Extended Kalman Filter Lecture Notes

EEE 581-Spring 1997
Darryl Morrell

1 Introduction

In this lecture note, we extend the Kalman Filter to non-linear system models to obtain an approximate filter–the Extended Kalman Filter. We will do this by finding an approximate error system that is linear, and applying the Kalman filter to this error system. Because the EKF is obtained using a linear approximation of a non-linear system, it offers no guarantees of optimality in a mean squared error sense (or in any other sense). However, for many systems, the EKF has proven to be a useful method of obtaining good estimates of the system state.

2 Discrete/Discrete EKF

The system model that we use is the following discrete / discrete model:

$$\mathbf{X}_{k+1} = \mathbf{f}\left(\mathbf{X}_k, k\right) + G\mathbf{W}_k \tag{1}$$

$$\mathbf{Z}_k = \mathbf{h}\left(\mathbf{X}_k, k\right) + \mathbf{V}_k \tag{2}$$

In this model, $\mathbf{W}_k$ is a discrete-time white noise process with mean zero and covariance matrix Q, $\mathbf{V}_k$ is a discrete-time white noise process with mean zero and covariance matrix R, and $\mathbf{W}_j$, $\mathbf{V}_k$, and $\mathbf{X}_0$ are uncorrelated for all j and k.

We develop the Extended Kalman filter by starting with a *nominal reference trajectory* denoted $\mathbf{x}_k^R$; $\mathbf{x}_k^R$ is obtained as the solution to the difference equation (1) without the process noise $\mathbf{W}_k$:

$$\mathbf{x}_{k+1}^R = \mathbf{f}\left(\mathbf{x}_k^R, k\right) \tag{3}$$

This difference equation has an initial condition $\mathbf{x}_0^R$. We define the error between $\mathbf{X}_k$ and $\mathbf{x}_k^R$ as

$$\boldsymbol{\delta}_k = \mathbf{X}_k - \mathbf{x}_k^R \tag{4}$$

We now find an approximate linear model that describes the dynamics of $\boldsymbol{\delta}_k$:

$$\begin{aligned}
\boldsymbol{\delta}_{k+1} &= \mathbf{X}_{k+1} - \mathbf{x}_{k+1}^R \\
&= \mathbf{f}\left(\mathbf{X}_k, k\right) + G\mathbf{W}_k - \mathbf{f}\left(\mathbf{x}_k^R, k\right)
\end{aligned}$$

To obtain a linear approximation of this equation, we make a Taylor series expansion[1] of $\mathbf{f}(\mathbf{X}_k, k)$ about the value $\mathbf{x}_k^R$ and drop all but the constant and linear terms:

$$\mathbf{f}(\mathbf{X}_k, k) \approx \mathbf{f}\left(\mathbf{x}_k^R, k\right) + \left.\frac{\partial \mathbf{f}(\mathbf{x}, k)}{\partial \mathbf{x}}\right|_{\mathbf{x}=\mathbf{x}_k^R} \left(\mathbf{X}_k - \mathbf{x}_k^R\right)$$

We make the definition

$$A_k = \left.\frac{\partial \mathbf{f}(\mathbf{x}, t)}{\partial \mathbf{x}}\right|_{\mathbf{x}=\mathbf{x}_k^R},$$

and note that A_k is an $n \times n$ matrix of the following form:

$$A_k = \left[\begin{array}{ccc} \frac{\partial f_1(\mathbf{X},k)}{\partial x_1} & \dots & \frac{\partial f_1(\mathbf{X},k)}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_n(\mathbf{X},k)}{\partial x_1} & \dots & \frac{\partial f_n(\mathbf{X},k)}{\partial x_n} \end{array}\right]_{\mathbf{x}=\mathbf{x}_k^R}$$

With this definition,

$$\begin{aligned} \boldsymbol{\delta}_{k+1} &\approx \mathbf{f}\left(\mathbf{x}_k^R, k\right) + A_k\left(\mathbf{X}_k - \mathbf{x}_k^R\right) + G\mathbf{W}_k - \mathbf{f}\left(\mathbf{x}_k^R, k\right) \\ &= A_k \boldsymbol{\delta}_k + G\mathbf{W}_k \end{aligned} \quad (5)$$

Note that (5) is a linear difference equation.

