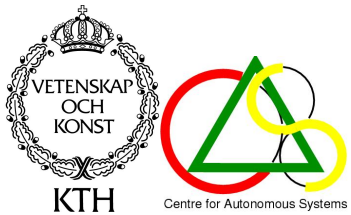
EL2320 - Other EKF Variations

John Folkesson

Other EKF Variations
(Chap 3.4, 7.6, 7.7 in Thrun)



Multi-Hypothesis Tracking - Mixture of Gaussians

The Kalman Filter and the EKF suffer from having to make a single yes no decsion regarding the data association.

This is particularly, problematic for EKF Localization and normally leads to failure if the association is wrong.

One option to fix this is to go from a single to a few data associations, running parallel EKF filters for each.

This leads to a mixture of Gaussians for the distribution function.

Mixture of Gaussians

We have a new form for the $bel(\mathbf{x}_t)$:

$$p(\mathbf{x}_t|\mathbf{z}_{1:t}, \mathbf{u}_{1:t}) = \frac{1}{\sum_i \psi_{t,i}} \sum_i \psi_{t,i} G(\mathbf{x}_t, \mu_{t,i}, \Sigma_{t,i})$$

$$= \frac{1}{\sum_i \psi_{t,i}} \sum_i \psi_{t,i} det(2\pi\Sigma_{t,i})^{-\frac{1}{2}} \exp -\frac{1}{2}(\mathbf{x}_t - \mu_{t,i})^T \Sigma_{t,i}^{-1}(\mathbf{x}_t - \mu_{t,i})$$

So i runs over the hypothesis set.

$\psi_{t,i}$ is the Mixture weight.

Mixture of Gaussians

$$p(\mathbf{x}_t|\mathbf{z}_{1:t}, \mathbf{u}_{1:t}) = \frac{1}{\sum_i \psi_{t,i}} \sum_i \psi_{t,i} G(\mathbf{x}_t, \mu_{t,i}, \Sigma_{t,i})$$

Each Gaussian in the sum is computed exactly as the $bel(\mathbf{x}_t)$ for the EKF.

They differ in the assumed data association with each i having some measurements attributed to features differently than the others for some time $1..t$.

Using the book notatation for an association vectors $\mathbf{c}_{1:t}$ we can say

$$\begin{aligned} bel_i(\mathbf{x}_t) &= p(\mathbf{x}_t|\mathbf{z}_{1:t}, \mathbf{u}_{1:t}, \mathbf{c}_{1:t,i}) \\ &= G(\mathbf{x}_t, \mu_{t,i}, \Sigma_{t,i}) \end{aligned}$$

Mixture of Gaussians

Predict phase:

$$\bar{\Sigma}_t = Q_t + G_t \Sigma_{t-1} G_t^T$$

$$\bar{\mu}_\mathbf{t} = g(\mu_{t-1}, \mathbf{u}_t)$$

and update:

$$K_t = \bar{\Sigma}_t H_t^T (H_t \bar{\Sigma}_t H_t^T + R_t)^{-1}$$

$$\Sigma_t = \bar{\Sigma}_t - K_t H_t \bar{\Sigma}_t$$

$$\mu_t = \bar{\mu}_\mathbf{t} + K_t(\mathbf{z}_t - \mathbf{h}(\bar{\mu}_t))$$

Mixture of Gaussians

$$p(\mathbf{x}_t|\mathbf{z}_{1:t}, \mathbf{u}_{1:t}) = \frac{1}{\sum_i \psi_{t,i}} \sum_i \psi_{t,i} G(\mathbf{x}_t, \mu_{t,i}, \Sigma_{t,i})$$

So really the only new thing here is these mixture weights $\psi_{t,i}$.

