10. Kalman Filtering for Interval Systems

If some system parameters such as certain elements of the system matrix are not precisely known or gradually change with time, then the Kalman filtering algorithm cannot be directly applied. In this case, robust Kalman filtering that has the ability of handling uncertainty is needed. In this chapter we introduce one of such robust Kalman filtering algorithms.

Consider the nominal system

$$\begin{cases} \mathbf{x}_{k+1} = A_k \mathbf{x}_k + \Gamma_k \underline{\xi}_k \,, \\ \mathbf{v}_k = C_k \mathbf{x}_k + \underline{\eta}_k \,, \end{cases} \tag{10.1}$$

where A_k, Γ_k and C_k are known $n \times n$, $n \times p$ and $q \times n$ matrices, respectively, with $1 \leq p,\, q \leq n$, and where

$$\begin{aligned} &E(\underline{\xi}_k) = 0\,, && E(\underline{\xi}_k \underline{\xi}_\ell^\top) = Q_k \delta_{k\ell}\,, \\ &E(\underline{\eta}_k) = 0\,, && E(\underline{\eta}_k \underline{\eta}_\ell^\top) = R_k \delta_{k\ell}\,, \\ &E(\underline{\xi}_k \underline{\eta}_\ell^\top) = 0\,, && E(\mathbf{x}_0 \underline{\xi}_k) = 0\,, \quad E(\mathbf{x}_0 \underline{\eta}_k) = 0\,, \end{aligned}$$

for all $k, \ell = 0, 1, \cdots$, with Q_k and R_k being positive definite and symmetric matrices.

If all the constant matrices, A_k, Γ_k, and C_k, are known, then the Kalman filter can be applied to the nominal system (10.1), which yields optimal estimates $\{\hat{\mathbf{x}}_k\}$ of the unknown state vectors $\{\mathbf{x}_k\}$ using the measurement data $\{\mathbf{v}_k\}$ in a recursive scheme. However, if some of the elements of these system matrices are unknown or uncertain, modification of the entire setting for filtering is necessary. Suppose that the uncertain parameters are only known to be bounded. Then we can write

$$\begin{aligned} A_k^I &= A_k + \Delta A_k = [A_k - |\Delta A_k|, A_k + |\Delta A_k|]\,, \\ \Gamma_k^I &= \Gamma_k + \Delta \Gamma_k = [\Gamma_k - |\Delta \Gamma_k|, \Gamma_k + |\Delta \Gamma_k|]\,, \\ C_k^I &= C_k + \Delta C_k = [C_k - |\Delta C_k|, C_k + |\Delta C_k|]\,, \end{aligned}$$

$k = 0, 1, \cdots$, where $|\Delta A_k|$, $|\Delta \Gamma_k|$, and $|\Delta C_k|$ are constant bounds for the unknowns. The corresponding system

$$\begin{cases} \mathbf{x}_{k+1} = A_k^I \mathbf{x}_k + \Gamma_k^I \underline{\xi}_k\,, \\ \mathbf{v}_k = C_k^I \mathbf{x}_k + \underline{\eta}_k\,, \end{cases} \tag{10.2}$$

$k = 0, 1, \cdots$, is then called an *interval system*.

Under this framework, how is the original Kalman filtering algorithm modified and applied to the interval system (10.2)? This question is to be addressed in this chapter.

10.1 Interval Mathematics

In this section, we first provide some preliminary results on interval arithmetic and interval analysis that are needed throughout the chapter.

10.1.1 Intervals and Their Properties

A closed and bounded subset $[\underline{x}, \overline{x}]$ in $R = (-\infty, \infty)$ is referred to as an interval. In particular, a single point $x \in R$ is considered as a degenerate interval with $\underline{x} = \overline{x} = x$.

Some useful concepts and properties of intervals are:

(a) *Equality*: Two intervals, $[\underline{x}_1, \overline{x}_1]$ and $[\underline{x}_2, \overline{x}_2]$, are said to be *equal*, and denoted by

$$[\underline{x}_1, \overline{x}_1] = [\underline{x}_2, \overline{x}_2]\,,$$

if and only if $\underline{x}_1 = \underline{x}_2$ and $\overline{x}_1 = \overline{x}_2$.

(b) *Intersection*: The intersection of two intervals, $[\underline{x}_1, \overline{x}_1]$ and $[\underline{x}_2, \overline{x}_2]$, is defined to be

$$[\underline{x}_1, \overline{x}_1] \cap [\underline{x}_2, \overline{x}_2] = [max\{\underline{x}_1, \underline{x}_2\}, min\{\overline{x}_1, \overline{x}_2\}]\,.$$

Furthermore, these two intervals are said to be *disjoint*, and denoted by

$$[\underline{x}_1, \overline{x}_1] \cap [\underline{x}_2, \overline{x}_2] = \phi\,,$$

if and only if $\underline{x}_1 > \overline{x}_2$ or $\underline{x}_2 > \overline{x}_1$.

(c) *Union*: The union of two non-disjoint intervals, $[\underline{x}_1, \overline{x}_1]$ and $[\underline{x}_2, \overline{x}_2]$, is defined to be

$$[\underline{x}_1, \overline{x}_1] \cup [\underline{x}_2, \overline{x}_2] = [min\{\underline{x}_1, \underline{x}_2\}, max\{\overline{x}_1, \overline{x}_2\}]\,.$$

Note that the union is defined only if the two intervals are not disjoint, i.e.,

$$[\underline{x}_1, \overline{x}_1] \cap [\underline{x}_2, \overline{x}_2] \neq \phi\,;$$

otherwise, it is undefined since the result is not an interval.

(d) *Inequality*: The interval $[\underline{x}_1, \overline{x}_1]$ is said to be *less than* (resp., *greater than*) the interval $[\underline{x}_2, \overline{x}_2]$, denoted by

$$[\underline{x}_1, \overline{x}_1] < [\underline{x}_2, \overline{x}_2] \qquad (resp., \quad [\underline{x}_1, \overline{x}_1] > [\underline{x}_2, \overline{x}_2])$$

if and only if $\overline{x}_1 < \underline{x}_2$ (resp., $\underline{x}_1 > \overline{x}_2$); otherwise, they cannot be compared. Note that the relations $\leq$ and $\geq$ are not defined for intervals.

(e) *Inclusion*: The interval $[\underline{x}_1, \overline{x}_1]$ is said to be *included* in $[\underline{x}_2, \overline{x}_2]$, denoted

$$[\underline{x}_1, \overline{x}_1] \subseteq [\underline{x}_2, \overline{x}_2]$$

if and only if $\underline{x}_2 \leq \underline{x}_1$ and $\overline{x}_1 \leq \overline{x}_2$; namely, if and only if $[\underline{x}_1, \overline{x}_1]$ is a subset (subinterval) of $[\underline{x}_2, \overline{x}_2]$.

