Lecture Notes: Extended Kalman filter

During the last lecture we derived the foundation for the extended Kalman filter (EKF). Now we look at the actual implementation. The neat thing is that the EKF works the same as the KF, but with slightly modified equations.

The model information brought to bear on a problem in the KF is represented in the observation matrix M and the state transition matrix Θ. Places in the KF equations that use these matrices need to be adjusted to use the matrices of nonlinear equations in g and f, and the related Jacobians.

Like the KF, the EKF is in a continuous cycle of predict-update. The following lists the equations for the original KF and the related equations for the EKF:

1. predict next state

$$\text{KF}: \qquad X_{t,t-1} = \Phi X_{t-1,t-1} \tag{1}$$

$$\text{EKF}: \qquad X_{t,t-1} = f(X_{t-1,t-1}, 0) \tag{2}$$

   where $f(X_{t-1,t-1}, 0)$ is the approximated state $\tilde{x}_t$.

2. predict next state covariance

$$\text{KF}: \qquad S_{t,t-1} = \Phi S_{t-1,t-1} \Phi^T + Q \tag{3}$$

$$\text{EKF}: \qquad S_{t,t-1} = \left(\frac{\partial f}{\partial x}\right) S_{t-1,t-1} \left(\frac{\partial f}{\partial x}\right)^T + \left(\frac{\partial f}{\partial a}\right) Q \left(\frac{\partial f}{\partial a}\right)^T \tag{4}$$

   where $\left(\frac{\partial f}{\partial x}\right)$ and $\left(\frac{\partial f}{\partial a}\right)$ are the Jacobians of the state transition equations. The notation $(...)^T$ indicates matrix transpose.

3. obtain measurement(s) Y_t

4. calculate the Kalman gain (weights)

$$\text{KF}: \qquad K_t = S_{t,t-1} M^T [M S_{t,t-1} M^T + R]^{-1} \tag{5}$$

$$\text{EKF}: \qquad K_t = S_{t,t-1} \left(\frac{\partial g}{\partial x}\right)^T \left[\left(\frac{\partial g}{\partial x}\right) S_{t,t-1} \left(\frac{\partial g}{\partial x}\right)^T + \left(\frac{\partial g}{\partial n}\right) R \left(\frac{\partial g}{\partial n}\right)^T \right]^{-1} \tag{6}$$

   where $\left(\frac{\partial g}{\partial x}\right)$ and $\left(\frac{\partial g}{\partial n}\right)$ are the Jacobians of the measurement equations.

5. update state

$$\text{KF}: \quad X_{t,t} = X_{t,t-1} + K_t[Y_t - MX_{t,t-1}] \tag{7}$$
$$\text{EKF}: \quad X_{t,t} = X_{t,t-1} + K_t[Y_t - g(\tilde{x}_t, 0)] \tag{8}$$

where $g(\tilde{x}_t, 0)$ is the ideal (noiseless) measurement of the approximated state from above.

6. update state covariance

$$\text{KF}: \quad S_{t,t} = [I - K_t M]S_{t,t-1} \tag{9}$$
$$\text{EKF}: \quad S_{t,t} = \left[I - K_t\left(\frac{\partial g}{\partial x}\right)\right] S_{t,t-1} \tag{10}$$

7. loop (now t becomes $t+1$)

In practice, the main difference between the KF and EKF is that the values in the Jacobian matrices must be calculated every iteration. In order to understand this concept, we will look at an example.

Consider tracking an object in 2D using a constant velocity model. Thus, the state variables are X and Y position and velocity:

$$X_t = \begin{bmatrix} x_t \\ \dot{x}_t \\ y_t \\ \dot{y}_t \end{bmatrix} \tag{11}$$

The state transition equations for this model are:

$$f(x_t, a_t) = \begin{bmatrix} x_{t+1} = x_t + T\dot{x}_t + 0 \\ \dot{x}_{t+1} = \dot{x}_t + u_1 \\ y_{t+1} = y_t + T\dot{y}_t + 0 \\ \dot{y}_{t+1} = \dot{y}_t + u_2 \end{bmatrix} \tag{12}$$

where u_1 and u_2 are random samples drawn from $N(0, \sigma_a^2)$ representing an unknown acceleration.

For observations, consider using a sensor that operates on polar coordinates, providing an R and Θ measurement:

$$Y_t = \begin{bmatrix} r_t \\ \theta_t \end{bmatrix} \tag{13}$$

The observation equations for this model are:

$$g(x_t, n_t) = \begin{bmatrix} r_t = \sqrt{x_t^2 + y_t^2} + n_1 \\ \theta_t = \tan^{-1} \frac{y_t}{x_t} + n_2 \end{bmatrix} \tag{14}$$

where n_1 is a random sample drawn from $N(0, \sigma^2_{\text{dist}})$ and n_2 is a random sample drawn from $N(0, \sigma^2_{\text{dir}})$ representing noises on the measured distance and direction, respectively.

