An Adaptive Unscented Kalman Filter

as an alternative to simplified EKF

Han The

given by Jelena Bojarova (theory) and Nils Gustafsson (practical application)
Article:

An Adaptive UKF Algorithm for the State and Parameter Estimations of a Mobile robot.
Song Qi, Han Jian-Da.

*Acta Automatica Sinica, 34, 72-79. (2008)*
Deterministic data assimilation

\[ x_t = f(x_{t-1}) + \eta_{t-1} \]
\[ y_t = h(x_t) + \epsilon_t \]
\[ Y_t = \{y_1, y_2, \ldots, y_t\} \]

Non-linear models

Extended Kalman Filter

\[ x^e_t = E(x_t|Y_t); \quad B^e_t = \text{var}(x_t|Y_t) \]
\[ x^a_t = E(x_t|Y_t); \quad B^a_t = \text{var}(x_t|Y_t) \]

Dynamic update

\[ B_f^t = \text{var}(f(x_{t-1})|Y_{t-1}) + Q_{t-1} \]
\[ x^f_t = x^f_{t-1} + K_t (y_t - h(x^f_t)) = x^f_{t-1} + \text{cov}(x_t, h(x_t)) \text{var}(y_t - h(x^f_t))^{-1} (y_t - h(x^f_t)) \]
\[ B^a_t = B^f_t - K_t (F_t)^{-1} (K_t)^T \]

Unscented Kalman Filter

The first-order approximation

\[ x_t = x_t^f + \delta x^a_{t-1} \approx f(x^a_{t-1}) + \frac{\partial f(x^a_{t-1})}{\partial x_{t-1}} \delta x^a_{t-1} \]
\[ h(x_t) = h(x_t^f) + \frac{\partial h(x^f_t)}{\partial x_t} \delta x^f_t \]

\[ B^f_t = FB^a_{t-1}F^T + Q; \]
\[ \text{var}(h(x_t)) = HB_f^eH^T \]
\[ \text{cov}(x_t, h(x_t)) = B_f^eH^T \]

The second-order approximation

\[ B^a_t \rightarrow \text{select} \rightarrow \text{apply} \rightarrow B^f_t \]
\[ Z^a_{t-1,i} = X_{t-1,i} - x^a_{t-1} \rightarrow \text{apply} \rightarrow f(X_{t-1,i}) \]
\[ Z^f_{t,i} = X_{t,i} - x^f_t \rightarrow \text{apply} \rightarrow h(X_{t,i}) \]

\[ B^a_{t-1} = Z^a_{t-1,i} (Z^a_{t-1,i})^T \]
\[ B^f_t = Z^f_t (Z^f_t)^T \]
\[ Z^a_t = Z^a_{t-1,i} \]

Ensemble Transform Kalman Filter
An Adaptive Unscented Kalman Filter

**Initialisation**

\[
\begin{align*}
    \mathbf{x}_0^a &= \mathbb{E}(\mathbf{X}_0) \\
    \mathbf{B}_0 &= \mathbb{E}((\mathbf{X}_0 - \mathbf{x}_0^a)(\mathbf{X}_0 - \mathbf{x}_0^a)^T)
\end{align*}
\]

**Forecast**

\[
\begin{align*}
    \mathbf{X}_{t,i}^{f*} &= f(\mathbf{x}_{t-1}^a \pm ((n+\lambda)\mathbf{B}_{t-1}^a)^{1/2}) \\
    \mathbf{x}_t^f &= \sum_{0 \leq i \leq 2n} w_i^m \mathbf{X}_{t,i}^{f*} \\
    \mathbf{B}_t^f &= \sum_{0 \leq i \leq 2n} w_i^c (\mathbf{X}_{t,i}^{f*} - \mathbf{x}_t^f)(\mathbf{X}_{t,i}^{f*} - \mathbf{x}_t^f)^T + \mathbf{Q}_t \\
    \mathbf{X}_{t,i}^f &= f(\mathbf{x}_t^f \pm ((n+\lambda)\mathbf{B}_t^f)^{1/2})
\end{align*}
\]

**Analysis**

\[
\begin{align*}
    \mathbf{y}_t^f &= \sum_{0 \leq i \leq 2n} w_i^m \mathbf{y}_{t,i} \\
    \mathbf{B}_{yy,t} &= \sum_{0 \leq i \leq 2n} w_i^c (\mathbf{y}_{t,i} - \mathbf{y}_t^f)(\mathbf{y}_{t,i} - \mathbf{y}_t^f)^T + \mathbf{R} \\
    \mathbf{B}_{xy,t} &= \sum_{0 \leq i \leq 2n} w_i^c (\mathbf{X}_{t,i}^f - \mathbf{x}_t^f)(\mathbf{y}_{t,i} - \mathbf{y}_t^f)^T \\
    \mathbf{K}_t &= \mathbf{B}_{xy,t} (\mathbf{B}_{yy,t}^{-1})^{1/2} \\
    \mathbf{B}_t^a &= \mathbf{B}_t^f - \mathbf{K}_t (\mathbf{B}_{yy,t}^{-1})(\mathbf{K}_t)^T \\
    \mathbf{x}_t^a &= \mathbf{x}_t^f + \mathbf{K}_t (\mathbf{y}_t - \mathbf{y}_t^f)
\end{align*}
\]
An adaptive filtering

- Kalman filter is optimal only in the case of **linear Gaussian systems**
- Stable linear Gaussian system converge to a **steady state solution**
- At steady state solution innovation variance \( (B_{yy,t} + R) \) is constant provided that the observation error variance \( R \) is constant

One way to optimize behavior of the system is to force the innovation variance vary less

\[
v_t = B_{yy,t} = (H(B^f_t + Q_t)H^T + R
\]
In the adaptive unscented Kalman filter scheme, all diagonal entries of $Q_t$ are corrected each assimilation step to minimize

$$V_t = \text{trace}(\Delta S_t) = \text{tr}((S_t - S^o_t))^2$$

$$S^o_t = \sum_{t-N+1 \leq i \leq t} v_i v_i^T$$ (long term innovation variance),

$$S_t = B_{yy,t}$$ (estimated innovation variance from the ensemble spread)

$$\dot{q}^m_t = -\eta \frac{\partial V}{\partial q^m_t}$$

Result: innovation is too large in comparison to the spread of the ensemble; the model does not predict the observation well; model is wrong; increase model error and come closer to the next observation;

innovation is small in comparison to the model spread; the model predict the observation well; model is right; reduce model error and filter away observational noise stronger.