Journal of Global Positioning Systems (2022)
Vol. 18, No. 1: 56-69
DOI: 10.5081/jgps.18.1.56

Innovative Formulation in Discrete Kalman Filtering with Constraints - A Generic Framework for Comprehensive Error Analysis

Jianguo Wang, Benjamin Brunson, Baoxin Hu

Department of Earth and Space Science and Engineering
Lassonde School of Engineering, York University, Toronto, Canada
Emails: jgwang@yorku.ca, brunson@yorku.ca and baoxin@yorku.ca

Corresponding author: Benjamin Brunson

Abstract: This manuscript establishes a generic framework for comprehensive error analysis in discrete Kalman filtering with constraints, which systematically provides a complete set of algorithmic formulas along with demonstrating an alternative process of theoretical analytics of discrete Kalman filter. This constructive work aims extensively to standardize the formulation of Kalman filter with constraints. In analogy to the similar framework for standard discrete Kalman filter (without any constraints), the proposed framework specifically considers: model formulation vs. the error sources, the solution of the state and process noise vectors, the residuals for the process noise vector and the measurement noise vector, the redundancy contribution of the predicted state vector, process noise vector and measurement vector, and other relevant essential aspects, of which some of the features are essential to comprehensive error analysis, but are nonexistent yet in the primary algorithm in Kalman filtering with constraints. Besides, the algorithmic form of the Extended Kalman filter with constraints is also provided for practical purpose. At the end, specific remarks about the developed framework are given to emphasize on its usage to a certain extend.

KEY WORDS: Kalman filter, state constraint, error analysis, generic framework, redundancy contribution.

1. INTRODUCTION

The Kalman filter is a recursive estimator that provides estimates of a group of selected states on the ground of a specific system model and measurements that are acquired over time. Its applications have steadily expanded in sciences and engineering since the 1960s.

Usually, the Kalman filter consists of a system model associated with its modeling errors as process noises and a measurement model associated with measurement noise. However, there are also many circumstances under which *a priori* knowledge of a dynamic system leads to equality constraints that may be imposed on the system states in Kalman filtering. Examples of this include path-constrained motion along roadways [Yang et al, 2005; Hasberg et al 2012] and constant velocity motion of tracking targets [Alouani and Blair, 1991]. In multisensor integrated navigation, the states representing the attitude commonly involve specific constraints, e.g., the elements of the direction cosine matrix have to conform to orthonormality conditions and the elements in a quaternion vector or rotation vector have to be in unit norm. Apparently, the formulation of indirect observation (Least Squares) adjustment with constraints in Geodesy and Geomatics has been generally standardized [Mikhail, 1970; Rao and Toutenburg, 1999; Wang, et al, 2019]. By contrast, the formulation on states-constrained Kalman filter is far from being standardized to the same degree.

Constrained Kalman filtering by augmentation was first proven by Doran [1992], which has been considered as a seminal paper on the subject [Pizzinga, 2012]. There exist several dominant strategies to impose constraints on the system states in Kalman filtering, which are generally divided into three categories [Simon, 2010; Khabbazi and Esfanjani, 2014]:

Reparameterization: this technique incorporates any system state constraints by reducing the parameterization of the system, through which the physical meaning of the system states may be lost [Simon, 2010].

"Perfect" Observations: this technique treats the system state constraints as pseudo-observations with zero variance. Without further simplification, it may cause numerical instability [Doran, 1992; Alouani and Blair, 1991].

Projection: this technique transforms the estimate of the system states onto a constraint surface [Khabbazi and Esfanjani, 2014]. Such transformation may be accomplished through projection of the system state estimate [Simon and Chia, 2002], projection of the system itself [Ko and Bitmead, 2007], or projection of the Kalman gain matrix [Teixeira et al, 2008]. State projection is the most commonly used method of imposing constraints on the system states in Kalman filtering [Khabbazi and Esfanjani, 2014]. The Kalman gain projection has been generalized for non-linear constraints [Xu et al, 2017]. These techniques may also apply their constraints less strictly by taking a weighted average between the constrained and the unconstrained solution [Baker and Thennadil, 2019], or by taking model uncertainty into account in the gain projection approach [Khabbazi and Esfanjani, 2015].

Besides, some other techniques have also been used to impose equality constraints in Kalman filtering that do not fit under the above mentioned three broad categories. Xu et al [2013] considered constraints *a priori* information that should also be incorporated into a system's dynamic models. Ghanbarpourasl and Zobar [2022] utilized singular value decomposition to separate the system state into a deterministic (i.e. fully constrained) and a stochastic component. Pizzinga [2012] framed the constrained Kalman filter as a recursive least-squares problem.

Unfortunately, there is still a lack of generic algorithmic formulas directly for the standard form of the discrete Kalman filter with constraints in literature for conducting comprehensive error analysis. This motivates the authors to develop a complete set of the generic formulas for it, so that one can easily adapt to theoretical development and practical implementation.

Following this introduction, this manuscript first summarizes the innovative alternate formulation of standard Kalman filter originally deduced by Wang [1997] and also specifically detailed and applied in [Caspary and Wang, 1998; Wang, 1997; Wang, 2008, 2009; Wang et al, 2009; Wang et al, 2009; Gopaul et al, 2010; Wang et al, 2010; Qian, 2017; Qian, et al, 2015, 2016; Wang et al, 2015, 2021; Zhang et al, 2017]. Then, as the core of this manuscript, Section 3 systematically develops the theoretical aspects and practical algorithm in discrete Kalman filtering with constraints, which innovatively promote the comprehensive error analysis. Section 4 further delivers the proposed algorithm in the form of Extended Kalman filter with constraints. The manuscript ends with concluding remarks in Section 5.

2. ALGORITHMIC FORMULATIONS OF STANDARD KALMAN FILTER

In general, a Kalman filter estimates the state vector by minimizing its mean squared errors after the minimum variance principle or equivalently its weighted sum of the residuals squared after the Principle of Least Squares, on the basis of operating system and measurement models recursively.

2.1 Standard form of Discrete Kalman filter

Let us define the standard form of Kalman filter first. Consider a linear or linearized system described in state space and the data are made available over a discrete time series $t_0, t_1, ..., t_k, ..., t_N$, of which each time instant corresponds to an observation epoch and is simply depicted as $0, 1, ..., k, ..., N$. Without loss of generality, the formulation here omits the deterministic system input.

At an arbitrary observation epoch k ($1 \le k \le N$), the system and measurement models are given as follows [Wang et al, 2021]:

$$\boldsymbol{x}(k) = \boldsymbol{A}(k, k-1)\boldsymbol{x}(k-1) + \boldsymbol{B}(k, k-1)\boldsymbol{w}(k) \quad (2.1)$$

(or simply $\boldsymbol{x}(k) = \boldsymbol{A}(k)\boldsymbol{x}(k-1) + \boldsymbol{B}(k)\boldsymbol{w}(k)$ (2.1a))

$$\boldsymbol{z}(k) = \boldsymbol{C}(k)\boldsymbol{x}(k) + \boldsymbol{\Delta}(k) \quad (2.2)$$

wherein $\boldsymbol{x}(k)$, $\boldsymbol{z}(k)$, $\boldsymbol{w}(k)$, and $\Delta(k)$ are the *n*-dimensional state-vector, the *p*-dimensional observation vector, the *m*-dimensional process noise vector, and the *p*-dimensional measurement noise vector, respectively, while $\boldsymbol{A}(k, k-1)$, $\boldsymbol{B}(k, k-1)$, and $\boldsymbol{C}(k)$ are the $n \times n$ coefficient matrix of $\boldsymbol{x}(k)$, the $n \times m$ coefficient matrix of $\boldsymbol{w}(k)$, and the $p \times n$ coefficient matrix of $\boldsymbol{z}(k)$, respectively. About the relevant stochastic information, $\boldsymbol{w}(k) \sim N(\boldsymbol{o}, \boldsymbol{Q}(k))$ and $\boldsymbol{\Delta}(k) \sim N(\boldsymbol{o}, \boldsymbol{R}(k))$ are assumed, where $N(a, b)$ represents a normal distribution with *a* and *b* as its expectation (vector) and variance (matrix). Between two different observation epochs, it is presumed to have $Cov(\boldsymbol{w}(i), \boldsymbol{w}(j)) = \boldsymbol{O}$ and $Cov(\boldsymbol{\Delta}(i), \boldsymbol{\Delta}(j)) = \boldsymbol{O}$ for ($i \ne j$), and $Cov(\boldsymbol{w}(i), \boldsymbol{\Delta}(j)) = \boldsymbol{O}$ for any *i* and *j*.

Besides, the initial state vector is given as $\boldsymbol{x}(0)$ with its variance matrix $\boldsymbol{D}_{xx}(0)$ and is independent of $\boldsymbol{w}(k)$ and $\Delta(k)$ for any *k*, i.e., $Cov(\boldsymbol{w}(k), \boldsymbol{x}(0)) = \boldsymbol{O}$ and $Cov(\boldsymbol{\Delta}(k), \boldsymbol{x}(0)) = \boldsymbol{O}$.

2.2 The Solution after Minimum Variance Principle

Without any unnecessary repetition of the solution derivation, the Kalman filtering algorithm at *k* from *k*-1 upon the definition in Section 2.1 after the minimum variance principle is directly summarized below:

$$\hat{\boldsymbol{x}}(k) = \hat{\boldsymbol{x}}(k/k-1) + \boldsymbol{G}(k)\boldsymbol{d}(k) \tag{2.3}$$

with its associated variance matrix

$$\begin{aligned}\boldsymbol{D}_{xx}(k) &= [I - \boldsymbol{G}(k)\boldsymbol{C}(k)]\boldsymbol{D}_{xx}(k/k-1) \\ &\cdot [I - \boldsymbol{G}(k)\boldsymbol{C}(k)]^T + \boldsymbol{G}(k)\boldsymbol{R}(k)\boldsymbol{G}^T(k)\end{aligned} \tag{2.4}$$

wherein $\boldsymbol{I}$ is a *n*x*n* identity matrix and $\boldsymbol{G}(k)$ is a *n*x*p* Kalman gain matrix:

$$\boldsymbol{G}(k) = \boldsymbol{D}_{xx}(k/k-1)\boldsymbol{C}^T(k)\boldsymbol{D}_{dd}^{-1}(k) \tag{2.5}$$

The predicted state vector (from the time update) and its variance matrix are as follows:

$$\hat{\boldsymbol{x}}(k/k-1) = \boldsymbol{A}(k)\hat{\boldsymbol{x}}(k-1/k-1) \tag{2.6}$$

$$\begin{aligned}\boldsymbol{D}_{xx}(k/k-1) &= \boldsymbol{A}(k)\boldsymbol{D}_{xx}(k-1/k-1)\boldsymbol{A}^T(k) \\ &+ \boldsymbol{B}(k)\boldsymbol{Q}(k)\boldsymbol{B}^T(k)\end{aligned} \tag{2.7}$$

The system innovation vector and its variance matrix are computed after:

$$\boldsymbol{d}(k) = \boldsymbol{z}(k) - \boldsymbol{C}(k)\hat{\boldsymbol{x}}(k/k-1) \tag{2.8}$$

$$\boldsymbol{D}_{dd}(k) = \boldsymbol{C}(k)\boldsymbol{D}_{xx}(k-1/k-1)\boldsymbol{C}^T(k) + \boldsymbol{R}(k) \tag{2.9}$$

Essentially, the system innovation vectors: $\boldsymbol{d}(1), \boldsymbol{d}(2), ..., \boldsymbol{d}(k), ...$ are independent of each other [Chui & Chen, 1987], i.e., $Cov(\boldsymbol{d}(i), \boldsymbol{d}(j)) = \boldsymbol{O}$ ($i \neq j$). However, the elements in $\boldsymbol{d}(k)$ at epoch *k* are not only correlated, but also blend all of the separate error sources. Traditionally, the error analysis has been centered on the system innovation series. In addition, it is proved that the estimate of the state vector $\boldsymbol{x}(k)$ and the system innovation vector $\boldsymbol{d}(k)$ are independent of each other based on (2.3) and (2.8), i.e.,

$$\boldsymbol{D}_{xd}(k) = O \tag{2.10}$$

2.3 Alternate Formulation for Comprehensive Error Analysis

Obviously, $\boldsymbol{d}(k)$ is originated from the process noise series $\boldsymbol{w}(1), ..., \boldsymbol{w}(k), ...$, the measurement noise series $\Delta(1), ..., \Delta(k), ...$ along with the initial state vector $\boldsymbol{x}(0)$. Therefore, as a matter of fact, the system and measurement models in (2.1) and (2.2) are associated with three groups of independent stochastic information that is propagated into the state solution from time to time. Specifically at *k*, the system is contaminated by (i) the measurement noise vector $\Delta(k)$, (ii) the process noise vector $\boldsymbol{w}(k)$, and (iii) the noise associated with the predicted state vector from $\boldsymbol{A}(k, k-1)\boldsymbol{x}(k-1)$, into which $\Delta(1), ...,$ $\Delta(k-1)$ and $\boldsymbol{w}(1), \ ..., \boldsymbol{w}(k-1)$ starting with $\boldsymbol{x}(0)$ are propagated through the recursive mechanism as in (2.1) and (2.2) from the past.

Along two different paths, either after the Minimum Variance Principle or Least Squares Principle, the Kalman filtering algorithm is equivalently derived. A widely repeated derivation is to deliver the equivalent solution on the ground of the predicted state vector $\boldsymbol{x}(k/k-1)$, as a pseudo-measurement vector by merging (ii) and (iii) as in (2.1) in Least Squares approach, and the measurement vector $\boldsymbol{z}(k)$ from (i). An apparent drawback to this formulation is that two groups of the independent stochastic information in (ii) and (iii) are blended into $\boldsymbol{x}(k/k-1)$ and are no more separable in error analysis.