We follow a similar approach to obtain an (approximate) linear relationship from (2). We expand $\mathbf{h}(\mathbf{X}_k, k)$ in a Taylor series about the nominal trajectory $\mathbf{x}_k^R$:

$$\mathbf{h}(\mathbf{X}_k, k) \approx \mathbf{h}(\mathbf{x}_k^R, k) + \left.\frac{\partial \mathbf{h}(\mathbf{x}, k)}{\partial \mathbf{x}}\right|_{\mathbf{x}=\mathbf{x}_k^R} \left(\mathbf{X}_k - \mathbf{x}_k^R\right)$$

We define H_k as

$$H_k = \left.\frac{\partial \mathbf{h}(\mathbf{x}, k)}{\partial \mathbf{x}}\right|_{\mathbf{x}=\mathbf{x}_k^R}$$

Then we can write (2) as

$$\mathbf{Z}_k \approx \mathbf{h}(\mathbf{x}_k^R, k) + H_k\left(\mathbf{X}_k - \mathbf{x}_k^R\right) + \mathbf{V}_k.$$

Define $\boldsymbol{\nu}_k$ as

$$\begin{aligned} \boldsymbol{\nu}_k &= \mathbf{Z}_k - \mathbf{h}(\mathbf{x}_k^R, k) \\ &= \mathbf{h}(\mathbf{X}_k, k) + \mathbf{V}_k - \mathbf{h}(\mathbf{x}_k^R, k) \\ &\approx H_k\left(\mathbf{X}_k - \mathbf{x}_k^R\right) + \mathbf{V}_k \\ &= H_k \boldsymbol{\delta}_k + \mathbf{V}_k \end{aligned} \quad (6)$$

[1] A Taylor series expansion of a vector function $\mathbf{f}(\mathbf{x})$ about the point $\mathbf{x}_0$ is the following:

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{x}_0) + \left.\frac{\partial \mathbf{f}(\mathbf{x})}{\partial \mathbf{x}}\right|_{\mathbf{x}=\mathbf{x}_0} (\mathbf{x} - \mathbf{x}_0) + \text{Higher Order Terms}$$

So we have an (approximate) linear relationship between the quantity $\boldsymbol{\nu}_k$ (which can be computed from $\mathbf{Z}_k$) and $\boldsymbol{\delta}_k$.

Assuming that $\boldsymbol{\delta}_k$ obeys the difference equation (5) and the observation equation (6), we can use a discrete / discrete Kalman filter to compute $\hat{\boldsymbol{\delta}}_{k+1|k}$, $P_{k+1|k}$, $\hat{\boldsymbol{\delta}}_{k+1|k+1}$, and $P_{k+1|k+1}$. Note that the actual error between $\mathbf{X}_k$ and $\mathbf{x}_k^R$ does not obey the linear equations (5) and (6). Thus, $\hat{\boldsymbol{\delta}}_{k+1|k}$ and $\hat{\boldsymbol{\delta}}_{k+1|k+1}$ are not optimal MMSE estimates of the actual error; we hope, however, that they are still good estimates of this error. How do we obtain an estimate of $\mathbf{X}_{k+1}$ from $\hat{\boldsymbol{\delta}}_{k+1|k}$ or $\hat{\boldsymbol{\delta}}_{k+1|k}$? Since $\hat{\boldsymbol{\delta}}_{k+1|k}$ and $\hat{\boldsymbol{\delta}}_{k+1|k}$ are estimates of the error between $\mathbf{X}_k$ and $\mathbf{x}_k^R$, reasonable estimates of $\mathbf{X}_{k+1}$ are

$$\hat{\mathbf{X}}_{k+1|k} = \mathbf{x}_{k+1}^R + \hat{\boldsymbol{\delta}}_{k+1|k}$$

$$\hat{\mathbf{X}}_{k+1|k+1} = \mathbf{x}_{k+1}^R + \hat{\boldsymbol{\delta}}_{k+1|k+1}$$

