AN5018

Basic Kalman Filter Theory

Rev. 2.0 — 21 June 2016 Application note

Document information

Info	Content
Abstract	This document derives the standard Kalman filter equations.

Revision history

Document ID	Release date	Supercedes
AN5018 v2.0	20160621	AN5018 v1.0
Modifications:	• Minor changes • The format of this document has been redesigned to comply with the new identity guidelines of NXP Semiconductors. Legal texts have been adapted to the new company name where appropriate.	
AN5018 v1.0	2015 September	—

Contact information

For more information, please visit: http://www.nxp.com

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 "*Sensor Fusion Kalman Filters*" which describes the more specialized indirect complementary Kalman filter used for the fusion of accelerometer, magnetometer and gyroscope data in the NXP 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. $\boldsymbol{x}_k = \boldsymbol{A}_k \boldsymbol{x}_{k-1} + \boldsymbol{w}_k$ $\widehat{\boldsymbol{x}}_k^- = \boldsymbol{A}_k \widehat{\boldsymbol{x}}_{k-1}^+$
$\boldsymbol{C}_k$	The measurement matrix relating $\boldsymbol{z}_k$ to $\boldsymbol{x}_k$ at sample k. $\boldsymbol{z}_k = \boldsymbol{C}_k \boldsymbol{x}_k + \boldsymbol{v}_k$
$E[]$	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 $\widehat{\boldsymbol{x}}_{\varepsilon,k}^-$ at sample k. $\boldsymbol{P}_k^- = E\left[\widehat{\boldsymbol{x}}_{\varepsilon,k}^- {\widehat{\boldsymbol{x}}_{\varepsilon,k}^-}^T\right]$
$\boldsymbol{P}_k^+$	The *a posteriori* covariance matrix of the Kalman (*a posteriori*) error $\widehat{\boldsymbol{x}}_{\varepsilon,k}^+$ at sample k. $\boldsymbol{P}_k^+ = E\left[\widehat{\boldsymbol{x}}_{\varepsilon,k}^+ {\widehat{\boldsymbol{x}}_{\varepsilon,k}^+}^T\right]$
$\boldsymbol{Q}_{w,k}$	The covariance matrix of the additive noise $\boldsymbol{w}_k$ in the process $\boldsymbol{x}_k$. $\boldsymbol{Q}_{w,k} = E[\boldsymbol{w}_k {\boldsymbol{w}_k}^T]$
$\boldsymbol{Q}_{v,k}$	The covariance matrix of the additive noise $\boldsymbol{v}_k$ in the measured process $\boldsymbol{z}_k$. $\boldsymbol{Q}_{v,k} = E[\boldsymbol{v}_k {\boldsymbol{v}_k}^T]$
$V[]$	Variance operator

Symbol	Definition
$\boldsymbol{v}_k$	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 $\boldsymbol{x}_k$. $\boldsymbol{x}_k = \boldsymbol{A}_k \boldsymbol{x}_{k-1} + \boldsymbol{w}_k$
$\hat{\boldsymbol{x}}_k^-$	The linear prediction (*a priori*) estimate of the process $\boldsymbol{x}_k$ at sample k. $\hat{\boldsymbol{x}}_k^- = \boldsymbol{A}_k \hat{\boldsymbol{x}}_{k-1}^+$
$\hat{\boldsymbol{x}}_k^+$	The Kalman filter (*a posteriori*) estimate of the process $\boldsymbol{x}_k$ at sample k. $\hat{\boldsymbol{x}}_k^+ = (\boldsymbol{I} - \boldsymbol{K}_k \boldsymbol{C}_k)\hat{\boldsymbol{x}}_k^- + \boldsymbol{K}_k \boldsymbol{z}_k = (\boldsymbol{I} - \boldsymbol{K}_k \boldsymbol{C}_k)\boldsymbol{A}_k \hat{\boldsymbol{x}}_{k-1}^+ + \boldsymbol{K}_k \boldsymbol{z}_k$
$\hat{\boldsymbol{x}}_{\varepsilon,k}^-$	The error in the linear prediction (*a priori*) estimate of the process $\boldsymbol{x}_k$. $\hat{\boldsymbol{x}}_{\varepsilon,k}^- = \hat{\boldsymbol{x}}_k^- - \boldsymbol{x}_k$
$\hat{\boldsymbol{x}}_{\varepsilon,k}^+$	The error in the *a posteriori* Kalman filter estimate of the process $\boldsymbol{x}_k$. $\hat{\boldsymbol{x}}_{\varepsilon,k}^+ = \hat{\boldsymbol{x}}_k^+ - \boldsymbol{x}_k$
$\boldsymbol{z}_k$	The measured process at sample k. $\boldsymbol{z}_k = \boldsymbol{C}_k \boldsymbol{x}_k + \boldsymbol{v}_k$
$\delta_{k,j}$	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. The proof is trivial.