To enhance the error analysis in discrete Kalman filtering, Wang [1997] proposed an innovative alternate formulation. Innovatively, the system state prediction in (2.1) was further split into two pseudo-measurement vectors:

$$\boldsymbol{l}_x(k) = \boldsymbol{A}(k)\hat{\boldsymbol{x}}(k-1) = \hat{\boldsymbol{x}}(k/k-1) \quad \boldsymbol{D}_{l_x l_x}(k) \tag{2.11}$$

$$\boldsymbol{l}_w(k) = \boldsymbol{w}_0(k) \qquad \boldsymbol{Q}(k) \tag{2.12}$$

with $\boldsymbol{w}_0(k) = \boldsymbol{o}$ (zero mean presumed) and

$$\boldsymbol{D}_{l_x l_x}(k) = \boldsymbol{A}(k)\boldsymbol{D}_{xx}(k-1)\boldsymbol{A}^T(k) \tag{2.13}$$

The real measurement vector $\boldsymbol{z}(k)$ remains as in (2.2) and denoted by $\boldsymbol{l}_z(k) = \boldsymbol{z}(k)$.

The residual equations corresponding to (2.11), (2.12) and (2.2) are as follows:

$$\boldsymbol{v}_{l_x}(k) = \quad \hat{\boldsymbol{x}}(k) - \boldsymbol{B}(k)\hat{\boldsymbol{w}}(k) - \boldsymbol{l}_x(k) \tag{2.14}$$

$$\boldsymbol{v}_{l_w}(k) = \quad \hat{\boldsymbol{w}}(k) - \boldsymbol{l}_w(k) \tag{2.15}$$

$$\boldsymbol{v}_{l_z}(k) = \boldsymbol{C}(k)\hat{\boldsymbol{x}}(k) \quad - \boldsymbol{l}_z(k) \tag{2.16}$$

with $\boldsymbol{D}_{l_x l_x}(k)$, $\boldsymbol{Q}(k)$ and $\boldsymbol{R}(k)$ as their measurement variance matrices, respectively, in which the state vector is extended to include the process noise vector $\boldsymbol{w}(k)$ being estimated together with $\boldsymbol{x}(k)$.

In seeking for a Least Squares solution for $\boldsymbol{x}(k)$ and $\boldsymbol{w}(k)$, the cost function is constructed

$$\begin{aligned}\boldsymbol{min}: g(k) &= \boldsymbol{v}_{l_x}^T(k)\boldsymbol{D}_{l_x l_x}^{-1}(k)\boldsymbol{v}_{l_x}(k) + \\ &\boldsymbol{v}_{l_w}^T(k)\boldsymbol{Q}^{-1}(k)\boldsymbol{v}_{l_w}(k) + \boldsymbol{v}_z^T(k)\boldsymbol{R}^{-1}(k)\boldsymbol{v}_z(k)\end{aligned} \tag{2.17}$$

In (2.14), (2.15) and (2.16), there are (*n* + *m*) states and (*n* + *m* +*p*) measurements. The number of the redundant measurements remains unchanged, namely, *p*. It is not in question about the identity of $\boldsymbol{x}(k)$ derived after (2.17) and the one in (2.3) [Wang, 1997]. The beauty of this formulation lies in the feasibility for the direct analysis of the three original error sources at any epoch *k*. Especially, it allows for reliability analysis in discrete Kalman filtering [Wang, 1997, 2009]. For the benefit of the presentation in next section, the outcome from this alternate formulation is summarized below:

1) The solution of the state vector

First, equations (2.3) – (2.9) in Section 2.2 remain unchanged to form the basis of the solution. Alternatively, (2.3) is also given as follows

$$\begin{aligned} \boldsymbol{x}(k/k) &= \boldsymbol{l}_x(k) + \boldsymbol{B}(k)\boldsymbol{l}_w(k) \\ &+ \boldsymbol{K}(k)\{\boldsymbol{l}_z(k) - \boldsymbol{C}(k)[\boldsymbol{l}_x(k) - \boldsymbol{B}(k)\boldsymbol{l}_w(k)]\} \end{aligned} \quad (2.18)$$

Importantly, the process noise vector is estimated by

$$\begin{aligned} \hat{\boldsymbol{w}}(k) &= \boldsymbol{w}_0(k) \\ &+ \boldsymbol{Q}(k)\boldsymbol{B}^T(k)\boldsymbol{D}_{xx}^{-1}(k/k-1)\boldsymbol{K}(k)\boldsymbol{d}(k) \end{aligned} \quad (2.19)$$

with its variance matrix

$$\begin{aligned} \boldsymbol{D}_{ww}(k) &= \boldsymbol{Q}(k) \\ &- \boldsymbol{Q}(k)\boldsymbol{B}^T(k)\boldsymbol{C}^T(k)\boldsymbol{D}_{dd}^{-1}(k)\boldsymbol{C}(k)\boldsymbol{B}(k)\boldsymbol{Q}(k) \end{aligned} \quad (2.20)$$

and its covariance matrix with the estimated state vector

$$\begin{aligned} \boldsymbol{D}_{xw}(k) &= \boldsymbol{B}(k)\boldsymbol{Q}(k) \\ &- \boldsymbol{D}_{xx}(k/k-1)\boldsymbol{C}^T(k)\boldsymbol{D}_{dd}^{-1}(k)\boldsymbol{C}(k)\boldsymbol{B}(k)\boldsymbol{Q}(k) \end{aligned} \quad (2.21)$$

2) The residual vectors

$$\boldsymbol{v}_{l_x}(k) = \boldsymbol{D}_{l_x l_x}(k)\boldsymbol{D}_{xx}^{-1}(k/k-1)\boldsymbol{G}(k)\boldsymbol{d}(k) \quad (2.22)$$

$$\boldsymbol{v}_w(k) = \boldsymbol{Q}(k)\boldsymbol{B}^T(k)\boldsymbol{D}_{xx}^{-1}(k/k-1)\boldsymbol{G}(k)\boldsymbol{d}(k) \quad (2.23)$$

$$\boldsymbol{v}_z(k) = [\boldsymbol{C}(k)\boldsymbol{G}(k) - \boldsymbol{I}]\boldsymbol{d}(k) \quad (2.24)$$

with ***their variance matrices***

$$\boldsymbol{D}_{v_{l_x}v_{l_x}}(k) = \boldsymbol{D}_{l_x l_x}(k)\boldsymbol{C}^T(k)\boldsymbol{D}_{dd}^{-1}(k-1)\boldsymbol{C}(k)\boldsymbol{D}_{l_x l_x}(k) \quad (2.25)$$

$$\boldsymbol{D}_{v_w v_w}(k) = \boldsymbol{Q}(k)\boldsymbol{B}^T(k)\boldsymbol{C}^T(k)\boldsymbol{D}_{dd}^{-1}(k)\boldsymbol{C}(k)\boldsymbol{B}(k)\boldsymbol{Q}(k) \quad (2.26)$$

$$\boldsymbol{D}_{v_z v_z}(k) = [\boldsymbol{I} - \boldsymbol{C}(k)\boldsymbol{G}(k)]\boldsymbol{R}(k) \quad (2.27)$$

3) ***The redundancy contributions*** in measurement groups corresponding to (2.11), (2.12) and 2.2):

$$r_{l_x}(k) = tr[\boldsymbol{A}(k)\boldsymbol{D}_{xx}(k-1)\boldsymbol{A}^T(k)\boldsymbol{C}^T(k)\boldsymbol{D}_{dd}^{-1}(k)\boldsymbol{C}(k)] \quad (2.28)$$

$$r_{l_w}(k) = tr[\boldsymbol{Q}(k)\boldsymbol{B}^T(k)\boldsymbol{C}^T(k)\boldsymbol{D}_{dd}^{-1}(k-1)\boldsymbol{C}(k)\boldsymbol{B}(k) \quad (2.29)$$

$$r_z(k) = tr[\boldsymbol{I} - \boldsymbol{C}(k)\boldsymbol{G}(k)] \quad (2.30)$$

For the entire system either after (2.1) and (2.2), or after (2.11), (2.12) and (2.2), the total redundancy number at epoch *k* satisfies [Wang, 1997; 2009, 2021; etc]

$$r(k) = r_{l_x}(k) + r_{l_w}(k) + r_z(k) = p(k) \quad (2.31)$$

wherein $p(k)$ is the number of the real measurements or the dimension of $\boldsymbol{z}(k)$.

4) The individual redundancy indexes

In practice, $\boldsymbol{Q}(k)$ and $\boldsymbol{R}(k)$ are commonly diagonal so that the individual redundancy indexes in components for the process noise vector are

$$r_{w_i}(k) = [\boldsymbol{Q}(k)\boldsymbol{B}^T(k)\boldsymbol{C}^T(k)\boldsymbol{D}_{dd}^{-1}(k)\boldsymbol{C}(k)\boldsymbol{B}(k)]_{ii}$$

$$(i = 1, 2, ..., m(k)) \quad (2.32)$$

and for the measurement vector

$$r_{z_i}(k) = [\boldsymbol{I} - \boldsymbol{C}(k)\boldsymbol{G}(k)]_{ii} \quad (i = 1, 2, ..., p(k)) \quad (2.33)$$

Indeed, as $\boldsymbol{D}_{l_x l_x}(k)$ in (2.11) is not a diagonal matrix in general, no individual redundancy indexes in components become meaningful here for $\boldsymbol{l}_x(k)$.

5) ***The variance of unit weight*** (the variance factor)

$$\hat{\sigma}_0^2(k) = \boldsymbol{d}^T(k)\boldsymbol{D}_{dd}^{-1}(k)\boldsymbol{d}(k)/p(k) \quad (2.34)$$

or

$$\begin{aligned} \hat{\sigma}_0^2(k) = [&\boldsymbol{v}_{l_x}^T(k)\boldsymbol{D}_{l_x l_x}^{-1}(k)\boldsymbol{v}_{l_x}(k) + \boldsymbol{v}_{l_w}^T(k)\boldsymbol{Q}^{-1}(k)\boldsymbol{v}_{l_w}(k) + \\ &+ \boldsymbol{v}_z^T(k)\boldsymbol{R}^{-1}(k)\boldsymbol{v}_z(k)]/p(k) \end{aligned} \quad (2.35)$$

6) ***The posteriori variance matrix*** of the estimated state vector

$$\hat{\boldsymbol{D}}_{xx}(k) = \hat{\sigma}_0^2(k)\boldsymbol{D}_{xx}(k) \quad (2.36)$$

which directly reflects the latest available residuals due to the modeling and measurement errors. For the usage in (2.36), one can apply the epochwise variance factor as in (2.34) or (2.35), a regional variance factor, i.e., an average over a specific time period, or even a global variance factor from the entire data period [Wang, 1997, 2009; Wang et al, 2021]. ***However, it is noticed that plenty of the applications with applying Kalman filter have inappropriately considered (2.4), instead of (2.36), as their posteriori state variance matrix***.

Refer to [Wang, 1997, 2008, 2009; Caspary and Wang, 1998; Wang et al, 2021] for more details about this alternate formulation and its advantages for error analysis in discrete Kalman filtering.

3. GENERIC FORMULATION OF DISCRETE KALMAN FILTER WITH CONSTRAINTS

This section provides readers with our original development of a generic formula set, which meaningfully serves as an innovative framework for comprehensive error analysis in discrete Kalman filtering with constraints in parallel with the one summarized in Section 2.3, and also describes their connections. In this work, the constraints are restricted to the equality constraints,

3.1 The Functional and Stochastic Models

Upon the models of the standard Kalman filter defined in Section 2.1, a Kalman filter with constraints indicates that there exist the following additional constraints among the states

$$\boldsymbol{H}^T(k)\boldsymbol{x}(k) - \boldsymbol{h} = \boldsymbol{o} \quad (3.1)$$

wherein $\boldsymbol{H}(k)$ is a $n\times h$-dimensional coefficient matrix and $tr[\boldsymbol{H}(k)] = h$ ($h < n$), which is either originally linear or linearized from nonlinear constraints and $\boldsymbol{h}$ is the *h*-dimensional constant vector. Hence, the equations (2.1), (2.2) and (3.1) together represent the system model, the measurement model, and the constraints among the states in discrete Kalman filtering.

In analogy to the alternate formulation summarized in Section 2.3, the Principle of Least Squares is straightforwardly applied epochwise hereinafter to result the solution for discrete Kalman filter with constraints. To demonstrate the flexibility in dealing with the available functional and stochastic models, three different ways that deliver an identical estimate of the state vector $\boldsymbol{x}(k)$ are introduced in Sections 3.2, 3.3 and 3.4, respectively, of which Section 3.4 is the focus of attention of this manuscript.

3.2 Approach One

The measurement equation system is here structured as follows

1) ***a pseudo-measurement vector*** $\boldsymbol{l}'_x(k)$ is given directly by using the solution of the state vector from the standard Kalman filter (without any constraints) as in Section 2, which establishes the following residual equation:

$$\boldsymbol{v}_x(k) = \boldsymbol{x}_h(k) - \boldsymbol{l}'_x(k) \quad (3.2)$$

wherein $\boldsymbol{x}_h(k)$ is the state estimate subject to the constraints as in (3.1) while the pseudo-measurement vector and its variance matrix are:

$$\boldsymbol{l}'_x(k) = \boldsymbol{x}(k/k) \quad \text{(refer to (2.3))} \quad (3.3)$$

$$\boldsymbol{D}_{l'_x l'_x}(k) = \boldsymbol{D}_{xx}(k) \quad \text{(refer to (2.4))} \quad (3.4)$$

2) ***a group of*** *h* ***linear constraints*** as in (3.1) are applied.

