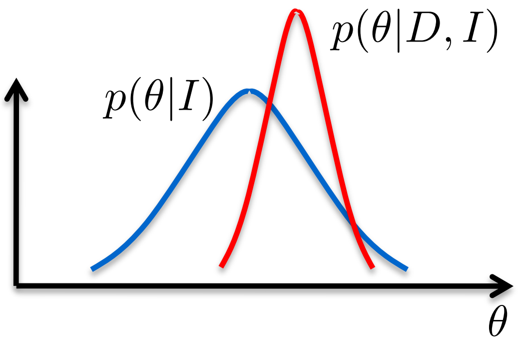
Kalman filter extensions



Dr. Chris Renton
School of Engineering

THE UNIVERSITY OF NEWCASTLE AUSTRALIA
FACULTY OF ENGINEERING AND BUILT ENVIRONMENT
www.newcastle.edu.au

Extensions to nonlinear Kalman filtering

Recall the model we've used so far,

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

We consider the following extensions by transforming back to the above form:

- Asynchronous measurement updates
- Correlated noise
- Non-additive noise
- Uncertain inputs
- Delayed/out-of-order measurements
- Merging multiple estimators

These extensions can be used in any combination.

Asynchronous measurement updates

Non-uniform measurement periods

- Sensors with nonuniform sampling rates (e.g., encoders)
- Fuse sensors with different uniform sampling rates (e.g., GPS + IMU)
- Need state estimates between measurement event times (e.g., control loop)

To handle these cases

- Decouple time update (prediction) and measurement update (correction) steps, so that these steps may be applied in any order.
- Model the problem in continuous time and integrate over the sampling period T_k at each time step using Euler integration or an ODE solver. Be sure to discretise the noise appropriately.

$$\mathbf{x}_k = \mathbf{x}[k] = \mathbf{x}(t_k), \qquad t_k = t_0 + \sum_{j=1}^{k} T_j, \quad k = 1, 2, \ldots$$

Every time a measurement is received, predict forward to the time stamp of the measurement and then do the measurement correction.

To obtain an estimate of the state between measurements, predict forward to the desired time from the most recent measurement.

Correlated noise

Consider the case where we have correlated process and measurement noise,

$$\begin{aligned} \mathbf{x}_{k+1} &= \mathbf{f}_k(\mathbf{x}_k) + \mathbf{w}_k, \\ \mathbf{y}_k &= \mathbf{h}_k(\mathbf{x}_k) + \mathbf{v}_k, \end{aligned} \qquad \begin{bmatrix} \mathbf{w}_k \\ \mathbf{v}_k \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \mathbf{Q}_k & \mathbf{S}_k \\ \mathbf{S}_k^\mathsf{T} & \mathbf{R}_k \end{bmatrix} \right)$$

We can write the joint future state and current measurement density as

$$p(\mathbf{x}_{k+1}, \mathbf{y}_k | \mathbf{x}_k) = \mathcal{N}\left(\begin{bmatrix} \mathbf{x}_{k+1} \\ \mathbf{y}_k \end{bmatrix}; \begin{bmatrix} \mathbf{f}_k(\mathbf{x}_k) \\ \mathbf{h}_k(\mathbf{x}_k) \end{bmatrix}, \begin{bmatrix} \mathbf{Q}_k & \mathbf{S}_k \\ \mathbf{S}_k^\mathsf{T} & \mathbf{R}_k \end{bmatrix} \right)$$

The state transition likelihood is given by conditioning on $\mathbf{y}_k$

$$p(\mathbf{x}_{k+1} | \mathbf{x}_k, \mathbf{y}_k) = \mathcal{N}\Big(\mathbf{x}_{k+1}; \underbrace{\mathbf{f}_k(\mathbf{x}_k) + \mathbf{S}_k \mathbf{R}_k^{-1}(\mathbf{y}_k - \mathbf{h}_k(\mathbf{x}_k))}_{\tilde{\mathbf{f}}_k(\mathbf{x}_k)}, \underbrace{\mathbf{Q}_k - \mathbf{S}_k \mathbf{R}_k^{-1} \mathbf{S}_k^\mathsf{T}}_{\tilde{\mathbf{Q}}_k}\Big)$$