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

2.2 Lemma 2

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

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

Proof: 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:

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

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

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

By inspection:

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

Substituting equation (5) into equation (4) completes the proof:

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

2.3 Lemma 3

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

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

Proof: 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 $\hat{\boldsymbol{x}}_k^+$ of the underlying process $\boldsymbol{x}_k$ from i) extrapolation from the previous iteration's *a posteriori* estimate $\hat{\boldsymbol{x}}_{k-1}^+$ and ii) from the current measurement $\boldsymbol{z}_k$:

$$\hat{\boldsymbol{x}}_k^+ = \boldsymbol{K}_k'\hat{\boldsymbol{x}}_{k-1}^+ + \boldsymbol{K}_k\boldsymbol{z}_k \quad (22)$$

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 $\boldsymbol{K}_k'\hat{\boldsymbol{x}}_{k-1}^+$ and vice versa. The Kalman filter is, therefore, a time varying, recursive filter.

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

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

$$E\left[\hat{\boldsymbol{x}}_{\varepsilon,k}^+\right] = E\left[\hat{\boldsymbol{x}}_k^+ - \boldsymbol{x}_k\right] = \boldsymbol{0} \quad (23)$$

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

$$\hat{\boldsymbol{x}}_{\varepsilon,k}^+ = \hat{\boldsymbol{x}}_k^+ - \boldsymbol{x}_k = \boldsymbol{K}_k'\hat{\boldsymbol{x}}_{k-1}^+ + \boldsymbol{K}_k\boldsymbol{z}_k - \boldsymbol{x}_k \quad (24)$$

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

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

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

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

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

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

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

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:

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

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

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

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

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

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

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

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 $\hat{\boldsymbol{x}}_{k-1}^+$:

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

Definition of *a posteriori* estimate

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

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

An equivalent form is:

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

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 (34):

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

$\boldsymbol{P}_k^-$ 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:

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

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

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 $\boldsymbol{P}_{k-1}^+$:

$$\boldsymbol{P}_k^- = E[(\boldsymbol{A}_k\hat{\boldsymbol{x}}_{k-1}^+ - \boldsymbol{A}_k\boldsymbol{x}_{k-1} - \boldsymbol{w}_k)(\boldsymbol{A}_k\hat{\boldsymbol{x}}_{k-1}^+ - \boldsymbol{A}_k\boldsymbol{x}_{k-1} - \boldsymbol{w}_k)^T] \quad (39)$$

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

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

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

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 $\boldsymbol{P}_k^+$:

$$E\left[\hat{\boldsymbol{x}}_{\varepsilon,k}^+{}^T\hat{\boldsymbol{x}}_{\varepsilon,k}^+\right] = tr(\boldsymbol{P}_k^+) \quad (43)$$

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

$$\hat{\boldsymbol{x}}_k^+ = \hat{\boldsymbol{x}}_{\varepsilon,k}^+ + \boldsymbol{x}_k = (\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{C}_k)\left(\hat{\boldsymbol{x}}_{\varepsilon,k}^- + \boldsymbol{x}_k\right) + \boldsymbol{K}_k(\boldsymbol{C}_k\boldsymbol{x}_k + \boldsymbol{v}_k) \quad (44)$$

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

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

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

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

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

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

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

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):

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

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

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

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:

$$\frac{\partial\{tr(\boldsymbol{P}_k^- {\boldsymbol{C}_k}^T {\boldsymbol{K}_k}^T)\}}{\partial \boldsymbol{K}_k} = \frac{\partial\{tr(\boldsymbol{K}_k \boldsymbol{C}_k \boldsymbol{P}_k^-)\}}{\partial \boldsymbol{K}_k} = (\boldsymbol{C}_k \boldsymbol{P}_k^-)^T = \boldsymbol{P}_k^- {\boldsymbol{C}_k}^T \quad \textbf{(53)}$$

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

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

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

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

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

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

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

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

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

Rearranging equation (57) gives:

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

Substituting $\boldsymbol{K}_k\boldsymbol{Q}_{v,k}$ from equation (59) into equation (49) gives:

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

$$= \boldsymbol{P}_k^- - \boldsymbol{K}_k\boldsymbol{C}_k\boldsymbol{P}_k^- - \boldsymbol{P}_k^-\boldsymbol{C}_k{}^T\boldsymbol{K}_k{}^T + \boldsymbol{K}_k\boldsymbol{C}_k\boldsymbol{P}_k^-\boldsymbol{C}_k{}^T\boldsymbol{K}_k{}^T + \boldsymbol{P}_k^-\boldsymbol{C}_k{}^T\boldsymbol{K}_k{}^T - \boldsymbol{K}_k\boldsymbol{C}_k\boldsymbol{P}_k^-\boldsymbol{C}_k{}^T\boldsymbol{K}_k{}^T \quad (61)$$

$$\Rightarrow \boldsymbol{P}_k^+ = (\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{C}_k)\boldsymbol{P}_k^- \qquad \textbf{Kalman equation (E)} \quad (62)$$

This completes the derivation of the standard Kalman filter equations.