The equations (3.2) and (3.1) together compose the model in the form of indirect observations with constraints. So, the Principle of Least Squares is applied to the following cost function at epoch *k*:

$$\begin{aligned}&\boldsymbol{min}: g(\boldsymbol{x}_h(k)/\boldsymbol{l}'_x(k), \boldsymbol{z}(k)) = \\ &\boldsymbol{v}^T_{l'_x}(k)\boldsymbol{D}^{-1}_{l'_x l'_x}(k)\boldsymbol{v}_{l'_x}(k) + 2\boldsymbol{k}^T_h[\boldsymbol{H}^T(k)\boldsymbol{x}_h(k) - \boldsymbol{h}]\end{aligned} \quad (3.5)$$

which was called the Mean Square Method in [Simon and Chia, 2000]. To seek for the (minimum) extreme value of (3.5), its first order derivative with respect to $\boldsymbol{x}_h(k)$ is assigned to 0:

$$\begin{aligned}&\frac{\partial}{\partial \boldsymbol{x}_h(k)} g(\boldsymbol{x}_h(k)/\boldsymbol{l}'_x(k), \boldsymbol{z}(k)) \\ &= 2\boldsymbol{v}^T_{l'_x}(k)\boldsymbol{D}^{-1}_{xx}(k/k) + 2\boldsymbol{k}^T_h \boldsymbol{H}^T(k) = \boldsymbol{o}\end{aligned} \quad (3.6)$$

which yields

$$\boldsymbol{D}^{-1}_{xx}(k)\boldsymbol{x}_h(k) + \boldsymbol{H}(k)\boldsymbol{k}_h(k) - \boldsymbol{D}^{-1}_{xx}(k)\boldsymbol{l}'_x(k) = \boldsymbol{o} \quad (3.7)$$

The equations (3.7) and (3.1) together compose a normal equation system:

$$\begin{pmatrix}\boldsymbol{D}^{-1}_{xx}(k/k) & \boldsymbol{H}(k) \\ \boldsymbol{H}^T(k) & \boldsymbol{O}\end{pmatrix}\begin{pmatrix}\boldsymbol{x}_h(k) \\ \boldsymbol{k}_h\end{pmatrix} = \begin{pmatrix}\boldsymbol{D}^{-1}_{xx}(k/k)\boldsymbol{l}'_x(k) \\ \boldsymbol{h}\end{pmatrix} \quad (3.8)$$

which possesses two unknown parameter vectors: the state vector $\boldsymbol{x}_h(k)$ and the Lagrange multiplier vector $\boldsymbol{k}_h(k)$ brought by the constraints.

To solve (3.8), one can first derive $\boldsymbol{x}_h(k)$ from the first equation:

$$\begin{aligned}\boldsymbol{x}_h(k) &= \boldsymbol{D}_{xx}(k)[\boldsymbol{D}^{-1}_{xx}(k/k)\boldsymbol{l}'_x(k) - \boldsymbol{H}(k)\boldsymbol{k}_h(k)] \\ &= \boldsymbol{x}(k/k) - \boldsymbol{D}_{xx}(k/k)\boldsymbol{H}(k)\boldsymbol{k}_h(k)\end{aligned} \quad (3.9)$$

wherein $\boldsymbol{x}(k/k)$ is the minimum variance estimate of the state vector given in (2.3). Substituting (3.9) into the second equation of (3.8) delivers the Lagrange multiplier vector $\boldsymbol{k}_h(k)$:

$$\boldsymbol{k}_h(k) = \boldsymbol{N}^{-1}_{hh}(k)[\boldsymbol{H}^T(k)\boldsymbol{x}(k/k) - \boldsymbol{h}] \quad (3.10)$$

where a helping matrix $\boldsymbol{N}_{hh}(k)$ is defined to simplify the notation

$$\boldsymbol{N}_{hh}(k) = \boldsymbol{H}^T(k)\boldsymbol{D}_{xx}(k/k)\boldsymbol{H}(k) \quad (3.11)$$

The substitution of (3.10) into (3.9) gives $\boldsymbol{x}_h(k)$

$$\begin{aligned}&\hat{\boldsymbol{x}}_h(k) = \boldsymbol{x}(k/k) \\ &- \boldsymbol{D}_{xx}(k/k)\boldsymbol{H}(k)\boldsymbol{N}^{-1}_{hh}(k)\{\boldsymbol{H}^T(k)\boldsymbol{x}(k/k) - \boldsymbol{h}\}\end{aligned} \quad (3.12)$$

in which the overhead symbol ^ is commonly ignored wherever no confusion may occur.

The associated variance matrix with $\boldsymbol{x}_h(k)$ is derived based on (3.12):

$$\begin{aligned}&\boldsymbol{D}_{x_h x_h}(k) = \boldsymbol{D}_{xx}(k/k) \\ &- \boldsymbol{D}_{xx}(k/k)\boldsymbol{H}(k)\boldsymbol{N}^{-1}_{hh}(k)\boldsymbol{H}^T(k)\boldsymbol{D}_{xx}(k/k)\end{aligned} \quad (3.13)$$

wherein $\boldsymbol{D}_{xx}(k/k)$ is the variance matrix of $\boldsymbol{x}(k/k)$ in (2.4).

This solution is indeed identical with the one after Maximum Probability Method and Projection Method presented in Simon and Chia [2000].

3.3 Approach Two

Differently from Approach One in Section 3.2, the measurement equation system is here structured as follows

1) ***A pseudo-measurement vector*** $\boldsymbol{l}''_x(k)$ is defined by the predicted state vector $\boldsymbol{x}(k/k-1)$ from t_{k-1} to t_k (i.e., time update) from the standard Kalman filter as in Section 2, which establishes the following residual equation:

$$\boldsymbol{v}_{l''_x}(k) = \boldsymbol{x}_h(k) - \boldsymbol{l}''_x(k) \quad (3.14)$$

with

$$\boldsymbol{l}''_x(k) = \boldsymbol{x}(k/k-1) = \boldsymbol{A}(k)\boldsymbol{x}(k-1) \quad (3.15)$$

$$\begin{aligned} \boldsymbol{D}_{l''_x l''_x}(k) &= \boldsymbol{D}_{xx}(k/k-1) \\ &= \boldsymbol{A}(k)\boldsymbol{D}_{xx}(k-1)\boldsymbol{A}^T(k) + \boldsymbol{B}(k)\boldsymbol{Q}(k)\boldsymbol{B}^T(k) \end{aligned} \quad (3.16)$$

2) ***A measurement vector*** $\boldsymbol{l}_z(k)$ is adapted from the real measurement vector $\boldsymbol{z}(k)$ at t_k as in (2.2), which yields the following residual equation:

$$\boldsymbol{v}_{l_z}(k) = \boldsymbol{v}_z(k) = \boldsymbol{C}(k)\boldsymbol{x}_h(k) - \boldsymbol{l}_z(k) \quad (3.17)$$

with

$$\boldsymbol{l}_z(k) = \boldsymbol{z}(k) \quad (3.18)$$

$$\boldsymbol{D}_{l_z l_z}(k) = \boldsymbol{R}(k) \quad (3.19)$$

wherein $\boldsymbol{C}(k)$ is the same as in (2.2).

3) ***a group of*** *h* ***linear constraints*** as in (3.1) are applied, wherein ***h*** (**bold** and *italic*) is the constant vector in the constraints.

Now, the equations (3.14), (3.17) and (3.1) together compose another model in the form of indirect observations with constraints at epoch *k*. Accordingly, the cost function for applying the Principle of Least Squares is as below:

$$\begin{aligned} \min: & g(\boldsymbol{x}_h(k)/\boldsymbol{l}''_x(k), \boldsymbol{z}(k)) \\ &= \boldsymbol{v}^T_{l''_x}(k)\boldsymbol{D}^{-1}_{l''_x l''_x}(k)\boldsymbol{v}_{l''_x}(k) + \boldsymbol{v}^T_{l_z}(k)\boldsymbol{R}^{-1}(k)\boldsymbol{v}_{l_z}(k) \\ &\quad + 2\boldsymbol{k}^T_h(k)[\boldsymbol{H}^T(k)\boldsymbol{x}_h(k) - \boldsymbol{h}(k)] \end{aligned} \quad (3.20)$$

The same as with (3.5), the 1st order derivative of (3.20) with respect to $\boldsymbol{x}(k)$ is assigned to 0:

$$\begin{aligned} \frac{\partial}{\partial \boldsymbol{x}_h(k)} g(\boldsymbol{x}_h(k)/\boldsymbol{l}''_x(k), \boldsymbol{z}(k)) &= 2\boldsymbol{v}^T_{l''_x}(k)\boldsymbol{D}^{-1}_{l''_x l''_x}(k) \\ &\quad + 2\boldsymbol{v}^T_{l_z}(k)\boldsymbol{R}^{-1}(k)\boldsymbol{C}(k) + 2\boldsymbol{k}^T_h\boldsymbol{H}^T(k) = \boldsymbol{o} \end{aligned} \quad (3.21)$$

which gives

$$\begin{aligned} &\{\boldsymbol{D}^{-1}_{l_x l_x}(k) + \boldsymbol{C}^T(k)\boldsymbol{R}^{-1}(k)\boldsymbol{C}(k)\}\boldsymbol{x}_h(k) + \boldsymbol{H}(k)\boldsymbol{k}_h \\ &- \boldsymbol{D}^{-1}_{l_x l_x}(k)\boldsymbol{l}_x(k) - \boldsymbol{C}^T(k)\boldsymbol{R}^{-1}(k)\boldsymbol{z}(k) = \boldsymbol{o} \end{aligned} \quad (3.22)$$

From (3.22) and (3.1), the normal equation system goes as follows:

$$\begin{aligned} &\begin{pmatrix} \boldsymbol{D}^{-1}_{l''_x l''_x}(k) + \boldsymbol{C}^T(k)\boldsymbol{R}^{-1}(k)\boldsymbol{C}(k) & \boldsymbol{H}(k) \\ \boldsymbol{H}^T(k) & \boldsymbol{O} \end{pmatrix} \begin{pmatrix} \boldsymbol{x}_h(k) \\ \boldsymbol{k}_h \end{pmatrix} \\ &= \begin{pmatrix} \boldsymbol{D}^{-1}_{l''_x l''_x}(k)\boldsymbol{l}''_x(k) + \boldsymbol{C}^T(k)\boldsymbol{R}^{-1}(k)\boldsymbol{z}(k) \\ \boldsymbol{h} \end{pmatrix} \end{aligned} \quad (3.23)$$

which is identical to (3.8) because it can be proved

$$[\boldsymbol{D}^{-1}_{l''_x l''_x}(k) + \boldsymbol{C}^T(k)\boldsymbol{R}^{-1}(k)\boldsymbol{C}(k)]^{-1} = \boldsymbol{D}_{xx}(k/k) \quad (3.24)$$

and

$$\begin{aligned} &\boldsymbol{D}^{-1}_{l''_x l''_x}(k)\boldsymbol{l}''_x(k) + \boldsymbol{C}^T(k)\boldsymbol{R}^{-1}(k)\boldsymbol{z}(k) \\ &= \boldsymbol{D}^{-1}_{xx}(k/k-1)\boldsymbol{x}(k/k-1) + \boldsymbol{C}^T(k)\boldsymbol{R}^{-1}(k)\boldsymbol{z}(k) \\ &= \boldsymbol{D}^{-1}_{xx}(k)\boldsymbol{x}(k) \end{aligned} \quad (3.25)$$

This implies that (3.8) and (3.23) result in the identical solution for the state vector.

3.4 Approach Three

Furthermore, differently from Approaches One and Two, Approach Three here develops the proposed framework for comprehensive error analysis in discrete Kalman filtering with constraints, which is particularly an extension of (2.14) – (2.16) by adding the constraints among the states. The measurement equation system is hereto structured as follows:

1) ***The first pseudo-measurement vector*** $\boldsymbol{l}_x(k)$ is here defined by the predicted state vector exclusive of the effect of the process noise vector. Its residual equation is (refer to (2.14):

$$\boldsymbol{v}_{l_x}(k) = \boldsymbol{x}_h(k) - \boldsymbol{B}(k)\boldsymbol{w}_h(k) - \boldsymbol{l}_x(k) \quad (3.25)$$

$$\boldsymbol{l}_x(k) = \boldsymbol{A}(k)\boldsymbol{x}(k-1) \quad (3.26)$$

$$\boldsymbol{D}_{l_x l_x}(k) = \boldsymbol{A}(k)\boldsymbol{D}_{xx}(k-1)\boldsymbol{A}^T(k) \quad (3.27)$$

2) ***The second pseudo-measurement vector*** $\boldsymbol{l}_w(k)$ is defined by the process noise vector, which gives the residual equation below (refer to (2.15):

$$\boldsymbol{v}_{l_w}(k) = \boldsymbol{w}_h(k) - \boldsymbol{l}_w(k) \quad (3.28)$$

$$\boldsymbol{l}_w(k) = \boldsymbol{w}_0(k) \quad (\text{usually } \boldsymbol{w}_0(k) = \boldsymbol{o}) \quad (3.29)$$

$$\boldsymbol{D}_{l_w l_w}(k) = \boldsymbol{Q}(k) \quad (3.30)$$

3) ***A measurement vector*** $\boldsymbol{l}_z(k)$ is adapted from the real measurement vector $\boldsymbol{z}(k)$ at t_k as in (2.2). So, the residual equation is as (3.17) alongside with (3.18) and (3.19).

4) a group of *h* ***linear or linearized constraints*** are as in (3.1).