The state transition likelihood now depends on $\mathbf{y}_k$, which we treat like an input

$$\begin{aligned} \mathbf{x}_{k+1} &= \tilde{\mathbf{f}}_k(\mathbf{x}_k) + \tilde{\mathbf{w}}_k, \qquad & \tilde{\mathbf{w}}_k &\sim \mathcal{N}(\mathbf{0}, \tilde{\mathbf{Q}}_k) \\ \mathbf{y}_k &= \mathbf{h}_k(\mathbf{x}_k) + \mathbf{v}_k, \qquad & \mathbf{v}_k &\sim \mathcal{N}(\mathbf{0}, \mathbf{R}_k) \end{aligned}$$

Non-additive noise

Consider the case where the process or measurement noise is non-additive, and we may have correlated process and measurement noise,

$$\begin{aligned} \mathbf{x}_{k+1} &= \mathbf{f}_k(\mathbf{x}_k, \mathbf{w}_k), \\ \mathbf{y}_k &= \mathbf{h}_k(\mathbf{x}_k, \mathbf{v}_k), \end{aligned} \qquad \begin{bmatrix} \mathbf{w}_k \\ \mathbf{v}_k \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \mathbf{Q}_k & \mathbf{S}_k \\ \mathbf{S}_k^\mathsf{T} & \mathbf{R}_k \end{bmatrix} \right)$$

By augmenting the state *with the noise*, we can transform this system into one that has only additive noise,

$$\begin{aligned} \begin{bmatrix} \mathbf{x}_{k+1} \\ \mathbf{d}_{k+1} \\ \mathbf{e}_{k+1} \end{bmatrix} &= \begin{bmatrix} \mathbf{f}_k(\mathbf{x}_k, \mathbf{d}_k) \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \mathbf{w}_{k+1} \\ \mathbf{v}_{k+1} \end{bmatrix}, \qquad & \mathcal{Q} &= \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_{k+1} & \mathbf{S}_{k+1} \\ \mathbf{0} & \mathbf{S}_{k+1}^\mathsf{T} & \mathbf{R}_{k+1} \end{bmatrix}, \\ \mathbf{y}_k &= \mathbf{h}(\mathbf{x}_k, \mathbf{e}_k), & \mathcal{R} &= \mathbf{0}. \end{aligned}$$

This assumes we can compute the model noise covariance one-step ahead and we treat the zero measurement noise above with care.

Uncertain inputs

We can model input uncertainty as noise,

$$
\begin{aligned}
\mathbf{x}_{k+1} &= \mathbf{f}(\mathbf{x}_k, \mathbf{u}_k - \mathbf{e}_k) + \mathbf{w}_k, \qquad & \mathbf{e}_k &\sim \mathcal{N}(\mathbf{0}, \mathbf{E}_k), \quad \mathbf{w}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_k), \\
\mathbf{y}_k &= \mathbf{h}(\mathbf{x}_k, \mathbf{u}_k - \mathbf{e}_k) + \mathbf{v}_k, \qquad & \mathbf{v}_k &\sim \mathcal{N}(\mathbf{0}, \mathbf{R}_k).
\end{aligned}
$$

By augmenting the state *with the noise*, we can transform this system into one that has only additive noise,

$$
\begin{aligned}
\begin{bmatrix} \mathbf{x}_{k+1} \\ \mathbf{d}_{k+1} \end{bmatrix} &= \begin{bmatrix} \mathbf{f}(\mathbf{x}_k, \mathbf{u}_k - \mathbf{d}_k) \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{w}_k \\ \mathbf{e}_{k+1} \end{bmatrix}, \qquad & \mathcal{Q} &= \begin{bmatrix} \mathbf{Q}_k & \mathbf{0} \\ \mathbf{0} & \mathbf{E}_{k+1} \end{bmatrix}, \\
\mathbf{y}_k &= \mathbf{h}(\mathbf{x}_k, \mathbf{u}_k - \mathbf{d}_k) + \mathbf{v}_k, \qquad & \mathcal{R} &= \mathbf{R}_k.
\end{aligned}
$$

This assumes we can compute the input noise covariance one-step ahead.

