Box Math and KSM: Extending Sherman–Morrison to Functions of Interval Matrices

Ralph Kelsey
School of Electrical Engineering
and Computer Science
Ohio University
Athens, Ohio 45701
Email: kelsey@ohio.edu

Abstract—Certain advantages of midpoint/radius notation for boxes are well known. Box notation, a concise midpoint/radius scheme employing a ‘box operator’, \( \Box \), significantly simplifies box calculations. The box math approach to interval analysis, emphasizing ‘image-centered’ representations and assessment of quality of approximation, demonstrates the power of midpoint/radius methods. A prime example is the KSM method, extending complex analysis type expansions to functions of matrix boxes. This works particularly well for the inverse of a matrix box, leading to simple formulas for the hull of the solution of a linear box equation (interval matrix equation). An outline of key ideas of box math and proofs of some basic KSM theorems is presented below.

I. INTRODUCTION

Boxes (intervals, interval vectors, interval matrices, and similar sets) can be represented with endpoint notation, or in terms of midpoint and radius. While endpoint notation is more familiar, midpoint/radius representation is advantageous for some applications and computations. **Box notation** is a midpoint/radius notation scheme that extends standard set and linear algebra notation with a **box operator** denoted \( \Box \). A box is represented as \( B = M + R \Box \), leading to easy calculations and exposing the similarity of box calculations to the usual \( z = a + bi \) used in complex arithmetic. The \( \Box \) operator is defined below so that all math operations work correctly.

Box math refers to applications of boxes, the mathematics of boxes, equations involving boxes, quality of estimation by boxes, and related topics. Using box notation for calculations leads naturally to the concept of ‘image centered constructions’ for the image of functions of boxes, products of boxes, etc. Methods similar to those of complex analysis can sometimes be used to obtain simple approximating boxes. In the case of matrix boxes, Sherman Morrison like constructions prove to be useful. This method works particularly well for constructing the hull of the solution set for matrix interval equations. Several results of this approach, called KSM theory, are reported on below.

The major innovation of box math, the KSM method, is an approach to constructing box approximations for functions of matrix boxes. This was originally developed to provide simple formulas for boxes bounding the solution sets of interval matrix equations, etc. The basic KSM theorem is proven in detail. The method of proof can be extended to prove a wide variety of other formulas, some listed below.

The innovations of box math do not provide narrower bounds than existing methods, but rather extend the range of problems that can be treated, and provide simpler methods for standard problems. This report assumes the reader is familiar with standard interval analysis as presented in, e.g., [1].

II. BOX NOTATION

Box notation extends standard linear algebra and set math notation with one new symbol for a box operator.

A. Scalar field.

The usual notations \( \mathbb{R} \), \( \mathbb{C} \), and \( \mathbb{Q} \) are used for the standard scalar fields. Certain ‘near fields’, \( \mathbb{H} \) (quaternions), \( \mathbb{O} \) (octonions), and \( \mathbb{M} \) (IEEE 754 floating point numbers) can also be used for scalars. The simplest case, and the material presented below, uses \( \mathbb{R} \) for the scalar field. A few adjustments are needed for \( \mathbb{C} \). Techniques such as directed rounding can be incorporated when working with \( \mathbb{M} \). Lowercase Greek letters \( (\alpha, \beta, \gamma, \ldots) \) denote scalars. Sets of scalars are denoted by bold lowercase Greek letters. Distinguished sets of reals include the unit interval \( \omega \equiv [0,1] \) and the symmetric unit interval \( \omega \equiv [-1,1] \). In the complex case, \( \omega \) is the unit disk.

B. Matrices and vectors.

Uppercase and lowercase Roman letters denote matrices \( (A,B,C,\ldots) \) and vectors \( (a,b,c,\ldots) \), respectively. Thus, e.g., \( \alpha (Aa) = (\alpha A)a = A(\alpha a) \). The identity matrix is denoted \( I \). This common convention is extended by using uppercase Greek letters \( (A, B, \Gamma, \ldots) \) to denote either a scalar, a vector, or a matrix. Thus, the product \( \Gamma \Delta \) can represent \( GD, Gd, \gamma D, \gamma d, \text{ or } \gamma \delta \). This convention is slightly problematic in \( \mathbb{M} \times \mathbb{X} \), because the mathrm font used for uppercase Greek letters does not look italic in mathematical expressions. Any bold Roman or Greek letter indicates a set of matrices, vectors, or scalars.
C. Set math.

