Kalman Filter

This is to give a very brief introduction to Kalman Filter and its family which can be employed in control engineering, robotics, signal processing, etc.

Given the probabilistic State-space model in discrete time

$$x_k = f_k(x_{k-1}, w_{k-1})$$
$$y_k = h_k(x_k, v_n)$$

Where x is the state vector and w is the noise vector; v_n is the measurement noise.
If the process is Markovian, then we have

$$p(x_k|x_{1:k-1}, y_{1:k-1}) = p(x_k|x_{k-1})$$
$$p(x_{k-1}|x_{k:T}, y_{k:T}) = p(x_{k-1}|x_k)$$
$$p(y_k|x_{1:k-1}, y_{1:k-1}) = p(y_k|x_k)$$

where T is the lase sample.
The prior distribution $p(x_0)$, where x_0 is the state prior to the first measurement
The state-space model

$$x_k \sim p(x_k|x_{k-1})$$
$$y_k \sim p(y_k|x_k)$$

The measurement sequence $y_{1:k} = y_1, y_2, \ldots., y_k$
Computation is based on the recursion rule

$$p(x_{k-1}|y_{1:k-1}) \rightarrow p(x_k|y_{1:k})$$

This means we get the current state from the prior state and all the past measurements. Assume the posterior distribution of the previous step

$$p(x_{k-1}|y_{1:k-1})$$

The joint distribution of x_{k-1}, x_k given $y_{1:k-1}$ can be computed as

$$p(x_{k-1}, x_k \,|y_{1:k-1}) = p(x_k \,|x_{k-1}, y_{1:k-1})p(x_{k-1} \,|y_{1:k-1})$$
$$= p(x_k \,|x_{k-1})p(x_{k-1} \,|y_{1:k-1})$$

Integrating over x_{k-1} gives the prediction step of the optimal filter, which is the Chapman-Kolmogorov equation.

$$p(x_k \,|y_{1:k-1}) = \int p(x_k \,|x_{k-1})p(x_{k-1} \,|y_{1:k-1})\, dx_{k-1}$$

The measurement update state is found from Bayes' Rule.

$$p(x_k \,|y_{1:k}) = \frac{1}{C_k} p(y_k|x_k)p(x_k \,|y_{k-1})$$
$$C_k = \int p(y_k|x_k)p(x_k \,|y_{k-1})dx_k$$

C_k is the probability of the current measurement, given all past measurements.
If the noise is additive and Gaussian with covariance Q_n and zero mean, and the measurement covariance R_n, then we have

$$x_k = f_k(x_{k-1}) + w_{k-1}$$
$$y_k = h_k(x_k) + v_n$$

Given that Q is not time dependent, we can write

$$p(x_k \,|x_{k-1}, y_{1:k-1}) = N(x_k; f_k(x_{k-1}), Q)$$

Hence we write the prediction step as

$$p(x_k \,|y_{1:k-1}) = \int N(x_k; f_k(x_{k-1}), Q) p(x_{k-1} \,|y_{1:k-1}) dx_{k-1}$$

We need to find the first two moments of x_k.

$$E[x_k] = \int x_k p(x_k \,|y_{1:k-1})\, dx_k$$
$$E\left[x_k x_k^T\right] = \int x_k x_k^T p(x_k \,|y_{1:k-1})\, dx_k$$

The identity $E[x] = \int x N(x; f(s), \Sigma) dx = f(s)$

$$E[x_k] = \int x_k \int N(x_k; f_k(x_{k-1}), Q) p(x_{k-1} \,|y_{1:k-1}) dx_{k-1}\, dx_k$$
$$= \int \left[\int x_k N(x_k; f_k(x_{k-1}), Q) dx_k\right] p(x_{k-1} \,|y_{1:k-1}) dx_{k-1}$$
$$= \int f_k(x_{k-1}) p(x_{k-1} \,|y_{1:k-1}) dx_{k-1}$$