In order to use this model in the EKF, we must calculate the four Jacobians. The derivative of the state transition equations with respect to the state variables is:

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial(x, \dot{x}, y, \dot{y})} = \begin{bmatrix} 1 & T & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & T \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{15}$$

The derivative of the state transition equations with respect to the dynamic noises is:

$$\frac{\partial f}{\partial a} = \frac{\partial f}{\partial(0, u_1, 0, u_2)} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{16}$$

The Jacobians of the state transition equations both fairly simple because that portion of this model is linear. Therefore all the derivatives are constant.

The derivative of the observation equations with respect to the state variables is:

$$\begin{aligned} \frac{\partial g}{\partial x} = \frac{\partial g}{\partial(x, \dot{x}, y, \dot{y})} &= \begin{bmatrix} \frac{\partial}{\partial x}\sqrt{x^2+y^2} + n_1 & 0 & \frac{\partial}{\partial y}\sqrt{x^2+y^2} + n_1 & 0 \\ \frac{\partial}{\partial x}\tan^{-1}\frac{y}{x} + n_2 & 0 & \frac{\partial}{\partial y}\tan^{-1}\frac{y}{x} + n_2 & 0 \end{bmatrix} \\ &= \begin{bmatrix} x(x^2+y^2)^{-1/2} & 0 & y(x^2+y^2)^{-1/2} & 0 \\ \frac{-y}{x^2+y^2} & 0 & \frac{x}{x^2+y^2} & 0 \end{bmatrix} \end{aligned} \tag{17}$$

The equation makes use of the functions for the derivative of arctan, the power rule and the chain rule. The time subscripts are ommitted for clarity. However, in practice the values in the matrix $\frac{\partial g}{\partial x}$ must be calculated every iteration. They are calculated using values from the current filtered estimate of the state variables. The derivative of the observation equations with respect to the measurement noises is:

$$\frac{\partial g}{\partial n} = \frac{\partial g}{\partial(n_1, n_2)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{18}$$

To finish this example we must look at the covariances. The covariance of the dynamic noises is:

$$Q = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \sigma_a^2 & 0 & \text{small\#} \\ 0 & 0 & 0 & 0 \\ 0 & \text{small\#} & 0 & \sigma_a^2 \end{bmatrix} \tag{19}$$

where a small number allows for some covariance in practice. The covariance of the measurement noises is

$$R = \begin{bmatrix} \sigma_{\text{dist}}^2 & \text{small\#} \\ \text{small\#} & \sigma_{\text{dir}}^2 \end{bmatrix} \tag{20}$$

The covariance of the state, S, is a 4×4 matrix with the variances of $x, \dot{x}, y, \dot{y}$ along the diagonal and the (hopefully very small) covariances in the other elements.

Finally, in order to verify that the problem has been configured properly, one should go through all the EKF equations from above and make sure that the matrix sizes match. Doing this will reveal that the Kalman gain matrix K is 4×2, and everything fits.

$$\overbrace{X}^{4\times1} = \overbrace{f}^{4\times1} \tag{21}$$

$$\overbrace{S}^{4\times4} = \overbrace{\left(\frac{\partial f}{\partial x}\right)}^{4\times4} \overbrace{S}^{4\times4} \overbrace{\left(\frac{\partial f}{\partial x}\right)^T}^{4\times4} + \overbrace{\left(\frac{\partial f}{\partial a}\right)}^{4\times4} \overbrace{Q}^{4\times4} \overbrace{\left(\frac{\partial f}{\partial a}\right)^T}^{4\times4} \tag{22}$$

$$\overbrace{K}^{4\times2} = \overbrace{S}^{4\times4} \overbrace{\left(\frac{\partial g}{\partial x}\right)^T}^{4\times2} \left[\overbrace{\left(\frac{\partial g}{\partial x}\right)}^{2\times4} \overbrace{S}^{4\times4} \overbrace{\left(\frac{\partial g}{\partial x}\right)^T}^{4\times2} + \overbrace{\left(\frac{\partial g}{\partial n}\right)}^{2\times2} \overbrace{R}^{2\times2} \overbrace{\left(\frac{\partial g}{\partial n}\right)^T}^{2\times2} \right]^{-1} \tag{23}$$

$$\overbrace{X}^{4\times1} = \overbrace{X}^{4\times1} + \overbrace{K}^{4\times2} (\overbrace{Y}^{2\times1} - \overbrace{g}^{2\times1}) \tag{24}$$

$$\overbrace{S}^{4\times4} = \left[\overbrace{I}^{4\times4} - \overbrace{K}^{4\times2} \overbrace{\left(\frac{\partial g}{\partial x}\right)}^{2\times4} \right] \overbrace{S_{t,t-1}}^{4\times4} \tag{25}$$