How do we choose the reference trajectory? A reasonable value for $\mathbf{x}_0^R$ would be $\mathbf{m}_{X_0}$, the mean of the initial state. Using this value gives

$$\hat{\boldsymbol{\delta}}_{0|0} = E\left[\boldsymbol{\delta}_0\right] = E\left[\mathbf{X}_0 - \mathbf{x}_0^R\right] = 0$$

$$P_{0|0} = P_{X_0}$$

Thus,

$$\hat{\boldsymbol{\delta}}_{1|0} = A_0 \hat{\boldsymbol{\delta}}_{0|0} = 0$$

The error covariance for the estimate $\hat{\boldsymbol{\delta}}_{1|0}$ is

$$P_{1|0} = A_0 P_{0|0} A_0^T + GQG^T$$

The estimate for $\mathbf{X}_1$ is

$$\hat{\mathbf{X}}_{1|0} = \mathbf{x}_1^R + \hat{\boldsymbol{\delta}}_{k+1|k} = \mathbf{f}\left(\mathbf{m}_{X_0}, 0\right)$$

Note that $P_{1|0}$ is also approximately the error covariance of the estimate $\hat{\mathbf{X}}_{1|0}$:

$$\begin{aligned}
P_{1|0} &= E\left[\left(\boldsymbol{\delta}_1 - \hat{\boldsymbol{\delta}}_{1|0}\right)\left(\boldsymbol{\delta}_1 - \hat{\boldsymbol{\delta}}_{1|0}\right)^T\right] \\
&\approx E\left[\left(\mathbf{X}_1 - \mathbf{x}_1^R - \left(\hat{\mathbf{X}}_{1|0} - \mathbf{x}_1^R\right)\right)\left(\mathbf{X}_1 - \mathbf{x}_1^R - \left(\hat{\mathbf{X}}_{1|0} - \mathbf{x}_1^R\right)\right)^T\right] \\
&= E\left[\left(\mathbf{X}_1 - \hat{\mathbf{X}}_{1|0}\right)\left(\mathbf{X}_1 - \hat{\mathbf{X}}_{1|0}\right)^T\right] \qquad (7)
\end{aligned}$$

At time $k = 1$, we process the observation $\mathbf{Z}_1$ to obtain

$$K_1 = P_{1|0} H_1^T \left(H_1 P_{1|0} H_1^T + R\right)^{-1}$$

$$\begin{aligned}
\hat{\boldsymbol{\delta}}_{1|1} &= \hat{\boldsymbol{\delta}}_{1|0} + K_1\left(\boldsymbol{\nu}_1 - H_1 \hat{\boldsymbol{\delta}}_{1|0}\right) \\
&= K_1 \boldsymbol{\nu}_1,
\end{aligned}$$

where we have used the fact that $\hat{\boldsymbol{\delta}}_{1|0} = 0$.

$$P_{1|1} = (I - K_1 H_1) P_{1|0}.$$

The estimate $\hat{\mathbf{X}}_{1|1}$ is obtained as

$$\begin{aligned}
\hat{\mathbf{X}}_{1|1} &= \mathbf{x}_1^R + \hat{\boldsymbol{\delta}}_{1|1} \\
&= \mathbf{x}_1^R + K_1 \boldsymbol{\nu}_1 \\
&= \hat{\mathbf{X}}_{1|0} + K_1 \left[\mathbf{Z}_1 - \mathbf{h}\left(\mathbf{x}_1^R, 1\right)\right] \\
&= \hat{\mathbf{X}}_{1|0} + K_1 \left[\mathbf{Z}_1 - \mathbf{h}\left(\hat{\mathbf{X}}_{1|0}, 1\right)\right]
\end{aligned}$$

Note that $P_{1|1}$ is approximately the error covariance of $\hat{\mathbf{X}}_{1|1}$; this can be shown using a derivation similar to (7).

