

Basic Kalman Filter Theory

Technical Note

Document TBD

Version: Draft

Authors: Mark Pedley and Michael Stanley

Date: September 2014

Table of Contents

- 1 Introduction.....4**
- 2 Mathematical Lemmas.....5**
 - 2.1 Lemma 1.....5
 - 2.2 Lemma 2.....5
 - 2.3 Lemma 3.....6
- 3 Kalman Filter Derivation8**
 - 3.1 Process Model8
 - 3.2 Derivation8
 - 3.3 Standard Kalman Equations 11

Glossary

| | |
|--------------------------------------|---|
| A_k | <p>The linear prediction or state matrix at sample k.</p> $\mathbf{x}_k = A_k \mathbf{x}_{k-1} + \mathbf{w}_k$ $\hat{\mathbf{x}}_k^- = A_k \hat{\mathbf{x}}_{k-1}^+$ |
| C_k | <p>The measurement matrix relating \mathbf{x}_k to \mathbf{z}_k at sample k.</p> $\mathbf{z}_k = C_k \mathbf{x}_k + \mathbf{v}_k$ |
| $E[]$ | Expectation operator |
| K_k | The Kalman filter gain at sample k |
| P_k^- | <p>The <i>a priori</i> covariance matrix of the linear prediction (<i>a priori</i>) error $\hat{\mathbf{x}}_{\varepsilon,k}^-$ at sample k.</p> $P_k^- = E[\hat{\mathbf{x}}_{\varepsilon,k}^- \hat{\mathbf{x}}_{\varepsilon,k}^{-T}]$ |
| P_k^+ | <p>The <i>a posteriori</i> covariance matrix of the Kalman (<i>a posteriori</i>) error $\hat{\mathbf{x}}_{\varepsilon,k}^+$ at sample k.</p> $P_k^+ = E[\hat{\mathbf{x}}_{\varepsilon,k}^+ \hat{\mathbf{x}}_{\varepsilon,k}^{+T}]$ |
| $Q_{w,k}$ | <p>The covariance matrix of the additive noise \mathbf{w}_k on the process \mathbf{x}_k</p> $Q_{w,k} = E[\mathbf{w}_k \mathbf{w}_k^T]$ |
| $Q_{v,k}$ | <p>The covariance matrix of the additive noise \mathbf{v}_k on the measured process \mathbf{z}_k</p> $Q_{v,k} = E[\mathbf{v}_k \mathbf{v}_k^T]$ |
| $V[]$ | Variance operator |
| \mathbf{v}_k | The additive noise on the measured process \mathbf{z}_k at sample k |
| \mathbf{w}_k | The additive noise on the process of interest \mathbf{x}_k at sample k |
| \mathbf{x}_k | <p>The state vector at time sample k of the process of interest \mathbf{x}_k</p> $\mathbf{x}_k = A_k \mathbf{x}_{k-1} + \mathbf{w}_k$ |
| $\hat{\mathbf{x}}_k^-$ | <p>The linear prediction (<i>a priori</i>) estimate of the process \mathbf{x}_k at sample k.</p> $\hat{\mathbf{x}}_k^- = A_k \hat{\mathbf{x}}_{k-1}^+$ |
| $\hat{\mathbf{x}}_k^+$ | <p>The Kalman filter (<i>a posteriori</i>) estimate of the process \mathbf{x}_k at sample k.</p> $\hat{\mathbf{x}}_k^+ = (I - K_k C_k) \hat{\mathbf{x}}_k^- + K_k \mathbf{z}_k = (I - K_k C_k) A_k \hat{\mathbf{x}}_{k-1}^+ + K_k \mathbf{z}_k$ |
| $\hat{\mathbf{x}}_{\varepsilon,k}^-$ | <p>The error in the linear prediction (<i>a priori</i>) estimate of the process \mathbf{x}_k.</p> $\hat{\mathbf{x}}_{\varepsilon,k}^- = \hat{\mathbf{x}}_k^- - \mathbf{x}_k$ |
| $\hat{\mathbf{x}}_{\varepsilon,k}^+$ | <p>The error in the <i>a posteriori</i> Kalman filter estimate of the process \mathbf{x}_k.</p> $\hat{\mathbf{x}}_{\varepsilon,k}^+ = \hat{\mathbf{x}}_k^+ - \mathbf{x}_k$ |
| \mathbf{z}_k | <p>The measured process at sample k.</p> $\mathbf{z}_k = C_k \mathbf{x}_k + \mathbf{v}_k$ |
| $\delta_{k,j}$ | The Kronecker delta function. $\delta_{k,j} = 1$ for $k = j$ and zero otherwise. |

1 Introduction

This document describes the assumptions underlying the basic Kalman filter and derives the standard Kalman equations. It is intended as a primer that should be read before tackling the documentation for the more specialized indirect complementary Kalman filter used for the fusion of accelerometer, magnetometer and gyroscope data.

