_k \end{bmatrix} \right).$$

Conditioning on the observed measurement $Y_k = y_k$ yields

$$X_k \mid x_{k-1}^{(i)}, y_k \sim \mathcal{N}\left(m_k^- + U_k S_k^{-1}(y_k - \mu_k),\ C_k^- - U_k S_k^{-1} U_k^T \right).$$

Define

$$\boxed{m_k^{(i)} = m_k^- + U_k S_k^{-1}(y_k - \mu_k),} \qquad \boxed{C_k^{(i)} = C_k^- - U_k S_k^{-1} U_k^T.}$$

These moments define the Gaussian approximation of $\mathbb{P}(X_k \mid x_{k-1}^{(i)}, y_k)$ used for sampling under the optimal proposal.

3   Gaussian-Approximated Optimal Proposal: Complete Recursion

For each particle $x_{k-1}^{(i)}$, assume Gaussian process and measurement noise, and define the conditional prior

$$X_k \mid x_{k-1}^{(i)} \sim \mathcal{N}(m_k^-, C_k^-), \qquad m_k^- = \Phi(x_{k-1}^{(i)}), \quad C_k^- = \Sigma.$$

Predictive Measurement Moments

Define the predictive measurement mean

$$\mu_k = \int h(x_k)\, \mathcal{N}(x_k; m_k^-, C_k^-)\, dx_k,$$

and predictive measurement covariance

$$S_k = \int \big(h(x_k) - \mu_k\big)\big(h(x_k) - \mu_k\big)^T \mathcal{N}(x_k; m_k^-, C_k^-)\, dx_k + \Gamma.$$

Cross-Covariance

Define the state–measurement cross-covariance

$$U_k = \int (x_k - m_k^-)\big(h(x_k) - \mu_k\big)^T \mathcal{N}(x_k; m_k^-, C_k^-)\, dx_k.$$

Gaussian-Approximated Optimal Proposal

The Gaussian approximation of the optimal proposal is

$$X_k \mid x_{k-1}^{(i)}, y_k \sim \mathcal{N}(m_k^{(i)}, C_k^{(i)}),$$

with

$$m_k^{(i)} = m_k^- + U_k S_k^{-1}(y_k - \mu_k), \qquad C_k^{(i)} = C_k^- - U_k S_k^{-1} U_k^T.$$

Weight Update

The importance weights are updated according to

$$w_k^{(i)} = \bar{w}_{k-1}^{(i)} \, \mathbb{P}(y_k \mid x_{k-1}^{(i)}),$$

with the Gaussian approximation

$$\mathbb{P}(y_k \mid x_{k-1}^{(i)}) \approx \mathcal{N}(y_k; \mu_k, S_k).$$