IMPLEMENTATION OF INTERVAL VALUED OPTIMIZATION TECHNIQUES APPLIED TO PARAMETER ESTIMATION UNDER UNCERTAINTY

Mathematical Engineering

CB0441 - Research Practice 2

Progress Report

April 10th, 2015

Jose D. Gallego-Posada
jgalle29@eafit.edu.co

Maria E. Puerta-Yepes
mpuerta@eafit.edu.co

Functional Analysis and Applications Research Group

UNIVERSIDAD EAFIT®

UNIVERSIDAD EAFIT

Definition - Notation

Consider the closed interval denoted by $[a, b]$ which represents the set of real numbers given by

$$[a, b] = \{x \in (R) : a \leq x \leq b\}$$

Define $I(\mathbb{R}) := \{ [a, b] : a \leq b, a, b \in \mathbb{R}\}$ be the set of all closed intevals of $\mathbb{R}$. We say a interval $[a, b]$ is degenerate if $a = b$.

We adopt the *infimum-supremum* notation for intervals:

$$X = [X^L, X^U] \text{ with } X^L, X^U \in \mathbb{R}$$

$$X = Y \text{ if } X^L = Y^L \wedge X^U = Y^U$$

Relevance of Intersection

Intersection plays a key role in interval analysis. If we have two intervals containing a result of interest — regardless of how they were obtained — then the intersection, which may be narrower, also contains the result.

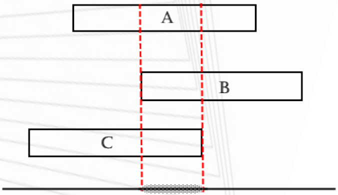

Figure 1 : Intersection of measurements.

Width, Absolute Value, Midpoint (I)

Length

$$l(X) := X^U - X^L$$

Absolute Value

$$|X| := \max\left\{\left|X^L\right|, \left|X^U\right|\right\}$$

Midpoint

$$m(X) := \frac{1}{2}(X^L + X^U)$$

Width, Absolute Value, Midpoint (II)

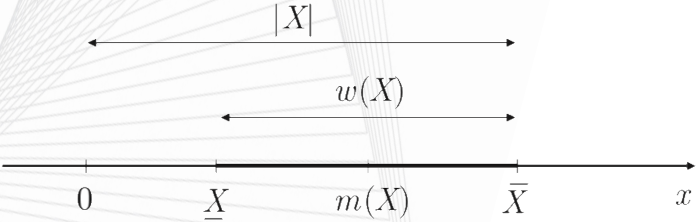

Figure 2 : Width, absolute value, and midpoint of an interval.

UNIVERSIDAD EAFIT

Definition of Arithmetic Operations

Let $\odot \in \{+, -, \cdot, /\}$ be a binary operation in the real numbers, e.g., addition, substraction, multiplication and division.

$$X \odot Y := \{x \odot y : x \in X,\, y \in Y\}$$

In order to simplify notation, the interval $[x, x]$ will be referred as the real number x itself, whenever the context is clear.

UNIVERSIDAD EAFIT

Endpoint Formulas for the Arithmetic Operations

Let X, $Y \in I(\mathbb{R})$. It can be shown that:

1. $X + Y = \left[X^L + Y^L, X^U + Y^U\right]$ Example
2. $-Y = \left[-Y^U, -Y^L\right]$ Example
3. $X - Y = X + (-Y) = \left[X^L - Y^U, X^U - Y^L\right]$ Example
4. $kX = \left[kX^L, kX^U\right]$ Example
5. $XY = [\min S, \max S]$, where
   $S = \left\{X^L Y^L, X^L Y^U, X^U Y^L, X^U Y^U\right\}$ Example
6. $1/Y = \left[1/Y^U, 1/Y^L\right]$ Example
7. $X/Y = X \cdot (1/Y)$ Example

