Extended Kalman Filter Tutorial

Gabriel A. Terejanu
Department of Computer Science and Engineering
University at Buffalo, Buffalo, NY 14260
terejanu@buffalo.edu

1 Dynamic process

Consider the following nonlinear system, described by the difference equation and the observation model with additive noise:

$$\mathbf{x}_k = \mathbf{f}(\mathbf{x}_{k-1}) + \mathbf{w}_{k-1} \tag{1}$$
$$\mathbf{z}_k = \mathbf{h}(\mathbf{x}_k) + \mathbf{v}_k \tag{2}$$

The **initial state** $\mathbf{x}_0$ is a random vector with known mean $\mu_0 = E[\mathbf{x}_0]$ and covariance $\mathbf{P}_0 = E[(\mathbf{x}_0 - \mu_0)(\mathbf{x}_0 - \mu_0)^T]$.

In the following we assume that the random vector $\mathbf{w}_k$ captures uncertainties in the model and $\mathbf{v}_k$ denotes the measurement noise. Both are temporally uncorrelated (white noise), zero-mean random sequences with known covariances and both of them are uncorrelated with the initial state $\mathbf{x}_0$.

$$E[\mathbf{w}_k] = 0 \qquad E[\mathbf{w}_k\mathbf{w}_k^T] = \mathbf{Q}_k \qquad E[\mathbf{w}_k\mathbf{w}_j^T] = 0 \text{ for } k \neq j \qquad E[\mathbf{w}_k\mathbf{x}_0^T] = 0 \text{ for all } k \tag{3}$$
$$E[\mathbf{v}_k] = 0 \qquad E[\mathbf{v}_k\mathbf{v}_k^T] = \mathbf{R}_k \qquad E[\mathbf{v}_k\mathbf{v}_j^T] = 0 \text{ for } k \neq j \qquad E[\mathbf{v}_k\mathbf{x}_0^T] = 0 \text{ for all } k \tag{4}$$

Also the two random vectors $\mathbf{w}_k$ and $\mathbf{v}_k$ are uncorrelated:

$$E[\mathbf{w}_k\mathbf{v}_j^T] = 0 \text{ for all } k \text{ and } j \tag{5}$$

Vectorial functions $\mathbf{f}(\cdot)$ and $\mathbf{h}(\cdot)$ are assumed to be C^1 functions (the function and its first derivative are continuous on the given domain).

Dimension and description of variables:

$\mathbf{x}_k$	$n \times 1$	– State vector
$\mathbf{w}_k$	$n \times 1$	– Process noise vector
$\mathbf{z}_k$	$m \times 1$	– Observation vector
$\mathbf{v}_k$	$m \times 1$	– Measurement noise vector
$\mathbf{f}(\cdot)$	$n \times 1$	– Process nonlinear vector function
$\mathbf{h}(\cdot)$	$m \times 1$	– Observation nonlinear vector function
$\mathbf{Q}_k$	$n \times n$	– Process noise covariance matrix
$\mathbf{R}_k$	$m \times m$	– Measurement noise covariance matrix

2 EKF derivation

Assuming the nonlinearities in the dynamic and the observation model are smooth, we can expand $\mathbf{f}(\mathbf{x}_k)$ and $\mathbf{h}(\mathbf{x}_k)$ in Taylor Series and approximate this way the forecast and the next estimate of $\mathbf{x}_k$.

