Interval analysis for the treatment of uncertainty in epidemiological models based on systems of ordinary differential equations

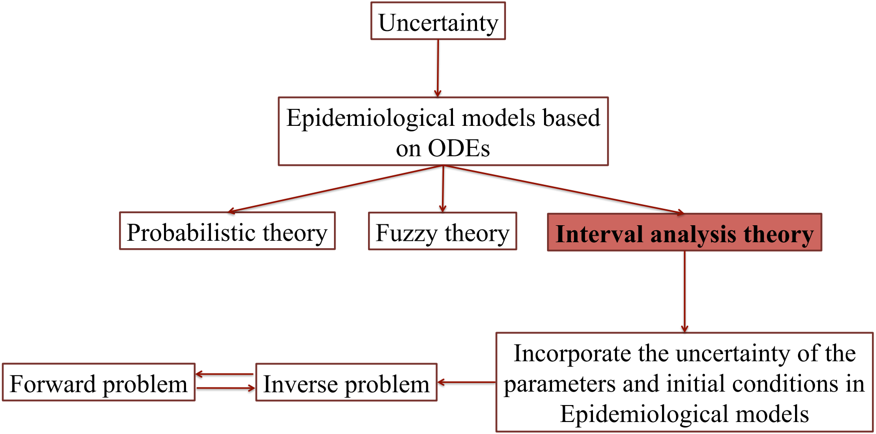
Paola Lizarralde-Bejarano
PhD Student in Mathematical Engineering

Supervisor: María Eugenia Puerta
Co-supervisor: Sair Arboleda-Sánchez

Doctoral Seminar 4
Universidad EAFIT
Medellín, Colombia
18 May, 2018

Objective

Review the algorithms available in the literature and determining a strategy to solve numerically epidemiological models based on ODEs where both the parameters and initial conditions are closed real intervals as a strategy to deal with the uncertainty of these.

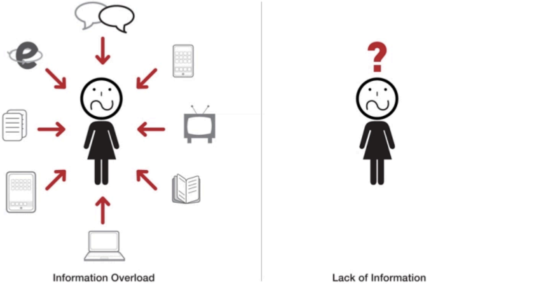

Uncertainty
Epidemiological models based on ODEs
Probabilistic theory
Fuzzy theory
Interval analysis theory
Incorporate the uncertainty of the parameters and initial conditions in Epidemiological models
Inverse problem
Forward problem

Uncertainty

Uncertainty is present in any process of measuring and obtaining information that is required to explain a real phenomenon. In many sciences it is possible to conduct experiments to obtain information and test hypotheses. Experiments with the spread of infectious diseases in human populations are often impossible, unethical or expensive (Hethcote, 2009).

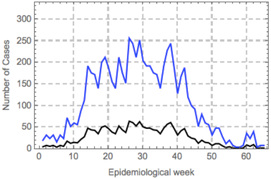
Causes of uncertainty

- Lack of information
- Conflict evidence
- Ambiguity
- Measurement
- Belief

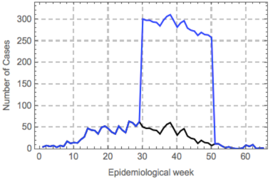


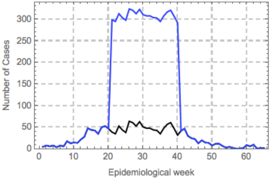
Uncertainty in epidemiological models

Uncertainty in experimental data



Uncertainty in dengue cases reported





Ross-McDonald model

$$\frac{dx}{dt} = \beta y(1-x) - \mu x$$
$$\frac{dy}{dt} = \alpha x(1-y) - \gamma y$$

SIR model

$$\frac{dS}{dt} = -\beta \frac{I}{H} S$$
$$\frac{dI}{dt} = \beta \frac{I}{H} S - \gamma I$$
$$\frac{dR}{dt} = \gamma I$$

