**Freescale Semiconductor**
APPLICATION NOTE

Document Number: AN5018
Rev. 1.0, 9/2015

# Basic Kalman Filter Theory

**Mark Pedley and Michael Stanley**

**Contents**



## 1 Introduction

This document derives the standard Kalman filter equations. It is intended as a primer that should be read before tackling Application Note AN5023, which describes the more specialized indirect complementary Kalman filter used for the fusion of accelerometer, magnetometer and gyroscope data in the *Freescale Sensor Fusion Library* software.

Section 2 calculates some mathematical results used in the derivation. The derivation itself is in Section 3.

# 1.1 Terminology

| Symbol | Definition |
|---|---|
| $\boldsymbol{A}_k$ | The linear prediction or state matrix at sample $k$.<br>$\boldsymbol{x}_k = \boldsymbol{A}_k \boldsymbol{x}_{k-1} + \boldsymbol{w}_k$<br>$\hat{\boldsymbol{x}}_k^- = \boldsymbol{A}_k \hat{\boldsymbol{x}}_{k-1}^+$ |
| $\boldsymbol{C}_k$ | The measurement matrix relating $\boldsymbol{z}_k$ **to** $\boldsymbol{x}_k$ at sample $k$.<br>$\boldsymbol{z}_k = \boldsymbol{C}_k \boldsymbol{x}_k + \boldsymbol{v}_k$ |
| $E[]$ | Expectation operator |
| $\boldsymbol{K}_k$ | The Kalman filter gain matrix at sample $k$ |
| $\boldsymbol{P}_k^-$ | The *a priori* covariance matrix of the linear prediction (*a priori*) error $\hat{\boldsymbol{x}}_{\varepsilon,k}^-$ at sample $k$.<br>$\boldsymbol{P}_k^- = E\left[\hat{\boldsymbol{x}}_{\varepsilon,k}^- {\hat{\boldsymbol{x}}_{\varepsilon,k}^-}^T\right]$ |
| $\boldsymbol{P}_k^+$ | The *a posteriori* covariance matrix of the Kalman (*a posteriori*) error $\hat{\boldsymbol{x}}_{\varepsilon,k}^+$ at sample $k$.<br>$\boldsymbol{P}_k^+ = E\left[\hat{\boldsymbol{x}}_{\varepsilon,k}^+ {\hat{\boldsymbol{x}}_{\varepsilon,k}^+}^T\right]$ |
| $\boldsymbol{Q}_{w,k}$ | The covariance matrix of the additive noise $\boldsymbol{w}_k$ in the process $\boldsymbol{x}_k$.<br>$\boldsymbol{Q}_{w,k} = E[\boldsymbol{w}_k {\boldsymbol{w}_k}^T]$ |
| $\boldsymbol{Q}_{v,k}$ | The covariance matrix of the additive noise $\boldsymbol{v}_k$ in the measured process $\boldsymbol{z}_k$.<br>$\boldsymbol{Q}_{v,k} = E[\boldsymbol{v}_k {\boldsymbol{v}_k}^T]$ |
| $V[]$ | Variance operator |
| $\boldsymbol{v}_k$ | The additive noise in the measured process $\boldsymbol{z}_k$ at sample $k$. |
| $\boldsymbol{w}_k$ | The additive noise in the process of interest $\boldsymbol{x}_k$ at sample $k$. |
| $\boldsymbol{x}_k$ | The state vector at time sample $k$ of the process $\boldsymbol{x}_k$.<br>$\boldsymbol{x}_k = \boldsymbol{A}_k \boldsymbol{x}_{k-1} + \boldsymbol{w}_k$ |
| $\hat{\boldsymbol{x}}_k^-$ | The linear prediction (*a priori*) estimate of the process $\boldsymbol{x}_k$ at sample $k$.<br>$\hat{\boldsymbol{x}}_k^- = \boldsymbol{A}_k \hat{\boldsymbol{x}}_{k-1}^+$ |
| $\hat{\boldsymbol{x}}_k^+$ | The Kalman filter (*a posteriori*) estimate of the process $\boldsymbol{x}_k$ at sample $k$.<br>$\hat{\boldsymbol{x}}_k^+ = (\boldsymbol{I} - \boldsymbol{K}_k \boldsymbol{C}_k)\hat{\boldsymbol{x}}_k^- + \boldsymbol{K}_k \boldsymbol{z}_k = (\boldsymbol{I} - \boldsymbol{K}_k \boldsymbol{C}_k)\boldsymbol{A}_k \hat{\boldsymbol{x}}_{k-1}^+ + \boldsymbol{K}_k \boldsymbol{z}_k$ |
| $\hat{\boldsymbol{x}}_{\varepsilon,k}^-$ | The error in the linear prediction (*a priori*) estimate of the process $\boldsymbol{x}_k$.<br>$\hat{\boldsymbol{x}}_{\varepsilon,k}^- = \hat{\boldsymbol{x}}_k^- - \boldsymbol{x}_k$ |
| $\hat{\boldsymbol{x}}_{\varepsilon,k}^+$ | The error in the *a posteriori* Kalman filter estimate of the process $\boldsymbol{x}_k$.<br>$\hat{\boldsymbol{x}}_{\varepsilon,k}^+ = \hat{\boldsymbol{x}}_k^+ - \boldsymbol{x}_k$ |
| $\boldsymbol{z}_k$ | The measured process at sample $k$.<br>$\boldsymbol{z}_k = \boldsymbol{C}_k \boldsymbol{x}_k + \boldsymbol{v}_k$ |
| $\delta_{k,j}$ | The Kronecker delta function. $\delta_{k,j} = 1$ for $k = j$ and zero otherwise. |

# 2 Mathematical Lemmas

## 2.1 Lemma 1

The trace of the sum of two square matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ equals the sum of the individual traces.