Model Forecast Step
Initially, since the only available information is the mean, μ_0, and the covariance, $\mathbf{P}_0$, of the initial state then the initial optimal estimate $\mathbf{x}_0^a$ and error covariance is:

$$\begin{aligned} \mathbf{x}_0^a &= \mu_0 = E[\mathbf{x}_0] && (6) \\ \mathbf{P}_0 &= E[(\mathbf{x}_0 - \mathbf{x}_0^a)(\mathbf{x}_0 - \mathbf{x}_0^a)^T] && (7) \end{aligned}$$

Assume now that we have an **optimal estimate** $\mathbf{x}_{k-1}^a \equiv E[\mathbf{x}_{k-1}|\mathbf{Z}_{k-1}]$ with $\mathbf{P}_{k-1}$ covariance at time $k-1$. The predictable part of $\mathbf{x}_k$ is given by:

$$\begin{aligned} \mathbf{x}_k^f &\equiv E[\mathbf{x}_k|\mathbf{Z}_{k-1}] && (8) \\ &= E[\mathbf{f}(\mathbf{x}_{k-1}) + \mathbf{w}_{k-1}|\mathbf{Z}_{k-1}] \\ &= E[\mathbf{f}(\mathbf{x}_{k-1})|\mathbf{Z}_{k-1}] \end{aligned}$$

Expanding $\mathbf{f}(\cdot)$ in Taylor Series about $\mathbf{x}_{k-1}^a$ we get:

$$\mathbf{f}(\mathbf{x}_{k-1}) \equiv \mathbf{f}(\mathbf{x}_{k-1}^a) + \mathbf{J_f}(\mathbf{x}_{k-1}^a)(\mathbf{x}_{k-1} - \mathbf{x}_{k-1}^a) + \mathbf{H.O.T.} \quad (9)$$

where $\mathbf{J_f}$ is the Jacobian of $\mathbf{f}(\cdot)$ and the higher order terms (**H.O.T.**) are considered negligible. Hence, the Extended Kalman Filter is also called the First-Order Filter. The Jacobian is defined as:

$$\mathbf{J_f} \equiv \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \cdots & \frac{\partial f_n}{\partial x_n} \end{bmatrix} \quad (10)$$

where $\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_n(\mathbf{x}))^T$ and $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T$. The eq.(9) becomes:

$$\mathbf{f}(\mathbf{x}_{k-1}) \approx \mathbf{f}(\mathbf{x}_{k-1}^a) + \mathbf{J_f}(\mathbf{x}_{k-1}^a)\mathbf{e}_{k-1} \quad (11)$$

where $\mathbf{e}_{k-1} \equiv \mathbf{x}_{k-1} - \mathbf{x}_{k-1}^a$. The expectated value of $\mathbf{f}(\mathbf{x}_{k-1})$ conditioned by $\mathbf{Z}_{k-1}$:

$$E[\mathbf{f}(\mathbf{x}_{k-1})|\mathbf{Z}_{k-1}] \approx \mathbf{f}(\mathbf{x}_{k-1}^a) + \mathbf{J_f}(\mathbf{x}_{k-1}^a)E[\mathbf{e}_{k-1}|\mathbf{Z}_{k-1}] \quad (12)$$

where $E[\mathbf{e}_{k-1}|\mathbf{Z}_{k-1}] = 0$. Thus the forecast value of $\mathbf{x}_k$ is:

$$\mathbf{x}_k^f \approx \mathbf{f}(\mathbf{x}_{k-1}^a) \quad (13)$$

Substituting (11) in the forecast error equation results:

$$\begin{aligned} \mathbf{e}_k^f &\equiv \mathbf{x}_k - \mathbf{x}_k^f && (14) \\ &= \mathbf{f}(\mathbf{x}_{k-1}) + \mathbf{w}_{k-1} - \mathbf{f}(\mathbf{x}_{k-1}^a) \\ &\approx \mathbf{J_f}(\mathbf{x}_{k-1}^a)\mathbf{e}_{k-1} + \mathbf{w}_{k-1} \end{aligned}$$

