A Small Theorem on the Interval Approach Criterion for Systems of Linear Equations.

Isaac T Yonemoto

April 10, 2017

1 Motivation

The development of Unums encourages reformulation of computational mathematics to approach numerical problems, keeping in mind traditional performance metrics, such as speed and energy consumption, while still yielding mathematically accurate solutions. Traditional numerical systems, such as IEEE floating points, abandon accuracy for performance. Symbolic systems incur high calculation costs and still may encounter situations without closed form solutions. Interval arithmetic can have exploding uncertainty under certain operations without clear guidance on how to resolve this. Unums, however, provide a natural solution to the exploding uncertainty, by way of methods such as the Ubox method.

For solving systems of linear equations, the naive solution is to first perform a gauss-jordan elimination and then whittle down the solution set using Uboxes. This is inefficient and cannot distinguish between a unit cell that contains a solution and a unit cell that merely happens to be at the intersection of all of the hyperplanes. Moreover, the search time proceeds with time O(something large)

2 The Interval Approach (uSlice) method

To address this issue, I developed a strategy called the interval approach method. The procedure is as follows: Select a dimension. Partition a superset of the solved interval in this dimension with equal width (simplest to choose a power of two), called "uslices". Apply the uslices as values to the system of linear equations and measure the maximum error from the constant vector of the linear equations. Because the equations are linear, conceptually the "outside" slices should have greater error than the "inside" slices, with a few exceptions that are otherwise easily solved (e.g. an irrelevant dimension that trivially reduces to a smaller system of linear equations).

3 Statement of the interval approach criterion

First, some definitions:

Let M be a matrix in $\mathbb{R}^{n \times n}$, and $\vec{v}, \vec{r} \in \mathbb{R}^n | M\vec{v} = \vec{r}$.

Let Π be a closed box (n-orthotrope) in $\mathbb{R}^n$, with an equal tripartition of closed boxes across one dimension $\Pi = \bigcup \Pi_1 \Pi_2 \Pi_3$. WOLOG, let the partitioning dimension be 1, so:

$$\vec{v} \in \Pi_1 = \begin{pmatrix} [x_1, x_1 + \Delta x_1] \\ [x_2, x_2 + \Delta x_2] \\ [x_3, x_3 + \Delta x_3] \end{pmatrix} ; \Pi_2 = \begin{pmatrix} [x_1 + \Delta x_1, x_1 + 2\Delta x_1] \\ [x_2, x_2 + \Delta x_2] \\ [x_3, x_3 + \Delta x_3] \end{pmatrix} ; \Pi_3 = \begin{pmatrix} [x_1 + 2\Delta x_1, x_1 + 3\Delta x_1] \\ [x_2, x_2 + \Delta x_2] \\ [x_3, x_3 + \Delta x_3] \end{pmatrix} ;$$

Define an "heuristic function" $d_{\alpha,\vec{r}} : \mathbb{R}^n \to \mathbb{R}^n$ for each dimension α, where $d_{\alpha,\vec{r}}(\vec{x}) = |(M\vec{x})_\alpha - r_\alpha|$. This is effectively the dot product between the vector $\vec{x}$ and the α'th row of the matrix, or, crudely, 'the error, across the partition, when attempting to solve the α'th linear equation'.

We would like a criterion that guarantees that the characteristic shape of applying the heuristic function to each of these intervals results in a diagnostic that can iteratively refine the partition Π. The interval approach theorem shows that the following condition:

$$\Delta x_1 > \frac{\sum\limits_{i=2}^{n} |m_{\alpha,i}| \Delta x_i}{|m_{\alpha,1}|}$$

guarantees that $max(d_{\alpha,\vec{r}}(M \cdot \Pi_1)) < max(d_{\alpha,\vec{r}}(M \cdot \Pi_3))$. As a result, any equal *4-partition* of a closed box that has $max(d_{\alpha,\vec{r}}(M \cdot \Pi_1)) > max(d_{\alpha,\vec{r}}(M \cdot \Pi_2))$ and $max(d_{\alpha,\vec{r}}(M \cdot \Pi_3)) < max(d_{\alpha,\vec{r}}(M \cdot \Pi_4))$ has a guarantee that $\vec{r}$ is in either Π_2 or Π_3, which iteratively reduces the bounds on the solution.