Essentially, one must give one's attention to (3.26), $\boldsymbol{l}_x(k) \neq \boldsymbol{x}(k/k-1)$ because

$$\begin{aligned} \boldsymbol{x}(k/k-1) &= \boldsymbol{l}_x(k) + \boldsymbol{B}(k)\boldsymbol{l}_w(k) \\ &= \boldsymbol{A}(k)\boldsymbol{x}(k-1) + \boldsymbol{B}(k)\boldsymbol{w}_0(k) \end{aligned} \quad (3.31)$$

Writing four equations (3.25), (3.28), (3.17), and (3.1) together gives the entire residual equation system with constraints as below:

$$\begin{pmatrix} \boldsymbol{v}^h_x(k) \\ \boldsymbol{v}^h_w(k) \\ \boldsymbol{v}^h_z(k) \end{pmatrix} = \begin{pmatrix} \boldsymbol{E}_x & -\boldsymbol{B}(k) \\ \boldsymbol{O} & \boldsymbol{E}_w \\ \boldsymbol{C}(k) & \boldsymbol{O} \end{pmatrix} \begin{pmatrix} \boldsymbol{x}_h(k) \\ \boldsymbol{w}_h(k) \end{pmatrix} - \begin{pmatrix} \boldsymbol{l}_x(k) \\ \boldsymbol{l}_w(k) \\ \boldsymbol{l}_z(k) \end{pmatrix} \quad (3.32)$$

$$\boldsymbol{H}^T(k)\boldsymbol{x}_h(k) - \boldsymbol{h} = \boldsymbol{o}$$

alongside with the blockwise covariance matrix of three independent measurement vectors $\boldsymbol{l}_x(k)$, $\boldsymbol{l}_w(k)$ and $\boldsymbol{z}(k)$ as in (3.27), (3.30), (3.19). The main difference of (3.32) from Approach One in Section 3.2 and Approach Two in Section 3.3 lies in directly modeling three originally independent random vectors as the measurement vectors. Accordingly, the unknown parameters have been extended from $\boldsymbol{x}_h(k)$ to both of $\boldsymbol{x}_h(k)$ and $\boldsymbol{w}_h(k)$, This modeling strategy allows estimating the process noise vector epochwise and also the residual vector of $\boldsymbol{l}_w(k)$, which has been of scarcely any mention in literature, except initially modeled in Wang [1997].

Frankly, (3.32) allows specifying the following cost function for applying the Principle of Least Squares:

$$\begin{aligned}&\boldsymbol{min}: g(\boldsymbol{x}_h(k),\boldsymbol{w}_h(k)/\boldsymbol{l}_x(k),\boldsymbol{l}_w(k),\boldsymbol{l}_z(k))\\&=\boldsymbol{v}_{l_x}^T(k)\boldsymbol{D}_{l_xl_x}^{-1}(k)\boldsymbol{v}_{l_x}(k)+\boldsymbol{v}_{l_w}^T(k)\boldsymbol{Q}^{-1}(k)\boldsymbol{v}_{l_w}(k)\\&\quad+\boldsymbol{v}_{l_z}^T(k)\boldsymbol{R}^{-1}(k)\boldsymbol{v}_{l_z}(k)\\&\quad+2\boldsymbol{k}_h^T(k)[\boldsymbol{H}^T(k)\boldsymbol{x}_h(k)-\boldsymbol{h}(k)]\end{aligned}\tag{3.33}$$

which yields two 1st order partial derivatives for $\boldsymbol{x}_h(k)$ and $\boldsymbol{w}_h(k)$, respectively:

$$\begin{aligned}&\frac{\partial g(\boldsymbol{x}_h(k),\boldsymbol{w}_h(k)/\boldsymbol{l}_x(k),\boldsymbol{l}_w(k),\boldsymbol{l}_z(k))}{\partial \boldsymbol{x}_h(k)}\\&=2\boldsymbol{v}_{l_x}^T(k)\boldsymbol{D}_{l_xl_x}^{-1}(k)\\&\quad+2\boldsymbol{v}_{l_z}^T(k)\boldsymbol{R}^{-1}(k)\boldsymbol{C}(k)+2\boldsymbol{k}_h^T(k)\boldsymbol{H}^T(k)=\boldsymbol{o}\end{aligned}\tag{3.34}$$

$$\begin{aligned}&\frac{\partial g(\boldsymbol{x}_h(k),\boldsymbol{w}_h(k)/\boldsymbol{l}_x(k),\boldsymbol{l}_w(k),\boldsymbol{l}_z(k))}{\partial \boldsymbol{w}(k)}\\&=-2\boldsymbol{v}_{l_x}^T(k)\boldsymbol{D}_{l_xl_x}^{-1}(k)\boldsymbol{B}(k)+2\boldsymbol{v}_{l_w}^T(k)\boldsymbol{Q}^{-1}(k)=\boldsymbol{o}\end{aligned}\tag{3.35}$$

Together with (3.1), (3.34) and (3.35) build up the corresponding normal equation system:

$$\begin{aligned}&\begin{pmatrix}\boldsymbol{D}_{l_xl_x}^{-1}(k)+\boldsymbol{C}^T(k)\boldsymbol{R}^{-1}(k)\boldsymbol{C}(k) & -\boldsymbol{D}_{l_xl_x}^{-1}(k)\boldsymbol{B}(k) & \boldsymbol{H}(k)\\ -\boldsymbol{B}^T(k)\boldsymbol{D}_{l_xl_x}^{-1}(k) & \boldsymbol{B}^T(k)\boldsymbol{D}_{l_xl_x}^{-1}(k)\boldsymbol{B}(k)+\boldsymbol{Q}^{-1}(k) & \boldsymbol{O}\\ \boldsymbol{H}^T(k) & \boldsymbol{O} & \boldsymbol{O}\end{pmatrix}\\&\cdot\begin{pmatrix}\boldsymbol{x}_h(k)\\ \boldsymbol{w}_h(k)\\ \boldsymbol{k}_h(k)\end{pmatrix}=\begin{pmatrix}\boldsymbol{D}_{l_xl_x}^{-1}(k)\boldsymbol{l}_x(k)+\boldsymbol{C}^T(k)\boldsymbol{R}^{-1}(k)\boldsymbol{l}_z(k)\\ -\boldsymbol{B}^T(k)\boldsymbol{D}_{l_xl_x}^{-1}(k)\boldsymbol{l}_x(k)+\boldsymbol{Q}^{-1}(k)\boldsymbol{l}_w(k)\\ \boldsymbol{h}(k)\end{pmatrix}\end{aligned}\tag{3.36}$$

Although deducing an explicit solution of (3.36) affirmatively seems complicated because the coefficient matrix of (3.36) is in the form of a 3×3 partitioned block matrix, we have successfully accomplished the algorithmic formulation of the solution for $\boldsymbol{x}_h(k)$, $\boldsymbol{w}_h(k)$ and $\boldsymbol{k}_h(k)$ inclusive of some further relevant contents, e.g., the residual vectors and redundancy contribution and redundant indexes of the measurements etc.

Before the solution is delivered, the equivalency of (3.36) to (3.8) and (3.23) is first proved. With the 2nd equation in (3.36), three specifics need readers' attention for the benefit of further derivation:

i) The coefficient matrix of $\boldsymbol{x}_h(k)$ in the 1st equation of (3.26) is $\boldsymbol{D}_{xx}^{-1}(k/k)$ (refer to (3.24)).

ii) The inverse of the coefficient matrix of $\boldsymbol{w}_h(k)$ in the 2nd equation of (3.26) gives

$$\begin{aligned}&[\boldsymbol{B}^T(k)\boldsymbol{D}_{l_xl_x}^{-1}(k)\boldsymbol{B}(k)+\boldsymbol{Q}^{-1}(k)]^{-1}\\&=\boldsymbol{Q}(k)-\boldsymbol{Q}(k)\boldsymbol{B}^T(k)\boldsymbol{D}_{xx}^{-1}(k/k-1)\boldsymbol{B}(k)\boldsymbol{Q}(k)\end{aligned}\tag{3.37}$$

iii) Solving for $\boldsymbol{w}_h(k)$ from the 2nd equation in (3.26) gives

$$\begin{aligned}\boldsymbol{w}_h(k)&=[\boldsymbol{Q}(k)-\boldsymbol{Q}(k)\boldsymbol{B}^T(k)\boldsymbol{D}_{xx}^{-1}(k/k-1)\boldsymbol{B}(k)\boldsymbol{Q}(k)]\cdot\\&\quad\cdot\boldsymbol{B}^T(k)\boldsymbol{D}_{l_xl_x}^{-1}(k)\boldsymbol{x}_h(k)\\&\quad+[\boldsymbol{Q}(k)-\boldsymbol{Q}(k)\boldsymbol{B}^T(k)\boldsymbol{D}_{xx}^{-1}(k/k-1)\boldsymbol{B}(k)\boldsymbol{Q}(k)]\cdot\\&\quad\cdot[-\boldsymbol{B}^T(k)\boldsymbol{D}_{l_xl_x}^{-1}(k)\boldsymbol{l}_x(k)+\boldsymbol{Q}^{-1}(k)\boldsymbol{l}_w(k)]\end{aligned}\tag{3.38}$$

Substituting (3.38) into the 1st equation of (3.36) eliminates $\boldsymbol{w}_h(k)$

$$\begin{aligned}&\begin{pmatrix}\boldsymbol{D}_{xx}^{-1}(k/k) & \boldsymbol{H}(k)\\ \boldsymbol{H}^T(k) & \boldsymbol{O}\end{pmatrix}\begin{pmatrix}\boldsymbol{x}_h(k)\\ \boldsymbol{k}_h(k)\end{pmatrix}=\\&\begin{pmatrix}\boldsymbol{D}_{xx}^{-1}(k/k-1)\boldsymbol{x}(k/k-1)+\boldsymbol{C}^T(k)\boldsymbol{R}^{-1}(k)\boldsymbol{z}(k)\\ \boldsymbol{h}\end{pmatrix}\end{aligned}\tag{3.39}$$

which proved that (3.36) is equivalent to (3.8) and (3.23) for $\boldsymbol{x}_h(k)$ and $\boldsymbol{k}_h(k)$ as

$$\begin{aligned}&\boldsymbol{D}_{xx}^{-1}(k/k-1)[\boldsymbol{l}_x(k)+\boldsymbol{B}(k)\boldsymbol{l}_w(k)]+\boldsymbol{C}^T(k)\boldsymbol{R}^{-1}(k)\boldsymbol{z}(k)\\&=\boldsymbol{D}_{xx}^{-1}(k/k-1)\boldsymbol{x}(k/k-1)+\boldsymbol{C}^T(k)\boldsymbol{R}^{-1}(k)\boldsymbol{z}(k)\\&=\boldsymbol{D}_{xx}^{-1}(k)\boldsymbol{x}(k)\end{aligned}\tag{3.40}$$

3.5 Solution

Now, without providing the lengthy intermediate steps, the solution of $\boldsymbol{x}_h(k)$, $\boldsymbol{w}_h(k)$ and $\boldsymbol{k}_h(k)$ is directly given below:

$$\begin{aligned}\boldsymbol{x}_h(k)&=\boldsymbol{x}(k/k-1)+\boldsymbol{K}(k)\boldsymbol{d}(k)\\&\quad-\boldsymbol{D}_{xx}(k/k)\boldsymbol{H}(k)\boldsymbol{k}_h(k)\end{aligned}\tag{3.41}$$

$$\begin{aligned}\boldsymbol{w}_h(k)&=\boldsymbol{l}_w(k)+\boldsymbol{Q}(k)\boldsymbol{B}^T(k)\boldsymbol{D}_{xx}^{-1}(k/k-1)\boldsymbol{K}(k)\boldsymbol{d}(k)\\&\quad-\boldsymbol{Q}(k)\boldsymbol{B}^T(k)\boldsymbol{D}_{xx}^{-1}(k/k-1)\boldsymbol{D}_{xx}(k/k)\cdot\\&\quad\cdot\boldsymbol{H}(k)\boldsymbol{N}_{hh}^{-1}(k)[\boldsymbol{H}^T(k)\boldsymbol{x}(k/k)-\boldsymbol{h}]\end{aligned}\tag{3.42}$$

$$\begin{aligned}\boldsymbol{k}_h(k)&=\boldsymbol{N}_{hh}^{-1}(k)\{\boldsymbol{H}^T(k)\boldsymbol{D}_{xx}(k/k)\cdot\\&\quad\cdot[\boldsymbol{D}_{l_xl_x}^{-1}(k)\boldsymbol{x}(k/k-1)+\boldsymbol{C}^T(k)\boldsymbol{R}^{-1}(k)\boldsymbol{z}(k)]-\boldsymbol{h}\}\end{aligned}\tag{3.43}$$

with the variance–covariance matrices of $\boldsymbol{x}_h(k)$ and $\boldsymbol{w}_h(k)$:

$$\begin{aligned}&\boldsymbol{D}_{x_hx_h}(k)=\boldsymbol{D}_{xx}(k/k-1)\\&-\boldsymbol{D}_{xx}(k/k-1)\boldsymbol{C}^T(k)\boldsymbol{D}_{dd}^{-1}(k)\boldsymbol{C}(k)\boldsymbol{D}_{xx}(k/k-1)\\&-\boldsymbol{D}_{xx}(k/k)\boldsymbol{H}(k)\boldsymbol{N}_{hh}^{-1}(k)\boldsymbol{H}^T(k)\boldsymbol{D}_{xx}(k/k)\end{aligned}\tag{3.44}$$

$$\boldsymbol{D}_{\boldsymbol{w}_h\boldsymbol{w}_h}(k)=\boldsymbol{Q}(k) -\boldsymbol{Q}(k)\boldsymbol{B}^T(k)\boldsymbol{C}^T(k)\boldsymbol{D}_{dd}^{-1}(k)\boldsymbol{C}(k)\boldsymbol{B}(k)\boldsymbol{Q}(k) -\boldsymbol{Q}(k)\boldsymbol{B}^T(k)[\boldsymbol{I}-\boldsymbol{K}(k)\boldsymbol{C}(k)]^T\cdot \cdot\boldsymbol{H}(k)\boldsymbol{N}_{hh}^{-1}(k)\boldsymbol{H}^T(k)[\boldsymbol{I}-\boldsymbol{K}(k)\boldsymbol{C}(k)]\boldsymbol{B}(k)\boldsymbol{Q}(k) \quad (3.45)$$

$$\boldsymbol{D}_{x_h\boldsymbol{w}_h}(k-1)=[\boldsymbol{I}-\boldsymbol{D}_{xx}(k/k)\boldsymbol{H}(k)\boldsymbol{N}_{hh}^{-1}(k)\boldsymbol{H}^T(k)]\cdot \cdot[\boldsymbol{I}-\boldsymbol{K}(k)\boldsymbol{C}(k)]\boldsymbol{B}(k)\boldsymbol{Q}(k) \quad (3.46)$$