The forecast error covariance is given by:

$$\begin{aligned}
\mathbf{P}_k^f &\equiv E[\mathbf{e}_k^f(\mathbf{e}_k^f)^T] \qquad (15)\\
&= \mathbf{J_f}(\mathbf{x}_{k-1}^a)E[\mathbf{e}_{k-1}\mathbf{e}_{k-1}^T]\mathbf{J_f}^T(\mathbf{x}_{k-1}^a) + E[\mathbf{w}_{k-1}\mathbf{w}_{k-1}^T]\\
&= \mathbf{J_f}(\mathbf{x}_{k-1}^a)\mathbf{P}_{k-1}\mathbf{J_f}^T(\mathbf{x}_{k-1}^a) + \mathbf{Q}_{k-1}
\end{aligned}$$

Data Assimilation Step

At time k we have two pieces of information: the forecast value $\mathbf{x}_k^f$ with the covariance $\mathbf{P}_k^f$ and the measurement $\mathbf{z}_k$ with the covariance $\mathbf{R}_k$. Our goal is to approximate the best unbiased estimate, in the least squares sense, $\mathbf{x}_k^a$ of $\mathbf{x}_k$. One way is to assume the estimate is a linear combination of both $\mathbf{x}_k^f$ and $\mathbf{z}_k$ [4]. Let:

$$\mathbf{x}_k^a = \mathbf{a} + \mathbf{K}_k\mathbf{z}_k \qquad (16)$$

From the unbiasedness condition:

$$\begin{aligned}
0 &= E[\mathbf{x}_k - \mathbf{x}_k^a|\mathbf{Z}_k] \qquad (17)\\
&= E[(\mathbf{x}_k^f + \mathbf{e}_k^f) - (\mathbf{a} + \mathbf{K}_k\mathbf{h}(\mathbf{x}_k) + \mathbf{K}_k\mathbf{v}_k)|\mathbf{Z}_k]\\
&= \mathbf{x}_k^f - \mathbf{a} - \mathbf{K}_kE[\mathbf{h}(\mathbf{x}_k)|\mathbf{Z}_k]\\
\mathbf{a} &= \mathbf{x}_k^f - \mathbf{K}_kE[\mathbf{h}(\mathbf{x}_k)|\mathbf{Z}_k] \qquad (18)
\end{aligned}$$

Substitute (18) in (16):

$$\mathbf{x}_k^a = \mathbf{x}_k^f + \mathbf{K}_k(\mathbf{z}_k - E[\mathbf{h}(\mathbf{x}_k)|\mathbf{Z}_k]) \qquad (19)$$

Following the same steps as in model forecast step, expanding $\mathbf{h}(\cdot)$ in Taylor Series about $\mathbf{x}_k^f$ we have:

$$\mathbf{h}(\mathbf{x}_k) \equiv \mathbf{h}(\mathbf{x}_k^f) + \mathbf{J_h}(\mathbf{x}_k^f)(\mathbf{x}_k - \mathbf{x}_k^f) + \mathbf{H.O.T.} \qquad (20)$$

where $\mathbf{J_h}$ is the Jacobian of $\mathbf{h}(\cdot)$ and the higher order terms (**H.O.T.**) are considered negligible. The Jacobian of $\mathbf{h}(\cdot)$ is defined as:

$$\mathbf{J_h} \equiv \begin{bmatrix} \frac{\partial h_1}{\partial x_1} & \frac{\partial h_1}{\partial x_2} & \cdots & \frac{\partial h_1}{\partial x_n} \\ \vdots & & \ddots & \vdots \\ \frac{\partial h_m}{\partial x_1} & \frac{\partial h_m}{\partial x_2} & \cdots & \frac{\partial h_m}{\partial x_n} \end{bmatrix} \qquad (21)$$

where $\mathbf{h}(\mathbf{x}) = (h_1(\mathbf{x}), h_2(\mathbf{x}), \ldots, h_m(\mathbf{x}))^T$ and $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T$. Taken the expectation on both sides of (20) conditioned by $\mathbf{Z}_k$:

$$E[\mathbf{h}(\mathbf{x}_k)|\mathbf{Z}_k] \approx \mathbf{h}(\mathbf{x}_k^f) + \mathbf{J_h}(\mathbf{x}_k^f)E[\mathbf{e}_k^f|\mathbf{Z}_k] \qquad (22)$$

where $E[\mathbf{e}_k^f|\mathbf{Z}_k] = 0$. Substitute in (19), the state estimate is:

$$\mathbf{x}_k^a \approx \mathbf{x}_k^f + \mathbf{K}_k(\mathbf{z}_k - \mathbf{h}(\mathbf{x}_k^f)) \qquad (23)$$