UNIVERSIDAD EAFIT

Hukuhara Difference

Difference

Let $X = [X^L, X^U]$ and $Y = [Y^L, Y^U]$ be two closed intervals in $\mathbb{R}$. If $X^L - Y^L \leq X^U - Y^U$, then the *Hukuhara difference* $Z = X \ominus Y$ exists and $Z = [Z^L, Z^U] = [X^L - Y^L, X^U - Y^U]$. Example

Note

The usual substraction and the Hukuhara difference betwen two intervals need not be the same:

$$[X^L - Y^U, X^U - Y^L] = X - Y \neq X \ominus Y = [X^L - Y^L, X^U - Y^U]$$

UNIVERSIDAD EAFIT

Hausdorff Metric

Let $X,\ Y \subseteq \mathbb{R}^n$. Then the Hausdorff metric between X and Y is defined by

$$d_H(X, Y) = \max\left\{\sup_{x\in X}\inf_{y\in Y}\|x-y\|, \sup_{y\in Y}\inf_{x\in X}\|x-y\|\right\}$$

where $\|\cdot\|$ is a norm in $\mathbb{R}^n$.

If $X = [X^L, X^U]$ and $Y = [Y^L, Y^U]$ are two closed intervals in $\mathbb{R}$, it is not hard to see that

$$d_H(X, Y) = \max\left\{|X^L - Y^L|, |X^U - Y^U|\right\}$$

Convergence

Convergence in $I(\mathbb{R})$

Let $\{X_n\}$ and $X \in I(\mathbb{R})$. We say that the sequence of intervals $\{X_n\}$ converges to X, denoted by $\lim_{n\to\infty} X_n = X$, if, for every $\epsilon > 0$, there exists $N \in \mathbb{N}$, such that, for $n \geq N$, we have $d_H(X_n, X) < \epsilon$.

Lemma

$$\lim_{n\to\infty} X_n = X \text{ if and only if } X_n^L \to X^L \wedge X_n^U \to X^U$$

Functions in $I(\mathbb{R})$ (I)

Interval-valued Function

The function $f : \mathbb{R}^n \to I(\mathbb{R})$ defined on an Eucliden space $\mathbb{R}^n$ is called an interval-valued function. This function can also be written as $f(\mathbb{x}) = [f^L(\mathbb{x}), f^U(\mathbb{x})]$, where f^L and f^U are real-valued functions defined on $\mathbb{R}^n$ and satisfy $f^L(\mathbb{x}) \leq f^U(\mathbb{x})$ for every $\mathbb{x} \in \mathbb{R}^n$.

Limit of a Function

For $\mathbb{c} \in \mathbb{R}^n$ we write $\lim_{\mathbb{x} \to \mathbb{c}} f(\mathbb{x}) = X$ if, for every $\epsilon > 0$, there exists $\delta > 0$ such that, for $\|\mathbb{x} - \mathbb{c}\| < \delta$, we have $d_H(f(\mathbb{x}), X) < \epsilon$.

Functions in $I(\mathbb{R})$ (II)

Lemma

Let f be an interval-valued function defined on $\mathbb{R}^n$ and $X = [X^L, X^U]$ be an interval in $\mathbb{R}$. Then $\lim_{\mathbb{x}\to\mathbb{c}} f(\mathbb{x}) = X$ if and only if $\lim_{\mathbb{x}\to\mathbb{c}} f^L(\mathbb{x}) = X^L$ and $\lim_{\mathbb{x}\to\mathbb{c}} f^U(\mathbb{x}) = X^U$.