$$tr(\boldsymbol{A}+\boldsymbol{B}) = \sum_{i=0}^{N-1} A_{ii}+B_{ii} = \sum_{i=0}^{N-1} A_{ii} + \sum_{i=0}^{N-1} B_{ii} = tr(\boldsymbol{A}) + tr(\boldsymbol{B}) \tag{1}$$

## 2.2 Lemma 2

The derivative with respect to $\boldsymbol{A}$ of the trace of the matrix product $\boldsymbol{C} = \boldsymbol{AB}$ equals $\boldsymbol{B}^T$.

$$\frac{\partial\{tr(\boldsymbol{C})\}}{\partial \boldsymbol{A}} = \frac{\partial\{tr(\boldsymbol{AB})\}}{\partial \boldsymbol{A}} = \begin{pmatrix} \left(\frac{\partial tr(\boldsymbol{AB})}{\partial A_{0,0}}\right) & \left(\frac{\partial tr(\boldsymbol{AB})}{\partial A_{0,1}}\right) & \cdots & \left(\frac{\partial tr(\boldsymbol{AB})}{\partial A_{0,N-1}}\right) \\ \left(\frac{\partial tr(\boldsymbol{AB})}{\partial A_{1,0}}\right) & \left(\frac{\partial tr(\boldsymbol{AB})}{\partial A_{1,1}}\right) & \cdots & \left(\frac{\partial tr(\boldsymbol{AB})}{\partial A_{1,N-1}}\right) \\ \cdots & \cdots & \cdots & \cdots \\ \left(\frac{\partial tr(\boldsymbol{AB})}{\partial A_{M-1,0}}\right) & \left(\frac{\partial tr(\boldsymbol{AB})}{\partial A_{M-1,1}}\right) & \cdots & \left(\frac{\partial tr(\boldsymbol{AB})}{\partial A_{M-1,N-1}}\right) \end{pmatrix} \tag{2}$$

If the matrix $\boldsymbol{A}$ has dimensions $M \times N$ and the matrix $\boldsymbol{B}$ has dimensions $N \times M$, then $\boldsymbol{C} = \boldsymbol{AB}$ has dimensions $M \times M$.

The element $C_{ij}$ of matrix $\boldsymbol{C}$ has value:

$$C_{ij} = \sum_{k=0}^{N-1} A_{ik}B_{kj} \Rightarrow tr(\boldsymbol{C}) = tr(\boldsymbol{AB}) = \sum_{i=0}^{M-1} C_{ii} = \sum_{i=0}^{M-1}\sum_{k=0}^{N-1} A_{ik}B_{ki} \tag{3}$$

Substituting equation (3) into equation (2) gives:

$$\frac{\partial\{tr(\boldsymbol{AB})\}}{\partial \boldsymbol{A}} = \begin{pmatrix} \left(\frac{\partial \sum_{i=0}^{M-1}\sum_{k=0}^{N-1} A_{ik}B_{ki}}{\partial A_{0,0}}\right) & \left(\frac{\partial \sum_{i=0}^{M-1}\sum_{k=0}^{N-1} A_{ik}B_{ki}}{\partial A_{0,1}}\right) & \cdots & \left(\frac{\partial \sum_{i=0}^{M-1}\sum_{k=0}^{N-1} A_{ik}B_{ki}}{\partial A_{0,N-1}}\right) \\ \left(\frac{\partial \sum_{i=0}^{M-1}\sum_{k=0}^{N-1} A_{ik}B_{ki}}{\partial A_{1,0}}\right) & \left(\frac{\partial \sum_{i=0}^{M-1}\sum_{k=0}^{N-1} A_{ik}B_{ki}}{\partial A_{1,1}}\right) & \cdots & \left(\frac{\partial \sum_{i=0}^{M-1}\sum_{k=0}^{N-1} A_{ik}B_{ki}}{\partial A_{1,N-1}}\right) \\ \cdots & \cdots & \cdots & \cdots \\ \left(\frac{\partial \sum_{i=0}^{M-1}\sum_{k=0}^{N-1} A_{ik}B_{ki}}{\partial A_{M-1,0}}\right) & \left(\frac{\partial \sum_{i=0}^{M-1}\sum_{k=0}^{N-1} A_{ik}B_{ki}}{\partial A_{M-1,1}}\right) & \cdots & \left(\frac{\partial \sum_{i=0}^{M-1}\sum_{k=0}^{N-1} A_{ik}B_{ki}}{\partial A_{M-1,N-1}}\right) \end{pmatrix} \tag{4}$$

By inspection:

$$\left(\frac{\partial \sum_{i=0}^{M-1} \sum_{k=0}^{N-1} A_{ik} B_{ki}}{\partial A_{lm}}\right) = B_{ml} \tag{5}$$

Substituting equation (5) into equation (4) gives the required proof:

$$\frac{\partial\{tr(\boldsymbol{AB})\}}{\partial \boldsymbol{A}} = \begin{pmatrix} B_{0,0} & B_{1,0} & \dots & B_{N-1,0} \\ B_{0,1} & B_{1,1} & \dots & B_{N-1,1} \\ \dots & \dots & \dots & \dots \\ B_{0,M-1} & B_{1,M-1} & \dots & B_{N-1,M-1} \end{pmatrix} = \boldsymbol{B}^T \tag{6}$$

## 2.3 Lemma 3

The derivative with respect to $\boldsymbol{A}$ of the trace of the matrix product $\boldsymbol{ABA}^T$ equals $\boldsymbol{A}(\boldsymbol{B} + \boldsymbol{B}^T)$.