For example, for three given intervals, $X_1 = [-1, 0]$, $X_2 = [-1, 2]$, and $X_3 = [2, 10]$, we have

$$\begin{aligned}
X_1 \cap X_2 &= [-1,0] \cap [-1,2] = [-1,0]\,,\\
X_1 \cap X_3 &= [-1,0] \cap [2,10] = \phi\,,\\
X_2 \cap X_3 &= [-1,2] \cap [2,10] = [2,2] = 2\,,\\
X_1 \cup X_2 &= [-1,0] \cup [-1,2] = [-1,2]\,,\\
X_1 \cup X_3 &= [-1,0] \cup [2,10] \quad is\ undefined\,,\\
X_2 \cup X_3 &= [-1,2] \cup [2,10] = [-1,10]\,,\\
X_1 &= [-1,0] < [2,10] = X_3\,,\\
X_1 &= [-1,0] \subset [-1,10] = X_2\,.
\end{aligned}$$

10.1.2 Interval Arithmetic

Let $[\underline{x}, \overline{x}]$, $[\underline{x}_1, \overline{x}_1]$, and $[\underline{x}_2, \overline{x}_2]$ be intervals. The basic arithmetic operations of intervals are defined as follows:

(a) *Addition*:

$$[\underline{x}_1, \overline{x}_1] + [\underline{x}_2, \overline{x}_2] = [\underline{x}_1 + \underline{x}_2, \overline{x}_1 + \overline{x}_2]\,.$$

(b) *Subtraction*:

$$[\underline{x}_1, \overline{x}_1] - [\underline{x}_2, \overline{x}_2] = [\underline{x}_1 - \overline{x}_2, \overline{x}_1 - \underline{x}_2] .$$

(c) *Reciprocal operation*:

$$If \quad 0 \notin [\underline{x}, \overline{x}] \quad then \quad [\underline{x}, \overline{x}]^{-1} = [1/\overline{x}, 1/\underline{x}] ;$$
$$If \quad 0 \in [\underline{x}, \overline{x}] \quad then \quad [\underline{x}, \overline{x}]^{-1} \quad is\ undefined .$$

(d) *Multiplication*:

$$[\underline{x}_1, \overline{x}_1] \cdot [\underline{x}_2, \overline{x}_2] = [\underline{y}, \overline{y}] ,$$

where

$$\underline{y} = min\,\{\underline{x}_1\underline{x}_2 \,,\, \underline{x}_1\overline{x}_2 \,,\, \overline{x}_1\underline{x}_2 \,,\, \overline{x}_1\overline{x}_2\} \ ,$$
$$\overline{y} = max\,\{\underline{x}_1\underline{x}_2 \,,\, \underline{x}_1\overline{x}_2 \,,\, \overline{x}_1\underline{x}_2 \,,\, \overline{x}_1\overline{x}_2\} \ .$$

(e) *Division*:

$$[\underline{x}_1, \overline{x}_1] \,/\, [\underline{x}_2, \overline{x}_2] = [\underline{x}_1, \overline{x}_1] \cdot [\underline{x}_2, \overline{x}_2]^{-1} ,$$

provided that $0 \notin [\underline{x}_2, \overline{x}_2]$; otherwise, it is undefined.

For three intervals, $X = [\underline{x}, \overline{x}]$, $Y = [\underline{y}, \overline{y}]$, and $Z = [\underline{z}, \overline{z}]$, consider the interval operations of addition (+), subtraction (−), multiplication (·), and division (/), namely,

$$Z = X * Y , \qquad * \in \{+, -, \cdot, /\} .$$

It is clear that $X * Y$ is also an interval. In other words, the family of intervals under the four operations $\{+, -, \cdot, /\}$ is algebraically closed. It is also clear that the real numbers $x, y, z, \cdots$ are isomorphic to degenerate intervals $[x, x]$, $[y, y]$, $[z, z]$, $\cdots$, so we will simply denote the point-interval operation $[x, x] * Y$ as $x * Y$. Moreover, the multiplication symbol "·" will often be dropped for notational convenience.

Similar to conventional arithmetic, the interval arithmetic has the following basic algebraic properties (cf. Exercise 10.1):

$$X + Y = Y + X ,$$
$$Z + (X + Y) = (Z + X) + Y ,$$
$$XY = YX ,$$
$$Z(XY) = (ZX)Y ,$$
$$X + 0 = 0 + X = X \quad and \quad X0 = 0X = 0 , \ where\ 0 = [0, 0] ,$$
$$XI = IX = X , \ where\ I = [1, 1] ,$$
$$Z(X + Y) \subseteq ZX + ZY , \ where\ = \ holds\ only\ if\ either$$

(*a*) $Z = [z, z]$,

(*b*) $X = Y = 0$, *or*

(*c*) $xy \geq 0$ *for all* $x \in X$ *and* $y \in Y$.

In addition, the following is an important property of interval operations, called the *monotonic inclusion* property.

Theorem 10.1. *Let X_1, X_2, Y_1, and Y_2 be intervals, with*

$$X_1 \subseteq Y_1 \qquad \text{and} \qquad X_2 \subseteq Y_2 .$$

Then for any operation $* \in \{+, -, \cdot, /\}$, it follows that

$$X_1 * X_2 \subseteq Y_1 * Y_2 .$$

This property is an immediate consequence of the relations $X_1 \subseteq Y_1$ and $X_2 \subseteq Y_2$, namely,

$$\begin{aligned} X_1 * X_2 &= \{x_1 * x_2 \,|\, x_1 \in X_1, x_2 \in X_2\} \\ &\subseteq \{y_1 * y_2 \,|\, y_1 \in Y_1, y_2 \in Y_2\} \\ &= Y_1 * Y_2 . \end{aligned}$$

Corollary 10.1. *Let X and Y be intervals and let $x \in X$ and $y \in Y$. Then,*

$$x * y \subseteq X * Y , \qquad \textit{for all} \;\; * \in \{+, -, \cdot, /\} .$$

What is seemingly counter-intuitive in the above theorem and corollary is that some operations such as reciprocal, subtraction, and division do not seem to satisfy such a monotonic inclusion property. However, despite the above proof, let us consider a simple example of two intervals, $X = [0.2, 0.4]$ and $Y = [0.1, 0.5]$. Clearly, $X \subseteq Y$. We first show that $I/X \subseteq I/Y$, where $I = [1.0, 1.0]$. Indeed,

$$\frac{I}{X} = \frac{[1.0, 1.0]}{[0.2, 0.4]} = [2.5, 5.0] \quad \textit{and} \quad \frac{I}{Y} = \frac{[1.0, 1.0]}{[0.1, 0.5]} = [2.0, 10.0] .$$