Continuity

Let f be an interval-valued function defined on $\mathbb{R}^n$. We say that f is continuous at $\mathbb{c} \in \mathbb{R}^n$ if

$$\lim_{\mathbb{x}\to\mathbb{c}} f(\mathbb{x}) = f(\mathbb{c})$$

Example

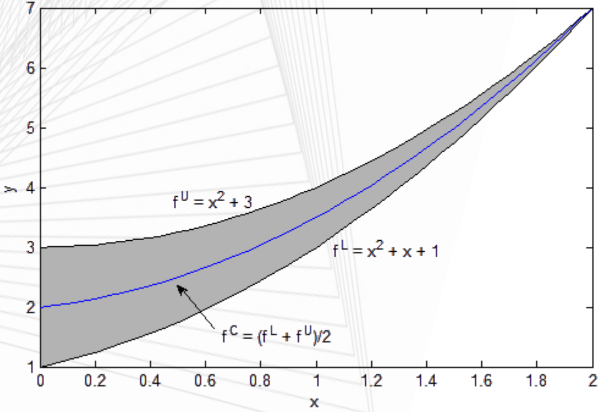

Figure 3 : Graphic representation $f(x) = [x^2 + x + 1,\ x^2 + 3]$.

Optimization Problems

Problem (RVOP)

$$\min \quad f(\mathbb{x}) = f(x_1, ..., x_n)$$
$$\text{subject to} \quad g_i(\mathbb{x}) \leq 0$$

Problem (IVOP)

$$\min \ f(\mathbb{x}) = [f^L(x_1, ..., x_n), f^U(x_1, ..., x_n)] = [f^L(\mathbb{x}), f^U(\mathbb{x})]$$
$$\text{subject to} \quad g_i(\mathbb{x}) \leq 0$$

Order Relations

Let $X = [X^L, X^U]$ and $Y = [Y^L, Y^U] \in I(\mathbb{R})$. It is possible to express X as a function of its center and width, as $X = \langle m(X), w(X) \rangle$.

Order Relations

$$X \preceq_{LU} Y \text{ if and only if } X^L \leq Y^L \text{ and } X^U \leq Y^U$$

$$X \preceq_{CW} Y \text{ if and only if } m(X) \leq m(Y) \text{ and } w(X) \leq w(Y)$$

$$X \preceq_{UC} Y \text{ if and only if } X^U \leq Y^U \text{ and } m(X) \leq m(Y)$$

Solution Types

Type-I

Let $\mathbb{x}^*$ be a feasible solution, i.e., $\mathbb{x}^* \in X$. We say that $\mathbb{x}^*$ is a *type-I solution* of problem (IVOP) if there exists no $\overline{\mathbb{x}} \in X$ such that $f(\overline{\mathbb{x}}) \prec_{LU} f(\mathbb{x}^*)$.

Type-II

Let $\mathbb{x}^*$ be a feasible solution, i.e., $\mathbb{x}^* \in X$. We say that $\mathbb{x}^*$ is a *type-II solution* of problem (IVOP) if there exists no $\overline{\mathbb{x}} \in X$ such that $f(\overline{\mathbb{x}}) \prec_{LU} f(\mathbb{x}^*)$ or $f(\overline{\mathbb{x}}) \prec_{CW} f(\mathbb{x}^*)$.

UNIVERSIDAD EAFIT

KKT Conditions - Real Case

Theorem

Assume that the constraint functions $g_i : \mathbb{R}^n \to \mathbb{R}$ are convex on $\mathbb{R}^n$ for $i = 1, \ldots, m$. Let $X = \{\mathbb{x} \in \mathbb{R}^n : g_i(\mathbb{x}) \leq 0, i = 1, \ldots, m$ be a feasible set and a point $\mathbb{x}^* \in X$. Suppose that the objective function $f : \mathbb{R}^n \to \mathbb{R}$ is convex at $\mathbb{x}^*$, and f, g_i, $i = 1, \ldots m$, are continuously differentiable at $\mathbb{x}^*$. If there exist (Langrange) multipliers $0 \leq \mu_i \in \mathbb{R}$, $i = 1, \ldots, m$, such that

1. $\nabla f(\mathbb{x}^*) + \sum_{i=1}^{m} \mu_i \nabla g_i(\mathbb{x}^*) = \mathbb{0}$
2. $\mu_i g_i(\mathbb{x}^*) = 0$ for all $i = 1, \ldots, m$.

then $\mathbb{x}^*$ is an optimal solution of problem (RVOP).