Metapopulation model

$$S_1' = -\beta_1 p_{11} S_1 \left[p11 \frac{I_1}{N_1} + p_{21} \frac{I_2}{N_2} \right] - \beta_2 p_{12} S_1 \left[p12 \frac{I_1}{N_1} + p_{22} \frac{I_2}{N_2} \right]$$
$$I_1' = \beta_1 p_{11} S_1 \left[p11 \frac{I_1}{N_1} + p_{21} \frac{I_2}{N_2} \right] + \beta_2 p_{12} S_1 \left[p12 \frac{I_1}{N_1} + p_{22} \frac{I_2}{N_2} \right] - \gamma_1 I_1$$
$$S_2' = -\beta_1 p_{21} S_2 \left[p11 \frac{I_1}{N_1} + p_{21} \frac{I_2}{N_2} \right] - \beta_2 p_{22} S_2 \left[p12 \frac{I_1}{N_1} + p_{22} \frac{I_2}{N_2} \right]$$
$$I_2' = \beta_1 p_{21} S_2 \left[p11 \frac{I_1}{N_1} + p_{21} \frac{I_2}{N_2} \right] + \beta_2 p_{22} S_2 \left[p12 \frac{I_1}{N_1} + p_{22} \frac{I_2}{N_2} \right] - \gamma_2 I_2$$

How has the uncertainty been considered in epidemiological models?

Probability theory

- In (Luz et al. 2003) assumed that parameters as duration of infectious period in humans, biting rate, mosquito to human transmission and human to mosquito transmission follows a uniform distribution, while the extrinsic incubation period follows a triangular distribution, this information was obtained from several works that conducted experiments with different vector populations.
- In (Britton and Lindenstrand, 2009) it is assumed that latent and infection periods are random and independent with the gamma distribution.

How has the uncertainty been considered in epidemiological models?

Fuzzy theory

In (Barros et al. 2001) an SIS model was formulated

$$\frac{dS}{dt} = -\beta SI + \gamma I$$
$$\frac{dI}{dt} = \beta SI - \gamma I$$

where β, and γ are given by functions that depend on the amount of virus, v.

$$\beta(v) = \begin{cases} 1, & \text{if } v_M < v \leq v_{max}, \\ \frac{v - v_{min}}{v_M - v_{min}}, & \text{if } v_{min} < v \leq v_M \\ 0, & \text{otherwise.} \end{cases}, \ \gamma(v) = \frac{(\gamma_0 - 1)}{v_{max}} v + 1$$

Motivation of interval arithmetic

- The notion of irrational number entails some process of approximation from above and below. Archimedes (287-212 BCE) was able to bracket π by taking a circle and considering inscribed and circumscribed polygons (Moore and Lodwick, 2003).
- The purpose of interval analysis is to provide upper and lower bounds on the effect all errors and uncertainties have on computed quantity (Hansen and Walster, 2003).
- Interval analysis began as a tool for bounding rounding errors.

Some history of interval analysis

Ramon E. Moore conceived interval arithmetic in 1957, while an employee of Lockheed Missiles and Space Co. Inc., as an approach to bound rounding errors in mathematical computation (Hijazi et al. 2008).

Ramon E. Moore (1929–2015)

Applications of interval analysis

- Practical application areas include chemical and structural engineering, economics, control circuitry design, beam physics, global optimization, constraint satisfaction, asteroid orbits, robotics, signal processing, computer graphics, and behavioral ecology (Moore and Lodwick, 2003).
- Interval analysis has been used in rigorous computer-assisted proofs, for example, Hales' proof on the Kepler conjecture.
- An interval Newton method has been developed for solving systems of nonlinear equations.
- Interval methods can bound the solutions of linear systems with inexact data.
- There are rigorous interval branch-and-bound methods for global optimization.