We also observe that $I - X \subseteq I - Y$, by noting that

$$I - X = [1.0, 1.0] - [0.2, 0.4] = [0.6, 0.8]$$

and

$$I - Y = [1.0, 1.0] - [0.1, 0.5] = [0.5, 0.9] .$$

Moreover, as a composition of these two operations, we again have $\frac{I}{I-X} \subseteq \frac{I}{I-Y}$, since

$$\frac{I}{I - X} = [5/4, 5/3] \qquad \textit{and} \qquad \frac{I}{I - Y} = [10/9, 2] .$$

We next extend the notion of intervals and interval arithmetic to include interval vectors and interval matrices. Interval vectors and interval matrices are similarly defined. For example,

$$A^I = \begin{bmatrix} [2,3] & [0,1] \\ [1,2] & [2,3] \end{bmatrix} \qquad \textit{and} \qquad \mathbf{b}^I = \begin{bmatrix} [0,10] \\ [-6,1] \end{bmatrix}$$

are an interval matrix and an interval vector, respectively.

Let $A^I = [a^I_{ij}]$ and $B^I = [b^I_{ij}]$ be $n \times m$ interval matrices. Then, A^I and B^I are said to be *equal* if $a^I_{ij} = b^I_{ij}$ for all $i = 1, \cdots, n$ and $j = 1, \cdots, m$; A^I is said to be *contained* in B^I, denoted $A^I \subseteq B^I$, if $a^I_{ij} \subseteq b^I_{ij}$ for all $i = 1, \cdots, n$ and $j = 1, \cdots, m$, where, in particular, if $A^I = A$ is an ordinary constant matrix, we write $A \in B^I$.

Fundamental operations of interval matrices include:

(a) *Addition* and *Subtraction*:

$$A^I \pm B^I = \left[a^I_{ij} \pm b^I_{ij}\right] .$$

(b) *Multiplication*: For two $n \times r$ and $r \times m$ interval matrices A^I and B^I,

$$A^I B^I = \left[\sum_{k=1}^{r} a^I_{ik} b^I_{kj} \right] .$$

(c) *Inversion*: For an $n \times n$ interval matrix A^I with $det\,[A^I] \neq 0$,

$$\left[A^I\right]^{-1} = \frac{adj\,[A^I]}{det\,[A^I]} .$$

For instance, if $A^I = \begin{bmatrix} [2,3] & [0,1] \\ [1,2] & [2,3] \end{bmatrix}$, then

$$\left[A^I\right]^{-1} = \frac{adj\,[A^I]}{det\,[A^I]} = \frac{\begin{bmatrix} [2,3] & -[0,1] \\ -[1,2] & [2,3] \end{bmatrix}}{[2,3]\,[2,3] - [0,1]\,[1,2]}$$
$$= \begin{bmatrix} [2/9, 3/2] & [-1/2, 0] \\ [-1, -1/9] & [2/9, 3/2] \end{bmatrix} .$$

Interval matrices (including vectors) obey many algebraic operational rules that are similar to those for intervals (cf. Exercise 10.2).

10.1.3 Rational Interval Functions

Let S_1 and S_2 be intervals in R and $f : S_1 \to S_2$ be an ordinary one-variable real-valued (i.e., point-to-point) function. Denote by Σ_{S_1} and Σ_{S_2} families of all subintervals of S_1 and S_2, respectively. The interval-to-interval function, $f^I : \Sigma_{S_1} \to \Sigma_{S_2}$, defined by

$$f^I(X) = \left\{ f(x) \in S_2 : \quad x \in X \,, \; X \in \Sigma_{S_1} \right\}$$

is called the *united extension* of the point-to-point function f on S_1. Obviously, its range is

$$f^I(X) = \bigcup_{x \in X} \{f(x)\} \,,$$

which is the union of all the subintervals of S_2 that contain the single point $f(x)$ for some $x \in X$.

The following property of the united extension $f^I : \Sigma_{S_1} \to \Sigma_{S_2}$ follows immediately from definition, namely,

$$X, Y \in \Sigma_{S_1} \quad \textit{and} \quad X \subseteq Y \quad \Longrightarrow \quad f^I(X) \subseteq f^I(Y) \,.$$

In general, an interval-to-interval function F of n-variables, $X_1, \cdots, X_n$, is said to have the *monotonic inclusion property*, if

$$X_i \subseteq Y_i \,, \quad \forall \; i = 1, \cdots, n \quad \Longrightarrow \quad F(X_1, \cdots, X_n) \subseteq F(Y_1, \cdots, Y_n) \,.$$

Note that not all interval-to-interval functions have this property.

However, all united extensions have the monotonic inclusion property. Since interval arithmetic functions are united extensions of the real arithmetic functions: addition, subtraction, multiplication and division $(+, -, \cdot, /)$, interval arithmetic has the monotonic inclusion property, as previously discussed (cf. Theorem 10.1 and Corollary 10.1).

An interval-to-interval function will be called an *interval function* for simplicity. *Interval vectors* and *interval matrices* are similarly defined. An interval function is said to be *rational*, and so is called a *rational interval function*, if its values are defined by a finite sequence of interval arithmetic operations. Examples of rational interval functions include $X + Y^2 + Z^3$ and $(X^2 + Y^2)/Z$, etc., for intervals X, Y and Z, provided that $0 \notin Z$ for the latter.

It follows from the transitivity of the partially ordered relation $\subseteq$ that all the rational interval functions have the monotonic inclusion property. This can be verified by mathematical induction.

Next, let $f = f(x_1, \cdots, x_n)$ be an ordinary n-variable real-valued function, and $X_1, \cdots, X_n$ be intervals. An interval function, $F = F(X_1, \cdots, X_n)$, is said to be an *interval extension* of f if

$$F(x_1, \cdots, x_n) = f(x_1, \cdots, x_n), \qquad \forall\ x_i \in X_i\,,\ i = 1, \cdots, n\,.$$

Note also that not all the interval extensions have the monotonic inclusion property.

The following result can be established (cf. Exercise 10.3):

Theorem 10.2. *If F is an interval extension of f with the monotonic inclusion property, then the united extension f^I of f satisfies*

$$f^I(X_1, \cdots, X_n) \subseteq F(X_1, \cdots, X_n)\,.$$

Since rational interval functions have the monotonic inclusion property, we have the following

Corollary 10.2. *If F is a rational interval function and is an interval extension of f, then*

$$f^I(X_1, \cdots, X_n) \subseteq F(X_1, \cdots, X_n)\,.$$

This corollary provides a means of finite evaluation of upper and lower bounds on the value-range of an ordinary rational function over an n-dimensional rectangular domain in R^n.