Now, to compute the estimate for $k = 2$, we could continue to use the reference trajectory that we used for $k = 1$; however, we now have a (hopefully) better estimate of the state at time 1, namely $\hat{\mathbf{X}}_{1|1}$. Thus, to get the reference trajectory for $k = 2$, we set $\mathbf{x}_1^R = \hat{\mathbf{X}}_{1|1}$. With this new reference trajectory, we have $\boldsymbol{\delta}_1^* = \mathbf{X}_1 - \hat{\mathbf{X}}_{1|1}$, where $\boldsymbol{\delta}_1^*$ represents the error with the new reference trajectory. Assuming that $\boldsymbol{\delta}_1^*$ has a mean of zero [2], the estimate of $\boldsymbol{\delta}_1^*$ is $\hat{\boldsymbol{\delta}^*}_{1|1} = 0$ and

$$\hat{\boldsymbol{\delta}}_{2|1} = A_1 \hat{\boldsymbol{\delta}^*}_{1|1} = 0$$

Thus,

$$\hat{\mathbf{X}}_{2|1} = \mathbf{x}_2^R + \hat{\boldsymbol{\delta}}_{2|1} = \mathbf{f}\left(\hat{\mathbf{X}}_{1|1}, 1\right)$$

The approximate error covariance for this estimate is

$$P_{2|1} = A_1 P_{1|1} A_1^T + GQG^T$$

At $k = 2$, we process the observation $\mathbf{Z}_2$ in the same way as we processed the observation $\mathbf{Z}_1$. Further prediction and filter updates follow this same pattern.

To summarize, the Extended Kalman Filter equations are the following:

Predictor Update:

$$\hat{\mathbf{X}}_{k+1|k} = \mathbf{f}\left(\hat{\mathbf{X}}_{k|k}, k\right)$$

$$P_{k+1|k} = A_k P_{k|k} A_k^T + GQG^T$$

$$A_k = \left.\frac{\partial \mathbf{f}(\mathbf{x}, t)}{\partial \mathbf{x}}\right|_{\mathbf{x}=\hat{\mathbf{X}}_{k|k}}.$$

[2] This assumption would be true if the system model were linear; since the system model is not linear, we hope that the assumption is approximately true and that the mean of $\boldsymbol{\delta}_1^*$ is approximately zero.

Filter Update:

$$H_{k+1} = \left.\frac{\partial \mathbf{h}(\mathbf{x}, t_k)}{\partial \mathbf{x}}\right|_{\mathbf{x}=\hat{\mathbf{X}}_{k+1|k}}$$

$$K_{k+1} = P_{k+1|k} H_{k+1}^T \left(H_{k+1} P_{k+1|k} H_{k+1}^T + R\right)^{-1}$$

$$\hat{\mathbf{X}}_{k+1|k+1} = \hat{\mathbf{X}}_{k+1|k} + K_{k+1}\left[\mathbf{Z}_{k+1} - \mathbf{h}\left(\hat{\mathbf{X}}_{k+1|k}, k+1\right)\right].$$

$$P_{k+1|k+1} = (I - K_{k+1} H_{k+1}) P_{k+1|k}$$

Note that in the filter update equation for $\hat{\mathbf{X}}_{k+1|k+1}$, the residual $\mathbf{Z}_{k+1} - \mathbf{h}\left(\hat{\mathbf{X}}_{k+1|k}, k+1\right)$ plays the same role as the innovations in the standard Kalman filter. This residual sequence can be monitored to verify correct filter operation, much in the same way that the innovation sequence can be monitored in the standard Kalman filter.

3 Continuous/Discrete EKF

The development of the continuous/discrete EKF is very similar to the development of the discrete/discrete EKF. The system model that we use is the following continuous/discrete model:

$$\dot{\mathbf{X}}(t) = \mathbf{f}(\mathbf{X}(t), t) + G\mathbf{W}(t) \tag{8}$$

$$\mathbf{Z}(t_k) = \mathbf{h}(\mathbf{X}(t_k), t_k) + \mathbf{V}_k \tag{9}$$

In this model, $\mathbf{W}(t)$ is a continuous-time white noise process with mean zero and intensity Q, $\mathbf{V}_k$ is a discrete-time white noise process with covariance R, and $\mathbf{W}(t)$, $\mathbf{V}_k$, and $\mathbf{X}(t_0)$ are uncorrelated for all t and k.