KKT Conditions - Interval-valued Case

Theorem

Suppose that the the real-valued constraint functions g_i, $i = 1, \ldots, m$, of problem (IVOP) satisfy the KKT assumptions at $\mathbb{x}^*$ and the interval-valued objective function $f : \mathbb{R}^n \to I(\mathbb{R})$ is LU-convex and weakly continuously differentiable at $\mathbb{x}^*$. If there exist (Langrange) multipliers $0 < \lambda^L, \lambda^U \in R$ and $0 \leq \mu_i \in \mathbb{R}$, $i = 1, \ldots, m$, such that

1. $\lambda^L \nabla f^L(\mathbb{x}^*) + \lambda^U \nabla f^U(\mathbb{x}^*) + \sum_{i=1}^{m} \mu_i \nabla g_i(\mathbb{x}^*) = \mathbb{0}$
2. $\mu_i g_i(\mathbb{x}^*) = 0$ for all $i = 1, \ldots, m$.

then $\mathbb{x}^*$ is a type-I and type-II solution of problem (IVOP).

Interval-Valued Polynomial

Interval-Valued Polynomial

Let $c_i = [c_i^L, c_i^U] \in I(\mathbb{R})$ for $i \in \mathbb{N}$. We say $p(x)$ is an interval-valued polynomial if it can be expressed in the form

$$p(x) = \sum_{i=0}^{n} c_i \cdot x^i = \sum_{i=0}^{n} [c_i^L, c_i^U] \cdot x^i$$

UNIVERSIDAD EAFIT

Matrix Representation

Vandermonde Matrix

Let $c_i = [c_i^L, c_i^U] \in I(\mathbb{R})$ for $i \in \{1, ..., n\}$.

$$\begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} = \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^n \\ 1 & x_2 & x_2^2 & \cdots & x_2^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_m & x_m^2 & \cdots & x_m^n \end{bmatrix} \begin{bmatrix} c_0 \\ \vdots \\ c_n \end{bmatrix} + \begin{bmatrix} \varepsilon_0 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

$$\mathbb{Y} = \mathbb{V}\mathbb{C} + \mathbb{E}$$

In this case, $\mathbb{V}$ is called a Vandermonde matrix.

Example

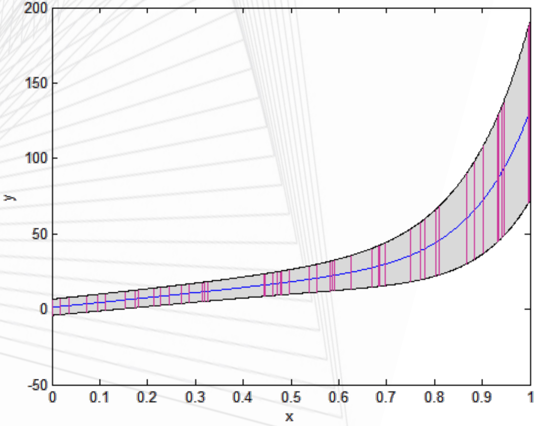

Figure 4 : Interval-valued polynomial graphic.

UNIVERSIDAD EAFIT

What do we look for?

In a nutshell

Find a parameter configuration that reduces at most as possible the discrepancies between the observed data and the information provided by the model proposed.