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

2 Mathematical Lemmas

2.1 Lemma 1

The trace of the sum of two matrices equals the sum of the individual traces.

Proof

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

Eq 2.1.1

2.2 Lemma 2

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

Proof

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

Eq 2.2.1

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

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

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

Eq 2.2.2

Substituting gives:

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

Eq 2.2.3

By inspection:

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

Eq 2.2.4

Substituting back gives:

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

Eq 2.2.5

2.3 Lemma 3

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

Proof

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

Eq 2.3.1

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

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

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

Eq 2.3.2

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

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

Eq 2.3.3

The trace of matrix \mathbf{D} has value:

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

Eq 2.3.4

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

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

Eq 2.3.5

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

Eq 2.3.6

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

Eq 2.3.7

If \mathbf{B} is also symmetric then:

$$\frac{\partial \{tr(\mathbf{A}\mathbf{B}\mathbf{A}^T)\}}{\partial \mathbf{A}} = 2\mathbf{A}\mathbf{B} \text{ if } \mathbf{B} = \mathbf{B}^T$$

Eq 2.3.8

3 Kalman Filter Derivation

3.1 Process Model

The Kalman filter models the vector process of interest x_k as linear and recursive:

$$x_k = A_k x_{k-1} + w_k \quad \text{Eq 3.1.1}$$

If x_k has N degrees of freedom then A_k is an $N \times N$ linear prediction matrix (possibly time varying but assumed known) and w_k is an $N \times 1$ noise vector.

The process x_k is assumed to be not directly measurable and must be estimated from a process z_k which can be measured. z_k is modeled as being linearly related to x_k with additive noise v_k .

$$z_k = C_k x_k + v_k \quad \text{Eq 3.1.2}$$

z_k is an $N \times 1$ vector, C_k is an $N \times N$ matrix (possibly time varying but assumed known) and v_k is an $N \times 1$ noise vector.

The noise vectors w_k and v_k are assumed to be zero mean white processes:

$$E[w_k] = 0 \quad \text{Eq 3.1.3}$$

$$E[v_k] = 0 \quad \text{Eq 3.1.4}$$

$$\text{cov}\{w_k, w_j\} = E[w_k w_j^T] = Q_{w,k} \delta_{kj} \quad \text{Eq 3.1.5}$$

$$\text{cov}\{v_k, v_j\} = E[v_k v_j^T] = Q_{v,k} \delta_{kj} \quad \text{Eq 3.1.6}$$

By definition, covariance matrices are symmetric.

$$Q_{w,k}^T = \{E[w_k w_k^T]\}^T = E[(w_k w_k^T)^T] = E[w_k w_k^T] = Q_{w,k} \quad \text{Eq 3.1.7}$$

3.2 Derivation

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

$$\hat{x}_k^+ = K'_k \hat{x}_{k-1}^+ + K_k z_k \quad \text{Eq 3.2.1}$$

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

Unbiased estimate constraint (determines K'_k)

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

$$E[\hat{x}_{\varepsilon,k}^+] = E[\hat{x}_k^+ - x_k] = 0 \quad \text{Eq 3.2.2}$$

Subtracting \mathbf{x}_k from equation 3.2.1 gives:

$$\hat{\mathbf{x}}_{\varepsilon,k}^+ = \hat{\mathbf{x}}_k^+ - \mathbf{x}_k = \mathbf{K}'_k \hat{\mathbf{x}}_{k-1}^+ + \mathbf{K}_k \mathbf{z}_k - \mathbf{x}_k \quad \text{Eq 3.2.3}$$