The standard set math notation for functions of sets, \( f(A) = \{ f(A) : A \in A \} \), and operations on sets, \( A \oplus B = \{ A \oplus B : A \in A, B \in B \} \) are used. For example, \( \omega = v - v \), and \( v^{-1} = [1, \infty] \). A nontrivial aspect of set math is subdistributivity: \( A(B + \Gamma) \subseteq AB + A\Gamma \), but usually \( A(B + \Gamma) \neq AB + A\Gamma \).

D. Box operator.

A simple notation for linear algebra with sets is enabled by the box operator, \( \varpi \), defined by \( A \varpi \equiv \{ B : |B| \leq |A| \} \). The \( \varpi \) operator is postfix (to enhance the analogy with complex numbers) and multiplies each component of the operand with an independent copy of \( \omega \); e.g.,

\[
\begin{bmatrix}
1 & 2 \\
3 & 4 \\
\end{bmatrix}
+ \begin{bmatrix}
5 & 4 \\
2 & 0 \\
\end{bmatrix}
\varpi = \begin{bmatrix}
1 + 5\omega & 2 + 4\omega \\
3 + 2\omega & 4 + 0\omega \\
\end{bmatrix} = \begin{bmatrix}
[1-5,6] & [-2,6] \\
[1,5] & [4,4] \\
\end{bmatrix}.
\]

The symbol \( \varpi \) is a bold, lowercase rendering of the obscure Greek letter pomegna, also called variant pi. It was chosen because its only common usage is for solid radian, and it looks rather like an operator version of \( \omega \).

A box (interval matrix, interval vector, or interval) is represented as \( B = M + R \varpi \), \( b = m + r \varpi \), or \( \beta = \mu + r \varpi \), respectively. For a box \( B = M + R \varpi \), the midpoint is \( M = \text{mid}(B) \) and the radius is \( R = \text{rad}(B) \).

For matrices and vectors, the pencil \( M + R \omega \) is a proper subset of the box \( M + R \varpi \). For scalars, \( \mu + r \omega = \mu + r \varpi \).

In all cases, \( A \varpi = |A| \varpi \), \( (AB) \varpi \subseteq A(B \varpi) = (|A||B|) \varpi = (A \varpi)B \), and \( (A + B \varpi) \pm (\Gamma + \Delta \varpi) = A \pm B + (|B| + \Gamma) \varpi \). Furthermore, \( A(AB) \varpi = A \varpi \), so \( \varpi^2 = \varpi \), i.e., \( \varpi \) is idempotent.

III. Box Math

Box math is an approach to interval analysis (IA) emphasizing the use of boxes to represent and approximate sets of matrices, vectors, or scalars. Such representation and approximation is essential in both modeling (most problems do not start with parameters defined by boxes) and calculations (the result of many operations on boxes are not boxes). Categorization and assessment of approximations is a major chapter in box math.

Any math operations indicated on boxes assumes the boxes are compatible. For example, \( Mv \), indicates that \( M \) is a set of \( n \times m \) matrices and \( v \) a set of \( m \) vectors. Any intersection of boxes is empty or a box. The intersection \( B \) of all boxes containing a nonempty set \( A \) is the hull of \( A \); \( B = \text{hull}(A) \).

The major innovation of the box math approach is the KSM method. The KSM formula for the inverse of a matrix box, and the kraal product (Section V) for matrix boxes, leads to simple approximations for the solutions for interval matrix equations, as developed in Section VIII.

IV. Representation and approximation of sets

A central task in many technical problems is representing and manipulating sets of matrices, vectors, or scalars. It is often useful or essential to approximate a set \( \Sigma \) with a simple set (a set that is significantly simpler to represent or manipulate).