As in the derivation of the discrete/discrete Kalman filter, we develop the continuous/discrete Extended Kalman filter by starting with a nominal reference trajectory denoted $\mathbf{x}^R(t)$; $\mathbf{x}^R(t)$ is obtained as the solution to the differential equation (8) without the process noise $\mathbf{W}(t)$:

$$\dot{\mathbf{x}}^R(t) = \mathbf{f}\left(\mathbf{x}^R(t), t\right) \tag{10}$$

This differential equation has some initial condition $\mathbf{x}^R(t_0)$. We denote the error between $\mathbf{X}(t)$ and $\mathbf{x}^R(t)$ as

$$\boldsymbol{\delta}(t) = \mathbf{X}(t) - \mathbf{x}^R(t) \tag{11}$$

We now find an approximate linear model that describes the dynamics of $\boldsymbol{\delta}(t)$ by taking the derivative of (11):

$$\begin{aligned}\dot{\boldsymbol{\delta}}(t) &= \dot{\mathbf{X}}(t) - \dot{\mathbf{x}}^R(t) \\ &= \mathbf{f}(\mathbf{X}(t), t) + G\mathbf{W}(t) - \mathbf{f}\left(\mathbf{x}^R(t), t\right)\end{aligned}$$

To obtain a linear approximation of this equation, we make a Taylor series expansion of $\mathbf{f}\,(\mathbf{X}(t), t)$ about the value $\mathbf{x}^R\,(t)$ and drop all but the constant and linear terms:

$$\mathbf{f}\,(\mathbf{X}(t), t) \approx \mathbf{f}\left(\mathbf{x}^R(t), t\right) + \left.\frac{\partial \mathbf{f}\,(\mathbf{x}, t)}{\partial \mathbf{x}}\right|_{\mathbf{x}=\mathbf{x}^R(t)} \left(\mathbf{X}(t) - \mathbf{x}^R\,(t)\right)$$

Denote

$$A(t) = \left.\frac{\partial \mathbf{f}\,(\mathbf{x}, t)}{\partial \mathbf{x}}\right|_{\mathbf{x}=\mathbf{x}^R(t)},$$

and note that $A(t)$ is an $n \times n$ matrix of the following form:

$$A(t) = \begin{bmatrix} \frac{\partial f_1(\mathbf{X},t)}{\partial x_1} & \cdots & \frac{\partial f_1(\mathbf{X},t)}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_n(\mathbf{X},t)}{\partial x_1} & \cdots & \frac{\partial f_n(\mathbf{X},t)}{\partial x_n} \end{bmatrix}_{\mathbf{x}=\mathbf{x}^R(t)}$$

With this definition,

$$\begin{aligned} \dot{\boldsymbol{\delta}}\,(t) &\approx \mathbf{f}\left(\mathbf{x}^R(t), t\right) + A(t)\left(\mathbf{X}(t) - \mathbf{x}^R\,(t)\right) + G\mathbf{W}(t) - \mathbf{f}\left(\mathbf{x}^R(t), t\right) \\ &= A(t)\boldsymbol{\delta}\,(t) + G\mathbf{W}(t) \end{aligned} \tag{12}$$

Note that (12) is now a linear differential equation.