3.6 Solutions with and without Constraints

The solution of $\boldsymbol{x}_h(k)$ and $\boldsymbol{w}_h(k)$ in discrete Kalman filtering with constraints is connected to the solution of $\boldsymbol{x}(k)$ and $\boldsymbol{w}(k)$ (without constraints) given in Section 2.3 as follows:

$$\boldsymbol{x}_h(k)=\boldsymbol{x}(k/k) -\boldsymbol{D}_{xx}(k/k)\boldsymbol{H}(k)\boldsymbol{N}_{hh}^{-1}(k)[\boldsymbol{H}^T(k)\boldsymbol{x}(k/k)-\boldsymbol{h}] \quad (3.47)$$

$$\boldsymbol{w}_h(k)=\boldsymbol{w}(k)-\boldsymbol{Q}(k)\boldsymbol{B}^T(k)\boldsymbol{D}_{xx}^{-1}(k/k-1)\boldsymbol{D}_{xx}(k/k)\cdot \cdot\boldsymbol{H}(k)\boldsymbol{N}_{hh}^{-1}(k)[\boldsymbol{H}^T(k)\boldsymbol{x}(k/k)-\boldsymbol{h}] \quad (3.48)$$

$$\boldsymbol{D}_{x_hx_h}(k)=\boldsymbol{D}_{xx}(k/k) -\boldsymbol{D}_{xx}(k/k)\boldsymbol{H}(k)\boldsymbol{N}_{hh}^{-1}(k)\boldsymbol{H}^T(k)\boldsymbol{D}_{xx}(k/k) \quad (3.49)$$

$$\boldsymbol{D}_{\boldsymbol{w}_h\boldsymbol{w}_h}(k-1)=\boldsymbol{D}_{ww}(k) -\boldsymbol{Q}(k)\boldsymbol{B}^T(k)\boldsymbol{D}_{xx}^{-1}(k/k-1)\boldsymbol{D}_{xx}(k/k)\boldsymbol{H}(k)\boldsymbol{N}_{hh}^{-1}(k)\cdot \cdot\boldsymbol{H}^T(k)\boldsymbol{D}_{xx}(k/k)\boldsymbol{D}_{xx}^{-1}(k/k-1)\boldsymbol{B}(k)\boldsymbol{Q}(k) \quad (3.50)$$

$$\boldsymbol{D}_{x_h\boldsymbol{w}_h}(k-1)=\boldsymbol{D}_{xw}(k)-\boldsymbol{D}_{xx}(k/k)\boldsymbol{H}(k)\boldsymbol{N}_{hh}^{-1}(k)\cdot \cdot\boldsymbol{H}^T(k)[\boldsymbol{I}-\boldsymbol{K}(k)\boldsymbol{C}(k)]\boldsymbol{B}(k)\boldsymbol{Q}(k) \quad (3.51)$$

This group of formulas provides the opportunity to obtain the solution with constraints directly upon the solution from the standard Kalman filtering described in Section 2. A hard-won advantage of the solution expressions from (3.47) to (3.51) lies in first obtaining the solution after (2.3) (or (2.18)), (2.19), (2.4), (2.20) and (2.21) without considering the constraints and then utilizing $\boldsymbol{x}(k/k)$ to linearize the constraints, when they are nonlinear, and apply them towards the solution with constraints.

3.7 Residual Vectors and their Variance Matrices

For error analysis in Kalman filtering, $\boldsymbol{v}_{x_h}(k)$, $\boldsymbol{v}_{w_h}(k)$ (when $\boldsymbol{w}_0(k)=\boldsymbol{o}$) and $\boldsymbol{v}_{\boldsymbol{z}_h}(k)$ with their associated covariance matrices are further derived below.

In general, they can directly be calculated according to (3.32) or individually after (3.25), (3.28) and (3.17). However, they are further detailed.

First, with the residual vector $\boldsymbol{v}_x^h(k)$ of $\boldsymbol{l}_x(k)$ in (3.25), substituting (3.41) and (3.42) or (3.47) and (3.48) into (3.25) gives

$$\boldsymbol{v}_{l_x}^h(k)=\boldsymbol{x}(k/k)-\boldsymbol{l}_x(k)-\boldsymbol{B}(k)\boldsymbol{w}(k) -\boldsymbol{D}_{xx}(k/k)\boldsymbol{H}(k)\boldsymbol{N}_{hh}^{-1}(k)[\boldsymbol{H}^T(k)\boldsymbol{x}(k/k)-\boldsymbol{h}] +\boldsymbol{B}(k)\boldsymbol{Q}(k)\boldsymbol{B}^T(k)\boldsymbol{D}_{xx}^{-1}(k/k-1)\boldsymbol{D}_{xx}(k/k)\cdot \cdot\boldsymbol{H}(k)\boldsymbol{N}_{hh}^{-1}(k)[\boldsymbol{H}^T(k)\boldsymbol{x}(k/k)-\boldsymbol{h}] \quad (3.52)$$

or

$$\boldsymbol{v}_{l_x}^h(k)=\boldsymbol{D}_{l_xl_x}(k)\boldsymbol{D}_{xx}^{-1}(k/k-1)\boldsymbol{K}(k)\boldsymbol{d}(k) -\boldsymbol{D}_{l_xl_x}(k)[\boldsymbol{I}-\boldsymbol{K}(k)\boldsymbol{C}(k)]^T\cdot \cdot\boldsymbol{H}(k)\boldsymbol{N}_{hh}^{-1}(k)[\boldsymbol{H}^T(k)\boldsymbol{x}(k/k)-\boldsymbol{h}] \quad (3.52a)$$

Based on (2.22), (3.52) is further simplified to

$$\boldsymbol{v}_{l_x}^h(k)=\boldsymbol{v}_{l_x}(k)-[\boldsymbol{I}+\boldsymbol{B}(k)\boldsymbol{Q}(k)\boldsymbol{B}^T(k)\boldsymbol{D}_{xx}^{-1}(k/k-1)] \cdot\boldsymbol{D}_{xx}(k/k)\boldsymbol{H}(k)\boldsymbol{N}_{hh}^{-1}(k)[\boldsymbol{H}^T(k)\boldsymbol{x}(k/k)-\boldsymbol{h}] \quad (3.53)$$

Second, with the residual vector $\boldsymbol{v}_{w_h}(k)$ of $\boldsymbol{l}_w(k)$ in (3.28), the substitution of (3.42) or (3.48) yields

$$\boldsymbol{v}_w^h(k)=\boldsymbol{w}(k)-\boldsymbol{Q}(k)\boldsymbol{B}^T(k)\boldsymbol{D}_{xx}^{-1}(k/k-1)\boldsymbol{D}_{xx}(k/k) \cdot\boldsymbol{H}(k)\boldsymbol{N}_{hh}^{-1}(k)[\boldsymbol{H}^T(k)\boldsymbol{x}(k/k)-\boldsymbol{h}]-\boldsymbol{l}_w(k) \quad (3.54)$$

or

$$\boldsymbol{v}_{l_w}^h(k)=\boldsymbol{Q}(k)\boldsymbol{B}^T(k)\boldsymbol{D}_{xx}^{-1}(k/k-1)\boldsymbol{K}(k)\boldsymbol{d}(k) -\boldsymbol{Q}(k)\boldsymbol{B}^T(k)[\boldsymbol{I}-\boldsymbol{K}(k)\boldsymbol{C}(k)]^T \cdot\boldsymbol{H}(k)\boldsymbol{N}_{hh}^{-1}(k)[\boldsymbol{H}^T(k)\boldsymbol{x}(k/k)-\boldsymbol{h}] \quad (3.54a)$$

After (2.23), (3.54) is further reformed to

$$\boldsymbol{v}_{l_w}^h(k)=\boldsymbol{v}_{l_w}(k)-\boldsymbol{Q}(k)\boldsymbol{B}^T(k)\boldsymbol{D}_{xx}^{-1}(k/k-1)\boldsymbol{D}_{xx}(k/k) \cdot\boldsymbol{H}(k)\boldsymbol{N}_{hh}^{-1}(k)[\boldsymbol{H}^T(k)\boldsymbol{x}(k/k)-\boldsymbol{h}] \quad (3.55)$$

Because the initial value of $\boldsymbol{l}_w(k)$ is usually assumed to be: $\boldsymbol{w}_0(k)=\boldsymbol{o}$ in practice, (3.54) becomes

$$\boldsymbol{v}_{l_w}^h(k)=\boldsymbol{w}(k)-\boldsymbol{Q}(k)\boldsymbol{B}^T(k)\boldsymbol{D}_{xx}^{-1}(k/k-1)\boldsymbol{D}_{xx}(k/k) \cdot\boldsymbol{H}(k)\boldsymbol{N}_{hh}^{-1}(k)[\boldsymbol{H}^T(k)\boldsymbol{x}(k/k)-\boldsymbol{h}] \quad (3.56)$$

Third, with the residual vector $\boldsymbol{v}_z^h(k)$ of $\boldsymbol{l}_z(k)=\boldsymbol{z}(k)$, substituting (3.41) or (3.47) into (3.17) delivers:

$$\boldsymbol{v}_z^h(k)=\boldsymbol{C}(k)\{\boldsymbol{x}(k/k)-\boldsymbol{D}_{xx}(k/k)\boldsymbol{H}(k)\boldsymbol{N}_{hh}^{-1}(k) \cdot[\boldsymbol{H}^T(k)\boldsymbol{x}(k/k)-\boldsymbol{h}]\}-\boldsymbol{l}_z(k) \quad (3.57)$$

or

$$\boldsymbol{v}_z^h(k)=[\boldsymbol{C}(k)\boldsymbol{K}(k)-\boldsymbol{I}]\boldsymbol{d}(k)-\boldsymbol{C}(k)\boldsymbol{D}_{xx}(k/k) \cdot\boldsymbol{H}(k)\boldsymbol{N}_{hh}^{-1}(k)[\boldsymbol{H}^T(k)\boldsymbol{x}(k/k)-\boldsymbol{h}] \quad (3.58)$$

According to (2.24), (3.58) is simplified to

$$\boldsymbol{v}_z^h(k)=\boldsymbol{v}_z(k)-\boldsymbol{C}(k)\boldsymbol{D}_{xx}(k/k) \cdot\boldsymbol{H}(k)\boldsymbol{N}_{hh}^{-1}(k)[\boldsymbol{H}^T(k)\boldsymbol{x}(k/k)-\boldsymbol{h}] \quad (3.59)$$

The covariance matrix of the residual vectors for each of $\boldsymbol{v}_{x_h}(k)$, $\boldsymbol{v}_{w_h}(k)$ and $\boldsymbol{v}_{\boldsymbol{z}_h}(k)$ are derived as follows:

(1) $\boldsymbol{D}_{v_{l_x}^h v_{l_x}^h}(k)$ is derived by applying the law of variance propagation to (3.52a)

$$\boldsymbol{D}_{\boldsymbol{v}_x^h\boldsymbol{v}_x^h}(k)=\boldsymbol{D}_{l_x l_x}(k)\boldsymbol{C}^T(k)\boldsymbol{D}_{dd}^{-1}(k)\boldsymbol{C}(k)\boldsymbol{D}_{l_x l_x}(k) + \boldsymbol{D}_{l_x l_x}(k)[\boldsymbol{I}-\boldsymbol{K}(k)\boldsymbol{C}(k)]^T\boldsymbol{H}(k)\boldsymbol{N}_{hh}^{-1}(k) \cdot\boldsymbol{H}^T(k)[\boldsymbol{I}-\boldsymbol{K}(k)\boldsymbol{C}(k)]\boldsymbol{D}_{l_x l_x}(k) \quad (3.60)$$

as $\boldsymbol{D}_{xd}(k)=\boldsymbol{O}$ in (2.10). Under the consideration of (2.25), (3.60) becomes

$$\boldsymbol{D}_{\boldsymbol{v}_x^h\boldsymbol{v}_x^h}(k)=\boldsymbol{D}_{\boldsymbol{v}_x\boldsymbol{v}_x}(k)+\boldsymbol{D}_{l_x l_x}(k)\boldsymbol{D}_{xx}^{-1}(k/k-1)\boldsymbol{D}_{xx}(k/k) \cdot\boldsymbol{H}(k)\boldsymbol{N}_{hh}^{-1}(k)\boldsymbol{H}^T(k)\boldsymbol{D}_{xx}(k/k)\boldsymbol{D}_{xx}^{-1}(k/k-1)\boldsymbol{D}_{l_x l_x}(k) \quad (3.61)$$