Meritorious attributes that a category of approximating sets might exhibit include: (1) simplicity and economy of representation; (2) ease of use in practical applications; (3) simplicity, ease, and speed of calculation; (4) closeness of approximation (fit, or tightness); (5) closure with respect to operations such as translation, rotation, intersection, addition, subtraction, 

\[ \text{matrix multiplication} \]  

and solution of sets of linear equations; and (6) minimal discordance with tradition. Boxes (and unions of boxes) are clearly the best category of approximating sets, with ellipsoids a distant second.

For a given set \( \Sigma \), the sets \( R \) and \( T \) are an inner approximation and an outer approximation if \( R \subseteq \Sigma \subseteq T \). A variety of other terms are in use for these bounds, [5]. An outer approximation provide guaranteed containment. A variety of ‘best fit’ or ‘likely’ approximations can also be defined. For a given category of outer approximations (say, boxes), an approximation is tight if no smaller approximation exists. For example, the intersection of all outer approximating boxes, the box hull, or simply the hull, is the tightest box containing \( \Sigma \). Similarly, an inner approximation is tight if there is no larger inner approximation in its category.

An important property of boxes (and ellipsoids) is that the midpoint is an obvious centrally located distinguished point. Boxes that approximate functions of boxes, or the result of math operations on boxes, are called image centered if the midpoint of the box approximating the image is the image of the midpoint(s) of the defining box(es).

An image centered box is a core if it is an inner approximation, and a kraal if it is an outer approximation. The intersection of all kralas, the optimal kraal, is tight. A sufficient condition for a kraal to be optimal is that the approximated set ‘touches a corner’. A major achievement of box math are methods for obtaining simple formulas for (often optimal) kralas in important situations, such as the solution of a linear box equation.

The closeness of approximation of a set \( \Sigma \) by a simple set \( T \) can be measured by a fitting function of form \( \phi(\Sigma, T) \to [1, \infty] \), with a smaller value indicating a tighter approximation. Meritorious properties of fitting functions can be defined in terms of sets \( R, \Sigma, T \), and \( Y \), with \( R \subseteq \Sigma \subseteq T \), and point \( X \). Then \( \phi \) is reflective if \( \phi(\Sigma, \Sigma) = 1 \), symmetric if \( \phi(\Sigma, Y) = \phi(Y, \Sigma) \), transitive if \( \phi(R, \Sigma) \phi(\Sigma, T) = \phi(R, T) \), monotone if \( \phi(R, \Sigma) \leq \phi(R, T) \) and \( \phi(\Sigma, T) \leq \phi(R, T) \), translation invariant if \( \phi(\Sigma, Y) = \phi(\Sigma + X, Y + X) \), and scaling invariant if \( \alpha R \subseteq \beta T \) implies \( \phi(\alpha R, \beta T) = (\beta/\alpha) \phi(R, T) \).

Several fitting functions have been studied, including the
volumetric fit,
\[
\phi_{\text{vol}}(\Sigma, T) = \begin{cases} 
\sqrt{\frac{\text{vol}(\Sigma \cup T)}{\text{vol}(\Sigma \cap T)}}, & \text{if } \text{vol}(\Sigma \cap T) \neq 0, \\
1, & \text{if } \text{vol}(\Sigma \cup T) = 0, \\
\infty, & \text{otherwise}.
\end{cases}
\]

Here \( r = 1 \) if the sets are scalars, \( r = n \) if the sets are \( n \) vectors, and \( r = nm \) if the sets are \( n \) by \( m \) matrices.

