Gaussian filters and moment matching

Sensor fusion & nonlinear filtering

Lars Hammarstrand

GAUSSIAN FILTERING

- We mainly consider models of the type

$$
\begin{aligned}
\mathbf{x}_k &= \mathbf{f}_{k-1}(\mathbf{x}_{k-1}) + \mathbf{q}_{k-1}, & \mathbf{q}_{k-1} &\sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_{k-1}) \\
\mathbf{y}_k &= \mathbf{h}_k(\mathbf{x}_k) + \mathbf{r}_k, & \mathbf{r}_k &\sim \mathcal{N}(\mathbf{0}, \mathbf{R}_k).
\end{aligned}
$$

Gaussian filtering solution

- Approximate distributions as Gaussian in the prediction and update steps:

Prediction:

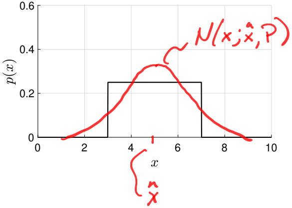
$$
\begin{aligned}
p(\mathbf{x}_k \mid \mathbf{y}_{1:k-1}) &= \int p(\mathbf{x}_k \mid \mathbf{x}_{k-1}) p(\mathbf{x}_{k-1} \mid \mathbf{y}_{1:k-1})\, d\mathbf{x}_{k-1} \\
&\approx \mathcal{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k-1}, \mathbf{P}_{k|k-1})
\end{aligned}
$$

Update:

$$
\begin{aligned}
p(\mathbf{x}_k \mid \mathbf{y}_{1:k}) &= \frac{p(\mathbf{y}_k \mid \mathbf{x}_k) p(\mathbf{x}_k \mid \mathbf{y}_{1:k-1})}{\int p(\mathbf{y}_k \mid \mathbf{x}_k) p(\mathbf{x}_k \mid \mathbf{y}_{1:k-1})\, d\mathbf{x}_k} \\
&\approx \mathcal{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k}, \mathbf{P}_{k|k})
\end{aligned}
$$

GAUSSIAN APPROXIMATIONS BY MOMENT MATCHING

- Suppose that we are given a non-Gaussian density $p(\mathbf{x})$.
- Task: find $\hat{\mathbf{x}}$ and $\mathbf{P}$ such that $p(\mathbf{x}) \approx \mathcal{N}(\mathbf{x}; \hat{\mathbf{x}}, \mathbf{P})$.



Moment matching

- One strategy is to select $\hat{\mathbf{x}}$ and $\mathbf{P}$ to match the moments of $p(\mathbf{x})$:

$$\hat{\mathbf{x}} = \mathbb{E}_{p(\mathbf{x})}\left\{\mathbf{x}\right\} = \int \mathbf{x} p(\mathbf{x})\, d\mathbf{x}$$

$$\mathbf{P} = \mathrm{Cov}_{p(\mathbf{x})}\{\mathbf{x}\} = \int (\mathbf{x} - \hat{\mathbf{x}})(\mathbf{x} - \hat{\mathbf{x}})^T p(\mathbf{x})\, d\mathbf{x}$$

- One can show that moment matching minimizes the Kullback-Leibler divergence

$$\mathrm{KL}\left(p(\mathbf{x}) \middle| \mathcal{N}(\mathbf{x}; \hat{\mathbf{x}}, \mathbf{P})\right) = \int p(\mathbf{x}) \log \frac{p(\mathbf{x})}{\mathcal{N}(\mathbf{x}; \hat{\mathbf{x}}, \mathbf{P})}\, d\mathbf{x}.$$

GAUSSIAN PREDICTION BY MOMENT MATCHING

- Given

$$\begin{cases} \mathbf{x}_{k-1} \big| \mathbf{y}_{1:k-1} \sim \mathcal{N}(\hat{\mathbf{x}}_{k-1|k-1}, \mathbf{P}_{k-1|k-1}) \\ \mathbf{x}_k = f(\mathbf{x}_{k-1}) + \mathbf{q}_{k-1} \end{cases}$$