Substituting equation 3.1.2 for \mathbf{z}_k gives:

$$\hat{\mathbf{x}}_{\varepsilon,k}^+ = \mathbf{K}'_k \hat{\mathbf{x}}_{k-1}^+ + \mathbf{K}_k (\mathbf{C}_k \mathbf{x}_k + \mathbf{v}_k) - \mathbf{x}_k \quad \text{Eq 3.2.4}$$

Substituting for \mathbf{x}_k from equation 3.1.1 and re-arranging gives:

$$\hat{\mathbf{x}}_{\varepsilon,k}^+ = \mathbf{K}'_k (\hat{\mathbf{x}}_{\varepsilon,k-1}^+ + \mathbf{x}_{k-1}) + \mathbf{K}_k \{\mathbf{C}_k (\mathbf{A}_k \mathbf{x}_{k-1} + \mathbf{w}_k) + \mathbf{v}_k\} - (\mathbf{A}_k \mathbf{x}_{k-1} + \mathbf{w}_k) \quad \text{Eq 3.2.5}$$

$$= \mathbf{K}'_k \hat{\mathbf{x}}_{\varepsilon,k-1}^+ + (\mathbf{K}_k \mathbf{C}_k \mathbf{A}_k - \mathbf{A}_k + \mathbf{K}'_k) \mathbf{x}_{k-1} + (\mathbf{K}_k \mathbf{C}_k - \mathbf{I}) \mathbf{w}_k + \mathbf{K}_k \mathbf{v}_k \quad \text{Eq 3.2.6}$$

Taking the expectation value of equation 3.2.6 and applying the unbiased estimate constraint gives:

$$E[\hat{\mathbf{x}}_{\varepsilon,k}^+] = E[\mathbf{K}'_k \hat{\mathbf{x}}_{\varepsilon,k-1}^+] + E[(\mathbf{K}_k \mathbf{C}_k \mathbf{A}_k - \mathbf{A}_k + \mathbf{K}'_k) \mathbf{x}_{k-1}] + E[(\mathbf{K}_k \mathbf{C}_k - \mathbf{I}) \mathbf{w}_k] + E[\mathbf{K}_k \mathbf{v}_k] = \mathbf{0} \quad \text{Eq 3.2.7}$$

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

$$E[(\mathbf{K}_k \mathbf{C}_k - \mathbf{I}) \mathbf{w}_k] = E[\mathbf{K}_k \mathbf{v}_k] = \mathbf{0} \quad \text{Eq 3.2.8}$$

With the additional assumption that the process \mathbf{x}_{k-1} is independent of the Kalman matrices at iteration k :

$$E[(\mathbf{K}_k \mathbf{C}_k \mathbf{A}_k - \mathbf{A}_k + \mathbf{K}'_k) \mathbf{x}_{k-1}] = (\mathbf{K}_k \mathbf{C}_k \mathbf{A}_k - \mathbf{A}_k + \mathbf{K}'_k) E[\mathbf{x}_{k-1}] = \mathbf{0} \quad \text{Eq 3.2.9}$$

Since \mathbf{x}_k is not, in general, a zero mean process:

$$\mathbf{K}_k \mathbf{C}_k \mathbf{A}_k - \mathbf{A}_k + \mathbf{K}'_k = \mathbf{0} \Rightarrow \mathbf{K}'_k = \mathbf{A}_k - \mathbf{K}_k \mathbf{C}_k \mathbf{A}_k = (\mathbf{I} - \mathbf{K}_k \mathbf{C}_k) \mathbf{A}_k \quad \text{Eq 3.2.10}$$

Eliminating \mathbf{K}'_k in equation 3.2.1 gives:

$$\hat{\mathbf{x}}_k^+ = (\mathbf{I} - \mathbf{K}_k \mathbf{C}_k) \mathbf{A}_k \hat{\mathbf{x}}_{k-1}^+ + \mathbf{K}_k \mathbf{z}_k \quad \text{Eq 3.2.11}$$

A priori estimate

The *a priori* Kalman filter estimate $\hat{\mathbf{x}}_k^-$ is defined as resulting from the application of the linear prediction matrix \mathbf{A}_k to the previous iteration's *a posteriori* estimate $\hat{\mathbf{x}}_{k-1}^+$:

$$\hat{\mathbf{x}}_k^- = \mathbf{A}_k \hat{\mathbf{x}}_{k-1}^+ \quad \text{Kalman equation 1} \quad \text{Eq 3.2.12}$$