Delayed/out-of-order measurements

We can account for delayed measurements up to N steps in the past by augmenting the state with delayed versions of itself.

Let $\mathbf{z}_{n,k} \triangleq \mathbf{x}_{k-n}$ for $0 \leq n \leq N$. Then,

$$\begin{bmatrix} \mathbf{z}_{0,k+1} \\ \mathbf{z}_{1,k+1} \\ \mathbf{z}_{2,k+1} \\ \vdots \\ \mathbf{z}_{N,k+1} \end{bmatrix} = \begin{bmatrix} \mathbf{f}(\mathbf{z}_{0,k}) \\ \mathbf{z}_{0,k} \\ \mathbf{z}_{1,k} \\ \vdots \\ \mathbf{z}_{N-1,k} \end{bmatrix} + \begin{bmatrix} \mathbf{w}_k \\ \mathbf{0} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix}, \qquad \mathcal{Q} = \begin{bmatrix} \mathbf{Q}_k & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \end{bmatrix},$$

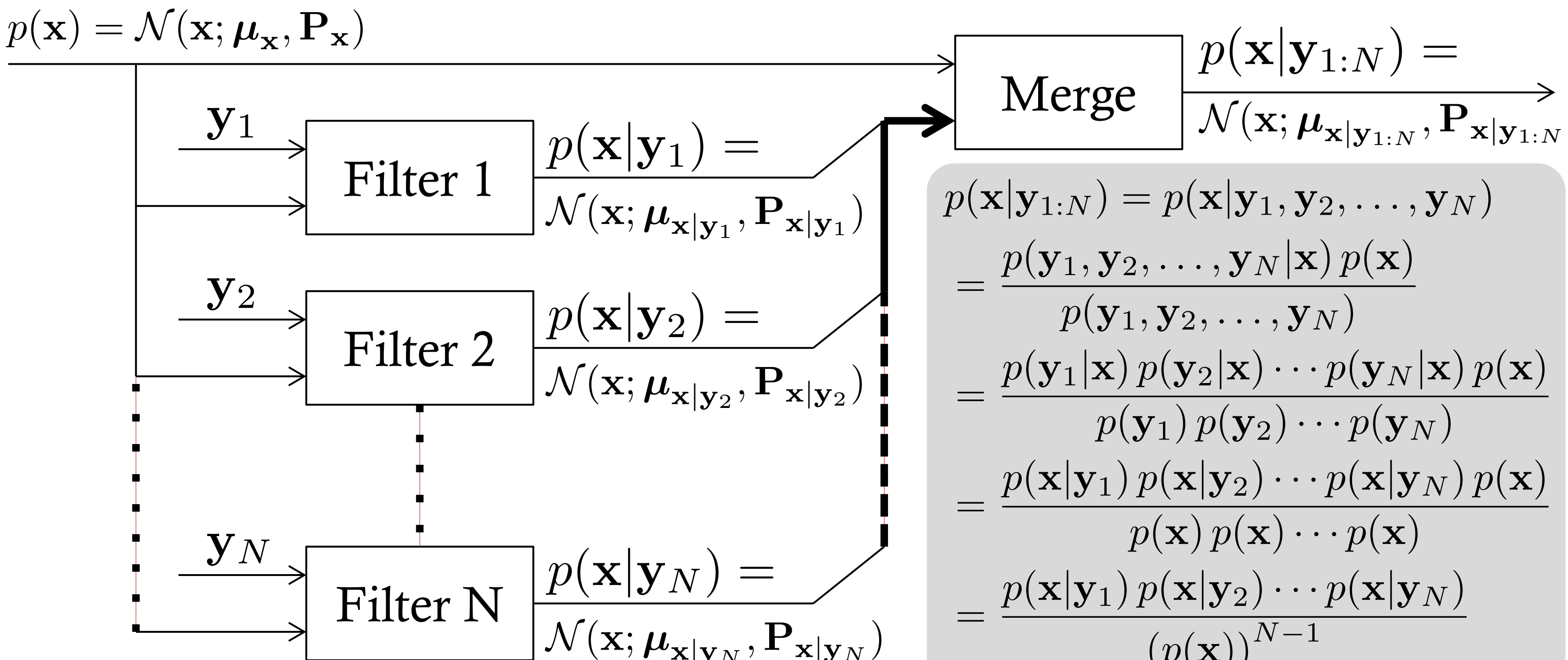
$$\mathbf{y}_{k-n} = \mathbf{h}_{k-n}(\mathbf{z}_{n,k}) + \mathbf{v}_{k-n}, \qquad \mathcal{R} = \mathbf{R}_{k-n}.$$