$$\frac{\partial\{tr(\boldsymbol{ABA}^T)\}}{\partial \boldsymbol{A}} = \begin{pmatrix} \left(\frac{\partial tr(\boldsymbol{ABA}^T)}{\partial A_{0,0}}\right) & \left(\frac{\partial tr(\boldsymbol{ABA}^T)}{\partial A_{0,1}}\right) & \dots & \left(\frac{\partial tr(\boldsymbol{ABA}^T)}{\partial A_{0,N-1}}\right) \\ \left(\frac{\partial tr(\boldsymbol{ABA}^T)}{\partial A_{1,0}}\right) & \left(\frac{\partial tr(\boldsymbol{ABA}^T)}{\partial A_{1,1}}\right) & \dots & \left(\frac{\partial tr(\boldsymbol{ABA}^T)}{\partial A_{1,N-1}}\right) \\ \dots & \dots & \dots & \dots \\ \left(\frac{\partial tr(\boldsymbol{ABA}^T)}{\partial A_{M-1,0}}\right) & \left(\frac{\partial tr(\boldsymbol{ABA}^T)}{\partial A_{M-1,1}}\right) & \dots & \left(\frac{\partial tr(\boldsymbol{ABA}^T)}{\partial A_{M-1,N-1}}\right) \end{pmatrix} \tag{7}$$

If the matrix $\boldsymbol{A}$ has dimensions $M \times N$, then the matrix $\boldsymbol{B}$ must be square with dimensions $N \times N$ in order for the product $\boldsymbol{ABA}^T$ to exist. The product $\boldsymbol{ABA}^T$ is always square with dimensions $M \times M$.

The element $C_{ij}$ of the matrix $\boldsymbol{C} = \boldsymbol{AB}$ has value:

$$C_{ij} = \sum_{k=0}^{N-1} A_{ik} B_{kj} \tag{8}$$

The element $D_{il}$ of matrix $\boldsymbol{D} = \boldsymbol{ABA}^T = \boldsymbol{CA}^T$ has value:

$$D_{il} = \sum_{j=0}^{N-1} C_{ij} A_{lj} = \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} A_{ik} B_{kj} A_{lj} \tag{9}$$

The trace of matrix $\boldsymbol{D}$ then equals:

$$tr(\boldsymbol{D}) = \sum_{i=0}^{N-1} D_{ii} = \sum_{i=0}^{N-1}\sum_{j=0}^{N-1}\sum_{k=0}^{N-1} A_{ik}B_{kj}A_{ij} \tag{10}$$

The derivative of $tr(\boldsymbol{D})$ with respect to $A_{lm}$ is:

$$\left(\frac{\partial tr(\boldsymbol{D})}{\partial A_{lm}}\right) = \left(\frac{\partial \sum_{i=0}^{N-1}\sum_{j=0}^{N-1}\sum_{k=0}^{N-1} A_{ik}B_{kj}A_{ij}}{\partial A_{lm}}\right) = \left(\frac{\partial \sum_{j=0}^{N-1}\sum_{k=0}^{N-1} A_{lk}B_{kj}A_{lj}}{\partial A_{lm}}\right) \tag{11}$$

$$= \sum_{j=0}^{N-1} A_{lj}B_{mj} + \sum_{j=0}^{N-1} A_{lj}B_{jm} = (\boldsymbol{AB}^T)_{lm} + (\boldsymbol{AB})_{lm} \tag{12}$$

$$\Rightarrow \frac{\partial\{tr(\boldsymbol{ABA}^T)\}}{\partial \boldsymbol{A}} = \boldsymbol{A}(\boldsymbol{B} + \boldsymbol{B}^T) \tag{13}$$

If $\boldsymbol{B}$ is also symmetric, then:

$$\frac{\partial\{tr(\boldsymbol{ABA}^T)\}}{\partial \boldsymbol{A}} = 2\boldsymbol{AB}\ if\ \boldsymbol{B} = \boldsymbol{B}^T \tag{14}$$

# 3 Kalman Filter Derivation

## 3.1 Process Model

The Kalman filter models the vector process of interest $\boldsymbol{x}_k$ with the linear and recursive model:

$$\boldsymbol{x}_k = \boldsymbol{A}_k\boldsymbol{x}_{k-1} + \boldsymbol{w}_k \tag{15}$$

If $\boldsymbol{x}_k$ has $N$ degrees of freedom, then $\boldsymbol{A}_k$ is an $N \times N$ linear prediction matrix (possibly time varying but assumed known) and $\boldsymbol{w}_k$ is an $N \times 1$ zero mean white noise vector.

The process $\boldsymbol{x}_k$ is assumed to be not directly measurable and must be estimated from a process $\boldsymbol{z}_k$, which can be measured. $\boldsymbol{z}_k$ is modeled as being linearly related to $\boldsymbol{x}_k$ with additive zero mean white noise $\boldsymbol{v}_k$.

$$\boldsymbol{z}_k = \boldsymbol{C}_k\boldsymbol{x}_k + \boldsymbol{v}_k \tag{16}$$

$\boldsymbol{z}_k$ is an $M \times 1$ vector, $\boldsymbol{C}_k$ is an $M \times N$ matrix (possibly time varying but assumed known) and $\boldsymbol{v}_k$ is an $M \times 1$ noise vector.