We consider the posterori over both $\mathbf{x}_t$ and $\mathbf{c}_{t,i}$:

$$\begin{aligned}
p(\mathbf{x}_t|\mathbf{z}_{1:t}, \mathbf{u}_{1:t}) &= \sum_i p(\mathbf{x}_t|\mathbf{z}_{1:t}, \mathbf{u}_{1:t}, \mathbf{c}_{1:t,i}) p(\mathbf{c}_{1:t,i}|\mathbf{z}_{1:t}, \mathbf{u}_{1:t}) \\
\psi_{t,i} &\propto p(\mathbf{c}_{1:t,i}|\mathbf{z}_{1:t}, \mathbf{u}_{1:t}) \\
&\propto \frac{p(\mathbf{z}_t|\mathbf{z}_{1:t-1}, \mathbf{u}_{1:t}, \mathbf{c}_{1:t,i}) p(\mathbf{c}_{1:t,i}|\mathbf{z}_{1:t-1}, \mathbf{u}_{1:t})}{p(\mathbf{z}_t|\mathbf{z}_{1:t-1}, \mathbf{u}_{1:t})} \\
&\propto p(\mathbf{z}_t|\mathbf{z}_{1:t-1}, \mathbf{u}_{1:t}, \mathbf{c}_{1:t,i}) p(\mathbf{c}_{1:t-1,i}|\mathbf{z}_{1:t-1}, \mathbf{u}_{1:t-1}) \\
&\propto p(\mathbf{z}_t|\mathbf{z}_{1:t-1}, \mathbf{u}_{1:t}, \mathbf{c}_{1:t,i}) \psi_{t-1,i}
\end{aligned}$$

Mixture of Gaussians

$$p(\mathbf{x}_t|\mathbf{z}_{1:t}, \mathbf{u}_{1:t}) = \frac{1}{\sum_i \psi_{t,i}} \sum_i \psi_{t,i} G(\mathbf{x}_t, \mu_{t,i}, \Sigma_{t,i})$$

So we generate correspondence $\mathbf{c}_{1:t,i}$ from an earlier hypothesis at $t-1$ with $\mathbf{c}_{1:t-1,i}$

$$\begin{aligned} \psi_{t,i} &= p(\mathbf{z}_t|\mathbf{z}_{1:t}, \mathbf{u}_{1:t}, \mathbf{c}_{1:t-1,i}, \mathbf{c}_{t,i})\psi_{t,i} \\ p(\mathbf{z}_t|\mathbf{z}_{1:t}, \mathbf{u}_{1:t}, \mathbf{c}_{1:t-1,i}, \mathbf{c}_{t,i}) &= G(\mathbf{z}_t, \mathbf{h}(\bar{\mu}_{t,i}, \mathbf{c}_{t,i}), S_t^i) \\ S_t^i &= H_t^i \bar{\Sigma}_{t,i} (H_t^i)^T + R_t \end{aligned}$$

This is exactly the same quantity you computed in Lab 1 for the liklihood of a data association.

Mixture of Gaussians

Of course the full mixture over all possible $\mathbf{c}$ vectors at each time t would be intractable.

This is handled by 'pruning' the set based on a criteria:

$$\frac{\psi_{t,i}}{\sum_i \psi_{t,i}} \geq \psi_{min}$$

In practice only a few hypothesis are maintained.

If one tries to use this for Multi-Hypothesis SLAM rather than Tracking one is faced with having to compute the determinate of a large matrix for each component. On the face and $O(n^3)$ operation.

This can be done in $O(n^2)$ however due to the sparseness of the measurement Jacobians.

Iterative Extended Kalman Filter iEKF

Deriving the Iterative Extended Kalman Filter from Bayes' Rule, DeansIEKFDerivation.pdf

The EKF update rule is a relatively crude attempt at finding the state that maximizes the posteori.

$$
\begin{aligned}
\mu_t &= \arg\max_{\mathbf{x}_t} p(\mathbf{x}_t | \mathbf{z}_{1:t}, \mathbf{u}_{1:t}) \\
&= \arg\max_{\mathbf{x}_t} G(\mathbf{z}_t, h(\mathbf{x}_t), R_t) G(\mathbf{x}_t, \bar{\mu}_t, \bar{\Sigma}_t) \\
&= \arg\min_{\mathbf{x}_t} (\mathbf{z}_t - h(\mathbf{x}_t))^T R_t^{-1} (\mathbf{z}_t - h(\mathbf{x}_t)) + (\mathbf{x}_t - \bar{\mu}_t)^T \bar{\Sigma}_t^{-1} (\mathbf{x}_t - \bar{\mu}_t) \\
&= \arg\min_{\mathbf{x}_t} f(\mathbf{x}_t)
\end{aligned}
$$