For concordance with prior usage, if \( \Sigma \subseteq T \), a fit \( \phi(\Sigma, T) \) is called an overestimation and denoted \( \rho(\Sigma, T) \). For any norm \( \| \cdot \|_p \), define the radius of a set by \( \text{rad}_p(\Sigma) \equiv (1/2) \max \{ \| A - B \|_p : A, B \in T \} \). An important overestimation is then defined by
\[
\rho_p(\Sigma, T) = \begin{cases} 
\frac{\text{rad}_p(\Sigma)}{\text{rad}_p(T)}, & \text{if } \text{rad}_p(T) \neq 0, \\
1, & \text{if } \text{rad}_p(T) = 0.
\end{cases}
\]

The overestimation used by Rump in [4] is then \( \rho(\Sigma, T) = \rho_\infty(\Sigma, T) \), where the infinity, or sup, norm is given by \( \| \Sigma \|_\infty = \sup \{ \| \Sigma_i \| : i \text{ an index} \} \).

It is common in KSM theory to have inner and outer approximations to a set \( \Sigma \) available, \( R \subseteq \Sigma \subseteq T \), but no way to calculate \( \phi(R, \Sigma) \) or \( \phi(\Sigma, T) \). For a transitive fit \( \phi \), the kappa estimate \( \kappa_\phi \) (or, \( \kappa \)),
\[
\kappa_\phi = \sqrt{\phi(R, T)},
\]
is a reasonable estimate for both. If furthermore, \( R \) and \( T \) are boxes with the same shape \( \Upsilon \); \( R = M + \alpha \Upsilon \varpi \) and \( T = N + \beta \Upsilon \varpi \); and \( \phi \) is an invariant fit, then \( \kappa_\phi = \sqrt{\beta/\alpha} \).

### V. Products of boxes

Products of matrix and vector boxes are an important and nontrivial topic. A consistent system for denoting various approximations is outlined in this section. For boxes \( A = M + R \varpi \) and \( B = N + T \varpi \), the product \( AB \equiv \{ AB : A \in A, B \in B \} \) is rarely a box, but more like ‘a box with some corners sliced off.’ To distinguish it from approximations, the product \( AB \) is also called the set product and denoted \( A \otimes B \). The set product is frequently an object of interest, but in general there is no easy way to compute it. Several box products, denoted \( \varpi X \) for various \( X \), are useful for approximating the set product. The Interval Analysis community often uses the IA product, \( A \boxtimes I \, A B \), calculated with the usual linear algebra rule, but with the inner products replaced by interval products. An important approximation is the hull product, \( A \boxtimes H B \equiv \text{hull}(AB) \). It is not difficult to show that \( A \boxtimes H B = A \boxtimes I A B \), so the hull product can be calculated, and obviously \( A \otimes B \subseteq A \boxtimes H B \).

Rump has long advocated use of the mid/rad product, \( A \boxtimes K B \equiv MN + ([M]T + |R|N + |R||T|) \varpi \). In box math this is referred to as the kraal product because it is a prime example of an optimal kraal. It is not hard to show that \( A \boxtimes H B \subseteq A \boxtimes K B \).

The kraal product is the simplest and fastest product to calculate. The kraal product always ‘touches’ the hull product in a corner, indicating it is the optimal kraal for the set product. If \( \epsilon \) is an appropriately defined measure of the relative width of the boxes, then \( \rho(A \boxtimes H B, A \boxtimes K B) \leq 1 + \epsilon \). Rump has proven in [4] a worst case bound on the overestimation of the kraal product with respect to the hull product,
\[
\rho(A \boxtimes H B, A \boxtimes K B) \leq 1.5.
\]

The kraal product, used with KSM kraals, results in some particularly simple formulas for important sets, see Section VIII.

### VI. Set and box equations I

The culmination of most mathematical models is a set equation that can be expressed as \( \Phi(X, \Pi_0, \Pi_1, \ldots) = 0 \), where \( X \) is the solution, and \( \Pi_i \) are sets of parameters. Solving the model consists of obtaining information about the solution set as a function of the parameters, \( X = X(\Pi_0, \Pi_1, \ldots) \). If the \( \Pi_i \) are boxes (or approximated by boxes), the equation is a box equation, with linear box equations (interval matrix equations) of particular interest.