Interval arithmetic

The set of intervals on the real line $\mathbb{R}$ is defined by

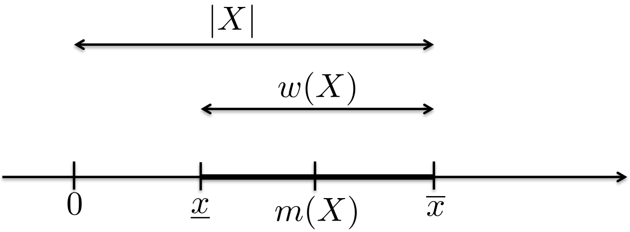
$$\mathcal{I} = \{\, X = [\underline{x}, \overline{x}] \,|\, \underline{x}, \overline{x} \in \mathbb{R},\ \underline{x} \leq \overline{x} \,\}. \tag{1}$$

Observations

- We say an interval X is degenerate if $\underline{x} = \overline{x}$.
- If $\underline{x} = -\overline{x}$ then X is symmetric.
- Two intervals X and Y are equal if $\underline{x} = \underline{y}$ and $\overline{x} = \overline{y}$.

Let $[X] = [\underline{x}, \overline{x}]$. We define the following quantities for intervals:

- **Width:** $w(X) = \overline{x} - \underline{x}$
- **Magnitude:** $|X| = \max\{|\underline{x}|, |\overline{x}|\}$
- **Midpoint:** $m(X) = \frac{1}{2}(\overline{x} + \underline{x})$



Interval-arithmetic operations

Let $X = [\underline{x}, \overline{x}]$ and $Y = [\underline{y}, \overline{y}]$

1. $X + Y = [\underline{x} + \underline{y}, \overline{x} + \overline{y}]$
2. $-X = [-\overline{x}, -\underline{x}]$
3. $X - Y = X + (-Y) = [\underline{x} - \overline{y}, \overline{x} - \underline{y}]$
4. $X \cdot Y = [\min\{S\}, \max\{S\}]$, where $S = \{\underline{x}\underline{y}, \underline{x}\overline{y}, \overline{x}\underline{y}, \overline{x}\overline{y}\}$
5. $1/Y = [1/\overline{y}, 1/\underline{y}]$, $0 \notin Y$
6. $X/Y = X \cdot 1/Y = \{x/y \,|\, x \in X,\ y \in Y\}$, $0 \notin Y$.

Examples

Consider $X = [-2, 1]$, $Y = [1, 3]$ and $Z = [2, 5]$

1. $X + Y = [-2 + 1, 1 + 3] = [-1, 4]$
2. $-Y = [-3, -1]$
3. $X - Y = [-2, 1] + [-3, -1] = [-5, 0]$
4. $X \cdot Y = [-6, 3]$, where $S = \{-2, -6, 1, 3\}$
5. $1/Z = \left[\frac{1}{5}, \frac{1}{2}\right]$
6. $Y/Z = Y \cdot 1/Z = [\frac{1}{5}, \frac{3}{2}]$, where $S = \{\frac{1}{5}, \frac{3}{5}, \frac{1}{2}, \frac{3}{2}\}$

Properties of interval-arithmetic operations

- Commutativity and asociativity of the sum operation and multiplication.
- Identity element of the sum $[0, 0]$ y $[1, 1]$ for multiplication.
- There is no inverse elements for both operations.
- Subdistributy property, $X(Y + Z) \subseteq XY + XZ$

Interval vectors and matrices

An *interval vector* is a vector with interval components. An *interval matrix* is a matrix with interval components. Let be $A \in \mathcal{I}^{n \times n}$ an interval matrix with elements A_{ij}.

- **Matrix norm:** $||A|| = \max_i \sum_j |A_{ij}|$.
- **Width:** $w(A) = \max_{i,j} w(A_{ij})$.
- **Midpoint:** $(m(A))_{ij} = m(A_{ij})$.

Hausdorff metric

Let $X,\ Y \in \mathbb{R}^n$. 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 $X = [\underline{x}, \overline{x}]$, and $Y = [\underline{y}, \overline{y}]$ the distance between two intervals is given by

$$d_H(X, Y) = \max\{|\underline{x} - \underline{y}|, |\overline{x} - \overline{y}|\} \tag{2}$$

Finally, if $X,\ Y \in \mathcal{I}^n$ are two interval vectors, the distance is given by

$$d_H(X, Y) = \max_{1\le i\le n}\{\, d_H(X_i,\ Y_i)\,\} \tag{3}$$

Interval-valued function

Given a function $f = f(x_1, \ldots, x_n)$ of several variables, we will wish to find the image set

$$f(X_1, \ldots, X_n) = \{f(x_1, \ldots, x_n) \,|\, x_1 \in X_1, \ldots, x_n \in X_n\},$$

where $X_1, \ldots, X_n$ are specified intervals.