Since the noise vectors $\boldsymbol{w}_k$ and $\boldsymbol{v}_k$ are zero-mean white noise processes their expectation vector is zero and their covariance matrices are uncorrelated at different times $j$ and $k$:

$$E[\boldsymbol{w}_k] = \boldsymbol{0} \tag{17}$$

$$E[\boldsymbol{v}_k] = \boldsymbol{0} \tag{18}$$

$$cov\{\boldsymbol{w}_k, \boldsymbol{w}_j\} = E\left[\boldsymbol{w}_k \boldsymbol{w}_j{}^T\right] = \boldsymbol{Q}_{w,k}\delta_{kj} \tag{19}$$

$$cov\{\boldsymbol{v}_k, \boldsymbol{v}_j\} = E\left[\boldsymbol{v}_k \boldsymbol{v}_j{}^T\right] = \boldsymbol{Q}_{v,k}\delta_{kj} \tag{20}$$

Covariance matrices are, by definition, symmetric but not necessarily diagonal:

$$\boldsymbol{Q}_{w,k}{}^T = \{E[\boldsymbol{w}_k \boldsymbol{w}_k{}^T]\}^T = E[(\boldsymbol{w}_k \boldsymbol{w}_k{}^T)^T] = E\left[\boldsymbol{w}_k \boldsymbol{w}_j{}^T\right] = \boldsymbol{Q}_{w,k} \tag{21}$$

The covariance matrices $\boldsymbol{Q}_{w,k}$ and $\boldsymbol{Q}_{v,k}$ need not be stationary and can, and generally will, vary with time.

## 3.2 Derivation

The objective of the Kalman filter is to compute an unbiased *a posterori* estimate $\widehat{\boldsymbol{x}}_k^+$ of the underlying process $\boldsymbol{x}_k$ from i) extrapolation from the previous iteration's *a posteriori* estimate $\widehat{\boldsymbol{x}}_{k-1}^+$ and ii) from the current measurement $\boldsymbol{z}_k$:

$$\widehat{\boldsymbol{x}}_k^+ = \boldsymbol{K}'_k \widehat{\boldsymbol{x}}_{k-1}^+ + \boldsymbol{K}_k \boldsymbol{z}_k \tag{22}$$

The time-varying Kalman gain matrices $\boldsymbol{K}'_k$ and $\boldsymbol{K}_k$ define the relative weightings given to the previous iteration's Kalman filter estimate $\boldsymbol{K}_k$ and to the current measurement $\boldsymbol{z}_k$. If the measurements $\boldsymbol{z}_k$ have low noise then the measurement term $\boldsymbol{K}_k \boldsymbol{z}_k$ will have a higher weighting compared to the extrapolated component $\boldsymbol{K}'_k \widehat{\boldsymbol{x}}_{k-1}^+$ and vice versa. The Kalman filter is, therefore, a time varying, recursive filter.

**Unbiased estimate constraint (determines $\boldsymbol{K}'_k$)**

For $\widehat{\boldsymbol{x}}_k^+$ to be an unbiased estimate of $\boldsymbol{x}_k$, the expectation value of the *a posteriori* Kalman filter error $\widehat{\boldsymbol{x}}_{\varepsilon,k}^+$ must be zero:

$$E\left[\widehat{\boldsymbol{x}}_{\varepsilon,k}^+\right] = E[\widehat{\boldsymbol{x}}_k^+ - \boldsymbol{x}_k] = \boldsymbol{0} \tag{23}$$

Subtracting $\boldsymbol{x}_k$ from equation (22) gives:

$$\widehat{\boldsymbol{x}}_{\varepsilon,k}^+ = \widehat{\boldsymbol{x}}_k^+ - \boldsymbol{x}_k = \boldsymbol{K}'_k \widehat{\boldsymbol{x}}_{k-1}^+ + \boldsymbol{K}_k \boldsymbol{z}_k - \boldsymbol{x}_k \tag{24}$$

Substituting equation (16) for the measurement $\boldsymbol{z}_k$ gives:

$$\widehat{\boldsymbol{x}}_{\varepsilon,k}^+ = \boldsymbol{K}'_k \widehat{\boldsymbol{x}}_{k-1}^+ + \boldsymbol{K}_k(\boldsymbol{C}_k \boldsymbol{x}_k + \boldsymbol{v}_k) - \boldsymbol{x}_k \tag{25}$$

Substituting for $\boldsymbol{x}_k$ from equation (15) and re-arranging gives:

$$\hat{\boldsymbol{x}}_{\varepsilon,k}^{+} = \boldsymbol{K}_k'\left(\hat{\boldsymbol{x}}_{\varepsilon,k-1}^{+} + \boldsymbol{x}_{k-1}\right) + \boldsymbol{K}_k\{\boldsymbol{C}_k(\boldsymbol{A}_k\boldsymbol{x}_{k-1} + \boldsymbol{w}_k) + \boldsymbol{v}_k\} - (\boldsymbol{A}_k\boldsymbol{x}_{k-1} + \boldsymbol{w}_k) \quad \textbf{(26)}$$

$$= \boldsymbol{K}_k'\hat{\boldsymbol{x}}_{\varepsilon,k-1}^{+} + (\boldsymbol{K}_k\boldsymbol{C}_k\boldsymbol{A}_k - \boldsymbol{A}_k + \boldsymbol{K}_k')\boldsymbol{x}_{k-1} + (\boldsymbol{K}_k\boldsymbol{C}_k - \boldsymbol{I})\boldsymbol{w}_k + \boldsymbol{K}_k\boldsymbol{v}_k \quad \textbf{(27)}$$

Taking the expected value of equation (27) and applying the unbiased estimate constraint gives:

$$E\left[\hat{\boldsymbol{x}}_{\varepsilon,k}^{+}\right] = E\left[\boldsymbol{K}_k'\hat{\boldsymbol{x}}_{\varepsilon,k-1}^{+}\right] + E[(\boldsymbol{K}_k\boldsymbol{C}_k\boldsymbol{A}_k - \boldsymbol{A}_k + \boldsymbol{K}_k')\boldsymbol{x}_{k-1}] + E[(\boldsymbol{K}_k\boldsymbol{C}_k - \boldsymbol{I})\boldsymbol{w}_k] + E[\boldsymbol{K}_k\boldsymbol{v}_k] = \boldsymbol{0} \quad \textbf{(28)}$$

