2 The Extended Kalman Filter (EKF)

The Process to be Estimated

As described above in section , the Kalman filter addresses the general problem of trying to estimate the state $x \in \Re^n$ of a discrete-time controlled process that is governed by a linear stochastic difference equation. But what happens if the process to be estimated and (or) the measurement relationship to the process is non-linear? Some of the most interesting and successful applications of Kalman filtering have been such situations. A Kalman filter that linearizes about the current mean and covariance is referred to as an extended Kalman filter or EKF.

In something akin to a Taylor series, we can linearize the estimation around the current estimate using the partial derivatives of the process and measurement functions to compute estimates even in the face of non-linear relationships. To do so, we must begin by modifying some of the material presented in section . Let us assume that our process again has a state vector $x \in \Re^n$, but that the process is now governed by the non-linear stochastic difference equation

$$x_k = f(x_{k-1}, u_k, w_{k-1}) \text{ , (2.1)}$$

with a measurement $z \in \Re^m$ that is

$$z_k = h(x_k, v_k) \text{ , (2.2)}$$

where the random variables w_k and v_k again represent the process and measurement noise as in (1.3) and (1.4). In this case the non-linear function f in the difference equation (2.1) relates the state at the previous time step $k-1$ to the state at the current time step k . It includes as parameters any driving function uk and the zero-mean process noise wk. The non-linear function h in the measurement equation (2.2) relates the state x_k to the measurement z_k .

In practice of course one does not know the individual values of the noise w_k and v_k at each time step. However, one can approximate the state and measurement vector without them as

$$\tilde{x}_k = f(\hat{x}_{k-1}, u_k, 0) \text{ (2.3)}$$

and

$$\tilde{z}_k = h(\tilde{x}_k, 0) \text{ , (2.4)}$$

where $\hat{x}_k$ is some a posteriori estimate of the state (from a previous time step k).

It is important to note that a fundamental flaw of the EKF is that the distributions (or densities in the continuous case) of the various random variables are no longer normal after undergoing their respective nonlinear transformations. The EKF is simply an ad hoc state estimator that only approximates the optimality of Bayes' rule by linearization. Some interesting work has been done by Julier et al. in developing a variation to the EKF, using methods that preserve the normal distributions throughout the non-linear transformations [Julier96].

The Computational Origins of the Filter

To estimate a process with non-linear difference and measurement relationships, we begin by writing new governing equations that linearize an estimate about (2.3) and (2.4),

$$x_k = \tilde{x}_k + A(x_{k-1} - \hat{x}_{k-1}) + Ww_{k-1} \text{ , (2.5)}$$

$$z_k = \tilde{z}_k + H(x_k - \tilde{x}_k) + Vv_k \text{ . (2.6)}$$

where

x_k and z_k are the actual state and measurement vectors,

$\tilde{x}_k$ and $\tilde{z}_k$ are the approximate state and measurement vectors from (2.3) and (2.4),

$\hat{x}_k$ is an a posteriori estimate of the state at step k,

the random variables w_k and v_k represent the process and measurement noise as in (1.3) and (1.4).

A is the Jacobian matrix of partial derivatives of f with respect to x, that is

$$A_{[i,j]} = \frac{\partial f_{[i]}}{\partial x_{[j]}}(\hat{x}_k, u_k, 0) \text{ ,}$$

W is the Jacobian matrix of partial derivatives of f with respect to w,

$$W_{[i,j]} = \frac{\partial f_{[i]}}{\partial w_{[j]}}(\hat{x}_k, u_k, 0) \text{ ,}$$

H is the Jacobian matrix of partial derivatives of h with respect to x,

$$H_{[i,j]} = \frac{\partial h_{[i]}}{\partial x_{[j]}}(\tilde{x}_k, 0)$$,

V is the Jacobian matrix of partial derivatives of h with respect to v,

$$V_{[i,j]} = \frac{\partial h_{[i]}}{\partial v_{[j]}}(\tilde{x}_k, 0)$$.

Note that for simplicity in the notation we do not use the time step subscript k with the Jacobians A , W , H , and V , even though they are in fact different at each time step.

Now we define a new notation for the prediction error,

$$\tilde{e}_{x_k} \equiv x_k - \tilde{x}_k$$ **, (2.7)**

and the measurement residual,

$$\tilde{e}_{z_k} \equiv z_k - \tilde{z}_k$$ **. (2.8)**

Remember that in practice one does not have access to x_k in (2.7), it is the actual state vector, i.e. the quantity one is trying to estimate. On the other hand, one does have access to z_k in (2.8), it is the actual measurement that one is using to estimate x_k . Using (2.7) and (2.8) we can write governing equations for an error process as

$$\tilde{e}_{x_k} = A(x_{k-1} - \hat{x}_{k-1}) + \varepsilon_k$$ **, (2.9)**

$$\tilde{e}_{z_k} = H\tilde{e}_{x_k} + \eta_k$$ **, (2.10)**

where ε_k and η_k represent new independent random variables having zero mean and covariance matrices WQW^T and VRV^T , with Q and R as in (1.3) and (1.4) respectively.