As an example of the monotonic inclusion property of rational interval functions, consider calculating the function

$$f^I(X, A) = \frac{AX}{I - X}$$

for two cases: $X_1 = [2,3]$ with $A_1 = [0,2]$, and $X_2 = [2,4]$ with $A_2 = [0,3]$, respectively. Here, $X_1 \subset X_2$ and $A_1 \subset A_2$. A direct calculation yields

$$f_1^I(X_1, A_1) = \frac{[0,2] \cdot [2,3]}{[1,1] - [2,3]} = [-6, 0]$$

and

$$f_2^I(X_2, A_2) = \frac{[0,3] \cdot [2,4]}{[1,1] - [2,4]} = [-12, 0]\,.$$

Here, we do have $f_1^I(X_1, A_1) \subset f_2^I(X_2, A_2)$, as expected.

We finally note that based on Corollary 10.2, when we have interval division of the type X^I/X^I where X^I does not contain zero, we can first examine its corresponding ordinary function and operation to obtain $x/x = 1$, and then return to the interval setting for the final answer. Thus, symbolically, we may write $X^I/X^I = 1$ for an interval X^I not containing zero. This is indeed a convention in interval calculations.

10.1.4 Interval Expectation and Variance

Let $f(x)$ be an ordinary function defined on an interval X. If f satisfies the ordinary Lipschitz condition

$$|f(x) - f(y)| \leq L\,|x - y|$$

for some positive constant L which is independent of $x, y \in X$, then the united extension f^I of f is said to be a *Lipschitz interval extension* of f over X.

Let $B(X)$ be a class of functions defined on X that are most commonly used in numerical computation, such as the four arithmetic functions $(+, -, \cdot, /)$ and the elementary type of functions like $e^{(\cdot)}$, $ln(\cdot)$, $\sqrt{\cdot}$, etc. We will only use some of these commonly used functions throughout this chapter, so $B(X)$ is introduced only for notational convenience.

Let N be a positive integer. Subdivide an interval $[a, b] \subseteq X$ into N subintervals, $X_1 = [\underline{X}_1, \overline{X}_1], \cdots, X_N = [\underline{X}_N, \overline{X}_N]$, such that

$$a = \underline{X}_1 < \overline{X}_1 = \underline{X}_2 < \overline{X}_2 = \cdots = \underline{X}_N < \overline{X}_N = b\,.$$

Moreover, for any $f \in B(X)$, let F be a Lipschitz interval extension of f defined on all X_i, $i = 1, \cdots, N$. Assume that F satisfies the monotonic inclusion property. Using the notation

$$S_N(F; [a, b]) = \frac{b-a}{N} \sum_{i=1}^{N} F(X_i)\,,$$

we have

$$\int_a^b f(t)dt = \bigcap_{N=1}^{\infty} S_N(F; [a, b]) = \lim_{N\to\infty} S_N(F; [a, b])\,.$$

Note that if we recursively define

$$\begin{cases} Y_1 = S_1\,, \\ Y_{k+1} = S_{k+1} \cap Y_k\,, \quad k = 1, 2, \cdots\,, \end{cases}$$

where $S_k = S_k(F;[a,b])$, then $\{Y_k\}$ is a nested sequence of intervals that converges to the exact value of the integral $\int_a^b f(t)dt$.

Note also that a Lipschitz interval extension F used here has the property that $F(x)$ is a real number for any real number $x \in R$. However, for other interval functions that have the monotonic inclusion property but are not Lipschitz, the corresponding function $F(x)$ may have interval coefficients even if x is a real number.

Next, based on the interval mathematics introduced above, we introduce the following important concept.

Let X be an interval of real-valued random variables of interest, and let

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma_x} exp\left\{\frac{-(x-\mu_x)^2}{2\sigma_x^2}\right\}, \qquad x \in X,$$

be an ordinary Gaussian density function with known μ_x and $\sigma_x > 0$. Then $f(x)$ has a Lipschitz interval extension, so that the *interval expectation*

$$\begin{aligned} E(X) &= \int_{-\infty}^{\infty} xf(x)dx \\ &= \int_{-\infty}^{\infty} \frac{x}{\sqrt{2\pi}\sigma_x} exp\left\{\frac{-(x-\mu_x)^2}{2\sigma_x^2}\right\} dx, \quad x \in X, \end{aligned} \tag{10.3}$$

and the *interval variance*

$$\begin{aligned} Var(X) &= E([X - E(X)]^2) \\ &= \int_{-\infty}^{\infty} (x-\mu_x)^2 f(x)dx \\ &= \int_{-\infty}^{\infty} \frac{(x-\mu_x)^2}{\sqrt{2\pi}\sigma_x} exp\left\{\frac{-(x-\mu_x)^2}{2\sigma_x^2}\right\} dx, \quad x \in X, \end{aligned} \tag{10.4}$$

are both well defined. This can be easily verified based on the definite integral defined above, with $a \to -\infty$ and $b \to \infty$. Also, with respect to another real interval Y of real-valued random variables, the *conditional interval expectation*

$$\begin{aligned} E(X|\, y \in Y) &= \int_{-\infty}^{\infty} xf(x|y)dx \\ &= \int_{-\infty}^{\infty} x\frac{f(x,y)}{f(y)}dx \\ &= \int_{-\infty}^{\infty} \frac{x}{\sqrt{2\pi}\sigma_{xy}} exp\left\{\frac{-(x-\mu_{xy})^2}{2\sigma_{xy}^2}\right\} dx, \quad x \in X, \end{aligned} \tag{10.5}$$

and the *conditional variance*

$$
\begin{aligned}
& Var(X|\, y \in Y) \\
&= E\big((x-\mu_x)^2|\ y \in Y\big) \\
&= \int_{-\infty}^{\infty} \big[x - E(x|\ y \in Y)\big]^2 f(x|y)dx \\
&= \int_{-\infty}^{\infty} \big[x - E(x|\, y \in Y)\big]^2 \frac{f(x,y)}{f(y)} dx \\
&= \int_{-\infty}^{\infty} \frac{\big[x - E(x|\, y \in Y)\big]^2}{\sqrt{2\pi}\widetilde{\sigma}}\, exp\left\{\frac{-(x-\widetilde{\mu})^2}{2\widetilde{\sigma}^2}\right\} dx\,, \quad x \in X, \qquad (10.6)
\end{aligned}
$$

are both well defined. This can be verified based on the same reasoning and the well-defined interval division operation (note that zero is not contained in the denominator for a Gaussian density interval function). In the above,

$$
\widetilde{\mu} = \mu_x + \sigma_{xy}^2 (y - \mu_y)/\sigma_y^2 \qquad and \qquad \widetilde{\sigma}^2 = \sigma_x^2 - \sigma_{xy}^2 \sigma_{yx}^2 / \sigma_y^2\,,
$$

with

$$
\sigma_{xy}^2 = \sigma_{yx}^2 = E(XY) - E(X)E(Y) = E(xy) - E(x)E(y)\,, \quad x \in X\,.
$$

Moreover, it can be verified (cf. Exercise 10.4) that

$$
E(X|\, y \in Y) = E(x) + \sigma_{xy}^2 \big[y - E(y)\big]/\sigma_y^2\,, \qquad x \in X\,, \qquad (10.7)
$$

and

$$
Var(X|\, y \in Y) = Var(x) - \sigma_{xy}^2 \sigma_{yx}^2 / \sigma_y^2\,, \qquad x \in X\,. \qquad (10.8)
$$

Finally, we note that all these quantities are well-defined rational interval functions, so that Corollary 10.2 can be applied to them.