4 Proof, n = 2

4.1 case $m_{\alpha,:} > 0$
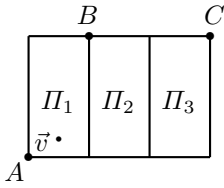
Lemma: $max(d_{\alpha,\vec{r}}(\Pi_1))$ **must be** $d_{\alpha,\vec{r}}(\vec{A})$ **or** $d_{\alpha,\vec{r}}(\vec{B})$. WOLOG, let's take $\alpha = 1$ and label the row $(a\ b) \equiv (m_{1,1}\ m_{1,2})$. For simplicity, we will identify $d = d_{\alpha,\vec{r}}$

Examine all points $\vec{x}$ where $v_1 < x_1 < B_1$, and $v_2 < x_2 < B_2$. Crudely, this is the "upper right sub-rectangle" created within Π_1 by $\vec{v}$. Compare $d(\vec{x})$ and $d(\vec{B})$.

$d(\vec{x}) \equiv |(M\vec{x} - \vec{r})_1| = |(M(\vec{v} + \delta\vec{x}) - M\vec{v})_1|$ for some positive vector $\delta\vec{x}$. By linearity, this $d(\vec{x}) = |(M\delta\vec{x})_1| = a(\delta x)_1 + b(\delta x)_2$

$d(\vec{A}) \equiv |(M\vec{A} - \vec{r})_1| = |(M(\vec{v} + \delta\vec{x} + \delta'\vec{x}) - M\vec{v})_1|$ for some positive or zero vector $\delta'\vec{x}$. By linearity, this $d(\vec{A}) = |(M\delta\vec{x} + M\delta'\vec{x})_1| = a(\delta x)_1 + b(\delta x)_2 + a(\delta'x)_1 + b(\delta'x)_2$. Since $\delta'\vec{x}$ is positive, $d(\vec{A}) \geq d(\vec{x})$ for all x in this upper right sub-rectangle, and so point B yields a maximum over this region.

Examine all points $\vec{x}$ where $A_1 < x_1 < v_1$, and $A_2 < x_2 < v_2$. Crudely, this is the "lower left sub-rectangle" created within Π_1 by $\vec{v}$. Compare $d(\vec{x})$ and $d(\vec{A})$.

$d(\vec{x}) \equiv |(M\vec{x} - \vec{r})_1| = |(M(\vec{v} - \delta\vec{x}) - M\vec{v})_1|$ for some positive vector $\delta\vec{x}$. By linearity, this $d(\vec{x}) = |-(M\delta\vec{x})_1| = |-a(\delta x)_1 - b(\delta x)_2| = a(\delta x)_1 + b(\delta x)_2$

$d(\vec{A}) \equiv |(M\vec{A} - \vec{r})_1| = |(M(\vec{v} - \delta\vec{x} - \delta'\vec{x}) - M\vec{v})_1|$ for some positive or zero vector $\delta'\vec{x}$. By linearity, this $d(\vec{A}) = |-(M\delta\vec{x} - M\delta'\vec{x})_1| = |-a(\delta x)_1 - b(\delta x)_2 - a(\delta'x)_1 - b(\delta'x)_2| a(\delta x)_1 + b(\delta x)_2 + a(\delta'x)_1 + b(\delta'x)_2$. Since $\delta'\vec{x}$ is positive, $d(\vec{A}) \geq d(\vec{x})$ for all x in this upper right sub-rectangle, and so point B yields a maximum over this region.