The important case of matrix/vector box equations (LBE), given as
\[
(M + R \varpi)x = m + r \varpi,
\]
can be regarded as a set of inequalities, complicated by the fact that the inequalities change in each orthant of \( x \) space. The Oettli/Prager procedure obtains the solution \( x(M, R, m, r) \) by systematically working with these inequalities. As is well known, the solution set is often surprisingly complex, indeed sometimes star-like. Obtaining the solution hull, \( h(M, R, m, r) = \text{hull}(x(M, R, m, r)) \) is known to be an NP-complete problem. A simple method for obtaining a kraal containing the solution set is given in Section VIII.

### VII. Basic Kral Sherman–Morrison theory

A matrix \( Q \) is rank 1 if it is an outer product of vectors \( a \) and \( b \); \( Q = ab^T \). Let \( \gamma = \text{tr}(Q) \) be the trace of \( A \). Then \( \gamma = a^T b \), \( Q^n = \gamma^{n-1} Q \) for integer \( n > 1 \), the operator norm is \( \| Q \| = \gamma \), and \( Qc = a(b^T c) = (b^T) a \), so \( Qc \) is a scalar times \( a \). The Sherman–Morrison formula states that if \( I + Q \) is not singular, then \( (I + Q)^{-1} = I - \frac{1}{1+\gamma} Q \), easily confirmed by calculation. Generalizations of the SM formula are found in [2].

The KSM (kraal Sherman–Morrison) approach provides a method for expanding functions of small matrix boxes. This method works particularly well for the inverse function, producing inverse kraals, boxes that contain the inverse of a matrix box. The KSM idea can be seen in an approximate calculation using box math; \( (I + Q \varpi)^{-1} = (I - Q \varpi)^{-1} = I + \sum_{i=1}^{\infty} (Q \varpi)^i = I + \sum_{i=1}^{\infty} (Q \varpi)^i = I + \sum_{i=0}^{\infty} \gamma^i Q \varpi = I + \frac{\gamma}{1-\gamma} Q \varpi \). While this argument is not rigorous due to concerns about factoring, subdistributivity, and convergence, it forms the basis of a proof.
\textbf{Theorem 1:} Let $Q$ be a rank 1 matrix with $\gamma = \text{tr}(Q) < 1$. Then
\[ (I + Q\varpi)^{-1} \subseteq I + \frac{1}{1-\gamma}Q\varpi. \]

\textbf{Proof.} Let $A \in (I + Q\varpi)^{-1}$. Then $\exists B$ such that $|B| \leq Q$ and $A = (I - B)^{-1} = I + (I - B)^{-1} - I$. Furthermore, $\|B\| \leq \|Q\| = \gamma < 1$, so the following expansion is valid: $|(I - B)^{-1} - I| = \sum_{i=1}^{\infty}|B|^i \leq \sum_{i=1}^{\infty}|Q|^i$. Using the fact that $Q^i = \gamma^i Q$, we have $|(I - B)^{-1} - I| \leq \sum_{i=1}^{\infty}\gamma^i |Q| = \frac{1}{1-\gamma}Q$. Thus, $(I - B)^{-1} - I \in \frac{1}{1-\gamma}Q\varpi$, so $A \in I + \frac{1}{1-\gamma}Q\varpi$. \hfill \Box

The symmetry of the inverse function leads to an inverse core inside the inverse of the matrix box. Thus $(I + Q\varpi)^{-1}$ has in inner-outper approximation pair.

\textbf{Theorem 2:} Let $Q$ be a rank 1 matrix with $\gamma = \text{tr}(Q) < 1$. Then
\[ I + \frac{1}{1+\gamma}Q\varpi \subseteq (I + Q\varpi)^{-1} \subseteq I + \frac{1}{1-\gamma}Q\varpi. \]

\textbf{Proof.} Apply the basic theorem to the left most set. \hfill \Box

The ‘corners’ of the approximating boxes, $I + \frac{1}{1-\gamma}Q$ and $I - \frac{1}{1+\gamma}Q$, are in $(I + Q\varpi)^{-1}$, so the approximating boxes are tight. The kappa estimate for this approximation is $\kappa = \sqrt{(1+\gamma)/(1-\gamma)} = 1 + \gamma + O(\gamma^2)$.