The measurement correction now correctly propagates old information at step k–n into the current time.

THE UNIVERSITY OF NEWCASTLE AUSTRALIA

Merging multiple estimators

Consider case of distributed estimators using independent measurements



$$
\begin{aligned}
p(\mathbf{x}|\mathbf{y}_{1:N}) &= p(\mathbf{x}|\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_N) \\
&= \frac{p(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_N|\mathbf{x})\, p(\mathbf{x})}{p(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_N)} \\
&= \frac{p(\mathbf{y}_1|\mathbf{x})\, p(\mathbf{y}_2|\mathbf{x}) \cdots p(\mathbf{y}_N|\mathbf{x})\, p(\mathbf{x})}{p(\mathbf{y}_1)\, p(\mathbf{y}_2) \cdots p(\mathbf{y}_N)} \\
&= \frac{p(\mathbf{x}|\mathbf{y}_1)\, p(\mathbf{x}|\mathbf{y}_2) \cdots p(\mathbf{x}|\mathbf{y}_N)\, p(\mathbf{x})}{p(\mathbf{x})\, p(\mathbf{x}) \cdots p(\mathbf{x})} \\
&= \frac{p(\mathbf{x}|\mathbf{y}_1)\, p(\mathbf{x}|\mathbf{y}_2) \cdots p(\mathbf{x}|\mathbf{y}_N)}{(p(\mathbf{x}))^{N-1}}
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{P}_{\mathbf{x}|\mathbf{y}_{1:N}} &= (\mathbf{P}^{-1}_{\mathbf{x}|y_1} + \mathbf{P}^{-1}_{\mathbf{x}|y_2} + \cdots + \mathbf{P}^{-1}_{\mathbf{x}|y_N} - (N-1)\mathbf{P}^{-1}_{\mathbf{x}})^{-1} \\
\boldsymbol{\mu}_{\mathbf{x}|\mathbf{y}_{1:N}} &= \mathbf{P}_{\mathbf{x}|\mathbf{y}_{1:N}}(\mathbf{P}^{-1}_{\mathbf{x}|\mathbf{y}_1}\boldsymbol{\mu}_{\mathbf{x}|\mathbf{y}_1} + \mathbf{P}^{-1}_{\mathbf{x}|\mathbf{y}_2}\boldsymbol{\mu}_{\mathbf{x}|\mathbf{y}_2} + \cdots + \mathbf{P}^{-1}_{\mathbf{x}|\mathbf{y}_N}\boldsymbol{\mu}_{\mathbf{x}|\mathbf{y}_N} - (N-1)\mathbf{P}^{-1}_{\mathbf{x}}\boldsymbol{\mu}_{\mathbf{x}})
\end{aligned}
$$