(2) $\boldsymbol{D}_{v_{l_w}^h v_{l_w}^h}(k)$ is developed similarly by applying the law of variance propagation to (3.54a) :

$$\boldsymbol{D}_{\boldsymbol{v}_w^h\boldsymbol{v}_w^h}(k)=\boldsymbol{Q}(k)\boldsymbol{B}^T(k)\boldsymbol{C}^T(k)\boldsymbol{D}_{dd}^{-1}(k)\boldsymbol{C}(k)\boldsymbol{B}(k)\boldsymbol{Q}(k) +\boldsymbol{Q}(k)\boldsymbol{B}^T(k)[\boldsymbol{I}-\boldsymbol{K}(k)\boldsymbol{C}(k)]^T\boldsymbol{H}(k)\boldsymbol{N}_{hh}^{-1}(k) \cdot\boldsymbol{H}^T(k)[\boldsymbol{I}-\boldsymbol{K}(k)\boldsymbol{C}(k)]\boldsymbol{B}(k)\boldsymbol{Q}(k) \quad (3.62)$$

and further, based on (2.26),

$$\boldsymbol{D}_{\boldsymbol{v}_w^h\boldsymbol{v}_w^h}(k)=\boldsymbol{D}_{\boldsymbol{v}_w\boldsymbol{v}_w}(k) +\boldsymbol{Q}(k)\boldsymbol{B}^T(k)\boldsymbol{D}_{xx}^{-1}(k/k-1)\boldsymbol{D}_{xx}(k/k)\boldsymbol{H}(k)\boldsymbol{N}_{hh}^{-1}(k) \cdot\boldsymbol{H}^T(k)\boldsymbol{D}_{xx}(k/k)\boldsymbol{D}_{xx}^{-1}(k/k-1)\boldsymbol{B}(k)\boldsymbol{Q}(k) \quad (3.63)$$

(3) $\boldsymbol{D}_{v_z^h v_z^h}(k)$ is given by applying the law of variance propagation to (3.58)

$$\boldsymbol{D}_{\boldsymbol{v}_z^h\boldsymbol{v}_z^h}(k)=[\boldsymbol{I}-\boldsymbol{C}(k)\boldsymbol{K}(k)]\boldsymbol{R}(k)+\boldsymbol{C}(k)\boldsymbol{D}_{xx}(k/k) \cdot\boldsymbol{H}(k)\boldsymbol{N}_{hh}^{-1}(k)\boldsymbol{H}^T(k)\boldsymbol{D}_{xx}(k/k)\boldsymbol{C}^T(k) \quad (3.64)$$

and further according to (2.24),

$$\boldsymbol{D}_{\boldsymbol{v}_z^h\boldsymbol{v}_z^h}(k)=\boldsymbol{D}_{\boldsymbol{v}_z\boldsymbol{v}_z}(k)+\boldsymbol{C}(k)\boldsymbol{D}_{xx}(k/k) \cdot\boldsymbol{H}(k)\boldsymbol{N}_{hh}^{-1}(k)\boldsymbol{H}^T(k)\boldsymbol{D}_{xx}(k/k)\boldsymbol{C}^T(k) \quad (3.65)$$

3.8 Redundancy Contribution of Measurements

There are two levels of redundancy contribution: the total redundancy contribution of $\boldsymbol{l}_x(k)$, $\boldsymbol{l}_w(k)$ and $\boldsymbol{z}(k)$ together as well as the subtotal redundancy contribution of each of the groups, and the individual redundant indexes associated with each element in a group of the independent measurements, here specifically $\boldsymbol{l}_w(k)$ and $\boldsymbol{z}(k)$ because $\boldsymbol{Q}(k)$ and $\boldsymbol{R}(k)$ are commonly diagonal in practice. The following discusses the redundancy contributions of $\boldsymbol{l}_x(k)$, $\boldsymbol{l}_w(k)$ and $\boldsymbol{l}_z(k)$ one by one:

(1) The redundancy contribution $r_{l_x}(k)$ of $\boldsymbol{l}_x(k)$

$$r_{l_x}(k)=tr\{\boldsymbol{D}_{\boldsymbol{v}_x^h\boldsymbol{v}_x^h}(k)\boldsymbol{D}_{\boldsymbol{l}_x\boldsymbol{l}_x}^{-1}(k)\} =tr\{\boldsymbol{D}_{l_x l_x}(k)\boldsymbol{C}^T(k)\boldsymbol{D}_{dd}^{-1}(k)\boldsymbol{C}(k)\} +tr\{\boldsymbol{D}_{l_x l_x}(k)\boldsymbol{D}_{xx}^{-1}(k/k-1)\boldsymbol{D}_{xx}(k/k)\boldsymbol{H}(k)\boldsymbol{N}_{hh}^{-1}(k) \cdot\boldsymbol{H}^T(k)\boldsymbol{D}_{xx}(k/k)\boldsymbol{D}_{xx}^{-1}(k/k-1)\} \quad (3.66)$$

However, no individual redundant indexes will have the usual meaning for $\boldsymbol{l}_x(k)$ as its variance matrix of $\boldsymbol{D}_{\boldsymbol{l}_x\boldsymbol{l}_x}(k)=\boldsymbol{A}(k,k-1)\boldsymbol{D}_{xx}(k-1)\boldsymbol{A}^T(k,k-1)$ will not be possibly diagonal in reality.

(2) The redundancy contribution $r_{l_w}(k)$ of $\boldsymbol{l}_w(k)$

$$r_{l_w}(k)=tr\{\boldsymbol{D}_{\boldsymbol{v}_w^h\boldsymbol{v}_w^h}(k)\boldsymbol{D}_{l_w l_w}^{-1}(k)\} =tr\{\boldsymbol{D}_{\boldsymbol{v}_w\boldsymbol{v}_w}(k)\boldsymbol{Q}^{-1}(k)\} +tr\{\boldsymbol{Q}(k)\boldsymbol{B}^T(k)\boldsymbol{D}_{xx}^{-1}(k/k-1)\boldsymbol{D}_{xx}(k/k)\boldsymbol{H}(k)\boldsymbol{N}_{hh}^{-1}(k) \cdot\boldsymbol{H}^T(k)\boldsymbol{D}_{xx}(k/k)\boldsymbol{D}_{xx}^{-1}(k/k-1)\boldsymbol{B}(k)\} \quad (3.67)$$

or

$$r_{l_w}(k)=tr\{\boldsymbol{Q}(k)\boldsymbol{B}^T(k)\boldsymbol{C}^T(k)\boldsymbol{D}_{dd}^{-1}(k)\boldsymbol{C}(k)\boldsymbol{B}(k)\} +tr\{\boldsymbol{Q}(k)\boldsymbol{B}^T(k)\boldsymbol{D}_{xx}^{-1}(k/k-1)\boldsymbol{D}_{xx}(k/k)\boldsymbol{H}(k)\boldsymbol{N}_{hh}^{-1}(k) \cdot\boldsymbol{H}^T(k)\boldsymbol{D}_{xx}(k/k)\boldsymbol{D}_{xx}^{-1}(k/k-1)\boldsymbol{B}(k)\} \quad (3.68)$$

Besides, the individual redundant index associated with each component in $\boldsymbol{l}_w(k)$, when $\boldsymbol{Q}(k)$ is diagonal, is derived as follows

$$r_{l_w}^i(k)=\{\boldsymbol{Q}(k)\boldsymbol{B}^T(k)\boldsymbol{C}^T(k)\boldsymbol{D}_{dd}^{-1}(k)\boldsymbol{C}(k)\boldsymbol{B}(k)\}_{ii} +\{\boldsymbol{Q}(k)\boldsymbol{B}^T(k)\boldsymbol{D}_{xx}^{-1}(k/k-1)\boldsymbol{D}_{xx}(k/k)\boldsymbol{H}(k)\boldsymbol{N}_{hh}^{-1}(k) \cdot\boldsymbol{H}^T(k)\boldsymbol{D}_{xx}(k/k)\boldsymbol{D}_{xx}^{-1}(k/k-1)\boldsymbol{B}(k)\}_{ii} \quad (i=1,2,...,m) \quad (3.69)$$

(3) The redundancy contribution $r_{l_z}(k)$ (or $r_z(k)$) for $\boldsymbol{l}_z(k)$ (or $\boldsymbol{z}(k)$)

$$r_z(k)=trace\{\boldsymbol{D}_{\boldsymbol{v}_z^h\boldsymbol{v}_z^h}(k)\boldsymbol{D}_{l_z l_z}^{-1}(k)\} =tr\{\boldsymbol{D}_{\boldsymbol{v}_z\boldsymbol{v}_z}(k)\boldsymbol{R}^{-1}(k)\} +tr\{\boldsymbol{C}(k)\boldsymbol{D}_{xx}(k/k)\boldsymbol{H}(k)\boldsymbol{N}_{hh}^{-1}(k) \cdot\boldsymbol{H}^T(k)\boldsymbol{D}_{xx}(k/k)\boldsymbol{C}^T(k)\boldsymbol{R}^{-1}(k)\} \quad (3.70)$$

in which the first item is

$$tr\{\boldsymbol{D}_{\boldsymbol{v}_z\boldsymbol{v}_z}(k)\boldsymbol{R}^{-1}(k)\}=tr\{[\boldsymbol{I}-\boldsymbol{C}(k)\boldsymbol{K}(k)]\boldsymbol{R}(k)\boldsymbol{R}^{-1}(k)\} =tr[\boldsymbol{I}_{p\times p}-\boldsymbol{C}(k)\boldsymbol{K}(k)] = p(k)-tr[\boldsymbol{C}(k)\boldsymbol{K}(k)] \quad (3.71)$$

When $\boldsymbol{R}(k)$ is diagonal, the individual redundant index with each component in $\boldsymbol{l}_z(k)$ is

$$r_z^i(k)=\{\boldsymbol{D}_{\boldsymbol{v}_z\boldsymbol{v}_z}(k)\boldsymbol{R}^{-1}(k)\}_{ii}+\{\boldsymbol{C}(k)\boldsymbol{D}_{xx}(k/k)\boldsymbol{H}(k)\boldsymbol{N}_{hh}^{-1}(k) \cdot\boldsymbol{H}^T(k)\boldsymbol{D}_{xx}(k/k)\boldsymbol{C}^T(k)\boldsymbol{R}^{-1}(k)\}_{ii} \quad (3.72)$$

and further

$$r_z^i(k)=[\boldsymbol{I}-\boldsymbol{C}(k)\boldsymbol{K}(k)]_{ii}+[\boldsymbol{C}(k)\boldsymbol{D}_{xx}(k/k)\boldsymbol{H}(k)\boldsymbol{N}_{hh}^{-1}(k)\cdot\boldsymbol{H}^T(k)\boldsymbol{D}_{xx}(k/k)\boldsymbol{C}^T(k)\boldsymbol{R}^{-1}(k)]_{ii} \quad (i=1,2,...,p) \tag{3.73}$$

Finally, the total redundancy contribution of the three independent observation vectors together at epoch k, i.e., total redundancy number of $\boldsymbol{l}_x(k)$, $\boldsymbol{l}_w(k)$ and $\boldsymbol{l}_z(k)$ together is equal to

$$r(k)=r_{l_x}(k)+r_{l_w}(k)+r_z(k) \tag{3.74}$$

with the following specific detail,

$$\begin{aligned}r(k)=&tr\{\boldsymbol{D}_{l_xl_x}(k)\boldsymbol{C}^T(k)\boldsymbol{D}_{dd}^{-1}(k)\boldsymbol{C}(k)\}\\&+tr\{\boldsymbol{D}_{l_xl_x}(k)\boldsymbol{D}_{xx}^{-1}(k/k-1)\boldsymbol{D}_{xx}(k/k)\boldsymbol{H}(k)\boldsymbol{N}_{hh}^{-1}(k)\\&\quad\cdot\boldsymbol{H}^T(k)\boldsymbol{D}_{xx}(k/k)\boldsymbol{D}_{xx}^{-1}(k/k-1)\}\\&+tr\{\boldsymbol{Q}(k)\boldsymbol{B}^T(k)\boldsymbol{C}^T(k)\boldsymbol{D}_{dd}^{-1}(k)\boldsymbol{C}(k)\boldsymbol{B}(k)\}\\&+tr\{\boldsymbol{Q}(k)\boldsymbol{B}^T(k)\boldsymbol{D}_{xx}^{-1}(k/k-1)\boldsymbol{D}_{xx}(k/k)\boldsymbol{H}(k)\\&\quad\cdot\boldsymbol{N}_{hh}^{-1}(k)\boldsymbol{H}^T(k)\boldsymbol{D}_{xx}(k/k)\boldsymbol{D}_{xx}^{-1}(k/k-1)\boldsymbol{B}(k)\}\\&+tr\{\boldsymbol{I}-\boldsymbol{C}(k)\boldsymbol{K}(k)\}\\&+tr\{\boldsymbol{C}(k)\boldsymbol{D}_{xx}(k/k)\boldsymbol{H}(k)\boldsymbol{N}_{hh}^{-1}(k)\boldsymbol{H}^T(k)\boldsymbol{D}_{xx}(k/k)\\&\quad\cdot\boldsymbol{C}^T(k)\boldsymbol{R}^{-1}(k)\}\end{aligned}$$

It can be proved that the total redundant index at epoch k is equal to

$$\begin{aligned}r(k)&=p(k)+tr\{\boldsymbol{N}_{hh}^{-1}(k)\boldsymbol{H}^T(k)\boldsymbol{D}_{xx}(k/k)\boldsymbol{H}(k)\}\\&=p(k)+tr\{\boldsymbol{I}_{h\times h}\}=p(k)+h(k)\end{aligned} \tag{3.75}$$

which means

$$r(k)=r_{l_x}(k)+r_{l_w}(k)+r_z(k)=p(k)+h(k) \tag{3.76}$$

with $p(k)$ and $h(k)$ being the number of the measurements in $\boldsymbol{z}(k)$ and the number of the constraints in (3.1).

3.9 Other Aspects

In addition, several algorithmic developments such as test statistics, variance factors and variance component estimation etc. may be further conducted, in analogy to the work in [Wang, 1997, 2008, 2009 etc.] and are excluded here due to the space restriction, except the following essential remark about the variance of unit weight:

(i) The variance of unit weight for Section 2 (standard discrete Kalman filter): the one in (2.34) is identical to the one in (2.35).