Definition

We say that F is an *interval extension* of f, if an interval-valued function F of n interval variables $X_1, \ldots, X_n$ such that for real arguments $x_1, \ldots, x_n$ we have

$$F(x_1, \ldots, x_n) = f(x_1, \ldots, x_n)$$

Interval-valued function

Definition

We say that $F = F(X_1, \ldots, X_n)$ is *inclusion isotonic* if

$$Y_i \subseteq X_i \text{ for } i = 1, \ldots, n \text{ then } F(Y_1, \ldots, Y_n) \subseteq F(X_1, \ldots X_n). \quad (4)$$

Theorem

If F is an inclusion isotonic interval extension of f, then

$$f(X_1, \ldots, X_n) \subseteq F(X_1, \ldots, X_n) \quad (5)$$

Integration

Definition

We define the interval integral

$$\int_{[a,b]} F(t)dt = \left[\int_{[a,b]} \underline{F}(t)dt,\ \int_{[a,b]} \overline{F}(t)dt,\right]$$

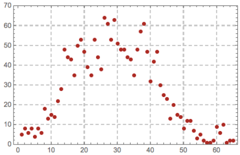
where $F(t) = [\,\underline{F}(t),\ \overline{F}(t)\,]$.

Inclusion property

If $F(t) \subseteq G(t)$ for all $t \in [a, b]$, then, $\int_{[a,b]} F(t)dt \subseteq \int_{[a,b]} G(t)dt.$

Inverse Problem

Observations of data



Quantitative model

$$\frac{dS}{dt} = \mu N - \beta S I - \mu S$$

$$\frac{dI}{dt} = \beta S I - (\gamma + \mu) I$$

$$\frac{dR}{dt} = \gamma I - \mu R$$

Estimates of model parameters

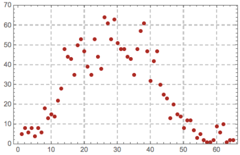
$\mu = ?$

$\beta = ?$

$\gamma = ?$

Inverse Problem

Observations of data



Quantitative model

$$\frac{dS}{dt} = \mu N - \beta SI - \mu S$$
$$\frac{dI}{dt} = \beta SI - (\gamma + \mu)I$$
$$\frac{dR}{dt} = \gamma I - \mu R$$

Estimates of model parameters

$$\mu = ?$$
$$\beta = ?$$
$$\gamma = ?$$

The optimization problem that we want to solve is given by

$$\min_{\mathbf{P} \in \mathcal{I}^k} \; d_H(x_i(t; \mathbf{P}), \; \mathbf{D}_i(t))$$
$$\text{s.t } \dot{x}(t) = F(x(t), \mathbf{P}, t), \quad t \in [t_0, T]$$
$$x(t_0) = \mathbf{X}_0$$

where $\mathbf{P}$ is the interval parameter vector, $\mathbf{X}_0$ is the interval vector of initial conditions, and $\mathbf{D}_i$ is the interval of data observed.

Forward Problem

Estimates of model parameters

$$\mu = c_1$$
$$\beta = c_2$$
$$\gamma = c_3$$

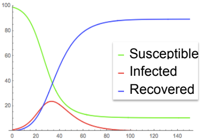
→

Quantitative model

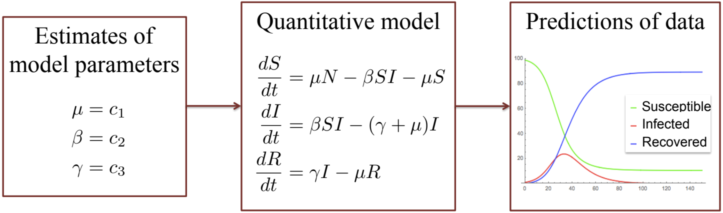
$$\frac{dS}{dt} = \mu N - \beta S I - \mu S$$
$$\frac{dI}{dt} = \beta S I - (\gamma + \mu) I$$
$$\frac{dR}{dt} = \gamma I - \mu R$$