Because the noise vectors $\boldsymbol{w}_k$ and $\boldsymbol{v}_k$ are zero mean and uncorrelated with the Kalman matrices for the same iteration, it follows that:

$$E[(\boldsymbol{K}_k\boldsymbol{C}_k - \boldsymbol{I})\boldsymbol{w}_k] = E[\boldsymbol{K}_k\boldsymbol{v}_k] = \boldsymbol{0} \quad \textbf{(29)}$$

With the additional assumption that the process $\boldsymbol{x}_{k-1}$ is independent of slowly varying matrices $\boldsymbol{A_k}$, $\boldsymbol{C_k}$, $\boldsymbol{K_k}$ and $\boldsymbol{K'_k}$ at iteration $k$:

$$E[(\boldsymbol{K}_k\boldsymbol{C}_k\boldsymbol{A}_k - \boldsymbol{A}_k + \boldsymbol{K}_k')\boldsymbol{x}_{k-1}] = (\boldsymbol{K}_k\boldsymbol{C}_k\boldsymbol{A}_k - \boldsymbol{A}_k + \boldsymbol{K}_k')E[\boldsymbol{x}_{k-1}] = \boldsymbol{0} \quad \textbf{(30)}$$

Because $\boldsymbol{x}_k$ is not, in general, a zero-mean process:

$$\boldsymbol{K}_k\boldsymbol{C}_k\boldsymbol{A}_k - \boldsymbol{A}_k + \boldsymbol{K}_k' = \boldsymbol{0} \Rightarrow \boldsymbol{K}_k' = \boldsymbol{A}_k - \boldsymbol{K}_k\boldsymbol{C}_k\boldsymbol{A}_k = (\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{C}_k)\boldsymbol{A}_k \quad \textbf{(31)}$$

Substituting for $\boldsymbol{K}_k'$ in equation (22) gives:

$$\hat{\boldsymbol{x}}_k^{+} = (\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{C}_k)\boldsymbol{A}_k\hat{\boldsymbol{x}}_{k-1}^{+} + \boldsymbol{K}_k\boldsymbol{z}_k \quad \textbf{(32)}$$

### *A priori* estimate

The *a priori* Kalman filter estimate $\hat{\boldsymbol{x}}_k^{-}$ is the result of applying the linear prediction matrix $\boldsymbol{A}_k$ to the previous iteration's *a posteriori* estimate $\hat{\boldsymbol{x}}_{k-1}^{+}$:

$$\hat{\boldsymbol{x}}_k^{-} = \boldsymbol{A}_k\hat{\boldsymbol{x}}_{k-1}^{+} \qquad \textbf{Kalman equation 1} \quad \textbf{(33)}$$

### Definition of *a posteriori* estimate

Substituting the *a priori* estimate $\hat{\boldsymbol{x}}_k^{-}$ from equation (33) into equation (32) gives:

$$\hat{\boldsymbol{x}}_k^{+} = (\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{C}_k)\hat{\boldsymbol{x}}_k^{-} + \boldsymbol{K}_k\boldsymbol{z}_k \qquad \textbf{Kalman equation 4} \quad \textbf{(34)}$$

An equivalent form is:

$$\hat{\boldsymbol{x}}_k^{+} = \hat{\boldsymbol{x}}_k^{-} + \boldsymbol{K}_k(\boldsymbol{z}_k - \boldsymbol{C}_k\hat{\boldsymbol{x}}_k^{-}) \quad \textbf{(35)}$$

From equation (16), the term $\boldsymbol{C}_k\hat{\boldsymbol{x}}_k^-$ can be interpreted as the *a priori* estimate $\hat{\boldsymbol{z}}_k^-$ of the measurement $\boldsymbol{z}_k$ giving another form of equation (35):

$$\hat{\boldsymbol{x}}_k^+ = \hat{\boldsymbol{x}}_k^- + \boldsymbol{K}_k(\boldsymbol{z}_k - \hat{\boldsymbol{z}}_k^-) \qquad \textbf{(36)}$$

### $\boldsymbol{P}_k^-$ as a function of $\boldsymbol{P}_{k-1}^+$

The *a priori* and *a posteriori* error covariance matrices $\boldsymbol{P}_k^-$ and $\boldsymbol{P}_k^+$ are defined as:

$$\boldsymbol{P}_k^- = cov\{\hat{\boldsymbol{x}}_{\varepsilon,k}^-, \hat{\boldsymbol{x}}_{\varepsilon,k}^-\} = E\left[\hat{\boldsymbol{x}}_{\varepsilon,k}^- {\hat{\boldsymbol{x}}_{\varepsilon,k}^-}^T\right] = E[(\hat{\boldsymbol{x}}_k^- - \boldsymbol{x}_k)(\hat{\boldsymbol{x}}_k^- - \boldsymbol{x}_k)^T] \qquad \textbf{(37)}$$

$$\boldsymbol{P}_k^+ = cov\{\hat{\boldsymbol{x}}_{\varepsilon,k}^+, \hat{\boldsymbol{x}}_{\varepsilon,k}^+\} = E\left[\hat{\boldsymbol{x}}_{\varepsilon,k}^+ {\hat{\boldsymbol{x}}_{\varepsilon,k}^+}^T\right] = E[(\hat{\boldsymbol{x}}_k^+ - \boldsymbol{x}_k)(\hat{\boldsymbol{x}}_k^+ - \boldsymbol{x}_k)^T] \qquad \textbf{(38)}$$