(ii) The variance of unit weight for Section 3 (discrete Kalman filter with constraints):

$$\hat{\sigma}_0^2=[\boldsymbol{v}_{l_x}^{hT}(k)\boldsymbol{D}_{l_xl_x}^{-1}(k)\boldsymbol{v}_{l_x}^h(k)+\boldsymbol{v}_{l_w}^{hT}(k)\boldsymbol{Q}^{-1}(k)\boldsymbol{v}_{l_w}^h(k)+\boldsymbol{v}_z^{hT}(k)\boldsymbol{R}^{-1}(k)\boldsymbol{v}_z^h(k)]/r(k) \tag{3.77}$$

as

$$\boldsymbol{d}^T(k)D_{dd}^{-1}(k)\boldsymbol{d}(k)\neq[\boldsymbol{v}_{l_x}^{hT}(k)\boldsymbol{D}_{l_xl_x}^{-1}(k)\boldsymbol{v}_{l_x}^h(k)+\boldsymbol{v}_{l_w}^{hT}(k)\boldsymbol{Q}^{-1}(k)\boldsymbol{v}_{l_w}^h(k)+\boldsymbol{v}_z^{hT}(k)\boldsymbol{R}^{-1}(k)\boldsymbol{v}_z^h(k)] \tag{3.78}$$

4. Algorithm in the Form of Extended Kalman Filter with Constraints

This section frames the relevant formulas in the form of Extended Kalman filter in accordance with the functional model defined in Section 3.1, but having them (i.e., (2.1), (2.2) and (3.1)) nonlinear.

The system model, the measurement model and constraint model appear nonlinear as follows:

$$\boldsymbol{x}(k)=\boldsymbol{A}(\boldsymbol{x}(k-1),k)+\boldsymbol{B}(k,k-1)\boldsymbol{w}(k-1) \tag{4.1}$$

$$\boldsymbol{z}(k)=\boldsymbol{C}(\boldsymbol{x}(k),k)+\Delta(k) \tag{4.2}$$

$$\boldsymbol{H}(\boldsymbol{x}(k),k)-\boldsymbol{h}=\boldsymbol{o} \tag{4.3}$$

As for the variance propagation, three Jacobian matrices are derived here,

$$\boldsymbol{A}(k,k-1)=\boldsymbol{A}(k)=\frac{\partial\boldsymbol{A}(\boldsymbol{x}(k-1),k)}{\partial\boldsymbol{x}(k-1)}=\begin{pmatrix}\frac{\partial\boldsymbol{A}_1(\boldsymbol{x}(k-1),k)}{\partial\boldsymbol{x}_1(k-1)}&\frac{\partial\boldsymbol{A}_1(\boldsymbol{x}(k-1),k)}{\partial\boldsymbol{x}_2(k-1)}&\cdots&\frac{\partial\boldsymbol{A}_1(\boldsymbol{x}(k-1),k)}{\partial\boldsymbol{x}_n(k-1)}\\\frac{\partial\boldsymbol{A}_2(\boldsymbol{x}(k-1),k)}{\partial\boldsymbol{x}_1(k-1)}&\frac{\partial\boldsymbol{A}_2(\boldsymbol{x}(k-1),k)}{\partial\boldsymbol{x}_2(k-1)}&\cdots&\frac{\partial\boldsymbol{A}_2(\boldsymbol{x}(k-1),k)}{\partial\boldsymbol{x}_n(k-1)}\\\vdots&\vdots&&\vdots\\\frac{\partial\boldsymbol{A}_n(\boldsymbol{x}(k-1),k)}{\partial\boldsymbol{x}_1(k-1)}&\frac{\partial\boldsymbol{A}_n(\boldsymbol{x}(k-1),k)}{\partial\boldsymbol{x}_2(k-1)}&\cdots&\frac{\partial\boldsymbol{A}_n(\boldsymbol{x}(k-1),k)}{\partial\boldsymbol{x}_n(k-1)}\end{pmatrix} \tag{4.4}$$

which is with respect to the estimated state vector $\boldsymbol{x}(k-1/k-1)$ at t_{k-1},

$$\boldsymbol{C}(k)=\frac{\partial\boldsymbol{C}(\boldsymbol{x}(k),k)}{\partial\boldsymbol{x}(k)}=\begin{pmatrix}\frac{\partial C_1(\boldsymbol{x}(k),k)}{\partial\boldsymbol{x}_1(k)}&\frac{\partial C_1(\boldsymbol{x}(k),k)}{\partial\boldsymbol{x}_2(k)}&\cdots&\frac{\partial C_1(\boldsymbol{x}(k),k)}{\partial\boldsymbol{x}_n(k)}\\\frac{\partial C_2(\boldsymbol{x}(k),k)}{\partial\boldsymbol{x}_1(k)}&\frac{\partial C_2(\boldsymbol{x}(k),k)}{\partial\boldsymbol{x}_2(k)}&\cdots&\frac{\partial C_2(\boldsymbol{x}(k),k)}{\partial\boldsymbol{x}_n(k)}\\\vdots&\vdots&&\vdots\\\frac{\partial C_p(\boldsymbol{x}(k),k)}{\partial\boldsymbol{x}_1(k)}&\frac{\partial C_p(\boldsymbol{x}(k),k)}{\partial\boldsymbol{x}_2(k)}&\cdots&\frac{\partial C_p(\boldsymbol{x}(k),k)}{\partial\boldsymbol{x}_n(k)}\end{pmatrix} \tag{4.5}$$

which is with respect to the predicted state vector $\boldsymbol{x}(k/k-1)$ through the time update from t_{k-1} to t_k, and

$$\boldsymbol{H}^T(k) = \frac{\partial \boldsymbol{H}(\boldsymbol{x}(k),k)}{\partial \boldsymbol{x}(k)} = \begin{pmatrix} \frac{\partial \boldsymbol{H}_1(\boldsymbol{x}(k),k)}{\partial \boldsymbol{x}_1(k)} & \frac{\partial \boldsymbol{H}_1(\boldsymbol{x}(k),k)}{\partial \boldsymbol{x}_2(k)} & \cdots & \frac{\partial \boldsymbol{H}_1(\boldsymbol{x}(k),k)}{\partial \boldsymbol{x}_n(k)} \\ \frac{\partial \boldsymbol{H}_2(\boldsymbol{x}(k),k)}{\partial \boldsymbol{x}_1(k)} & \frac{\partial \boldsymbol{H}_2(\boldsymbol{x}(k),k)}{\partial \boldsymbol{x}_2(k)} & \cdots & \frac{\partial \boldsymbol{H}_2(\boldsymbol{x}(k),k)}{\partial \boldsymbol{x}_n(k)} \\ \vdots & \vdots & & \vdots \\ \frac{\partial \boldsymbol{H}_h(\boldsymbol{x}(k),k)}{\partial \boldsymbol{x}_1(k)} & \frac{\partial \boldsymbol{H}_h(\boldsymbol{x}(k),k)}{\partial \boldsymbol{x}_2(k)} & \cdots & \frac{\partial \boldsymbol{H}_h(\boldsymbol{x}(k),k)}{\partial \boldsymbol{x}_n(k)} \end{pmatrix} \quad (4.6)$$

which is with respect to the estimated state vector $\boldsymbol{x}(k)$ through the measurement update before the constraints are applied at t_k.

The following gives the algorithm in the form of Extended Kalman filter by referring to Sections 3.4 and 3.5:

1) THE MEASUREMENT MODEL

The predicted state vector exclusive of the effect from the process noise vector:

$$\boldsymbol{v}_x(k) = \boldsymbol{x}_h(k) - \boldsymbol{B}(k,k-1)\boldsymbol{w}_h(k-1) - \boldsymbol{l}_x(k) \quad (3.25)$$

$$\boldsymbol{l}_x(k) = \boldsymbol{A}(\boldsymbol{x}_h(k-1),k,k-1) \text{ (vs. (3.26))} \quad (4.7)$$

$$\boldsymbol{D}_{l_x l_x}(k) = \boldsymbol{A}(k,k-1)\boldsymbol{D}_{xx}(k-1)\boldsymbol{A}^T(k,k-1) \quad (3.27)$$

The process noise vector as a group of the pseudo-measurements: the same as (3.28), (3.29) and (3.30).

A group of the measurements from the measurement vector $\boldsymbol{z}(k)$ at t_k:

$$\boldsymbol{v}_{l_z}(k) = \boldsymbol{C}(\boldsymbol{x}_h(k),k) - \boldsymbol{l}_z(k) \quad \text{(vs. (3.17))} \quad (4.8)$$

$$\boldsymbol{l}_z(k) = \boldsymbol{z}(k) \quad (3.18)$$

$$\boldsymbol{D}_{l_z l_z}(k) = \boldsymbol{R}(k) \quad (3.19)$$

A group of the constraints on the states:

$$\boldsymbol{H}(\boldsymbol{x}_h(k),k) - \boldsymbol{h} = \boldsymbol{o} \quad \text{(vs. (3.1))} \quad (4.9)$$

2) THE SOLUTION

The state vector, the process noise vector and the Lagrange multiplier vector:

$$\begin{aligned}\boldsymbol{x}_h(k/k) &= \boldsymbol{x}(k/k) \\ &- \boldsymbol{D}_{xx}(k/k)\boldsymbol{H}(k)\boldsymbol{N}_{hh}^{-1}(k)[\boldsymbol{H}(\boldsymbol{x}(k/k),k) - \boldsymbol{h}]\end{aligned} \quad (4.10)$$

$$\begin{aligned}\boldsymbol{w}_h(k-1) &= \boldsymbol{l}_w(k) \\ &+ \boldsymbol{Q}(k)\boldsymbol{B}^T(k)\boldsymbol{D}_{xx}^{-1}(k/k-1)\boldsymbol{K}(k)\boldsymbol{d}(k) \\ &- \boldsymbol{Q}(k)\boldsymbol{B}^T(k)\boldsymbol{D}_{xx}^{-1}(k/k-1)\boldsymbol{D}_{xx}(k/k) \\ &\cdot \boldsymbol{H}(k)\boldsymbol{N}_{hh}^{-1}(k)[\boldsymbol{H}(\boldsymbol{x}(k/k),k) - \boldsymbol{h}]\end{aligned} \quad (4.11)$$

$$\boldsymbol{k}_h(k) = \boldsymbol{N}_{hh}^{-1}(k)[\boldsymbol{H}(\boldsymbol{x}(k/k),k) - \boldsymbol{h}] \quad (4.12)$$

wherein

$$\boldsymbol{x}(k/k) = \boldsymbol{x}(k/k-1) + \boldsymbol{K}(k)\boldsymbol{d}(k) \quad (4.13)$$

$$\boldsymbol{x}(k/k-1) = \boldsymbol{A}(\boldsymbol{x}_h(k-1),k,k-1) + \boldsymbol{B}(k)\boldsymbol{w}_0(k) \quad (4.14)$$

$$\boldsymbol{d}(k) = \boldsymbol{z}(k) - \boldsymbol{A}(\boldsymbol{x}_h(k-1),k,k-1) - \boldsymbol{B}(k)\boldsymbol{w}_0(k) \quad (4.15)$$

The var-covariance matrices of the state vector and the process noise vector: the same as (3.49), (3.50) and (3.51).

3) THE MEASUREMENT RESIDUALS

The residual vectors:

$$\begin{aligned}\boldsymbol{v}_{l_x}^h(k) &= \boldsymbol{D}_{l_x l_x}(k)\boldsymbol{D}_{xx}^{-1}(k/k-1)\boldsymbol{K}(k)\boldsymbol{d}(k) \\ &- \boldsymbol{D}_{l_x l_x}(k)[\boldsymbol{I} - \boldsymbol{K}(k)\boldsymbol{C}(k)]^T \\ &\cdot \boldsymbol{H}(k)\boldsymbol{N}_{hh}^{-1}(k)[\boldsymbol{H}(\boldsymbol{x}(k/k),k) - \boldsymbol{h}]\end{aligned} \quad (4.16)$$

$$\begin{aligned}\boldsymbol{v}_{l_w}^h(k) &= \boldsymbol{Q}(k)\boldsymbol{B}^T(k)\boldsymbol{D}_{xx}^{-1}(k/k-1)\boldsymbol{K}(k)\boldsymbol{d}(k) \\ &- \boldsymbol{Q}(k)\boldsymbol{B}^T(k)[\boldsymbol{I} - \boldsymbol{K}(k)\boldsymbol{C}(k)]^T \\ &\cdot \boldsymbol{H}(k)\boldsymbol{N}_{hh}^{-1}(k)[\boldsymbol{H}(\boldsymbol{x}(k/k),k) - \boldsymbol{h}]\end{aligned} \quad (4.17)$$

$$\begin{aligned}\boldsymbol{v}_z^h(k) &= [\boldsymbol{C}(k)\boldsymbol{K}(k) - \boldsymbol{I}]\boldsymbol{d}(k) - \boldsymbol{C}(k)\boldsymbol{D}_{xx}(k/k) \\ &\cdot \boldsymbol{H}(k)\boldsymbol{N}_{hh}^{-1}(k)[\boldsymbol{H}(\boldsymbol{x}(k/k),k) - \boldsymbol{h}]\end{aligned} \quad (4.18)$$

The variance matrices of the residual vectors: are the same as (3.60) – (3.65).

4) **THE REDUNDANCY CONTRIBUTION**:

The same as in Section 3.8.

For the convenience of practical implementation and better understanding of the proposed framework, an algorithmic flow is suggested in Fig. 4.1.