The error in the estimate $\mathbf{x}_k^a$ is:

$$\begin{aligned}
\mathbf{e}_k &\equiv \mathbf{x}_k - \mathbf{x}_k^a \\
&= \mathbf{f}(\mathbf{x}_{k-1}) + \mathbf{w}_{k-1} - \mathbf{x}_k^f - \mathbf{K}_k(\mathbf{z}_k - \mathbf{h}(\mathbf{x}_k^f)) \\
&\approx \mathbf{f}(\mathbf{x}_{k-1}) - \mathbf{f}(\mathbf{x}_{k-1}^a) + \mathbf{w}_{k-1} - \mathbf{K}_k(\mathbf{h}(\mathbf{x}_k) - \mathbf{h}(\mathbf{x}_k^f) + \mathbf{v}_k) \\
&\approx \mathbf{J_f}(\mathbf{x}_{k-1}^a)\mathbf{e}_{k-1} + \mathbf{w}_{k-1} - \mathbf{K}_k(\mathbf{J_h}(\mathbf{x}_k^f)\mathbf{e}_k^f + \mathbf{v}_k) \\
&\approx \mathbf{J_f}(\mathbf{x}_{k-1}^a)\mathbf{e}_{k-1} + \mathbf{w}_{k-1} - \mathbf{K}_k\mathbf{J_h}(\mathbf{x}_k^f)(\mathbf{J_f}(\mathbf{x}_{k-1}^a)\mathbf{e}_{k-1} + \mathbf{w}_{k-1}) - \mathbf{K}_k\mathbf{v}_k \\
&\approx (\mathbf{I} - \mathbf{K}_k\mathbf{J_h}(\mathbf{x}_k^f))\mathbf{J_f}(\mathbf{x}_{k-1}^a)\mathbf{e}_{k-1} + (\mathbf{I} - \mathbf{K}_k\mathbf{J_h}(\mathbf{x}_k^f))\mathbf{w}_{k-1} - \mathbf{K}_k\mathbf{v}_k
\end{aligned} \tag{24}$$

Then, the posterior covariance of the new estimate is:

$$\begin{aligned}
\mathbf{P}_k &\equiv E[\mathbf{e}_k\mathbf{e}_k^T] \\
&= (\mathbf{I} - \mathbf{K}_k\mathbf{J_h}(\mathbf{x}_k^f))\mathbf{J_f}(\mathbf{x}_{k-1}^a)\mathbf{P}_{k-1}\mathbf{J_f}^T(\mathbf{x}_{k-1}^a)(\mathbf{I} - \mathbf{K}_k\mathbf{J_h}(\mathbf{x}_k^f))^T \\
&\quad + (\mathbf{I} - \mathbf{K}_k\mathbf{J_h}(\mathbf{x}_k^f))\mathbf{Q}_{k-1}(\mathbf{I} - \mathbf{K}_k\mathbf{J_h}(\mathbf{x}_k^f))^T + \mathbf{K}_k\mathbf{R}_k\mathbf{K}_k^T \\
&= (\mathbf{I} - \mathbf{K}_k\mathbf{J_h}(\mathbf{x}_k^f))\mathbf{P}_k^f(\mathbf{I} - \mathbf{K}_k\mathbf{J_h}(\mathbf{x}_k^f))^T + \mathbf{K}_k\mathbf{R}_k\mathbf{K}_k^T \\
&= \mathbf{P}_k^f - \mathbf{K}_k\mathbf{J_h}(\mathbf{x}_k^f)\mathbf{P}_k^f - \mathbf{P}_k^f\mathbf{J_h}^T(\mathbf{x}_k^f)\mathbf{K}_k^T + \mathbf{K}_k\mathbf{J_h}(\mathbf{x}_k^f)\mathbf{P}_k^f\mathbf{J_h}^T(\mathbf{x}_k^f)\mathbf{K}_k^T + \mathbf{K}_k\mathbf{R}_k\mathbf{K}_k^T
\end{aligned} \tag{25}$$