Substituting the definitions of $\hat{\boldsymbol{x}}_k^-$ and $\boldsymbol{x}_k$ into equation (37) gives an expression relating the current a priori error covariance $\boldsymbol{P}_k^-$ to the previous iteration's a posteriori error covariance estimate $\boldsymbol{P}_{k-1}^+$:

$$\boldsymbol{P}_k^- = E[(\boldsymbol{A}_k\hat{\boldsymbol{x}}_{k-1}^+ - \boldsymbol{A}_k\boldsymbol{x}_{k-1} - \boldsymbol{w}_k)(\boldsymbol{A}_k\hat{\boldsymbol{x}}_{k-1}^+ - \boldsymbol{A}_k\boldsymbol{x}_{k-1} - \boldsymbol{w}_k)^T] \qquad \textbf{(38)}$$

$$= E[\{\boldsymbol{A}_k(\hat{\boldsymbol{x}}_{k-1}^+ - \boldsymbol{x}_{k-1}) - \boldsymbol{w}_k\}\{\boldsymbol{A}_k(\hat{\boldsymbol{x}}_{k-1}^+ - \boldsymbol{x}_{k-1}) - \boldsymbol{w}_k\}^T] \qquad \textbf{(40)}$$

$$= \boldsymbol{A}_k E[(\hat{\boldsymbol{x}}_{k-1}^+ - \boldsymbol{x}_{k-1})(\hat{\boldsymbol{x}}_{k-1}^+ - \boldsymbol{x}_{k-1})^T]{\boldsymbol{A}_k}^T + \boldsymbol{Q}_{w,k} \qquad \textbf{(41)}$$

$$\Rightarrow \boldsymbol{P}_k^- = \boldsymbol{A}_k\boldsymbol{P}_{k-1}^+{\boldsymbol{A}_k}^T + \boldsymbol{Q}_{w,k} \qquad \textbf{Kalman Equation 2} \qquad \textbf{(42)}$$

### Minimum error covariance constraint (determines $\boldsymbol{K}_k$)

The Kalman gain matrix $\boldsymbol{K}_k$ minimizes the *a posteriori* error $\hat{\boldsymbol{x}}_{\varepsilon,k}^+$ variance via the trace of the *a posteriori* error covariance matrix $\boldsymbol{P}_k^+$:

$$E\left[{\hat{\boldsymbol{x}}_{\varepsilon,k}^+}^T \hat{\boldsymbol{x}}_{\varepsilon,k}^+\right] = tr(\boldsymbol{P}_k^+) \qquad \textbf{(39)}$$

Substituting equation (16) for $\boldsymbol{z}_k$ into equation (32) gives a relation between the *a posteriori* and *a priori* errors:

$$\hat{\boldsymbol{x}}_k^+ = \hat{\boldsymbol{x}}_{\varepsilon,k}^+ + \boldsymbol{x}_k = (\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{C}_k)\left(\hat{\boldsymbol{x}}_{\varepsilon,k}^- + \boldsymbol{x}_k\right) + \boldsymbol{K}_k(\boldsymbol{C}_k\boldsymbol{x}_k + \boldsymbol{v}_k) \qquad \textbf{(40)}$$

$$\Rightarrow \hat{\boldsymbol{x}}_{\varepsilon,k}^+ + \boldsymbol{x}_k = (\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{C}_k)\hat{\boldsymbol{x}}_{\varepsilon,k}^- + \boldsymbol{x}_k - \boldsymbol{K}_k\boldsymbol{C}_k\boldsymbol{x}_k + \boldsymbol{K}_k(\boldsymbol{C}_k\boldsymbol{x}_k + \boldsymbol{v}_k) \qquad \textbf{(45)}$$

$$\Rightarrow \hat{\boldsymbol{x}}_{\varepsilon,k}^+ = (\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{C}_k)\hat{\boldsymbol{x}}_{\varepsilon,k}^- + \boldsymbol{K}_k\boldsymbol{v}_k \qquad \textbf{(46)}$$

Substituting this result into the definition of the *a posteriori* covariance matrix $\boldsymbol{P}_k^+$ in equation (38) gives:

$$\boldsymbol{P}_k^+ = E\left[\left\{(\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{C}_k)\widehat{\boldsymbol{x}}_{\varepsilon,k}^- + \boldsymbol{K}_k\boldsymbol{v}_k\right\}\left\{(\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{C}_k)\widehat{\boldsymbol{x}}_{\varepsilon,k}^- + \boldsymbol{K}_k\boldsymbol{v}_k\right\}^T\right] \tag{47}$$

$$= (\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{C}_k)E\left[\widehat{\boldsymbol{x}}_{\varepsilon,k}^-\widehat{\boldsymbol{x}}_{\varepsilon,k}^{-\,T}\right](\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{C}_k)^T + \boldsymbol{K}_k E[\boldsymbol{v}_k\boldsymbol{v}_k{}^T]\boldsymbol{K}_k{}^T \tag{48}$$

$$= (\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{C}_k)\boldsymbol{P}_k^-(\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{C}_k)^T + \boldsymbol{K}_k\boldsymbol{Q}_{v,k}\boldsymbol{K}_k{}^T \tag{49}$$

$$= \boldsymbol{P}_k^- - \boldsymbol{P}_k^-\boldsymbol{C}_k{}^T\boldsymbol{K}_k{}^T - \boldsymbol{K}_k\boldsymbol{C}_k\boldsymbol{P}_k^- + \boldsymbol{K}_k\boldsymbol{C}_k\boldsymbol{P}_k^-\boldsymbol{C}_k{}^T\boldsymbol{K}_k{}^T + \boldsymbol{K}_k\boldsymbol{Q}_{v,k}\boldsymbol{K}_k{}^T \tag{50}$$