3.3 Standard Kalman Equations

This section simply re-lists the key Kalman filter equations derived in the previous section.

Kalman equation (A)

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 $\hat{\boldsymbol{x}}_{k-1}^+$.

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

Kalman equation (B)

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 $\boldsymbol{Q}_{w,k}$.

$$\boldsymbol{P}_k^- = \boldsymbol{A}_k\boldsymbol{P}_{k-1}^+\boldsymbol{A}_k{}^T + \boldsymbol{Q}_{w,k} \quad \textbf{(B1)}$$

Kalman equations (B) and (E) 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 (E):

$$\boldsymbol{P}_k^- = \boldsymbol{A}_k(\boldsymbol{I} - \boldsymbol{K}_{k-1}\boldsymbol{C}_{k-1})\boldsymbol{P}_{k-1}^-\boldsymbol{A}_k{}^T + \boldsymbol{Q}_{w,k} \quad \textbf{(B2)}$$

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

Kalman equation (C)

The Kalman filter gain matrix $\boldsymbol{K}_k$ is updated:

$$\boldsymbol{K}_k = \boldsymbol{P}_k^-\boldsymbol{C}_k{}^T\left(\boldsymbol{C}_k\boldsymbol{P}_k^-\boldsymbol{C}_k{}^T + \boldsymbol{Q}_{v,k}\right)^{-1} \quad \textbf{(C)}$$

Kalman equation (D)

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

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

Kalman equation (E)

The *a posteriori* Kalman error covariance matrix $\boldsymbol{P}_k^+$ is updated and ready for the next iteration. This equation can be skipped if $\boldsymbol{P}_k^-$ is updated recursively.

$$\boldsymbol{P}_k^+ = (\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{C}_k)\boldsymbol{P}_k^- \tag{E}$$

3.4 Limiting Cases

From equation (C), as the measurement noise covariance $\boldsymbol{Q}_{v,k}$ decreases relative to the process noise covariance $\boldsymbol{Q}_{w,k}$, the Kalman gain matrix $\boldsymbol{K}_k$ satisfies:

$$\boldsymbol{K}_k\boldsymbol{C}_k\boldsymbol{P}_k^-{\boldsymbol{C}_k}^T = \boldsymbol{P}_k^-{\boldsymbol{C}_k}^T \Rightarrow \boldsymbol{K}_k\boldsymbol{C}_k = \boldsymbol{I} \tag{63}$$

Using equation (D), the *a posteriori* estimate $\hat{\boldsymbol{x}}_k^+$ is then only dependent on the measurement $\boldsymbol{z}_k$:

$$\hat{\boldsymbol{x}}_k^+ = (\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{C}_k)\hat{\boldsymbol{x}}_k^- + \boldsymbol{K}_k\boldsymbol{z}_k = \boldsymbol{K}_k\boldsymbol{z}_k \tag{64}$$

As the measurement noise covariance $\boldsymbol{Q}_{v,k}$ increases relative to the process noise covariance $\boldsymbol{Q}_{w,k}$, the Kalman gain matrix $\boldsymbol{K}_k$ approaches zero:

$$\boldsymbol{K}_k = \boldsymbol{P}_k^-{\boldsymbol{C}_k}^T\left(\boldsymbol{Q}_{v,k}\right)^{-1} = \boldsymbol{0} \tag{65}$$

The *a posteriori* process estimate $\hat{\boldsymbol{x}}_k^+$ then approximates the *a priori* estimate $\hat{\boldsymbol{x}}_k^-$:

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

4. Legal information

4.1 Definitions

Draft — The document is a draft version only. The content is still under internal review and subject to formal approval, which may result in modifications or additions. NXP Semiconductors does not give any representations or warranties as to the accuracy or completeness of information included herein and shall have no liability for the consequences of use of such information.

4.2 Disclaimers

Information in this document is provided solely to enable system and software implementers to use NXP products. There are no express or implied copyright licenses granted hereunder to design or fabricate any integrated circuits based on the information in this document. NXP reserves the right to make changes without further notice to any products herein. NXP makes no warranty, representation, or guarantee regarding the suitability of its products for any particular purpose, nor does NXP assume any liability arising out of the application or use of any product or circuit, and specifically disclaims any and all liability, including without limitation consequential or incidental damages. “Typical” parameters that may be provided in NXP data sheets and/ or specifications can and do vary in different applications, and actual performance may vary over time. All operating parameters, including “typicals,” must be validated for each customer application by customer's technical experts. NXP does not convey any license under its patent rights nor the rights of others. NXP sells products pursuant to standard terms and conditions of sale, which can be found at the following address: nxp.com/salestermsandconditions.

4.3 Trademarks

Notice: All referenced brands, product names, service names and trademarks are property of their respective owners.

NXP, the NXP logo, Freescale, and the Freescale logo are trademarks of NXP B.V. ARM and Cortex are registered trademarks of ARM Limited (or its subsidiaries) in the EU and/or elsewhere.

5. Contents

Date of release: 21 June 2016

Document identifier: AN5018