→

Predictions of data



Forward Problem



Obtain a solution of the system of equations given by

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(t; \mathbf{x}; \mathbf{p}), \qquad \mathbf{x}(t_0) = \mathbf{x}_0, \qquad (6)$$

where $\mathbf{x} = (x_1, \ldots, x_n) \in \mathcal{I}^n$ is the vector of n variables, $\mathbf{p} = (p_1, \ldots, p_k) \in \mathcal{I}^k$ is the vector of k parameters, $\mathbf{x}_0 = (x_{1_0}, \ldots, x_{n_0})$ is the vector of initial conditions, $\mathbf{f}$ is an interval-valued vectorial function, and t represents the time.

Taylor coefficients

The ith Taylor coefficient of $u(t)$ evaluated at some point t_j is given by

$$(u_j)_i = \frac{u^{(i)}(t_j)}{i!} \tag{7}$$

where $u^{(i)}(t)$ is the ith derivative of $u(t)$. Consider the autonomous differential system

$$y'(t) = f(y), \;\; y(t_j) = y_j \tag{8}$$

The Taylor coefficients of $y(t)$ at t_j satisfy

$$\begin{aligned} (y_j)_0 &= y_j, \\ (y_j)_1 &= f^{[1]}(y_j) = f(y_j), \\ (y_j)_i &= \frac{1}{i}\left(\frac{\partial f^{[i-1]}}{\partial y} f\right)(y_j) \text{ for } i \geq 2 \end{aligned}$$

Taylor coefficients

In this way, the approximation of degree N by Taylor series for the function $y(t)$ is given by:

$$y(t) = \sum_{i=0}^{N-1} (y)_i (t - t_0)^i + R_N([\, t_0 \,, t \,]) \qquad (9)$$

where $R_N([\, t_0 \,, t \,]) = (y)_N(s)(t - t_0)^N$ for all $s \in [\, t_0 \,, t \,]$.

Taylor coefficients

If we consider the initial value problem given by an interval, Y_j

$$y'(t) = f(y), \qquad\qquad y(t_j) = y_j \in Y_j$$

Taylor coefficients would be given by:

$$\begin{aligned}
(Y_j)_0 &= Y_j \\
(Y_j)_1 &= f^{[1]}(Y_j) = f(Y_j) \\
(Y_j)_i &= f^{[i]}(Y_j) = \frac{1}{i}\left(\frac{\partial f^{(i-1)}}{\partial y} f\right)(Y_j) \qquad\qquad \text{for } i \geq 2
\end{aligned}$$

and therefore we could construct (9) for functions defined in $\mathcal{I}$.

ODEs and Initial Value Problem

Consider the integral equation

$$y(t) = y_0 + \int_0^t f(s, y(s))ds \tag{10}$$

which is formally equivalent to the *initial value problem* for the ODE

$$\begin{aligned} y'(t) &= f(t, y(t)) \\ y(0) &= y_0 \end{aligned} \tag{11}$$

We define an operator $p(y)(t) = y_0 + \int_0^t f(s, y(s))ds$. Let the *interval operator* $P : M \rightarrow M$ be defined on class M of interval enclosures of operator p.

ODEs and Initial Value Problem

Theorem

If P is an inclusion isotonic interval of p, and if $P(Y^{(0)}) \subseteq Y^{(0)}$, then the sequence defined by

$$Y^{(k+1)} = P(Y^{(k)}), \quad (k = 0, 1, 2, \dots) \tag{12}$$

has the following properties

(1) $Y^{(k+1)} \subseteq Y^{(k)}$, $k = 0, 1, 2, \dots$

(2) *For every $a \leq t \leq b$, the limit*

$$Y(t) = \bigcap_{k=0}^{\infty} Y^{(k)}(t) \tag{13}$$

exists as an interval $Y(t) \subseteq Y^{(k)}(t)$, $k = 0, 1, 2, \dots$.