The Kalman filter gain $\boldsymbol{K}_k$ is that which minimizes the trace of the *a posteriori* error covariance matrix $\boldsymbol{P}_k^+$ as in equation (43):

$$\frac{\partial}{\partial \boldsymbol{K}_k} tr(\boldsymbol{P}_k^+) = \frac{\partial}{\partial \boldsymbol{K}_k}\left\{tr(\boldsymbol{P}_k^-) - tr\left(\boldsymbol{P}_k^-\boldsymbol{C}_k{}^T\boldsymbol{K}_k{}^T\right) - tr(\boldsymbol{K}_k\boldsymbol{C}_k\boldsymbol{P}_k^-) + tr\left(\boldsymbol{K}_k\boldsymbol{C}_k\boldsymbol{P}_k^-\boldsymbol{C}_k{}^T\boldsymbol{K}_k{}^T\right) + tr\left(\boldsymbol{K}_k\boldsymbol{Q}_{v,k}\boldsymbol{K}_k{}^T\right)\right\} = 0 \tag{51}$$

The first term $tr(\boldsymbol{P}_k^-)$ has no dependence on $\boldsymbol{K}_k$ giving:

$$\frac{\partial\{tr(\boldsymbol{P}_k^-)\}}{\partial \boldsymbol{K}_k} = \frac{\partial\left\{tr(\boldsymbol{A}_k\boldsymbol{P}_{k-1}^+\boldsymbol{A}_k{}^T + \boldsymbol{Q}_{w,k})\right\}}{\partial \boldsymbol{K}_k} = 0 \tag{52}$$

Because a matrix trace is obviously unaffected by transposition, the second term of equation (51) can be transposed and simplified using equation (6) to give:

$$\frac{\partial\left\{tr(\boldsymbol{P}_k^-\boldsymbol{C}_k{}^T\boldsymbol{K}_k{}^T)\right\}}{\partial \boldsymbol{K}_k} = \frac{\partial\{tr(\boldsymbol{K}_k\boldsymbol{C}_k\boldsymbol{P}_k^-)\}}{\partial \boldsymbol{K}_k} = (\boldsymbol{C}_k\boldsymbol{P}_k^-)^T = \boldsymbol{P}_k^-\boldsymbol{C}_k{}^T \tag{53}$$

The fourth term can be simplified using equations (13) and (14) exploiting the fact that the covariance matrix $\boldsymbol{P}_k^-$ is symmetric:

$$\frac{\partial\left\{tr(\boldsymbol{K}_k\boldsymbol{C}_k\boldsymbol{P}_k^-\boldsymbol{C}_k{}^T\boldsymbol{K}_k{}^T)\right\}}{\partial \boldsymbol{K}_k} = \boldsymbol{K}_k\left\{\boldsymbol{C}_k\boldsymbol{P}_k^-\boldsymbol{C}_k{}^T + \left(\boldsymbol{C}_k\boldsymbol{P}_k^-\boldsymbol{C}_k{}^T\right)^T\right\} = 2\boldsymbol{K}_k\boldsymbol{C}_k\boldsymbol{P}_k^-\boldsymbol{C}_k{}^T \tag{54}$$

The final term can be simplified also using equations (13) and (14) and the symmetry of $\boldsymbol{Q}_{v,k}$ to give:

$$\frac{\partial\left\{tr(\boldsymbol{K}_k\boldsymbol{Q}_{v,k}\boldsymbol{K}_k{}^T)\right\}}{\partial \boldsymbol{K}_k} = 2\boldsymbol{K}_k\boldsymbol{Q}_{v,k} \tag{55}$$

Substituting equations (52) to (55) back into equation (51) gives an expression for the optimal Kalman filter gain matrix $\boldsymbol{K}_k$:

$$-2\boldsymbol{P}_k^-\boldsymbol{C}_k{}^T + 2\boldsymbol{K}_k\boldsymbol{C}_k\boldsymbol{P}_k^-\boldsymbol{C}_k{}^T + 2\boldsymbol{K}_k\boldsymbol{Q}_{v,k} = \boldsymbol{0} \tag{56}$$

$$\Rightarrow \boldsymbol{K}_k\left(\boldsymbol{C}_k\boldsymbol{P}_k^-{\boldsymbol{C}_k}^T + \boldsymbol{Q}_{v,k}\right) = \boldsymbol{P}_k^-{\boldsymbol{C}_k}^T \qquad \textbf{(57)}$$

$$\Rightarrow \boldsymbol{K}_k = \boldsymbol{P}_k^-{\boldsymbol{C}_k}^T\left(\boldsymbol{C}_k\boldsymbol{P}_k^-{\boldsymbol{C}_k}^T + \boldsymbol{Q}_{v,k}\right)^{-1} \qquad \textbf{Kalman equation 3} \qquad \textbf{(58)}$$

### $\boldsymbol{P}_k^+$ as a function of $\boldsymbol{P}_k^-$

Rearranging equation (57) gives:

$$\boldsymbol{K}_k\boldsymbol{Q}_{v,k} = \boldsymbol{P}_k^-{\boldsymbol{C}_k}^T - \boldsymbol{K}_k\boldsymbol{C}_k\boldsymbol{P}_k^-{\boldsymbol{C}_k}^T \qquad \textbf{(59)}$$