$\Rightarrow p(x_k | y_{1:k-1}) \approx \mathcal{N}(x_k; \hat{x}_{k|k-1}, P_{k|k-1})$

the first two predicted moments of $\mathbf{x}_k$ are

Prediction by moment matching

$$\begin{aligned} \hat{\mathbf{x}}_{k|k-1} &= \mathbb{E}\{\mathbf{x}_k \big| \mathbf{y}_{1:k-1}\} \\ &= \int f(\mathbf{x}_{k-1}) \mathcal{N}(\mathbf{x}_{k-1}; \hat{\mathbf{x}}_{k-1|k-1}, \mathbf{P}_{k-1|k-1})\, d\mathbf{x}_{k-1} \end{aligned}$$

$$\mathbf{P}_{k|k-1} = \mathrm{Cov}\{\mathbf{x}_k \big| \mathbf{y}_{1:k-1}\} = \mathbf{Q}_{k-1} +$$

$$\int (f(\mathbf{x}_{k-1}) - \hat{\mathbf{x}}_{k|k-1})(\cdot)^T \mathcal{N}(\mathbf{x}_{k-1}; \hat{\mathbf{x}}_{k-1|k-1}, \mathbf{P}_{k-1|k-1})\, d\mathbf{x}_{k-1}$$

GAUSSIAN UPDATE BY MOMENT MATCHING

- We are given

$$\begin{cases} \mathbf{x}_k \big| \mathbf{y}_{1:k-1} \sim \mathcal{N}(\hat{\mathbf{x}}_{k|k-1}, \mathbf{P}_{k|k-1}) \\ \mathbf{y}_k = h(\mathbf{x}_k) + \mathbf{r}_k. \end{cases}$$

- Ideal solution: set $\hat{\mathbf{x}}_{k|k}$ and $\mathbf{P}_{k|k}$ to the first two moments of

$$p(\mathbf{x}_k \big| \mathbf{y}_{1:k}) \propto \mathcal{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k-1}, \mathbf{P}_{k|k-1}) p(\mathbf{y}_k \big| \mathbf{x}_k).$$

- Unfortunately, it is difficult to efficiently compute the moments of $p(\mathbf{x}_k \big| \mathbf{y}_{1:k})$.

Alternative moment matching strategy

- Approximate $p(\mathbf{x}_k, \mathbf{y}_k \big| \mathbf{y}_{1:k-1})$ as Gaussian using moment matching.

  $\rightsquigarrow$ When $\mathbf{x}_k, \mathbf{y}_k \big| \mathbf{y}_{1:k-1}$ is Gaussian, we can find $p(\mathbf{x}_k \big| \mathbf{y}_k, \mathbf{y}_{1:k-1})$ analytically.

GAUSSIAN UPDATE BY MOMENT MATCHING

- We approximate $(\mathbf{x}_k, \mathbf{y}_k)$ as jointly Gaussian using moment matching:

$$\begin{bmatrix} \mathbf{x}_k \\ \mathbf{y}_k \end{bmatrix} \Bigg| \mathbf{y}_{1:k-1} \sim \mathcal{N}\left(\begin{bmatrix} \hat{\mathbf{x}}_{k|k-1} \\ \hat{\mathbf{y}}_{k|k-1} \end{bmatrix}, \begin{bmatrix} \mathbf{P}_{k|k-1} & \mathbf{P}_{xy} \\ \mathbf{P}_{yx} & \mathbf{S}_k \end{bmatrix} \right)$$