The posterior covariance formula holds for any $\mathbf{K}_k$. Like in the standard Kalman Filter we find out $\mathbf{K}_k$ by minimizing $tr(\mathbf{P}_k)$ w.r.t. $\mathbf{K}_k$.

$$\begin{aligned}
0 &= \frac{\partial tr(\mathbf{P}_k)}{\partial \mathbf{K}_k} \\
&= -(\mathbf{J_h}(\mathbf{x}_k^f)\mathbf{P}_k^f)^T - \mathbf{P}_k^f\mathbf{J_h}^T(\mathbf{x}_k^f) + 2\mathbf{K}_k\mathbf{J_h}(\mathbf{x}_k^f)\mathbf{P}_k^f\mathbf{J_h}^T(\mathbf{x}_k^f) + 2\mathbf{K}_k\mathbf{R}_k
\end{aligned} \tag{26}$$

Hence the Kalman gain is:

$$\mathbf{K}_k = \mathbf{P}_k^f\mathbf{J_h}^T(\mathbf{x}_k^f)\left(\mathbf{J_h}(\mathbf{x}_k^f)\mathbf{P}_k^f\mathbf{J_h}^T(\mathbf{x}_k^f) + \mathbf{R}_k\right)^{-1} \tag{27}$$

Substituting this back in (25) results:

$$\begin{aligned}
\mathbf{P}_k &= (\mathbf{I} - \mathbf{K}_k\mathbf{J_h}(\mathbf{x}_k^f))\mathbf{P}_k^f - (\mathbf{I} - \mathbf{K}_k\mathbf{J_h}(\mathbf{x}_k^f))\mathbf{P}_k^f\mathbf{J_h}^T(\mathbf{x}_k^f)\mathbf{K}_k^T + \mathbf{K}_k\mathbf{R}_k\mathbf{K}_k^T \\
&= (\mathbf{I} - \mathbf{K}_k\mathbf{J_h}(\mathbf{x}_k^f))\mathbf{P}_k^f - \left(\mathbf{P}_k^f\mathbf{J_h}^T(\mathbf{x}_k^f) - \mathbf{K}_k\mathbf{J_h}(\mathbf{x}_k^f)\mathbf{P}_k^f\mathbf{J_h}^T(\mathbf{x}_k^f) - \mathbf{K}_k\mathbf{R}_k\right)\mathbf{K}_k^T \\
&= (\mathbf{I} - \mathbf{K}_k\mathbf{J_h}(\mathbf{x}_k^f))\mathbf{P}_k^f - \left[\mathbf{P}_k^f\mathbf{J_h}^T(\mathbf{x}_k^f) - \mathbf{K}_k\left(\mathbf{J_h}(\mathbf{x}_k^f)\mathbf{P}_k^f\mathbf{J_h}^T(\mathbf{x}_k^f) + \mathbf{R}_k\right)\right]\mathbf{K}_k^T \\
&= (\mathbf{I} - \mathbf{K}_k\mathbf{J_h}(\mathbf{x}_k^f))\mathbf{P}_k^f - \left[\mathbf{P}_k^f\mathbf{J_h}^T(\mathbf{x}_k^f) - \mathbf{P}_k^f\mathbf{J_h}^T(\mathbf{x}_k^f)\right]\mathbf{K}_k^T \\
&= (\mathbf{I} - \mathbf{K}_k\mathbf{J_h}(\mathbf{x}_k^f))\mathbf{P}_k^f
\end{aligned} \tag{28}$$