Notice that the equations (2.9) and (2.10) are linear, and that they closely resemble the difference and measurement equations (1.1) and (1.2) from the discrete Kalman filter. This motivates us to use the actual measurement residual $\tilde{e}_{z_k}$ in (2.8) and a second (hypothetical) Kalman filter to estimate the prediction error $\tilde{e}_{x_k}$ given by (2.9). This estimate, call it $\hat{e}_k$, could then be used along with (2.7) to obtain the a posteriori state estimates for the original non-linear process as

$$\hat{x}_k = \tilde{x}_k + \hat{e}_k$$ **. (2.11)**

The random variables of (2.9) and (2.10) have approximately the following probability distributions (see the previous footnote):

$$p(\tilde{e}_{x_k}) \sim N(0, E[\tilde{e}_{x_k}\tilde{e}_{x_k}^T])$$
$$p(\varepsilon_k) \sim N(0, WQ_kW^T)$$
$$p(\eta_k) \sim N(0, VR_kV^T)$$

Given these approximations and letting the predicted value of $\hat{e}_k$ be zero, the Kalman filter equation used to estimate $\hat{e}_k$ is

$$\hat{e}_k = K_k\tilde{e}_{z_k} \text{ . } \textbf{(2.12)}$$

By substituting (2.12) back into (2.11) and making use of (2.8) we see that we do not actually need the second (hypothetical) Kalman filter:

$$\begin{aligned}\hat{x}_k &= \tilde{x}_k + K_k\tilde{e}_{z_k} \\ &= \tilde{x}_k + K_k(z_k - \tilde{z}_k) \quad \textbf{(2.13)}\end{aligned}$$

Equation (2.13) can now be used for the measurement update in the extended Kalman filter, with $\tilde{x}_k$ and $\tilde{z}_k$ coming from (2.3) and (2.4), and the Kalman gain K_k coming from (1.11) with the appropriate substitution for the measurement error covariance.

The complete set of EKF equations is shown below in Table 2-1 and Table 2-2. Note that we have substituted $\hat{x}_k^-$ for $\tilde{x}_k$ to remain consistent with the earlier "super minus" a priori notation, and that we now attach the subscript k to the Jacobians A , W , H , and V , to reinforce the notion that they are different at (and therefore must be recomputed at) each time step.

Table 2-1: EKF time update equations.

$\hat{x}_k^- = f(\hat{x}_{k-1}, u_k, 0)$ **(2.14)**
$P_k^- = A_kP_{k-1}A_k^T + W_kQ_{k-1}W_k^T$ **(2.15)**

As with the basic discrete Kalman filter, the time update equations in Table 2-1 project the state and covariance estimates from the previous time step $k-1$ to the current time step k . Again f in (2.14) comes from (2.3), A_k and W_k are the process Jacobians at step k, and Q_k is the process noise covariance (1.3) at step k.

Table 2-2: EKF measurement update equations.

$K_k = P_k^- H_k^T (H_k P_k^- H_k^T + V_k R_k V_k^T)^{-1}$	**(2.16)**
$\hat{x}_k = \hat{x}_k^- + K_k(z_k - h(\hat{x}_k^-, 0))$	**(2.17)**
$P_k = (I - K_k H_k) P_k^-$	**(2.18)**

As with the basic discrete Kalman filter, the measurement update equations in Table 2-2 correct the state and covariance estimates with the measurement z_k . Again h in (2.17) comes from (2.4), H_k and V are the measurement Jacobians at step k, and R_k is the measurement noise covariance (1.4) at step k. (Note we now subscript R allowing it to change with each measurement.)

The basic operation of the EKF is the same as the linear discrete Kalman filter as shown in Figure 1-1. Figure 2-1 below offers a complete picture of the operation of the EKF, combining the high-level diagram of Figure 1-1 with the equations from Table 2-1 and Table 2-2.

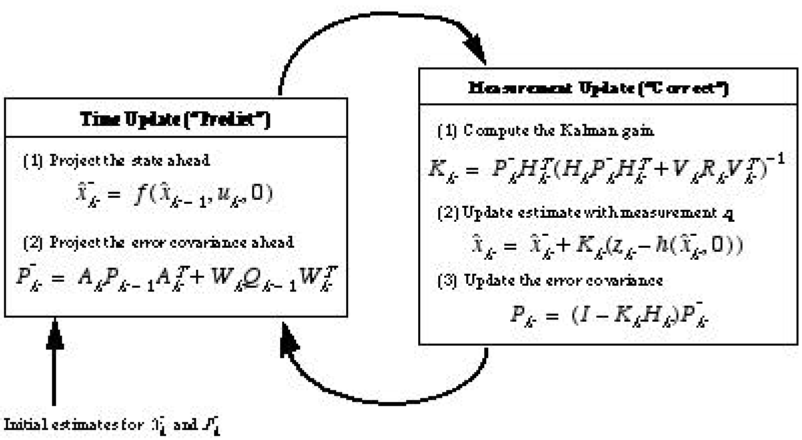

Figure 2-1. A complete picture of the operation of the extended Kalman filter, combining the hig

An important feature of the EKF is that the Jacobian H_k in the equation for the Kalman gain K_k serves to correctly propagate or "magnify" only the relevant component of the measurement information. For example, if there is not a one-to-one mapping between the measurement z_k and the state via h , the Jacobian H_k affects the Kalman gain so that it only magnifies the portion of the residual $z_k - h(\hat{x}_k^-, 0)$ that does affect the state. Of course if over all measurements there is not a one-to-one mapping between the measurement z_k and the state via h , then as you might expect the filter will quickly diverge. In this case the process is unobservable.