ODEs and Initial Value Problem

(3) Any solution of (10) which is in $Y^{(0)}$ is also in $Y^{(k)}$ for all k and in Y as well. That is, if $y(t) \in Y^{(0)}(t)$ for all $a \leq t \leq b$, then $y(t) \in Y^{(k)}(t)$ for all k and all $a \leq t \leq b$.

(4) If there is a real number c such that $0 \leq c < 1$, for which $X \subseteq Y_{(0)}$ implies

$$\sup_t w(P(X)(t)) \leq c \sup_t w(X(t)), \quad a \leq t \leq b. \tag{14}$$

for every $X \in M$, then (10) has the unique solution $Y(t)$ in $Y^{(0)}$ given by (13).

ODEs and Initial Value Problem

In (Nedialkov, Jackson, and Corliss, 1999) is considered the problem

$$\begin{aligned} y'(t) &= f(y) \\ y(t_0) &\in Y_0 \end{aligned} \tag{15}$$

where $t \in [t_0, T]$. We consider a grid $t_0 < t_1 < \cdots < t_m = T$, and denote the stepsize from t_j to t_{j+1} by h_j. We denote a solution of (15) with an initial condition y_j at t_j by $y(t; t_j, y_j)$. In most of methods to solve this problem, each integration step consists of two phases:

- Validating existence and uniqueness.
- Computing a tighter enclosure

ODEs and Initial Value Problem

In (Lin and Stadtherr, 2007) for the system

$$y_1' = \theta_1 y_1(1 - y_2), \quad y_1(0) = 1.2, \quad \theta_1 \in 3 + [-0.01, 0.01],$$
$$y_2' = \theta_2 y_2(y_1 - 1), \quad y_2(0) = 1.1, \quad \theta_2 \in 1 + [-0.01, 0.01],$$

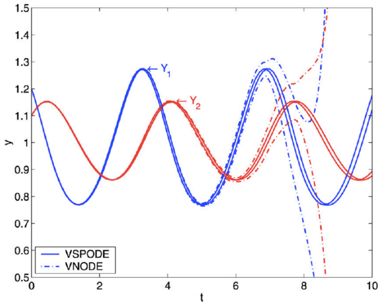


Fig. 1. Solution enclosure of Lotka–Volterra equations computed using VSPODE and VNODE. Solid curves show the enclosures of the state variables as determined by VSPODE, and the dashed curves show the enclosures determined by VNODE.

On the other hand, seeking to provide the space $\mathcal{I}$ for a richer structure in (Hukuhara, 1967), the following definition of the difference between intervals was formulated.

Definition

Let $X, Y \in \mathcal{I}$ such that $X = [\,\underline{x}, \, \overline{x}\,]$ and $Y = [\,\underline{y}, \, \overline{y}\,]$. If $\underline{x} - \underline{y} \leq \overline{x} - \overline{y}$, then the *Hukuhara difference* exists and is given by

$$Z = X \ominus Y = [\underline{x} - \underline{y}, \, \overline{x} - \overline{y}].$$

If we consider the space $\mathcal{I}$ with Hukuhara's difference, we would have the additive inverses for each interval, however this difference does not exist if the condition $\underline{x} - \underline{y} \leq \overline{x} - \overline{y}$ is not satisfy.

Definition

(Stefanini and Bede, 2009; Stefanini, 2010) Let $X, Y \in \mathcal{I}$ such that $X = [\underline{x}, \overline{x}]$ and $Y = [\underline{y}, \overline{y}]$ be two intervals; the gHdifference is

$$[\underline{x}, \overline{x}] \ominus_g [\underline{y}, \overline{y}] = [\underline{z}, \overline{z}] \Longleftrightarrow \begin{cases} (i) & \begin{cases} \underline{x} & = \underline{y} + \underline{z}, \\ \overline{x} & = \overline{y} + \overline{z} \end{cases} \\ \text{or } (ii) & \begin{cases} \underline{y} & = \underline{x} - \overline{z}, \\ \overline{y} & = \overline{x} - \underline{z}, \end{cases} \end{cases}$$

so that $[\underline{x}, \overline{x}] \ominus_g [\underline{y}, \overline{y}] = [\underline{z}, \overline{z}]$ is always defined by

$$\underline{c} = \min\{\underline{x} - \underline{y}, \overline{x} - \overline{y}\}, \qquad \overline{c} = \max\{\underline{x} - \underline{y}, \overline{x} - \overline{y}\},$$

i.e.