10.2 Interval Kalman Filtering

Now, return to the interval system (10.2). Observe that this system has an upper boundary system defined by all upper bounds of elements of its interval matrices:

$$\begin{cases} \mathbf{x}_{k+1} = [A_k + |\Delta A_k|]\mathbf{x}_k + [\Gamma_k + |\Delta\Gamma_k|]\underline{\xi}_k\,, \\ \quad \mathbf{v}_k = [C_k + |\Delta C_k|]\mathbf{x}_k + \underline{\eta}_k \end{cases} \tag{10.9}$$

and a lower boundary system using all lower bounds of the elements of its interval matrices:

$$\begin{cases} \mathbf{x}_{k+1} = [A_k - |\Delta A_k|]\mathbf{x}_k + [\Gamma_k - |\Delta\Gamma_k|]\underline{\xi}_k\,, \\ \quad \mathbf{v}_k = [C_k - |\Delta C_k|]\mathbf{x}_k + \underline{\eta}_k\,. \end{cases} \tag{10.10}$$

We first point out that by performing the standard Kalman filtering algorithm for these two boundary systems, the resulting two filtering trajectories do not encompass all possible optimal solutions of the interval system (10.2) (cf. Exercise 10.5). As a matter of fact, there is no specific relation between these two boundary trajectories and the entire family of optimal filtering solutions: the two boundary trajectories and their neighboring ones are generally intercrossing each other due to the noise perturbations. Therefore, a new filtering algorithm that can provide all-inclusive estimates for the interval system is needed. The interval Kalman filtering scheme derived below serves this purpose.

10.2.1 The Interval Kalman Filtering Scheme

Recall the derivation of the standard Kalman filtering algorithm given in Chapter 3, in which only matrix algebraic operations (additions, subtractions, multiplications, and inversions) and (conditional) expectations and variances are used. Since all these operations are well defined for interval matrices and rational interval functions, as discussed in the last section, the same derivation can be carried out for interval systems in exactly the same way to yield a Kalman filtering algorithm for the interval system (10.2). This interval Kalman filtering algorithm is simply summarized as follows:

The Interval Kalman Filtering Scheme

The main-process:

$$\begin{aligned}
&\hat{\mathbf{x}}_0^I = E(\mathbf{x}_0^I)\,,\\
&\hat{\mathbf{x}}_k^I = A_{k-1}^I\hat{\mathbf{x}}_{k-1}^I + G_k^I\Big[\mathbf{v}_k^I - C_k^I A_{k-1}^I\hat{\mathbf{x}}_{k-1}^I\Big]\,,\\
&k = 1,2,\cdots\,.
\end{aligned} \tag{10.11}$$

The co-process:

$$\begin{aligned}
&P_0^I = Var(\mathbf{x}_0^I)\,,\\
&M_{k-1}^I = A_{k-1}^I P_{k-1}^I\big[A_{k-1}^I\big]^\top + B_{k-1}^I Q_{k-1}\big[B_{k-1}^I\big]^\top\,,\\
&G_k^I = M_{k-1}^I\big[C_k^I\big]^\top\Big[\big[C_k^I\big]M_{k-1}^I\big[C_k^I\big]^\top + R_k\Big]^{-1}\,,\\
&P_k^I = \big[I - G_k^I C_k^I\big]M_{k-1}^I\big[I - G_k^I C_k^I\big]^\top + \big[G_k^I\big]R_k\big[G_k^I\big]^\top\,,\\
&k = 1,2,\cdots\,.
\end{aligned} \tag{10.12}$$

A comparison of this algorithm with the standard Kalman filtering scheme (3.25) reveals that they are exactly the same in form, except that all matrices and vectors in (10.11)–(10.12) are intervals. As a result, the interval estimate trajectory will diverge rather quickly. However, this is due to the conservative interval modeling but not the new filtering algorithm.

It should be noted that from the theory this interval Kalman filtering algorithm is optimal for the interval system (10.2), in the same sense as the standard Kalman filtering scheme, since no approximation is needed in its derivation. The filtering result produced by the interval Kalman filtering scheme is a sequence of interval estimates, $\{\hat{\mathbf{x}}_k^I\}$, that encompasses all possible optimal estimates $\{\hat{\mathbf{x}}_k\}$ of the state vectors $\{\mathbf{x}_k\}$ which the interval system may generate. Hence, the filtering result produced by this interval Kalman filtering scheme is inclusive but generally conservative in the sense that the range of interval estimates is often unnecessarily wide in order to include all possible optimal solutions.

It should also be remarked that just like the random vector (the measurement data) $\mathbf{v}_k$ in the ordinary case, the interval data vector $\mathbf{v}_k^I$ shown in the interval Kalman filtering scheme above is an uncertain interval vector before its realization (i.e., before the data actually being obtained), but will be an ordinary constant vector after it has been measured and obtained. This should avoid possible confusion in implementing the algorithm.

10.2.2 Suboptimal Interval Kalman Filter

To improve the computational efficiency, appropriate approximations of the interval Kalman filtering algorithm (10.11)–(10.12) may be applied. In this subsection, we suggest a suboptimal interval Kalman filtering scheme, by replacing its interval matrix inversion with its worst-case inversion, while keeping everything else unchanged.