We follow a similar approach to obtain an (approximate) linear relationship from (9). We expand $\mathbf{h}(\mathbf{X}(t_k), t_k)$ in a Taylor series about the nominal trajectory $\mathbf{x}^R\,(t_k)$:

$$\mathbf{h}(\mathbf{X}(t_k), t_k) \approx \mathbf{h}(\mathbf{x}^R\,(t_k)\,, t_k) + \left.\frac{\partial \mathbf{h}\,(\mathbf{x}, t_k)}{\partial \mathbf{x}}\right|_{\mathbf{x}=\mathbf{x}^R(t_k)} \left(\mathbf{X}(t_k) - \mathbf{x}^R\,(t_k)\right)$$

We define H_k as

$$H_k = \left.\frac{\partial \mathbf{h}\,(\mathbf{x}, t_k)}{\partial \mathbf{x}}\right|_{\mathbf{x}=\mathbf{x}^R(t_k)}$$

Then we can write (9) as

$$\mathbf{Z}\,(t_k) \approx \mathbf{h}(\mathbf{x}^R\,(t_k)\,, t_k) + H_k\left(\mathbf{X}(t_k) - \mathbf{x}^R\,(t_k)\right) + \mathbf{V}_k.$$

Define $\boldsymbol{\nu}(t_k)$ as

$$\begin{aligned} \boldsymbol{\nu}(t_k) &= \mathbf{Z}\,(t_k) - \mathbf{h}(\mathbf{x}^R\,(t_k)\,, t_k) \\ &\approx \mathbf{h}(\mathbf{x}^R\,(t_k)\,, t_k) + H_k\left(\mathbf{X}(t_k) - \mathbf{x}^R\,(t_k)\right) + \mathbf{V}_k - \mathbf{h}(\mathbf{x}^R\,(t_k)\,, t_k) \\ &= H_k\left(\mathbf{X}(t_k) - \mathbf{x}^R\,(t_k)\right) + \mathbf{V}_k \\ &= H_k\boldsymbol{\delta}(t_k) + \mathbf{V}_k \end{aligned} \tag{13}$$

So we have an (approximate) linear relationship between the quantity $\boldsymbol{\nu}(t_k)$ (which can be computed from $\mathbf{Z}_k$) and $\boldsymbol{\delta}(t_k)$.

We can now use a continuous/discrete Kalman filter to compute $\hat{\boldsymbol{\delta}}\,(t|k)$ and $P\,(t|k)$ using the linear dynamics equation (12) and the linear observation equation (13). An estimate of $\mathbf{X}(t)$ from $\hat{\boldsymbol{\delta}}\,(t|k)$ is obtained as

$$\hat{\mathbf{X}}\,(t|k) = \mathbf{x}^R\,(t) + \hat{\boldsymbol{\delta}}_X\,(t|k)\,.$$

How do we choose the reference trajectory? Let us first consider the interval $t_0 \le t \le t_1$; a reasonable value for $\mathbf{x}^R\,(t_0)$ would be $\mathbf{m}_X\,(t_0) = E\,[\mathbf{X}\,(t_0)]$. Using this value gives

$$E\,[\boldsymbol{\delta}\,(t_0)] = 0$$

$$P\,(t_0|0) = P_{X_0}$$

Thus, for $t_0 \le t \le t_1$, $\hat{\boldsymbol{\delta}}\,(t|0) = 0$. The estimate for $\mathbf{X}\,(t)$ is

$$\hat{\mathbf{X}}\,(t|0) = \mathbf{x}^R\,(t)\,,$$

where $\mathbf{x}^R\,(t)$ is the solution to the differential equation (10) with an initial condition of $\mathbf{m}_X\,(t_0)$. The error covariance is

$$P\,(t|0) = \phi(t,t_0)P\,(t_0|0)\,\phi^T(t,t_0) + \int_{t_0}^{t} \phi(t,\tau)GQG^T\phi^T(t,\tau)d\tau,$$

where $\phi(t,\tau)$ is the state transition matrix of the linear system model (12).