Substituting $\boldsymbol{K}_k\boldsymbol{Q}_{v,k}$ from equation (61) into equation (51) gives:

$$\boldsymbol{P}_k^+ = (\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{C}_k)\boldsymbol{P}_k^-\left(\boldsymbol{I} - {\boldsymbol{C}_k}^T{\boldsymbol{K}_k}^T\right) + (\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{C}_k)\boldsymbol{P}_k^-{\boldsymbol{C}_k}^T{\boldsymbol{K}_k}^T \qquad \textbf{(60)}$$

$$\boldsymbol{P}_k^+ = (\boldsymbol{P}_k^- - \boldsymbol{K}_k\boldsymbol{C}_k\boldsymbol{P}_k^-)\left(\boldsymbol{I} - {\boldsymbol{C}_k}^T{\boldsymbol{K}_k}^T\right) + (\boldsymbol{I}\boldsymbol{P}_k^- - \boldsymbol{K}_k\boldsymbol{C}_k\boldsymbol{P}_k^-){\boldsymbol{C}_k}^T{\boldsymbol{K}_k}^T$$

$$\boldsymbol{P}_k^+ = \boldsymbol{P}_k^- - \boldsymbol{K}_k\boldsymbol{C}_k\boldsymbol{P}_k^- - \boldsymbol{P}_k^-{\boldsymbol{C}_k}^T{\boldsymbol{K}_k}^T + \boldsymbol{K}_k\boldsymbol{C}_k\boldsymbol{P}_k^-{\boldsymbol{C}_k}^T{\boldsymbol{K}_k}^T + \boldsymbol{P}_k^-{\boldsymbol{C}_k}^T{\boldsymbol{K}_k}^T - \boldsymbol{K}_k\boldsymbol{C}_k\boldsymbol{P}_k^-{\boldsymbol{C}_k}^T{\boldsymbol{K}_k}^T$$

$$\boldsymbol{P}_k^+ = \boldsymbol{P}_k^- - \boldsymbol{K}_k\boldsymbol{C}_k\boldsymbol{P}_k^-$$

$$\Rightarrow \boldsymbol{P}_k^+ = (\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{C}_k)\boldsymbol{P}_k^- \qquad \textbf{Kalman equation 5} \qquad \textbf{(61)}$$

This completes the derivation of the standard Kalman filter equations.

## 3.3 Standard Kalman Equations

### Kalman equation 1

The linear prediction (*a priori*) estimate $\hat{\boldsymbol{x}}_k^-$is made by applying the linear prediction matrix $\boldsymbol{A}_k$ to the previous sample's Kalman (*a posteriori*) filter estimate $\hat{\boldsymbol{x}}_{k-1}^+$.

$$\hat{\boldsymbol{x}}_k^- = \boldsymbol{A}_k\hat{\boldsymbol{x}}_{k-1}^+ \qquad \textbf{(33)}$$

### Kalman equation 2

The *a priori* (linear extrapolation) error covariance matrix $\boldsymbol{P}_k^-$ is then updated using the model matrix $\boldsymbol{A}_k$ and the noise matrix $\boldsymbol{Q}_{w,k}$.

$$\boldsymbol{P}_k^- = \boldsymbol{A}_k\boldsymbol{P}_{k-1}^+{\boldsymbol{A}_k}^T + \boldsymbol{Q}_{w,k} \qquad \textbf{(42)}$$

Kalman equations 2 and 5 can be combined to give a recursive update of $\boldsymbol{P}_k^-$ without explicit calculation of the *a posteriori* error covariance matrix $\boldsymbol{P}_k^+$ in Kalman equation 5:

$$P_k^- = A_k(I - K_{k-1}C_{k-1})P_{k-1}^- {A_k}^T + Q_{w,k} \tag{62}$$

The only purpose of $P_k^-$ is to permit the calculation of the Kalman gain matrix $K_k$ for the determination of the a posteriori estimate $\hat{x}_k^+$.

**Kalman equation 3**

The Kalman filter gain matrix $K_k$ is updated:

$$K_k = P_k^- {C_k}^T\left(C_k P_k^- {C_k}^T + Q_{v,k}\right)^{-1} \tag{58}$$

**Kalman equation 4**

The Kalman filter (*a posteriori*) estimate $\hat{x}_k^+$ is computed from the current *a priori* estimate $\hat{x}_k^-$ and the current measurement $z_k$.

$$\hat{x}_k^+ = \hat{x}_k^- + K_k(z_k - C_k\hat{x}_k^-) = (I - K_kC_k)\hat{x}_k^- + K_kz_k \tag{34}$$

**Kalman equation 5**

The *a posteriori* Kalman error covariance matrix $P_k^+$ is updated and ready for the next iteration. This equation can be skipped if $P_k^-$ is updated recursively in terms of itself as in equation (62).

$$P_k^+ = (I - K_kC_k)P_k^- \tag{61}$$

## 3.4 Limiting Cases

From (59), we can see that as the measurement noise covariance $Q_{v,k}$ decreases relative to the process noise covariance $Q_{w,k}$, the Kalman gain matrix $K_k$ then satisfies:

$$K_kC_kP_k^-{C_k}^T = P_k^-{C_k}^T \Rightarrow K_kC_k = I \tag{63}$$

Using (34), we see that the a posteriori process estimate $\hat{x}_k^+$ is then only dependent on the measurement $z_k$:

$$\hat{x}_k^+ = (I - K_kC_k)\hat{x}_k^- + K_kz_k = K_kz_k \tag{64}$$

From (59), we can see that as the measurement noise covariance $Q_{v,k}$ increases relative to the process noise covariance $Q_{w,k}$, the Kalman gain matrix $K_k$ approaches zero:

$$K_k = P_k^-{C_k}^T\left(Q_{v,k}\right)^{-1} = 0 \tag{65}$$

## Kalman Filter Derivation

Using (34), the a posteriori process estimate $\hat{\boldsymbol{x}}_k^+$ then approximates the a priori estimate $\hat{\boldsymbol{x}}_k^-$:

$$\hat{\boldsymbol{x}}_k^+ = \hat{\boldsymbol{x}}_k^- + \boldsymbol{K}_k(\boldsymbol{z}_k - \boldsymbol{C}_k\hat{\boldsymbol{x}}_k^-) = \hat{\boldsymbol{x}}_k^- \tag{66}$$

*How to Reach Us:*

**Home Page:**
freescale.com

**Web Support:**
freescale.com/support