Definition of a posteriori estimate

Substituting the *a priori* estimate $\hat{\mathbf{x}}_k^-$ into equation 3.2.11 gives:

$$\hat{\mathbf{x}}_k^+ = (\mathbf{I} - \mathbf{K}_k \mathbf{C}_k) \hat{\mathbf{x}}_k^- + \mathbf{K}_k \mathbf{z}_k \quad \text{Kalman equation 4} \quad \text{Eq 3.2.13}$$

An equivalent form is:

$$\hat{\mathbf{x}}_k^+ = \hat{\mathbf{x}}_k^- + \mathbf{K}_k (\mathbf{z}_k - \mathbf{C}_k \hat{\mathbf{x}}_k^-) \quad \text{Eq 3.2.14}$$

P_k^- as a function of P_{k-1}^+

The *a priori* and *a posteriori* error covariance matrices P_k^- and P_k^+ are defined as:

$$P_k^- = cov\{\hat{x}_{\varepsilon,k}^-, \hat{x}_{\varepsilon,k}^-\} = E[\hat{x}_{\varepsilon,k}^- \hat{x}_{\varepsilon,k}^{-T}] = E[(\hat{x}_k^- - x_k)(\hat{x}_k^- - x_k)^T] \quad \text{Eq 3.2.15}$$

$$P_k^+ = cov\{\hat{x}_{\varepsilon,k}^+, \hat{x}_{\varepsilon,k}^+\} = E[\hat{x}_{\varepsilon,k}^+ \hat{x}_{\varepsilon,k}^{+T}] = E[(\hat{x}_k^+ - x_k)(\hat{x}_k^+ - x_k)^T] \quad \text{Eq 3.2.16}$$

Substituting the definitions of \hat{x}_k^- and x_k into equation 3.2.15 gives:

$$P_k^- = E[(A_k \hat{x}_{k-1}^+ - A_k x_{k-1} - w_k)(A_k \hat{x}_{k-1}^+ - A_k x_{k-1} - w_k)^T] \quad \text{Eq 3.2.17}$$

$$= E[\{A_k(\hat{x}_{k-1}^+ - x_{k-1}) - w_k\}\{A_k(\hat{x}_{k-1}^+ - x_{k-1}) - w_k\}^T] \quad \text{Eq 3.2.18}$$

$$= A_k E[(\hat{x}_{k-1}^+ - x_{k-1})(\hat{x}_{k-1}^+ - x_{k-1})^T] A_k^T + Q_{w,k} \quad \text{Eq 3.2.19}$$

$$\Rightarrow P_k^- = A_k P_{k-1}^+ A_k^T + Q_{w,k} \quad \text{Kalman equation 2} \quad \text{Eq 3.2.20}$$

Minimum error covariance constraint (determines K_k)

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

$$E[\hat{x}_{\varepsilon,k}^{+T} \hat{x}_{\varepsilon,k}^+] = tr(P_k^+) \quad \text{Eq 3.2.21}$$

Substituting equation 2.1.2 for z_k into equation 3.2.11 gives a relation between the *a posteriori* and *a priori* errors:

$$\hat{x}_k^+ = \hat{x}_{\varepsilon,k}^+ + x_k = (I - K_k C_k) \hat{x}_k^- + K_k z_k = (I - K_k C_k)(\hat{x}_{\varepsilon,k}^- + x_k) + K_k(C_k x_k + v_k) \quad \text{Eq 3.2.22}$$

$$\Rightarrow \hat{x}_{\varepsilon,k}^+ + x_k = (I - K_k C_k) \hat{x}_{\varepsilon,k}^- + x_k - K_k C_k x_k + K_k(C_k x_k + v_k) \quad \text{Eq 3.2.23}$$

$$\Rightarrow \hat{x}_{\varepsilon,k}^+ = (I - K_k C_k) \hat{x}_{\varepsilon,k}^- + K_k v_k \quad \text{Eq 3.2.24}$$