To do so, let

$$C_k^I = C_k + \Delta C_k \qquad \textit{and} \qquad M_{k-1}^I = M_{k-1} + \Delta M_{k-1}\,,$$

where C_k is the center point of C_k^I and M_{k-1} is center point of M_{k-1}^I (i.e., the nominal values of the interval matrices). Write

$$\begin{aligned}
&\left[[C_k^I]M_{k-1}^I[C_k^I]^\top + R_k\right]^{-1} \\
&= \left[[C_k + \Delta C_k]\,[M_{k-1} + \Delta M_{k-1}]\,[C_k + \Delta C_k]^\top + R_k\right]^{-1} \\
&= \left[C_k M_{k-1} C_k^\top + \Delta R_k\right]^{-1},
\end{aligned}$$

where

$$\begin{aligned}
\Delta R_k &= C_k M_{k-1}[\Delta C_k]^\top + C_k[\Delta M_{k-1}]\,C_k^\top + C_k[\Delta M_{k-1}]\,[\Delta C_k]^\top \\
&\quad + [\Delta C_k]M_{k-1}C_k^\top + [\Delta C_k]M_{k-1}[\Delta C_k]^\top + [\Delta C_k]\,[\Delta M_{k-1}]C_k^\top \\
&\quad + [\Delta C_k]\,[\Delta M_{k-1}]\,[\Delta C_k]^\top + R_k\,.
\end{aligned}$$

Then, in the algorithm (10.11)–(10.12), replace ΔR_k by its upper bound matrix, $|\Delta R_k|$, which consists of all the upper bounds of the interval elements of $\Delta R_k = [[-r_k(i,j), r_k(i,j)]]$, namely,

$$|\Delta R_k| = [r_{ij}]\,, \qquad r_k(i,j) \geq 0\,. \tag{10.13}$$

We should note that this $|\Delta R_k|$ is an ordinary (non-interval) matrix, so that when the ordinary inverse matrix $[C_k M_{k-1} C_k^\top + |\Delta R_k|]^{-1}$ is used to replace the interval matrix inverse $[[C_k^I]M_{k-1}^I [C_k^I]^\top + R_k]^{-1}$, the matrix inversion becomes much easier. More importantly, when the perturbation matrix $\Delta C_k = 0$ in (10.13), meaning that the measurement equation in system (10.2) is as accurate as the nominal system model (10.1), we have $|\Delta R_k| = R_k$.

Thus, by replacing ΔR_k with $|\Delta R_k|$, we obtain the following *suboptimal* interval Kalman filtering scheme.

A Suboptimal Interval Kalman Filtering Scheme

The main-process:

$$\begin{aligned}
&\hat{\mathbf{x}}_0^I = E(\mathbf{x}_0^I)\,,\\
&\hat{\mathbf{x}}_k^I = A_{k-1}^I \hat{\mathbf{x}}_{k-1}^I + G_k^I\big[\mathbf{v}_k^I - C_k^I A_{k-1}^I \hat{\mathbf{x}}_{k-1}^I\big]\,,\\
&k = 1, 2, \cdots .
\end{aligned} \tag{10.14}$$

The co-process:

$$\begin{aligned}
&P_0^I = Var(x_0^I)\,,\\
&M_{k-1}^I = A_{k-1}^I P_{k-1}^I \big[A_{k-1}^I\big]^\top + B_{k-1}^I Q_{k-1} \big[B_{k-1}^I\big]^\top\,,\\
&G_k^I = M_{k-1}^I \big[C_k^I\big]^\top \Big[C_k M_{k-1} C_k^\top + |\Delta R_k|\Big]^{-1}\,,\\
&P_k^I = \big[I - G_k^I C_k^I\big] M_{k-1}^I \big[I - G_k^I C_k^I\big]^\top + \big[G_k^I\big] R_k \big[G_k^I\big]^\top\,,\\
&k = 1, 2, \cdots .
\end{aligned} \tag{10.15}$$

Finally, we remark that the worst-case matrix $|\Delta R_k|$ given in (10.13) contains the largest possible perturbations and is in some sense the "best" matrix that yields a numerically stable inverse. Another possible approximation is, if ΔC_k is small, to simply use $|\Delta R_k| \approx R_k$. For some specific systems such as the radar tracking system to be discussed in the next subsection, special techniques are also possible to improve the speed and/or accuracy in performing suboptimal interval filtering.

10.2.3 An Example of Target Tracking

In this subsection, we show a computer simulation by comparing the interval Kalman filtering with the standard one, for a simplified version of the radar tracking system (3.26), see also (4.22), (5.22), and (6.43),

$$\begin{cases}
\mathbf{x}_{k+1} = \begin{bmatrix} 1 & h^I \\ 0 & 1 \end{bmatrix} \mathbf{x}_k + \xi_k\,,\\
v_k = \begin{bmatrix} 1 & 0 \end{bmatrix} x_k + \eta_k\,,
\end{cases} \tag{10.16}$$

where basic assumptions are the same as those stated for the system (10.2). Here, the system has uncertainty in an interval entry:

$$h^I = [h - \Delta h, h + \Delta h] = [0.01 - 0.001, 0.01 + 0.001] = [0.009, 0.011]\,,$$

in which the modeling error Δh was taken to be 10% of the nominal value of $h = 0.01$. Suppose that the other given data are:

$$E(x_0) = \begin{bmatrix} x_{01} \\ x_{02} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad Var(x_0) = \begin{bmatrix} P_{00} & P_{01} \\ P_{10} & P_{11} \end{bmatrix} = \begin{bmatrix} 0.5 & 0.0 \\ 0.0 & 0.5 \end{bmatrix},$$
$$Q_k = \begin{bmatrix} q & 0 \\ 0 & q \end{bmatrix} = \begin{bmatrix} 0.1 & 0.0 \\ 0.0 & 0.1 \end{bmatrix}, \qquad R_k = r = 0.1\,.$$

For this model, using the interval Kalman filtering algorithm (10.11)–(10.12), we have

$$\begin{aligned}
M_{k-1}^I &= \begin{bmatrix} \begin{matrix} h^I\left[2P_{k-1}^I(1,0) + h^I P_{k-1}^I(1,1)\right] \\ +P_{k-1}^I(0,0)+q \end{matrix} & P_{k-1}^I(0,1) + h^I P_{k-1}^I(1,1) \\ P_{k-1}^I(1,0) + h^I P_{k-1}^I(1,1) & P_{k-1}^I(1,1)+q \end{bmatrix} \\
&:= \begin{bmatrix} M_{k-1}^I(0,0) & M_{k-1}^I(0,1) \\ M_{k-1}^I(1,0) & M_{k-1}^I(1,1) \end{bmatrix} \\
G_k^I &= \begin{bmatrix} 1 - r/(M_{00}^I + r) \\ M_{10}^I/(M_{00}^I + r) \end{bmatrix} := \begin{bmatrix} G_{k,1}^I \\ G_{k,2}^I \end{bmatrix} \\
P_k^I &= \begin{bmatrix} rG_{k,1}^I & rG_{k,2}^I \\ rG_{k,2}^I & \begin{matrix} q + \Big[P_{k-1}^I(1,1)\left[P_{k-1}^I(0,0)+q+r\right] \\ -[P_{k-1}^I(0,1)]^2\Big]\Big/\left(M_{k-1}^I(0,0)+r\right) \end{matrix} \end{bmatrix} \\
&:= \begin{bmatrix} P_k^I(0,0) & P_k^I(0,1) \\ P_k^I(1,0) & P_k^I(1,1) \end{bmatrix}.
\end{aligned}$$