$$[a, b] \ominus_g [c, d] = [\min\{a - c, b - d\}, \max\{a - c, b - d\}]$$

In (Stefanini and Bede, 2009) is considered the problem

$$y' = f(x, y), \quad y(x_0) = y_0$$

where $f : [a, b] \times \mathcal{I} \to \mathcal{I}$, with $f(x, y) = [\underline{f}(x, y), \overline{f}(x, y)]$ for $y \in \mathcal{I}$, i.e $y = [\underline{y}, \overline{y}]$, and $y_0 = [\underline{y_0}, \overline{y_0}]$.

Theorem

The interval differential equation (36) *is equivalent to the integral equation*

$$y(x) \ominus_g y_0 = \int_{x_0}^{x} f(t, y(t))dt$$

on some interval $[x_0, x_0 + \delta]$.

Conclusions

- The analysis interval theory has taken boom for the handling of uncertainty.
- Several methodologies have been developed to solve differential equations where the initial conditions are defined as an interval. However, there are not many developments about how to solve systems of equations where both the parameters and the initial conditions belong to the space of the intervals.

References

Barros, L.C., R.C. Bassanezi, R. Zotin G. Oliveira, and M.B.F Leite (2001). A disease evolution model with uncertain parameters. In: IFSA World Congress and 20th NAFIPS International Conference, 2001. Joint 9th. Vol. 3. IEEE, pp. 1626–1630 (cit. on p. 10).

Britton, Tom and David Lindenstrand (2009). Epidemic modelling: aspects where stochasticity matters. Mathematical biosciences 222.2, pp. 109–116 (cit. on p. 9).

Hansen, Eldon and G William Walster (2003). Global optimization using interval analysis: revised and expanded. Vol. 264. CRC Press (cit. on p. 11).

Hethcote, Herbert W (2009). The basic epidemiology models: models, expressions for R_0, parameter estimation, and applications. In: Mathematical understanding of infectious disease dynamics. World Scientific, pp. 1–61 (cit. on p. 5).

Hijazi, Younis, Hans Hagen, Charles D Hansen, and Kenneth I Joy (2008). Why interval arithmetic is so useful. Visualization of large and unstructured data sets (cit. on p. 12).

Hukuhara, Masuo (1967). Integration des applications mesurables dont la valeur est un compact convexe. Funkcialaj Ekvacioj 10.3 (cit. on p. 36).

Lin, Youdong and Mark A Stadtherr (2007). Validated solutions of initial value problems for parametric ODEs. Applied Numerical Mathematics 57.10, pp. 1145–1162 (cit. on p. 35).

Luz, Paula Mendes, Cláudia Torres Codeço, Eduardo Massad, and Claudio José Struchiner (2003). Uncertainties regarding dengue modeling in Rio de Janeiro, Brazil. Memórias do Instituto Oswaldo Cruz 98.7, pp. 871–878 (cit. on p. 9).

Moore, Ramon and Weldon Lodwick (2003). Interval analysis and fuzzy set theory. Fuzzy sets and systems 135.1, pp. 5–9 (cit. on pp. 11, 13).

Nedialkov, Nedialko S, Kenneth R Jackson, and George F Corliss (1999). Validated solutions of initial value problems for ordinary differential equations. Applied Mathematics and Computation 105.1, pp. 21–68 (cit. on p. 34).

Stefanini, Luciano (2010). A generalization of Hukuhara difference and division for interval and fuzzy arithmetic. Fuzzy sets and systems 161.11, pp. 1564–1584 (cit. on p. 37).

Stefanini, Luciano and Barnabas Bede (2009). Generalized Hukuhara differentiability of interval-valued functions and interval differential equations. Nonlinear Analysis: Theory, Methods and Applications 71.3-4, pp. 1311–1328 (cit. on pp. 37, 38).