Assuming that $p(x_{k-1} \,|y_{1:k-1}) = N\left(x_{k-1}; \hat{x}_{k-1|k-1}, P^{xx}_{k-1|k-1}\right)$, we get

$$\hat{x}_{k|k-1} = \int f_k(x_{k-1}) N\left(x_{k-1}; \hat{x}_{k-1|k-1}, P^{xx}_{k-1|k-1}\right) dx_{k-1}$$

And the second moment is

$$E\left[x_k x_k^T\right] = \int \left[\int x_k x_k^T N(x_k; f_k(x_{k-1}), Q) dx_k\right] p(x_{k-1} \,|y_{1:k-1}) dx_{k-1}$$
$$P^{xx}_{k|k-1} = \mathrm{Q} + \int f_k(x_{k-1}) {f_k}^T(x_{k-1}) N\left(x_{k-1}; \hat{x}_{k-1|k-1}, P^{xx}_{k-1|k-1}\right) dx_{k-1} - \hat{x}^T_{k|k-1} \hat{x}_{k|k-1}$$

Then the Covariance for the initial state is Gaussian and is P^{xx}_0. The Kalman Filter is approximated by

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_n\left[y_k - \hat{y}_{k|k-1}\right]$$
$$P^{xx}_{k|k} = P^{xx}_{k|k-1} - K_n P^{xx}_{k|k-1} K_n^T$$

$$K_n = P^{xy}_{k|k-1}\left[P^{yy}_{k|k-1}\right]^{-1}$$

$$P^{yy}_{k|k-1} = R + \int h_k(x_k){h_k}^T(x_k)N\left(x_k;\hat{x}_{k|k-1},P^{xx}_{k|k-1}\right)dx_k - \hat{x}^T_{k|k-1}\hat{y}_{k|k-1}$$

$$P^{xy}_{k|k-1} = \int x_k{h_k}^T(x_k)N\left(x_k;\hat{x}_{k|k-1},P^{xx}_{k|k-1}\right)dx_k$$

$$\hat{y}_{k|k-1} = \int h_k(x_k)N\left(x_k;\hat{x}_{k|k-1},P^{xx}_{k|k-1}\right)dx_k$$

Conventional Kalman Filter

The state-space model in discrete time

$$x_k = A_{k-1}x_{k-1} + B_{k-1}u_{k-1} + \ q_{k-1}$$
$$y_k = H_kx_k + r_n$$

Measurements update

$$P^{yy}_{k|k-1} = H_kP^-_kH^T_k + R_k$$

$$P^{xy}_{k|k-1} = P^-_kH^T_k$$

$$P^{xx}_{k|k} = A_{k-1}P_{k-1}A^T_{k-1} + Q_{k-1}$$

$$\hat{x}_{k|k-1} = m^-_k$$

$$\hat{y}_{k|k-1} = H_km^-_k$$

The prediction step becomes

$$m^-_k = A_{k-1}m_{k-1}$$

$$P^-_k = A_{k-1}P_{k-1}A^T_{k-1} + Q_{k-1}$$

The update step is

$$v_k = y_k - H_km^-_k$$

$$S_k = H_kP^-_kH^T_k + R_k$$

$$K_k = P^-_kH^T_kS^{-1}_k$$

$$m_k = m^-_k + K_kv_k$$

$$P_k = P^-_k - K_kS_kK^T_k$$

Example

$$r_{k+1} = r_k + Tv_k + \frac{1}{2}T^2 a_k$$
$$v_{k+1} = v_k + Ta_k$$

In matrix form that is

$$\begin{bmatrix} r_{k+1} \\ v_{k+1} \end{bmatrix} = \begin{bmatrix} 1 & T \\ 0 & T \end{bmatrix} \begin{bmatrix} r_k \\ v_k \end{bmatrix} + \begin{bmatrix} \frac{1}{2}T^2 \\ T \end{bmatrix} a_k$$
$$x_{k+1} = fx_k + bu_k$$