In the above, the matrices M_{k-1}^I and P_k^I are both symmetrical. Hence, $M_{k-1}^I(0,1) = M_{k-1}^I(1,0)$ and $P_{k-1}^I(0,1) = P_{k-1}^I(1,0)$. It follows from the filtering algorithm that

$$\begin{bmatrix} \hat{x}_{k,1}^I \\ \hat{x}_{k,2}^I \end{bmatrix} = \begin{bmatrix} \left[r\left(\hat{x}_{k-1,1}^I + h^I \hat{x}_{k-1,2}^I\right) + M_{k-1}^I(0,0) y_k\right]/M_{k-1}^I(0,0) \\ \hat{x}_{k-1,2}^I + G_{k,1}^I\left(y_k - \hat{x}_{k-1,1}^I - h^I \hat{x}_{k-1,2}^I\right) \end{bmatrix}.$$

The simulation results for $\hat{x}_{k,1}$ of this interval Kalman filtering versus the standard Kalman filtering, where the latter used the nominal value of h, are shown and compared in both Figures 10.1 and 10.2. From these two figures we can see that the new scheme (10.11)–(10.12) produces the upper and lower boundaries of the single estimated curve obtained by the standard Kalman filtering algorithm (3.25), and these two boundaries encompass all possible optimal estimates of the interval system (10.2).

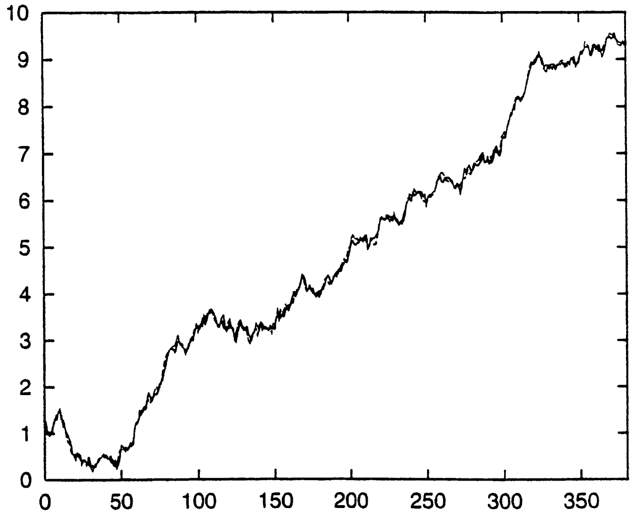

Fig. 10.1.

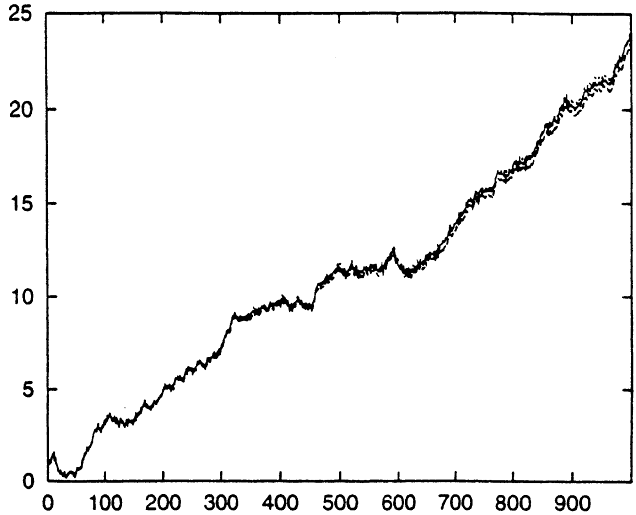

Fig. 10.2.

10.3 Weighted-Average Interval Kalman Filtering

As can be seen from Figures 10.1 and 10.2, the interval Kalman filtering scheme produces the upper and lower boundaries for all possible optimal trajectories obtained by using the standard Kalman filtering algorithm. It can also be seen that as the iterations continue, the two boundaries are expanding. Here, it should be emphasized once again that this seemingly divergent result is not caused by the filtering algorithm, but rather, by the iterations of the interval system model. That is, the upper and lower trajectory boundaries of the interval system keeps expanding by themselves even if there is no noise in the model and no filtering is performed. Hence, this phenomenon is inherent with interval systems, although it is a natural and convenient way for modeling uncertainties in dynamical systems.

To avoid this divergence while using interval system models, a practical approach is to use a weighted average of all possible optimal estimate trajectories encompassed by the two boundaries. An even more convenient way is to simply use a weighted average of the two boundary estimates. For instance, taking a certain weighted average of the two interval filtering trajectories in Figures 10.1 and 10.2 gives the the results shown in Figures 10.3 and 10.4, respectively.

Finally, it is very important to remark that this averaging is by nature different from the averaging of two standard Kalman filtering trajectories produced by using the two (upper and lower) boundary systems (10.9) and (10.10), respectively. The main reason is that the two boundaries of the filtering trajectories here, as shown in Figures 10.1 and 10.2, encompass all possible optimal estimates, but the standard Kalman filtering trajectories obtained from the two boundary systems do not cover all solutions (as pointed out above, cf. Exercise 10.5).

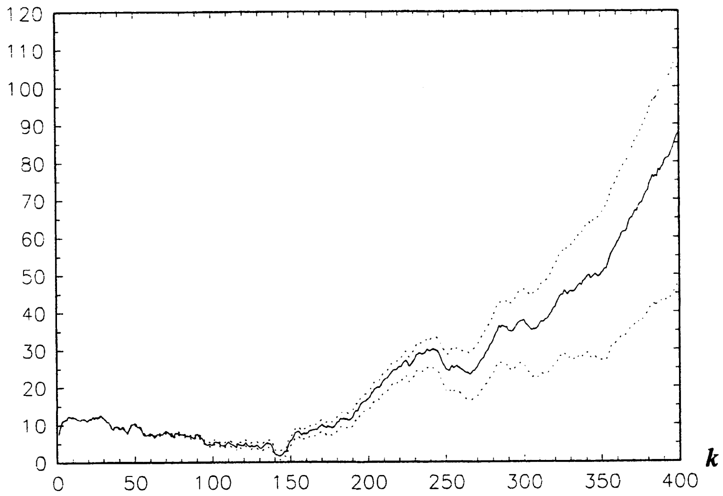

Fig. 10.3.