5. CONCLUDING REMARKS

This manuscript exhibited flexible algorithmic formulation for Kalman filtering with equality constraints on the system states, and practically developed an analytic framework for comprehensive error analysis accordingly. Specifically, this manuscript has:

(a) Developed a unique formula set as an innovative framework on the base of the three independent error sources that influence the system state estimate (Section 3.1-3.6);

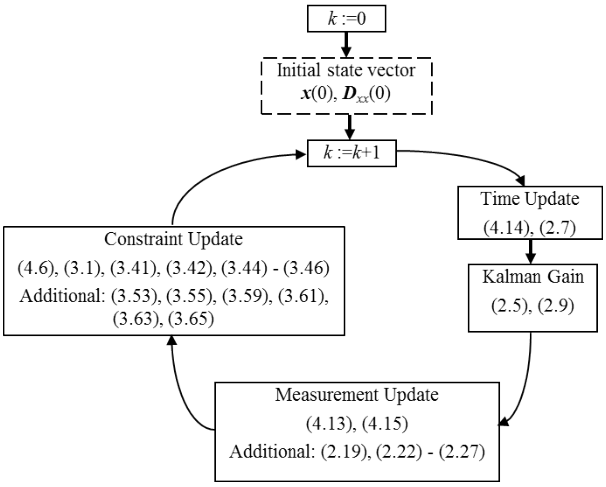


Fig. 4.1 an algorithmic flow of EKF with Constraints

(b) Specifically introduced the equation for the residual vector of the process noise vector , as well as their covariance matrices (Section 3.6);

(c) Made the reliability analysis feasible through parametrically introducing the redundancy contribution for the predicted state vector, process noise vector, and measurement vector, and the individual redundant indexes for the elements in the process noise and measurement vectors under the assumption of diagonal $\boldsymbol{Q}(k)$ and $\boldsymbol{R}(k)$ (Section 3.7); and

(d) Pointed out its essential potentials how further algorithmic extension may be accomplished from the proposed formulation (Sections 3.9).

This work took an important step towards a standardized generic approach to performing Kalman filtering with equality constraints that enables comprehensive and rigorous error analysis, which is particularly important for high accuracy applications, for instance, the centimeter level kinematic positioning and navigation using GNSS and/or multisensor-integrated systems in the modern direct-georeferencing technology, autonomous vehicle driving, and other robotic applications etc., wherever it is important to examine the sources of any deviations in the estimated system states. The issue of comprehensive error analysis in Kalman filtering has been addressed previously [Wang, 1997; Caspary and Wang, 1998; Wang, et al, 2021; etc.], but not yet in the context of a Kalman filter with equality constraints. It is the authors' hope that comprehensive error analysis becomes a necessary part of the estimation process in the constrained Kalman filtering as a result of this work.

Acknowledgements:

The authors would like to sincerely thank the Natural Sciences and Engineering Research Council of Canada (NSERC) for their financial support. We would also like to thank the editor-in-charge and anonymous reviewers for their valuable comments and suggestions.

References

Alouani, A. T. and Blair, W. D. (1991): Use of a kinematic constraint in tracking constant speed, maneuvering targets, *Proceedings of the 30th IEEE Conference on Decision and Control*, doi.org/10.1109/cdc.1991.261780.

Caspary, Wilhelm and Wang, Jianguo (1998): *Redundanzanteile und Varianzkomponentenim Kalman Filter*, Vol.123, No.4, 1998, pp.121-128.

Gopaul, Nilesh; Wang, Jianguo and and Jiming Guo (2010): Improving of GPS Observation Quality Assessment through Posteriori Variance-covariance Component Estimation, Proceedings of *CPGPS 2010 Navigation and Location Services: Emerging Industry and International Exchanges*, Scientific Research Publishing Inc., Shanghai, 2010.

Hasberg, C.; Hensel, S. and Stiller, C. (2012): Simultaneous localization and mapping for path-constrained motion. *IEEE Transactions on Intelligent Transportation Systems*, *13*(2), 541–552. doi.org/10.1109/tits.2011.2177522.

Khabbazi, M.-R. and Esfanjani, R. M. (2014): Constrained two-stage Kalman filter for Target Tracking, *4th International Conference on Computer and Knowledge Engineering, 2014*, doi.org/10.1109/iccke.2014.6993432.

Mikhail, Edward (1970): Parameter Constraints in Least Squares, Photogrammetric Engineering, 1970, pp. 1277 - 1291.

Pizzinga, Adrian (2012): Restricted Klaman Filtering – Theory, Methods, and Application, Springer Science+Business Media New York 2012.

Qian, Kun; Wang, Jianguo and Baoxin Hu (2016): A posteriori estimation of stochastic models for multi-sensor integrated inertial kinematic positioning and navigation on basis of variance component estimation, Journal of GPS, 2016, 14:5.

Qian, Kun (2017): Generic Multisensor Integration Strategy and Innovative Error Analysis for Integrated Navigation, PhD Dissertation, York University, Canada, 2017.

Qian, Kun; Wang, Jianguo and Baoxin Hu (2015): Novel Integration Strategy for GNSS-aided Inertial Integrated Navigation, No. 2, Vol. 69, Geomatica, pp. 217-230.

Rao, C. Radhakrishna and Toutenburg, Helge (1999): Linear Models: Least Squares and Alternatives, 2nd edition, Springer, ISBN 0-387-98848-3, 1999.

Simon, D. (2010): Kalman filtering with state constraints: A survey of linear and nonlinear algorithms, *IET Control Theory & Applications*, *4*(8), 1303–1318.

Simon, D. and Tien Li Chia (2002): Kalman filtering with state equality constraints. *IEEE Transactions on Aerospace and Electronic Systems*, *38*(1), 128–136. https://doi.org/10.1109/7.993234.

Teixeira, B., Chandrasekar, J., et al (2008) : Gain-constrained Kalman filtering for linear and nonlinear systems, *IEEE Transactions on Signal Processing*, *56*(9), 4113–4123, doi.org/10.1109/tsp.2008.926101.

Wang, Jianguo; Qiu, Weining; Yao, Yinbin and Wu, Yun (2019): Error Theory and Foundation of Surveying Adjustment, English Edition, Wuhan University Press, 2019.

Wang, Jianguo (1997): *Filtermethoden zur fehler-toleranten kinematischen Positionsbestimmung*, Schritenreihe Studiengang Vermessungswesen, Federal Arm-Forced University Munich, Germany, No. 52, Neubiberg, 1997.

Wang, Jianguo (2008): Test Statistics in Kalman Filtering. Vol. 7, No. 1, Journal of GPS, pp. 81-90.

Wang, Jianguo (2009): Reliability Analysis in Kalman Filtering, Journal of GPS, Vol. 8, No. 1, pp.101-111.

Wang, Jianguo; Gopaul, Nilesh and Scherzinger, Bruno (2009): Simplified Algorithms of Variance Compo-nent Estimation for Static and Kinematic GPS Single Point Positioning**,** J. of GPS,Vol. 8, No. 1:43-51.

Wang, Jianguo; Gopaul, Nilesh and Jiming Guo (2010): Adaptive Kalman Filter based on Posteriori Variance-covariance Component Estimation, Proceedings of *CPGPS 2010 Navigation and Location Services: Emerging Industry and International Exchanges*, Scientific Research Publishing Inc., Shanghai, 2010.

Wang, Jianguo; Kun Qian and Baoxin Hu (2015): An Unconventional Full Tightly-Coupled Multi-Sensor Integration for Kinematic Positioning and Navigation, CSNC 2015 Proceedings, Volume III, Vol. 7, No.1, 2008, pp. 81~90.

Wang, Jianguo; Aaron Boda and Baoxin Hu (2021): Comprehensive Error Analysis beyond System Innovations in Kalman Filtering, Chapter 3, Learning Control, Elservier, pages 59-92.

Yang, C.; Bakich, M. and Blasch, E. (2005): Nonlinear constrained tracking of targets on roads, *2005 7th International Conference on Information Fusion*, doi.org/10.1109/icif.2005.1591860.

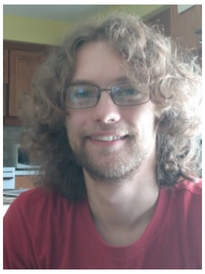
Appendix: Proof of (3.24) and (3.25)

$$
\begin{aligned}
&\{\boldsymbol{D}_{l''_x l''_x}^{-1}(k)+\boldsymbol{C}^T(k)\boldsymbol{R}^{-1}(k)\boldsymbol{C}(k)\}^{-1}\\
&=\{\boldsymbol{D}_{xx}^{-1}(k/k-1)+\boldsymbol{C}^T(k)\boldsymbol{R}^{-1}(k)\boldsymbol{C}(k)\}^{-1}\\
&=\boldsymbol{D}_{xx}(k/k-1)-\boldsymbol{D}_{xx}(k/k-1)\boldsymbol{C}^T(k)\{\boldsymbol{R}(k)+\boldsymbol{C}(k)\boldsymbol{D}_{l_x l_x}(k)\boldsymbol{C}^T(k)\}^{-1}\boldsymbol{C}(k)\boldsymbol{D}_{xx}(k/k-1)\\
&=\boldsymbol{D}_{xx}(k/k-1)-\boldsymbol{D}_{xx}(k/k-1)\boldsymbol{C}^T(k)\boldsymbol{D}_{dd}^{-1}(k)\boldsymbol{C}(k)\boldsymbol{D}_{xx}(k/k-1)\\
&=[\boldsymbol{E}-\boldsymbol{D}_{xx}(k/k-1)\boldsymbol{C}^T(k)\boldsymbol{D}_{dd}^{-1}(k)\boldsymbol{C}(k)]\boldsymbol{D}_{xx}(k/k-1)\\
&=[\boldsymbol{E}-\boldsymbol{K}(k)\boldsymbol{C}(k)]\boldsymbol{D}_{xx}(k/k-1)\\
&=\boldsymbol{D}_{xx}(k/k)
\end{aligned}
$$

$$
\begin{aligned}
&\boldsymbol{D}_{xx}^{-1}(k)\boldsymbol{x}(k)=\boldsymbol{D}_{xx}^{-1}(k)[(\boldsymbol{E}-\boldsymbol{K}(k)\boldsymbol{C}(k))\boldsymbol{x}(k/k-1)+\boldsymbol{K}(k)\boldsymbol{z}(k)]\\
&=[\boldsymbol{C}^T(k)\boldsymbol{R}^{-1}(k)\boldsymbol{C}(k)+\boldsymbol{D}_{xx}^{-1}(k/k-1)][(\boldsymbol{E}-\boldsymbol{K}(k)\boldsymbol{C}(k))\boldsymbol{x}(k/k-1)+\boldsymbol{K}(k)\boldsymbol{z}(k)]\\
&=\boldsymbol{C}^T(k)\boldsymbol{R}^{-1}(k)\boldsymbol{C}(k)\boldsymbol{x}(k/k-1)-\boldsymbol{C}^T(k)\boldsymbol{R}^{-1}(k)\boldsymbol{C}(k)\boldsymbol{K}(k)\boldsymbol{C}(k)\boldsymbol{x}(k/k-1)+\boldsymbol{D}_{xx}^{-1}(k/k-1)\boldsymbol{x}(k/k-1)\\
&\quad-\boldsymbol{D}_{xx}^{-1}(k/k-1)\boldsymbol{K}(k)\boldsymbol{C}(k)\boldsymbol{x}(k/k-1)+\boldsymbol{C}^T(k)\boldsymbol{R}^{-1}(k)\boldsymbol{C}(k)\boldsymbol{K}(k)\boldsymbol{z}(k)+\boldsymbol{D}_{xx}^{-1}(k/k-1)\boldsymbol{K}(k)\boldsymbol{z}(k)\\
&=\boldsymbol{D}_{xx}^{-1}(k/k-1)\boldsymbol{x}(k/k-1)+\boldsymbol{C}^T(k)\boldsymbol{R}^{-1}(k)\{\boldsymbol{C}(k)\boldsymbol{x}(k/k-1)\\
&\quad+[(\boldsymbol{D}_{dd}(k)-\boldsymbol{C}(k)\boldsymbol{D}_{xx}(k/k-1)\boldsymbol{C}^T(k))\boldsymbol{D}_{dd}^{-1}(k)+\boldsymbol{C}(k)\boldsymbol{K}(k)]\boldsymbol{d}(k)\}\\
&=\boldsymbol{D}_{xx}^{-1}(k/k-1)\boldsymbol{x}(k/k-1)+\boldsymbol{C}^T(k)\boldsymbol{R}^{-1}(k)\{\boldsymbol{C}(k)\boldsymbol{x}(k/k-1)+\boldsymbol{z}(k)-\boldsymbol{C}(k)\boldsymbol{x}(k/k-1)\}\\
&=\boldsymbol{D}_{xx}^{-1}(k/k-1)\boldsymbol{x}(k/k-1)+\boldsymbol{C}^T(k)\boldsymbol{R}^{-1}(k)\boldsymbol{z}(k)
\end{aligned}
$$

Authors

Dr.-Ing. Jianguo Wang is a professor in Geomatics Science and Engineering, Department of Earth and Space Science and Engineering, York University, Canada. He holds a Dr.-Ing. in Surveying Engineering from the University of German Federal Armed Forces Munich (Universität der Bundeswher München), a MEng. and a BEng. in Surveying Engineering from Wuhan Technical University of Surveying and Mapping, China. His current research is focused on novel multisensor integration strategy for kinematic positioning and navigation and advanced estimation algorithms.



Benjamin Brunson is currently a Doctorate student at the Department of Earth and Space Science and Engineering of York University. His research is focused on multisensor integrated kinematic positioning and navigation, especially on the usage of IMU Arrays and development of advanced filtering algorithms. Prior to his PhD studies, he also received his B. Eng. In Geomatics engineering in 2017 and MSc. focus on phase unwrapping for SAR interferometry in Earth and Space Science in 2019 from the same institution.

Dr. Baoxin Hu received PhD degree in Remote Sensing and GIS at Boston University, USA in 1998. She is currently a Professor in Geomatics Science and Engineering, Department of Earth and Space Science and Engineering, York University, Canada. Her research interests include vegetation characterization from multi-source remotely sensed data, 3D scene reconstruction, information fusion, and deep learning.