If $|A| \leq B$, then $B$ is said to majorize $A$. Rank 1 majorizing matrices play an important role in KSM theory, as in the following theorem.

\textbf{Theorem 3:} For a matrix $R$ and a nonsingular matrix $M$, suppose there is a rank 1 matrix $Q$ with $|M^{-1}| |R| \leq Q$ and $\gamma = \text{tr}(Q) < 1$, then
\[ (M + R\varpi)^{-1} \subseteq M^{-1} + \frac{1}{1-\gamma}(Q|M^{-1}|)\varpi. \]

\textbf{Proof.} $|R| = |MM^{-1}R| \leq |M| |Q|$, so $M + R\varpi \subseteq M(I + Q\varpi)$. Thus $(M + R\varpi)^{-1} \subseteq (I + \frac{1}{1+\gamma}Q\varpi)M^{-1} \subseteq M^{-1} + \frac{1}{1-\gamma}(Q|M^{-1}|)\varpi$. \hfill \Box

The bound established in Theorem 3 is usually fairly tight in the sense that corners of the solution set touch the krala. It is not as easy to find a core, or any simple set, inside the inverse of a generic matrix box.

Generalizations of these theorems producing inverse kralas are based on S–M formulas for the inverses of $I + (eI + Q)\varpi$, $D + Q\varpi$ ($D$ diagonal), $I + P\varpi$ (rank 2) and $M + UKV$ (Woodbury generalization of S–M). The S-M formula for $(D + Q\varpi)^{-1}$ is quite useful, and apparently new. Other dimensions of generalization are to functions of matrices other than the inverse, and to ellipsoids.

If the matrix box is not small, there might be no rank 1 $Q$ with $|M^{-1}| |R| \leq Q$ and $\gamma = \text{tr}(Q) < 1$. In this case, for tighter results, or otherwise, the problem can be treated in pieces, with $M + R\varpi \subseteq \bigcup_{i=1}^{n} (M_i + R_i\varpi)$ and $|M_i^{-1}| |R_i| \leq Q_i$.

For constructing minimal rank 1 majorizing matrices, the following functions are useful. For matrix $A$ and vector $a > 0$, define
\[
\text{majRow}(A, a) = \min \{ b : |A| \leq ab^T \}, \quad \text{majCol}(A, a) = \min \{ b : |A| \leq ba^T \},
\]
so that $|A| \leq \text{majCol}(A, a)a^T$ and $|A| \leq a \text{majRow}(A, a)^T$. These functions are used in examples treated below.

\section{Linear Box Equations}

The \textbf{midpoint equation} of $(M + R\varpi)x = m + r\varpi$ is $Mx = m$. Theorem 3 can be used to obtain a solution \textbf{krala} centered at the midpoint \textbf{solution}, $\hat{x} = M^{-1}m$.

\textbf{Theorem 4:} Suppose $(M + R\varpi)x = m + r\varpi$, with nonsingular $M$, and there is a rank 1 matrix $Q$ with $|M^{-1}| |R| \leq Q$ and $\gamma = \text{tr}(Q) < 1$, then
\[ x \subseteq \left( M^{-1} + \frac{1}{1-\gamma}(Q|M^{-1}|)\varpi \right) (m + r\varpi) \subseteq M^{-1}m + \left( \frac{1}{1-\gamma}(Q|M^{-1}|(|m| + |r|) + |M^{-1}| |r|) \right) \varpi \]

\textbf{Proof.} Theorem 3 gives the first inclusion, and the krala product gives the second. \hfill \Box

Similar results holds for the matrix-matrix equation, $(M + R\varpi)X = A + B\varpi$.

General results about the tightness of this approximation are not available. However, in many of the standard examples given in IA books, corners of the solution set $x$ touch, or nearly touch, a solution krala. This indicates that the solution krala is often near optimal.