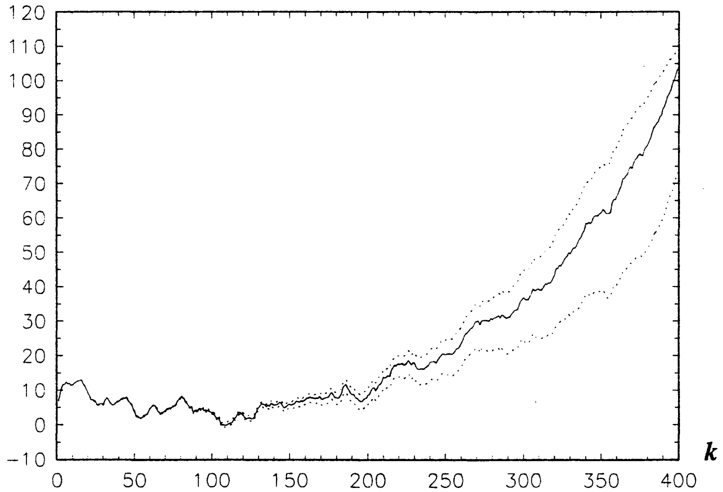

Fig. 10.4.

Exercises

10.1. For three intervals X, Y, and Z, verify that

$$X + Y = Y + X,$$
$$Z + (X + Y) = (Z + X) + Y,$$
$$XY = YX,$$
$$Z(XY) = (ZX)Y,$$
$$X + 0 = 0 + X = X \quad \textit{and} \quad X0 = 0X = 0, \ \textit{where } 0 = [0, 0],$$
$$XI = IX = X, \ \textit{where } I = [1, 1],$$
$$Z(X + Y) \subseteq ZX + ZY, \ \textit{where} \ = \ \textit{holds only if}$$
$$(a)\ Z = [z, z];$$
$$(b)\ X = Y = 0;$$
$$(c)\ xy \geq 0 \ \textit{for all } x \in X \textit{ and } y \in Y.$$

10.2. Let A, B, C be ordinary constant matrices and A^I, B^I, C^I be interval matrices, respectively, of appropriate dimensions. Show that

(a) $A^I \pm B^I = \{A \pm B \,|\, A \in A^I, B \in B^I\}$;
(b) $A^I B = \{AB \,|\, A \in A^I\}$;
(c) $A^I + B^I = B^I + A^I$;
(d) $A^I + (B^I + C^I) = (A^I + B^I) + C^I$;
(e) $A^I + 0 = 0 + A^I = A^I$;
(f) $A^I I = I A^I = A^I$;
(g) Subdistributive law:
(g.1) $(A^I + B^I)C^I \subseteq A^I C^I + B^I C^I$;
(g.2) $C^I(A^I + B^I) \subseteq C^I A^I + C^I B^I$;
(h) $(A^I + B^I)C = A^I C + B^I C$;
(i) $C(A^I + B^I) = CA^I + CB^I$;
(j) Associative and Subassociative laws:
(j.1) $A^I(BC) \subseteq (A^I B)C$;
(j.2) $(AB^I)C^I \subseteq A(B^I C^I)$ if $C^I = -C^I$;
(j.3) $A(B^I C) = (AB^I)C$;
(j.4) $A^I(B^I C^I) = (A^I B^I)C^I$, if $B^I = -B^I$ and $C^I = -C^I$.

10.3. Prove Theorem 10.2.

10.4. Verify formulas (10.7) and (10.8).

10.5. Carry out a simple one-dimensional computer simulation to show that by performing the standard Kalman filtering algorithm for the two boundary systems (10.9)–(10.10) of the interval system (10.2), the two resulting filtering trajectories do not encompass all possible optimal estimation solutions for the interval system.

10.6. Consider the target tracking problem of tracking an uncertain incoming ballistic missile. This physical problem is described by the following simplified interval model:

$$\begin{cases} \mathbf{x}_{k+1}^I = A_k^I \mathbf{x}_k^I + \underline{\xi}_k , \\ \mathbf{v}_k^I = C_k^I \mathbf{x}_k^I + \underline{\eta}_k , \end{cases}$$

where $\mathbf{x}^I = \begin{bmatrix} x_1^I & \cdots & x_7^I \end{bmatrix}^\top$,

$$\begin{aligned}
&A_{11} = A_{22} = A_{33} = 1 , \quad A_{44} = -\frac{1}{2}\, g\, x_7 \,\frac{z^2 + x_4^2}{z} , \\
&A_{45} = A_{54} = -\frac{1}{2}\, g \,\frac{x_7 x_4 x_5}{z} , \quad A_{46} = A_{64} = -\frac{1}{2}\, g \,\frac{x_7 x_4 x_6}{z} , \\
&A_{47} = -\frac{1}{2}\, g\, x_4 z , \quad A_{55} = -\frac{1}{2}\, g\, x_7 \,\frac{z^2 + x_5^2}{z} , \\
&A_{56} = A_{65} = -\frac{1}{2}\, g \,\frac{x_7 x_5 x_6}{z} , \quad A_{57} = -\frac{1}{2}\, g\, x_5 z , \\
&A_{66} = -\frac{1}{2}\, g\, x_7 \,\frac{z^2 + x_6^2}{z} , \quad A_{67} = -\frac{1}{2}\, g\, x_6 z , \\
&A_{76} = -K^I x_7 , \quad A_{77} = -K^I x_6 ,
\end{aligned}$$

with all other $A_{ij} = 0$, where K^I is an uncertain system parameter, g is the gravety constant, $z = \sqrt{x_4^2 + x_5^2 + x_6^2}$, and

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} .$$

Both the dynamical and measurement noise sequences are zero-mean Gaussian, mutually independent, with covariances $\{Q_k\}$ and $\{R_k\}$, respectively. Perform interval Kalman filtering on this model, using the following data set:

$$\begin{aligned}
&g = 0.981 , \\
&K^I = \left[2.3 \times 10^{-5}, 3.5 \times 10^{-5}\right] , \\
&\mathbf{x}_0^I = \big[3.2 \times 10^5, 3.2 \times 10^5, 2.1 \times 10^5, -1.5 \times 10^4, -1.5 \times 10^4, \\
&\qquad\quad - 8.1 \times 10^3, 5 \times 10^{-10}\big]^\top , \\
&P_0^I = diag\left\{10^6, 10^6, 10^6, 10^6, 10^6, 1.286 \times 10^{-13} exp\{-23.616\}\right\} , \\
&Q_k = \frac{1}{k+1}\; diag\left\{0, 0, 0, 100, 100, 100, 2.0 \times 10^{-18}\right\} , \\
&R_k = \frac{1}{k+1}\; diag\left\{150, 150, 150\right\} .
\end{aligned}$$