3 Summary of Extended Kalman Filter

Model and Observation:

$$\begin{aligned}
\mathbf{x}_k &= \mathbf{f}(\mathbf{x}_{k-1}) + \mathbf{w}_{k-1} \\
\mathbf{z}_k &= \mathbf{h}(\mathbf{x}_k) + \mathbf{v}_k
\end{aligned}$$

Initialization:

$$\mathbf{x}_0^a = \mu_0 \text{ with error covariance } \mathbf{P}_0$$

Model Forecast Step/Predictor:

$$\begin{aligned}
\mathbf{x}_k^f &\approx \mathbf{f}(\mathbf{x}_{k-1}^a) \\
\mathbf{P}_k^f &= \mathbf{J_f}(\mathbf{x}_{k-1}^a)\mathbf{P}_{k-1}\mathbf{J_f}^T(\mathbf{x}_{k-1}^a) + \mathbf{Q}_{k-1}
\end{aligned}$$

Data Assimilation Step/Corrector:

$$\begin{aligned}
\mathbf{x}_k^a &\approx \mathbf{x}_k^f + \mathbf{K}_k(\mathbf{z}_k - \mathbf{h}(\mathbf{x}_k^f)) \\
\mathbf{K}_k &= \mathbf{P}_k^f\mathbf{J_h}^T(\mathbf{x}_k^f)\left(\mathbf{J_h}(\mathbf{x}_k^f)\mathbf{P}_k^f\mathbf{J_h}^T(\mathbf{x}_k^f) + \mathbf{R}_k\right)^{-1} \\
\mathbf{P}_k &= \left(\mathbf{I} - \mathbf{K}_k\mathbf{J_h}(\mathbf{x}_k^f)\right)\mathbf{P}_k^f
\end{aligned}$$

4 Iterated Extended Kalman Filter

In the EKF, $\mathbf{h}(\cdot)$ is linearized about the predicted state estimate $\mathbf{x}_k^f$. The IEKF tries to linearize it about the most recent estimate, improving this way the accuracy [3, 1]. This is achieved by calculating $\mathbf{x}_k^a$, $\mathbf{K}_k$, $\mathbf{P}_k$ at each iteration.

Denote $\mathbf{x}_{k,i}^a$ the estimate at time k and ith iteration. The iteration process is initialized with $\mathbf{x}_{k,0}^a = \mathbf{x}_k^f$. Then the measurement update step becomes for each i:

$$\begin{aligned}
\mathbf{x}_{k,i}^a &\approx \mathbf{x}_k^f + \mathbf{K}_k(\mathbf{z}_k - \mathbf{h}(\mathbf{x}_{k,i}^a)) \\
\mathbf{K}_{k,i} &= \mathbf{P}_k^f\mathbf{J_h}^T(\hat{\mathbf{x}}_{k,i})\left(\mathbf{J_h}(\mathbf{x}_{k,i}^a)\mathbf{P}_k^f\mathbf{J_h}^T(\mathbf{x}_{k,i}^a) + \mathbf{R}_k\right)^{-1} \\
\mathbf{P}_{k,i} &= \left(\mathbf{I} - \mathbf{K}_{k,i}\mathbf{J_h}(\mathbf{x}_{k,i}^a)\right)\mathbf{P}_k^f
\end{aligned}$$

If there is little improvement between two consecutive iterations then the iterative process is stopped. The accuracy reached this way is achieved with higher computational time.