Important applications of interval methods include representing uncertainty due to imprecise measurements, etc., and studying the propagation of uncertainty in computations. In many situations, including these, the radius of the boxes are likely to be small, and might represent absolute uncertainty, relative uncertainty, or a combination of both. In this context, a linear box equation takes a simple form. Let $E$ denote the matrix of all 1’s, and $e$ for the vector of all 1’s. Then the relative width of the matrix box is given by $\delta |M|$, the absolute width by $eE$, and similarly $\mu|m|$ and $\nu e$ for the vector box width. The \textbf{uniform width linear box equation} (UWLBE), with $R = \delta |M| + eE$ and $r = \mu|m| + \nu e$, is
\[ (M + (\delta |M| + eE)\varpi)x = m + (\mu|m| + \nu e)\varpi. \]

The coefficients $\delta$, $e$, $\mu$, and $\nu$ are nonnegative, and assumed small enough that a majorizing rank 1 matrix $Q$ for use in Theorem 4 can be found.

Let $f = |M^{-1}| e$, so that $|M^{-1}| E = fe^T$. Vectors $u$ and $h$ such that $|M^{-1}| |M| \leq uh^T$ are needed. It often
turns out that $f$, the vector of row sums of $|M|^{-1}$, is a good choice for $u$. Define $h = \text{majRow}(|M|^{-1}|M|, f)$, so that $|M|^{-1}|M| \leq fh^T$. Then $Q = f (\delta h + \epsilon e)^T$ is rank 1, and $|R| \leq Q$. Define also $q = |M|^{-1}m$. The solution krall involves $|M|^{-1}|m| = q$ and $|M|^{-1}|r| = \mu|M|^{-1}|m| + \nu|M|^{-1}e = \mu q + \nu f$. The result can be expressed in terms of $\rho = h^T f$, $\sigma = e^T f$, $\tau = q^T q$, $\theta = e^T q$, and $\gamma = \delta \rho + \epsilon \sigma$.

Theorem 5: A solution krall for the UWLBFE is given by

$$M^{-1}m + \left(\mu q + \frac{\nu + (1 + \mu)(\delta \tau + \epsilon \theta)}{1 - (\delta \rho + \epsilon \sigma)} f\right) \omega.$$  

Proof. Theorem 4 with $Q = f (\delta h + \epsilon e)^T$ gives the kraal radius as

$$\frac{1}{\gamma} (Q|M|^{-1}|(m + |r|) + |M|^{-1}|r|)$$

$$= f (\delta h + \epsilon e)^T (q + \mu q + \nu f) + \mu q + \nu f$$

$$= \frac{\delta (1 + \mu) + \delta \nu + \epsilon (1 + \mu) \theta + \epsilon \nu \sigma}{1 - \delta \rho - \epsilon \sigma} f$$

$$+ \mu q = \frac{\delta (1 + \mu) + \epsilon (1 + \mu) \theta + \epsilon \nu \sigma}{1 - \delta \rho - \epsilon \sigma} f + \mu q.$$  

This expression shows the role of the equation width parameters; $\delta$, $\epsilon$, $\mu$, and $\nu$. Assuming these parameters are small, the kraal is approximately $M^{-1}m + (\mu q + [\nu + \delta(\rho + \tau) + \epsilon(\sigma + \theta) + (\delta \tau + \epsilon \theta) \mu + \delta \rho + \epsilon \sigma]) f) \omega$. Evidently $f$ plays a major role in the kraal.

IX. Extensions

Other aspects of box math include (1) KSM theory for ellipsoids, (2) software implementation, and (3) representation of uncertainty with by nested sequences of boxes. Additional topics are summarized next.

A. Generalized KSM.

The proof of Theorem 1 exploits the fact that all terms in $(1 - z)^{-1}$ are positive. It turns out that this method can be extended to any function with a power series that is eventually either all of one sign, or alternating. Several of the results that can be obtained are given in the next theorem.