$$\min \sum_{i=1}^{m} [m(y_i) - m(\widehat{y}_i)]^2$$

ℓ_2 Norm - Least Squares

$$\min \sum_{i=1}^{m} d_H(y_i, \widehat{y}_i)$$

ℓ_1 Norm - Least Absolute Values

ℓ_2 Norm - Least Squares Estimation

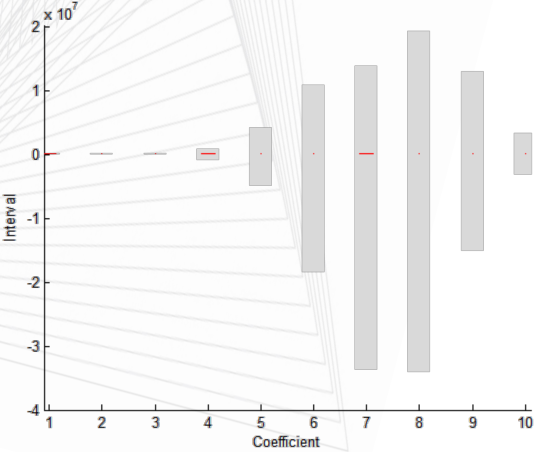

Figure 5 : Parameter estimation result using OLS.

ℓ_1 Norm - Heuristic

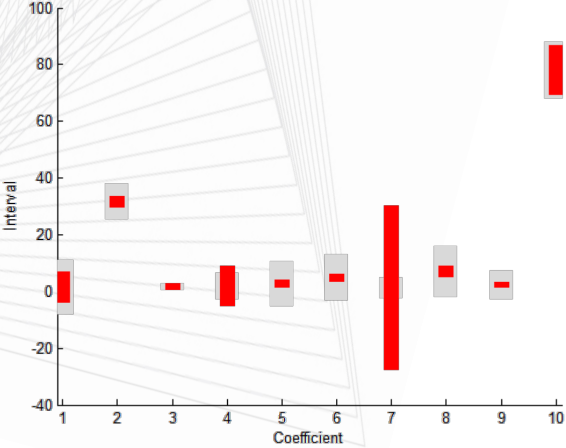

Figure 6 : Parameter estimation result using Differential Evolution.

ℓ_1 Norm - CVX

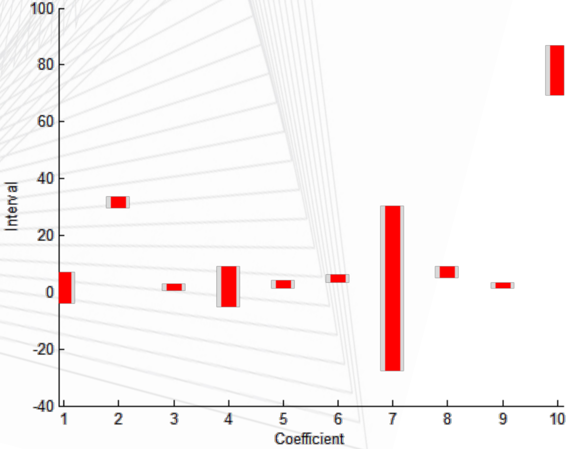

Figure 7 : Parameter estimation result using CVX .

Next Challenges

Experimental Application

$$\frac{\partial^2 u}{\partial t^2} = c^2 \nabla^2 ut$$

What is required?

- Experimental Data
- Numerical Methods - Computational Implementation
- Estimates Validation

References (I)

1 R. Moore, R. Kearfott and M. Cloud, *Introduction to interval analysis.* Philadelphia, PA: Society for Industrial and Applied Mathematics, 2009.

2 H. Wu, 'The Karush–Kuhn–Tucker optimality conditions in an optimization problem with interval-valued objective function', *European Journal of Operational Research*, vol. 176, no. 1, pp. 46-59, 2007.

3 A. Tarantola, *Inverse problem theory and methods for model parameter estimation.* Philadelphia, PA: Society for Industrial and Applied Mathematics, 2005.

UNIVERSIDAD EAFIT

References (II)

4 Michael Grant and Stephen Boyd. Graph implementations for nonsmooth convex programs, Recent Advances in Learning and Control (a tribute to M. Vidyasagar), V. Blondel, S. Boyd, and H. Kimura, editors, pages 95-110, Lecture Notes in Control and Information Sciences, Springer, 2008. http://stanford.edu/~boyd/graph_dcp.html.

5 Michael Grant and Stephen Boyd. CVX: Matlab software for disciplined convex programming, version 2.0 beta. http://cvxr.com/cvx, September 2013.