Substituting this result into the definition of the *a posteriori* covariance matrix P_k^+ gives:

$$P_k^+ = E[\{(I - K_k C_k) \hat{x}_{\varepsilon,k}^- + K_k v_k\} \{(I - K_k C_k) \hat{x}_{\varepsilon,k}^- + K_k v_k\}^T] \quad \text{Eq 3.2.25}$$

$$= (I - K_k C_k) E[\hat{x}_{\varepsilon,k}^- \hat{x}_{\varepsilon,k}^{-T}] (I - K_k C_k)^T + K_k E[v_k v_k^T] K_k^T \quad \text{Eq 3.2.26}$$

$$= (I - K_k C_k) P_k^- (I - K_k C_k)^T + K_k Q_{v,k} K_k^T \quad \text{Eq 3.2.27}$$

$$= P_k^- - P_k^- C_k^T K_k^T - K_k C_k P_k^- + K_k C_k P_k^- C_k^T K_k^T + K_k Q_{v,k} K_k^T \quad \text{Eq 3.2.28}$$

The Kalman filter gain K_k is that which minimizes the trace of the *a posteriori* error covariance matrix P_k^+ :

$$\frac{\partial}{\partial \mathbf{K}_k} \text{tr}(\mathbf{P}_k^+) = \frac{\partial}{\partial \mathbf{K}_k} \{ \text{tr}(\mathbf{P}_k^-) - \text{tr}(\mathbf{P}_k^- \mathbf{C}_k^T \mathbf{K}_k^T) - \text{tr}(\mathbf{K}_k \mathbf{C}_k \mathbf{P}_k^-) + \text{tr}(\mathbf{K}_k \mathbf{C}_k \mathbf{P}_k^- \mathbf{C}_k^T \mathbf{K}_k^T) + \text{tr}(\mathbf{K}_k \mathbf{Q}_{v,k} \mathbf{K}_k^T) \} = 0$$

Eq 3.2.29

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

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

Eq 3.2.30

Since the trace of a transposed matrix equals the trace of the original matrix and using equation 2.2.5 gives:

$$\frac{\partial \{ \text{tr}(\mathbf{P}_k^- \mathbf{C}_k^T \mathbf{K}_k^T) \}}{\partial \mathbf{K}_k} = \frac{\partial \{ \text{tr}(\mathbf{K}_k \mathbf{C}_k \mathbf{P}_k^-) \}}{\partial \mathbf{K}_k} = (\mathbf{C}_k \mathbf{P}_k^-)^T = \mathbf{P}_k^- \mathbf{C}_k^T$$

Eq 3.2.31

The third term can be simplified using equations 2.3.7 and 2.3.8 exploiting the fact that the covariance matrix is symmetric:

$$\frac{\partial \{ \text{tr}(\mathbf{K}_k \mathbf{C}_k \mathbf{P}_k^- \mathbf{C}_k^T \mathbf{K}_k^T) \}}{\partial \mathbf{K}_k} = \mathbf{K}_k \{ \mathbf{C}_k \mathbf{P}_k^- \mathbf{C}_k^T + (\mathbf{C}_k \mathbf{P}_k^- \mathbf{C}_k^T)^T \} = 2 \mathbf{K}_k \mathbf{C}_k \mathbf{P}_k^- \mathbf{C}_k^T$$

Eq 3.2.32

The final term can be simplified also using equations 2.3.7 and 2.3.8 to give:

$$\frac{\partial \{ \text{tr}(\mathbf{K}_k \mathbf{Q}_{v,k} \mathbf{K}_k^T) \}}{\partial \mathbf{K}_k} = 2 \mathbf{K}_k \mathbf{Q}_{v,k}$$

Eq 3.2.33

Substituting back into equation 2.2.29 gives the optimal Kalman filter gain matrix \mathbf{K}_k :

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

Eq 3.2.34

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

Eq 3.2.35

$$\Rightarrow \mathbf{K}_k = \mathbf{P}_k^- \mathbf{C}_k^T (\mathbf{C}_k \mathbf{P}_k^- \mathbf{C}_k^T + \mathbf{Q}_{v,k})^{-1}$$

Kalman equation 3
Eq 3.2.36

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

Rearranging equation 3.2.35 gives:

$$\mathbf{K}_k \mathbf{Q}_{v,k} = \mathbf{P}_k^- \mathbf{C}_k^T - \mathbf{K}_k \mathbf{C}_k \mathbf{P}_k^- \mathbf{C}_k^T$$

Eq 3.2.37

Substituting equation 3.2.37 into equation 3.2.27 gives:

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

Eq 3.2.38

$$\Rightarrow \mathbf{P}_k^+ = (\mathbf{I} - \mathbf{K}_k \mathbf{C}_k) \mathbf{P}_k^-$$

Kalman equation 5
Eq 3.2.39

This completes the derivation of the standard Kalman filter equations.