Extended Kalman Filter

It was developed to handle models with nonlinear dynamical model. Given a nonlinear model of the form

$$x_k = f(x_{k-1}, k-1) + q_{k-1}$$
$$y_k = h(x_k, k) + r_k$$

The prediction step is

$$m_k^- = f(m_{k-1}, k-1)$$

$$P_k^- = F_x(m_{k-1}, k-1)P_{k-1}F_x^T(m_{k-1}, k-1) + Q_{k-1}$$

The update step is

$$v_k = y_k - h\left(m_k^-, k\right)$$

$$S_k = H_x\left(m_k^-, k\right)P_k^- H_x^T\left(m_k^-, k\right) + R_k$$

$$K_k = P_k^- H_x\left(m_k^-, k\right)S_k^{-1}$$

$$m_k = m_k^- + K_k v_k$$

$$P_k = P_k^- - K_k S_k K_k^T$$

F_x and H_x are the Jacobians of the nonlinear functions f and h, respectively.

$$f = \begin{bmatrix} f_x(x, y) \\ f_y(x, y) \end{bmatrix}$$
$$F_k = \begin{bmatrix} \frac{\partial f_x(x_k, y_k)}{\partial x} & \frac{\partial f_x(x_k, y_k)}{\partial y} \\ \frac{\partial f_y(x_k, y_k)}{\partial x} & \frac{\partial f_y(x_k, y_k)}{\partial y} \end{bmatrix}$$

Unscented Kalman Filter

With the UKF we work with the nonlinear dynamical and measurement equations directly. The UKF is also known as a σ point filter because it simultaneously maintains models one sigma from the mean.

In the following sections we develop the equations for the nonaugmented kalman filter.

$$x_k = f(x_{k-1}, k-1) + q_{k-1}$$
$$y_k = h(x_k, k) + r_k$$

Define weights as

$$W_m^0 = \frac{\lambda}{n+\lambda}$$
$$W_c^0 = \frac{\lambda}{n+\lambda} + 1 - \alpha^2 - \beta$$
$$W_m^i = \frac{\lambda}{2(n+\lambda)}, i = 1,\ldots, 2n$$
$$W_c^i = \frac{\lambda}{2(n+\lambda)}, i = 1,\ldots, 2n$$

Note that $W_m^i = W_c^i$

$$\lambda = \alpha^2(n+\kappa) - n$$
$$c = \lambda + n = \alpha^2(n+\kappa)$$

α, β, $\kappa$are scaling constants.

α: 0 for state estimation, 3 minus the number of states for parameter estimation.

β: Determines spread of sigma points. Smaller means more closely spaced sigma points.

κ: Constant for prior knowledge. Set to 2 for Gaussian processes.

n is the order of the system. The weights can be put into matrix from.

$$w_m = \left[W_m^0 \ldots W_m^{2n}\right]^T$$

$$W = (I - [w_m \ldots w_m]) \begin{bmatrix} W_c^0 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & W_c^0 \end{bmatrix} (I - [w_m \ldots w_m])^T$$

I is the 2n+1 by 2n+1 identity matrix.

The prediction step is

$$X_{k-1} = [m_{k-1} \ldots m_{k-1}] + \sqrt{c}\left[0\ \sqrt{P_{k-1}} - \sqrt{P_{k-1}}\right]$$

$$\widehat{X}_k = f(X_{k-1}, k-1)$$

$$m_k^- = \widehat{X}_k w_m$$

$$P_k^- = \widehat{X}_k W \widehat{X}_{k-1}^T + Q_{k-1}$$

The update step is

$$X_k^- = \left[m_k^- \ldots m_k^-\right] + \sqrt{c}\left[0\ \sqrt{P_k^-} - \sqrt{P_k^-}\right]$$

$$Y_k^- = h(X_k^-, k)$$

$$\mu_k = Y_k^- w_m$$

$$S_k = Y_k^- W [Y_k^-]^T + R_k$$

$$C_k = X_k^- W [Y_k^-]^T$$

$$K_k = C_k S_k^{-1}$$

$$m_k = m_k^- + K_k (y_k - \mu_k)$$

$$P_k = P_k^- - K_k S_k K_k^T$$