Theorem 6: Let $Q$ be a rank 1 matrix with $\|Q\| \leq 1$ and $\gamma = \text{tr}(Q)$. Then the following inclusions hold:

1. For negative real $p$,

$$I + \frac{1}{\gamma} (1 - (1 - \gamma)^p) Q \omega \subseteq (I + Q \omega)^p$$

$$\subseteq I + \frac{1}{\gamma} (1 - (1 - \gamma)^p - 1) Q \omega.$$  

2. For real $p \in [0, 1]$,

$$(I + Q \omega)^p \subseteq I + \frac{1}{\gamma} (1 - (1 - \gamma)^p) Q \omega.$$  

3. For positive integer $m$,

$$I + \frac{1}{\gamma} [1 - (1 - \gamma)^m] Q \omega \subseteq (I + Q \omega)^m$$

$$\subseteq I + \frac{1}{\gamma} [1 - (1 - \gamma)^m - 1] Q \omega.$$  

4. 

$$I + \frac{1}{\gamma} \sqrt{1 + \gamma - 1} Q \omega \subseteq \sqrt{I + Q \omega}$$

$$\subseteq I + \frac{1}{\gamma} [1 - \sqrt{1 - \gamma}] Q \omega.$$  

5. 

$$I + \frac{1}{\gamma} (1 - \exp(-\gamma)) Q \omega \subseteq \exp(Q \omega)$$

$$\subseteq I + \frac{1}{\gamma} (\exp(\gamma) - 1) Q \omega.$$  

6. 

$$\frac{1}{\gamma} \ln(1 + \gamma) Q \omega \subseteq \ln(I + Q \omega)$$

$$\subseteq \frac{1}{\gamma} \ln(1 - \gamma) Q \omega.$$  

Tightness estimations are available for these and other set inclusions.

B. Optimal low rank majorizing matrices.

It is not unusual for a matrix box to be rank 1. e.g., the uniform absolute width case. Otherwise, a rank 1 majorizing matrix must be obtained. A variety of criteria for an optimal majorizer are useful. The majorizer of minimum trace is an obvious objective, leading to the following problem.

Minimum majorizer problem: Given an $n$ by $n$ matrix $R$, find $n$-vectors $a$ and $b$ that minimize $a^T b$ with $|R| \leq ab^T$.

This is a non-trivial optimization problem in $2n - 1$ real variables with $n^2$ constraints, most of which are slack. To the knowledge of the author, only the $n = 2$ case is fully solved. Other criteria, such as minimum $L_\infty$ norm are ever harder. R. Horn has stated that he does not think much is known about such minimum majorizer problems. For approximation where best guess, as opposed to absolute containment is the goal, a 'best fit' low rank matrix is needed, leading to the following problem.

Best rank $k$ approximation problem. Given an $n$ by $n$ matrix $R$ and an integer $k < n$, find the rank $k$ matrix $A$ that minimizes $\|R - A\|_2$.

This problem has an elegant solution given in [3]: Take the singular value decomposition of $|R|$, retain the first $k$ singular values and set the rest to zero to get $A$.  


X. NOTATION AND TERMINOLOGY

The usual notation $\mathbf{u}$ for $[0, 1]$ is modified to $\mathbf{v}$ to conform to notation conventions. The $\omega$ symbol was chosen because it suggests two adjacent $\mathbf{v}$’s; one each for $[-1, 0]$ and $[0, 1]$. The $\varpi$ symbol is a bold $\text{pomega}$, an obscure Greek letter (also called ‘variant pi’). It was chosen because it looks like an operator version of $\omega$, and the only other common usage of $\varpi$ is for solid angle.

An ‘image centered outer approximation’ is a central concept in box math, but an awkward expression. The Afrikaans word kraal for ‘An area containing some huts, surrounded by a fence’, concisely captures the concept.

XI. CONCLUSION

Box notation is concise, logical, simple, and close to common usage. The box math approach suggests new methods in interval analysis. The KSM method leads to many formulas useful in applications of higher dimensional interval analysis.

ACKNOWLEDGMENT

The author would like to thank Scott Ferson, Vladik Kreinovich, Baker Kearfott, David Juedes, Roger Horn, and the referee.

REFERENCES