$\hookrightarrow h(x_k) + r_k$

$= P_{xy}^T$

where

$$\hat{\mathbf{y}}_{k|k-1} = \mathbb{E}\{\mathbf{y}_k | \mathbf{y}_{1:k-1}\} = \mathbb{E}\{h(\mathbf{x}_k) | \mathbf{y}_{1:k-1}\} = \int h(\mathbf{x}_k) \mathcal{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k-1}, \mathbf{P}_{k|k-1})\, d\mathbf{x}_k$$

$$\begin{aligned} \mathbf{P}_{xy} &= \mathbb{E}\{(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})(\mathbf{y}_k - \hat{\mathbf{y}}_{k|k-1})^T | \mathbf{y}_{1:k-1}\} \\ &= \int (\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})(h(\mathbf{x}_k) - \hat{\mathbf{y}}_{k|k-1})^T \mathcal{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k-1}, \mathbf{P}_{k|k-1})\, d\mathbf{x}_k \end{aligned}$$

$$\begin{aligned} \mathbf{S}_k &= \mathrm{Cov}\{\mathbf{y}_k | \mathbf{y}_{1:k-1}\} = \mathbf{R}_k + \mathrm{Cov}\{h(\mathbf{x}_k) | \mathbf{y}_{1:k-1}\} \\ &= \mathbf{R}_k + \int (h(\mathbf{x}_k) - \hat{\mathbf{y}}_{k|k-1})(\cdot)^T \mathcal{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k-1}, \mathbf{P}_{k|k-1})\, d\mathbf{x}_k \end{aligned}$$

GAUSSIAN UPDATE BY MOMENT MATCHING

- If we know that

$$\begin{bmatrix} \mathbf{x}_k \\ \mathbf{y}_k \end{bmatrix} \Bigg| \mathbf{y}_{1:k-1} \sim \mathcal{N}\left(\begin{bmatrix} \hat{\mathbf{x}}_{k|k-1} \\ \hat{\mathbf{y}}_{k|k-1} \end{bmatrix}, \begin{bmatrix} \mathbf{P}_{k|k-1} & \mathbf{P}_{xy} \\ \mathbf{P}_{yx} & \mathbf{S}_k \end{bmatrix} \right)$$

  then it holds that

$$p(x_k|y_k, y_{1:k-1}): \left\{ \begin{aligned} \hat{\mathbf{x}}_{k|k} &= \hat{\mathbf{x}}_{k|k-1} + \overbrace{\mathbf{P}_{xy}\mathbf{S}_k^{-1}}^{K_k} \overbrace{(\mathbf{y}_k - \hat{\mathbf{y}}_{k|k-1})}^{v_k} \\ \mathbf{P}_{k|k} &= \mathbf{P}_{k|k-1} - \underbrace{\mathbf{P}_{xy}\mathbf{S}_k^{-1}\mathbf{P}_{xy}^T}_{K_k \cdot S_k K_k^T}. \end{aligned} \right.$$

- Key difference compared to the Kalman filter: we need to *approximate* $\hat{\mathbf{y}}_{k|k-1}$, $\mathbf{S}_k$ and $\mathbf{P}_{xy}$.
- Note: we find these components by solving integrals of the type $\int g(\mathbf{x})\mathcal{N}(\mathbf{x}; \hat{\mathbf{x}}, \mathbf{P})\, d\mathbf{x}$.

SELF ASSESSMENT

In Gaussian filtering using moment matching we need to solve several different integrals (two during the prediction and three during the update step). Among the following integrals, which ones do we need to solve?

- To compute $\mathbf{P}_{xy}$ in the update step: $\int f(\mathbf{x}_k)h(\mathbf{x}_k)\, d\mathbf{x}_k$.
- To compute $\mathbf{P}_{k|k-1}$ in the prediction step we take $\mathbf{Q}_{k-1}$ plus this integral: $\int (f(\mathbf{x}_{k-1}) - \hat{\mathbf{x}}_{k|k-1})(\cdot)^T \mathcal{N}(\mathbf{x}_{k-1}; \hat{\mathbf{x}}_{k-1|k-1}, \mathbf{P}_{k-1|k-1})\, d\mathbf{x}_{k-1}$
- To compute $\hat{\mathbf{y}}_{k|k-1}$ in the update step: $\int h(\mathbf{x}_k)\mathcal{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k-1}, \mathbf{P}_{k|k-1})\, d\mathbf{x}_k$
- To compute $\mathbf{P}_{k|k}$ in the update step: $\int (\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})(\cdot)^T \mathcal{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k}, \mathbf{P}_{k|k})\, d\mathbf{x}_k$