Iterative Extended Kalman Filter iEKF

$$\mu_t = \arg\min_{\mathbf{x}_t} f(\mathbf{x}_t)$$

A simple method to solve such non-linear minimization is to iteratively appoximate this as a parabolic hyper-bowl and move to the bottom of the bowl, (Newton-Raphson).

That is do a quadratic Taylor expansion around your current iterative estimate and then move to:

$$\mathbf{x}_t^i = \mathbf{x}_t^{i-1} - \left(\frac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j}\right)^{-1}\left(\bigtriangledown f(\mathbf{x})\right)$$

Iterative Extended Kalman Filter iEKF

We have this particular f to minimize:

$$2f(\mathbf{x}_t) = (\mathbf{z}_t - h(\mathbf{x}_t))^T R_t^{-1}(\mathbf{z}_t - h(\mathbf{x}_t)) + (\mathbf{x}_t - \bar{\mu}_t)^T \bar{\Sigma}_t^{-1}(\mathbf{x}_t - \bar{\mu}_t)$$

$$\bigtriangledown f(\mathbf{x}) = H^T(\mathbf{x}_t) R_t^{-1}(\mathbf{z}_t - h(\mathbf{x}_t)) + \bar{\Sigma}_t^{-1}(\mathbf{x}_t - \bar{\mu}_t)$$

$$\frac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j} = \frac{\partial^2 h(\mathbf{x})}{\partial x_i \partial x_j} R_t^{-1}(\mathbf{z}_t - h(\mathbf{x}_t)) + H^T(\mathbf{x}_t) R_t^{-1} H(\mathbf{x}_t) + \bar{\Sigma}_t^{-1}$$

At this point one wants very much to ignore the first term as it may lead to a local max or sadel point instead of a local min.

Luckily one can safely ignore it as it is typically small (...).

One defines the 'Covariance' as:

$$P_i^{-1} = H^T(\mathbf{x}_t^i) R_t^{-1} H(\mathbf{x}_t^i) + \bar{\Sigma}_t^{-1}$$

Iterative Extended Kalman Filter iEKF

Woddbury Matrix Identity:

$$\begin{aligned} P_i &= [H^T(\mathbf{x}_t^i)R_t^{-1}H(\mathbf{x}_t^i) + \bar{\Sigma}_t^{-1}]^{-1} \\ &= \bar{\Sigma}_t - \bar{\Sigma}_t H^T(\mathbf{x}_t^i)[R_t + H(\mathbf{x}_t^i)\bar{\Sigma}_t H^T(\mathbf{x}_t^i)]^{-1}H((\mathbf{x}_t^i)\bar{\Sigma}_t \end{aligned}$$

$$\mathbf{x}_t^{i+1} = \mathbf{x}_t^i - P_i[H^T(\mathbf{x}_t^i)R_t^{-1}(\mathbf{z}_t - h(\mathbf{x}_t^i)) + \bar{\Sigma}_t^{-1}(\mathbf{x}_t^i - \bar{\mu}_t)]$$

One begins with $\mathbf{x}_t^0 = \bar{\mu}_t$ then iterates for untill some stopping criteria is reached, such as

- $i = i_{max}$.
- magnitude of the change in f is below some threshold.
- $|\bigtriangledown f(\mathbf{x})|$ is below some threshold.
- ...

Iterative Extended Kalman Filter iEKF

The final updated $\mu_t = \mathbf{x}_t^i$ and $\Sigma_t = P_i$.

Thus we can avoid some of the linearization problems particularly when the measurement is relatively accurate and indicates a large innovation.