5 Stability

Since $\mathbf{Q}_k$ and $\mathbf{R}_k$ are symmetric positive definite matrices then we can write:

$$\mathbf{Q}_k = \mathbf{G}_k\mathbf{G}_k^T \tag{29}$$
$$\mathbf{R}_k = \mathbf{D}_k\mathbf{D}_k^T \tag{30}$$

Denote by φ and χ the high order terms resulted in the following subtractions:

$$\mathbf{f}(\mathbf{x}_k) - \mathbf{f}(\mathbf{x}_k^a) = \mathbf{J_f}(\mathbf{x}_k^a)\mathbf{e}_k + \varphi(\mathbf{x}_k, \mathbf{x}_k^a) \tag{31}$$
$$\mathbf{h}(\mathbf{x}_k) - \mathbf{h}(\mathbf{x}_k^a) = \mathbf{J_h}(\mathbf{x}_k^a)\mathbf{e}_k + \chi(\mathbf{x}_k, \mathbf{x}_k^a) \tag{32}$$

Konrad Reif showed in [2] that the estimation error remains bounded if the followings hold:

1. $\alpha, \beta, \gamma_1, \gamma_2 > 0$ are positive real numbers and for every k:

$$\begin{aligned} \|\mathbf{J_f}(\mathbf{x}_k^a)\| &\leq \alpha & (33) \\ \|\mathbf{J_h}(\mathbf{x}_k^a)\| &\leq \beta & (34) \\ \gamma_1 \mathbf{I} &\leq \mathbf{P}_k \leq \gamma_2 & (35) \end{aligned}$$

2. $\mathbf{J_f}$ is nonsingular for every k

3. There are positive real numbers $\epsilon_\varphi, \epsilon_\chi, \kappa_\varphi, \kappa_\chi > 0$ such that the nonlinear functions φ, χ are bounded via:

$$\begin{aligned} \|\varphi(\mathbf{x}_k, \mathbf{x}_k^a)\| &\leq \epsilon_\varphi \|\mathbf{x}_k - \mathbf{x}_k^a\|^2 \text{ with } \|\mathbf{x}_k - \mathbf{x}_k^a\| \leq \kappa_\varphi & (36) \\ \|\chi(\mathbf{x}_k, \mathbf{x}_k^a)\| &\leq \epsilon_\chi \|\mathbf{x}_k - \mathbf{x}_k^a\|^2 \text{ with } \|\mathbf{x}_k - \mathbf{x}_k^a\| \leq \kappa_\chi & (37) \end{aligned}$$

Then the estimation error $\mathbf{e}_k$ is exponentially bounded in mean square and bounded with probability one, provided that the initial estimation error satisfies:

$$\|\mathbf{e}_k\| \leq \epsilon \quad (38)$$

and the covariance matrices of the noise terms are bounded via:

$$\mathbf{G}_k \mathbf{G}_k^T \leq \delta \mathbf{I} \quad (39)$$

$$\mathbf{D}_k \mathbf{D}_k^T \leq \delta \mathbf{I} \quad (40)$$

for some $\epsilon, \delta > 0$.

6 Conclusion

In EKF the state distribution is propagated analytically through the first-order linearization of the nonlinear system. It does not take into account that $\mathbf{x}_k$ is a random variable with inherent uncertainty and it requires that the first two terms of the Taylor series to dominate the remaining terms.

Second-Order version exists [4, 5], but the computational complexity required makes it unfeasible for practical usage in cases of real time applications or high dimensional systems.

References

[1] Arthur Gelb. *Applied Optimal Estimation*. M.I.T. Press, 1974.

[2] K.Reif, S.Gunther, E.Yaz, and R.Unbehauen. Stochastic Stability of the Discrete-Time Extended Kalman Filter. *IEEE Trans.Automatic Control*, 1999.

[3] Tine Lefebvre and Herman Bruyninckx. Kalman Filters for Nonlinear Systems: A Comparison of Performance.

[4] John M. Lewis and S.Lakshmivarahan. *Dynamic Data Assimilation, a Least Squares Approach*. 2006.

[5] R. van der Merwe. Sigma-Point Kalman Filters for Probabilistic Inference in Dynamic State-Space Models. Technical report, 2003.