At time t_1, we process the observation $\mathbf{Z}\,(t_1)$ to obtain

$$K_1 = P\,(t_1|0)\,H_1^T\left(H_1 P\,(t_1|0)\,H_1^T + R\right)^{-1}$$

$$\begin{aligned}\hat{\boldsymbol{\delta}}\,(t_1|1) &= \hat{\boldsymbol{\delta}}\,(t_1|0) + K_1\left(\boldsymbol{\nu}(t_1) - H_1\hat{\boldsymbol{\delta}}\,(t_1|0)\right)\\ &= K_1\boldsymbol{\nu}(t_1),\end{aligned}$$

where we have used the fact that $\hat{\boldsymbol{\delta}}\,(t_1|0) = 0$.

$$P_\delta\,(t_1|1) = (I - K_1H_1)\,P_\delta\,(t_1|0)\,.$$

The estimate $\hat{\mathbf{X}}\,(t_1|1)$ is obtained as

$$\hat{\mathbf{X}}(t_1|1) = \mathbf{x}^R(t_1) + K_1\boldsymbol{\nu}(t_1) = \hat{\mathbf{X}}(t_1|0) + K_1\left(\mathbf{Z}\,(t_1) - \mathbf{h}\left(\hat{\mathbf{X}}\,(t_1|0)\,,t_1\right)\right).$$

Now, to obtain $\hat{\mathbf{X}}(t|1)$ for $t_1 \le t \le t_2$, we could continue to use the reference trajectory that we used for the interval $[t_0, t_1]$; however, we now have a (hopefully) better estimate of the state at time t_1, namely $\hat{\mathbf{X}}\,(t_1|1)$. Thus, to get the reference trajectory for $t_1 \le t \le t_2$, we find the solution of (10) with initial condition $\mathbf{x}^R\,(t_1) = \hat{\mathbf{X}}\,(t_1|1)$. If our dynamics and observation models for the linearized error system were exact, we would expect $E[\boldsymbol{\delta}(t)] = 0$; making the assumption that the mean of the error is small, we get

$$\hat{\mathbf{X}}(t|1) = \mathbf{x}^R(t),$$

where $\mathbf{x}^R(t)$ is the solution of the differential equation (10) with an initial condition of $\mathbf{x}^R(t_1) = \hat{\mathbf{X}}(t_1|1)$. The error covariance for this estimate is

$$P(t|1) = \phi(t,t_1)P(t_1|1)\,\phi^T(t,t_1) + \int_{t_1}^{t} \phi(t,\tau)GQG^T\phi^T(t,\tau)d\tau.$$

At t_2, we process the observation $\mathbf{Z}(t_2)$ in the same way as we processed the observation $\mathbf{Z}(t_1)$.

To summarize, the Extended Kalman Filter equations are the following:

Predictor Step: For $t_k \leq t \leq t_{k+1}$, $\hat{\mathbf{X}}(t|k)$ is the solution of (10) with $\hat{\mathbf{X}}(t_k|k)$ as the initial condition. The corresponding error covariance update is

$$P(t|k) = \phi(t,t_k)P(t_k|k)\,\phi^T(t,t_k) + \int_{t_k}^{t} \phi(t,\tau)GQG^T\phi^T(t,\tau)d\tau,$$

where $\phi(t,\tau)$ is the state transition matrix obtained from

$$A(t) = \left.\frac{\partial \mathbf{f}(\mathbf{x},t)}{\partial \mathbf{x}}\right|_{\mathbf{x}=\hat{\mathbf{X}}(t|k)}.$$

Filter Step At t_{k+1}, the filter step is the following:

$$H_{k+1} = \left.\frac{\partial \mathbf{h}(\mathbf{x},t_{k+1})}{\partial \mathbf{x}}\right|_{\mathbf{x}=\hat{\mathbf{X}}(t_{k+1}|k)}$$

$$K_{k+1} = P(t_{k+1}|k)\,H_{k+1}^T\left(H_{k+1}P(t_{k+1}|k)\,H_{k+1}^T + R\right)^{-1}$$

$$\hat{\mathbf{X}}(t_{k+1}|k+1) = \hat{\mathbf{X}}(t_{k+1}|k) + K_{k+1}\left[\mathbf{Z}(t_{k+1}) - \mathbf{h}\left(\hat{\mathbf{X}}(t_{k+1}|k)\right)\right].$$

$$(t_{k+1}|k+1) = (I - K_{k+1}H_{k+1})\,P(t_{k+1}|k)$$