3.3 Standard Kalman Equations

Kalman equation 1

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

$$\hat{\mathbf{x}}_k^- = \mathbf{A}_k \hat{\mathbf{x}}_{k-1}^+ \quad \text{Eq 3.3.1}$$

Kalman equation 2

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

$$\mathbf{P}_k^- = \mathbf{A}_k \mathbf{P}_{k-1}^+ \mathbf{A}_k^T + \mathbf{Q}_{w,k} \quad \text{Eq 3.3.2}$$

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

$$\mathbf{P}_k^- = \mathbf{A}_k (\mathbf{I} - \mathbf{K}_{k-1} \mathbf{C}_{k-1}) \mathbf{P}_{k-1}^- \mathbf{A}_k^T + \mathbf{Q}_{w,k} \quad \text{Eq 3.3.3}$$

Kalman equation 3

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

$$\mathbf{K}_k = \mathbf{P}_k^- \mathbf{C}_k^T (\mathbf{C}_k \mathbf{P}_k^- \mathbf{C}_k^T + \mathbf{Q}_{v,k})^{-1} \quad \text{Eq 3.3.4}$$

Kalman equation 4

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

$$\hat{\mathbf{x}}_k^+ = \hat{\mathbf{x}}_k^- + \mathbf{K}_k (\mathbf{z}_k - \mathbf{C}_k \hat{\mathbf{x}}_k^-) = (\mathbf{I} - \mathbf{K}_k \mathbf{C}_k) \hat{\mathbf{x}}_k^- + \mathbf{K}_k \mathbf{z}_k \quad \text{Eq 3.3.5}$$

Kalman equation 5

The *a posteriori* Kalman error covariance matrix \mathbf{P}_k^+ is updated ready for the next iteration. This equation can be skipped if \mathbf{P}_k^- is updated recursively in terms of itself as in equation 3.3.3.

$$\mathbf{P}_k^+ = (\mathbf{I} - \mathbf{K}_k \mathbf{C}_k) \mathbf{P}_k^- \quad \text{Eq 3.3.6}$$

Information in this document is provided solely to enable system and software implementers to use Freescale Semiconductors products. There are no express or implied copyright licenses granted hereunder to design or fabricate any integrated circuits or integrated circuits based on the information in this document.

Freescale Semiconductor reserves the right to make changes without further notice to any products herein. Freescale Semiconductor makes no warranty, representation, or guarantee regarding the suitability of its products for any particular purpose, nor does Freescale Semiconductor assume any liability arising out of the application or use of any product or circuit, and specifically disclaims any liability, including without limitation consequential or incidental damages. "Typical" parameters that may be provided in Freescale Semiconductor 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. Freescale Semiconductor does not convey any license under its patent rights nor the rights of others. Freescale Semiconductor products are not designed, intended, or authorized for use as components in systems intended for surgical implant into the body, or other applications intended to support or sustain life, or for any other application in which failure of the Freescale Semiconductor product could create a situation where personal injury or death may occur. Should Buyer purchase or use Freescale Semiconductor products for any such unintended or unauthorized application,

Buyer shall indemnify Freescale Semiconductor and its officers, employees, subsidiaries, affiliates, and distributors harmless against all claims, costs, damages, and expenses, and reasonable attorney fees arising out of, directly or indirectly, any claim of personal injury or death associated with such unintended or unauthorized use, even if such claims alleges that Freescale Semiconductor was negligent regarding the design or manufacture of the part.

RoHS-compliant and/or Pb-free versions of Freescale products have the functionality and electrical characteristics as their non-RoHS-complaint and/or non-Pb-free counterparts.

For further information, see <http://www.freescale.com> or contact your Freescale sales representative.

For information on Freescale's Environmental Products program, go to <http://www.freescale.com/epp>.

Freescale™ and the Freescale logo are trademarks of Freescale Semiconductor, Inc.

All other product or service names are the property of their respective owners.

© 2013 Freescale Semiconductor, Inc.



Basic Kalman Filter Theory by [Freescale Semiconductor](#) is licensed under a [Creative Commons Attribution 4.0 International License](#).