UNIVERSIDAD EAFIT

Addition - Example

Addition

$$X = [1, 2] \qquad Y = [-4, 5]$$

$$X + Y = [1,\ 2] + [-4,\ 5] = [1 + (-4),\ 2 + 5] = [-3,\ 7]$$

Back to Operations

Negative - Example

Negative

$$X = [-5, 2]$$

$$-X = [-2, -(-5)] = [-2, 5]$$

Back to Operations

Substraction - Example

Substraction

$$X = [-5, 2] \qquad Y = [-1, 9]$$

$$X + (-Y) = [-5,\ 2] + [-9,\ 1] = [-14,\ 3]$$

Back to Operations

Scalar Multiplication - Example

Scalar Multiplication (I)

$$X = [-5, 2] \qquad k = 3$$

$$3X = [-5 \cdot 3,\ 2 \cdot 3] = [-15,\ 6]$$

Back to Operations

Scalar Multiplication (II)

$$X = [-5, 2] \qquad k = -8$$

$$-8X = [2 \cdot -8,\ -5 \cdot -8] = [-16,\ 40]$$

Back to Operations

UNIVERSIDAD EAFIT

Product - Example

Product

$$X = [-5, 2] \qquad Y = [-1, 9]$$

$$S = \{(-5)(-1),\ (-5)(9),\ (2)(-1),\ (2)(9)\}$$

$$XY = [\min S,\ \max S] = [-45,\ 18]$$

Back to Operations

Multiplicative Inverse - Example

Multiplicative Inverse (I)

$$X = [2, 8] \to \frac{1}{X} = \left[\frac{1}{8}, \frac{1}{2}\right]$$

Back to Operations

Multiplicative Inverse (II)

$$X = [-1, 5]$$

$$\frac{1}{X} = \left\{\frac{1}{x} : x \in X\right\} = (-\infty, \infty)$$

Back to Operations

UNIVERSIDAD EAFIT

Division - Example

Division

$$X = [-5, 2] \qquad Y = [3, 7]$$

$$\frac{1}{Y} = \left[\frac{1}{7}, \frac{1}{3}\right]$$

$$S = \left\{(-5)\left(\frac{1}{7}\right), (-5)\left(\frac{1}{3}\right), (2)\left(\frac{1}{7}\right), (2)\left(\frac{1}{3}\right)\right\}$$

$$\frac{X}{Y} = [\min S, \max S] = \left[-\frac{5}{3}, \frac{2}{3}\right]$$

Back to Operations

Hukuhara Difference - Example

Hukuhara Difference

$$X = [-5, 2] \qquad Y = [-1, 3]$$

$$X \ominus Y = [-5,\ 2] \ominus [-1,\ 3] = [-5 - (-1),\ 2 - 3]$$

$$X \ominus Y = [-4,\ -1]$$

Back to H-Difference

Additive Inverses - Example

Additive Inverses

$$X = [-5, 2]$$

Usual Difference

$$X - X = [-5,\ 2] - [-5,\ 2] = [-5,\ 2] + [-2,\ 5]$$

$$X - X = [-7,\ 7] = 7[-1, 1] \ni [0, 0]$$

Hukuhara Difference

$$X \ominus X = [-5,\ 2] \ominus [-5,\ 2] = [-5 - (-5),\ 2 - 2] = [0, 0]$$

Back to main

Multiplicative Inverses - Example

Multiplicative Inverses

$$X = [3, 7]$$

$$\frac{1}{X} = \left[\frac{1}{7}, \frac{1}{3}\right]$$

$$S = \left\{(3)\left(\frac{1}{7}\right), (3)\left(\frac{1}{3}\right), (7)\left(\frac{1}{7}\right), (7)\left(\frac{1}{3}\right)\right\}$$

$$\frac{X}{X} = [\min S, \max S] = \left[\frac{3}{7}, \frac{7}{3}\right] \ni [1, 1]$$

Back to main