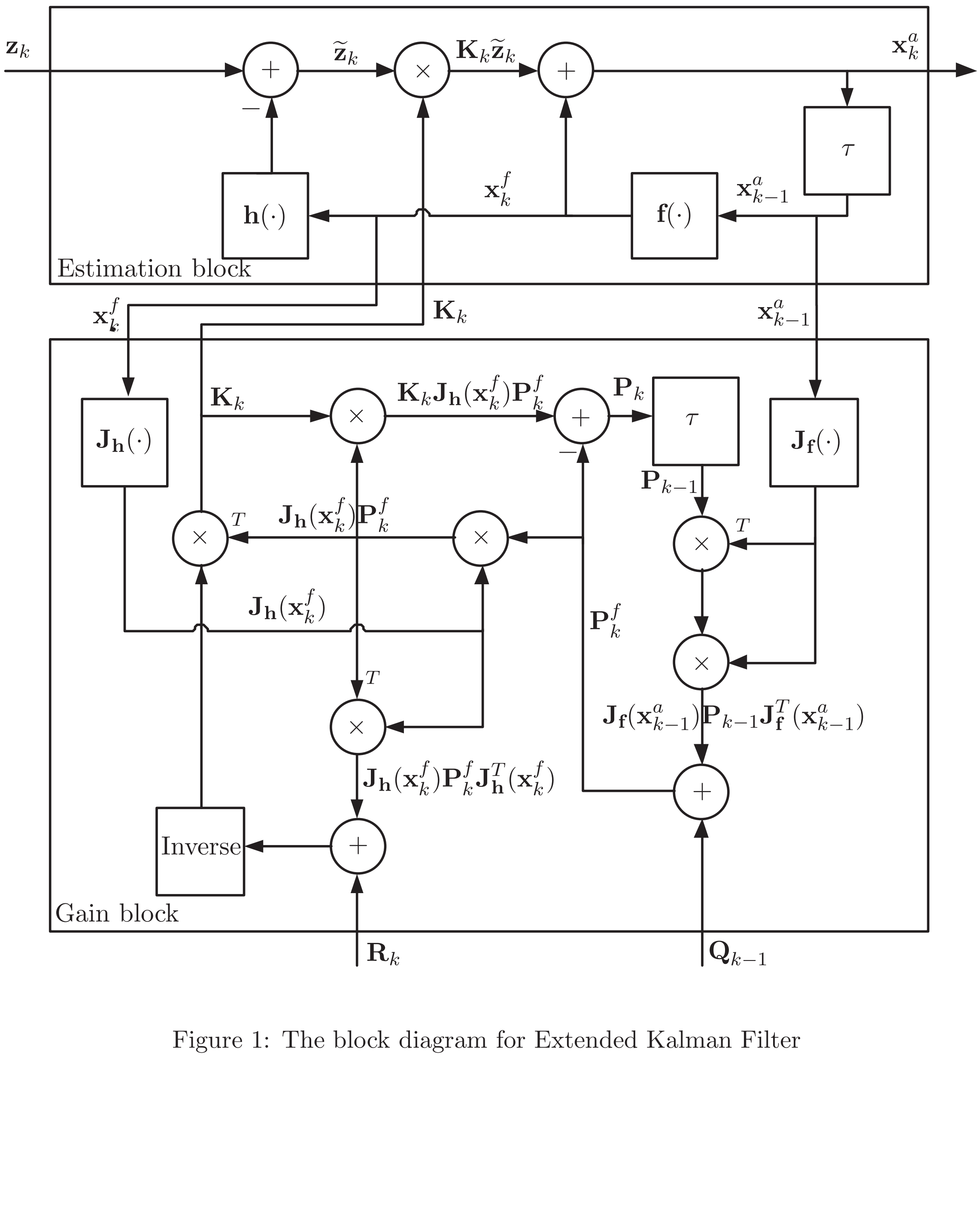

Figure 1: The block diagram for Extended Kalman Filter