Modeling Uncertainty with Emphasis on Non-Stochastic Approaches

Jan Chleboun

Faculty of Civil Engineering
Czech Technical University
Prague

ISUME 2011, Prague May 2, 2011

Outline

1. A broader context
   - Ultimate goal of modeling
   - Omnipresent uncertainty
2. Stochastic approaches
   - Transformation of a stochastic DE to a deterministic DE
3. Non-stochastic methods
   - The worst (case) scenario method (WSM)
   - Dempster-Shafer theory
   - Fuzzy set theory
4. DE-driven problems: uncertain functions
   - The worst scenario method
   - The WSM and fuzzy sets

Outline

1. A broader context
   - Ultimate goal of modeling
   - Omnipresent uncertainty
2. Stochastic approaches
   - Transformation of a stochastic DE to a deterministic DE
3. Non-stochastic methods
   - The worst (case) scenario method (WSM)
   - Dempster-Shafer theory
   - Fuzzy set theory
4. DE-driven problems: uncertain functions
   - The worst scenario method
   - The WSM and fuzzy sets

Ultimate goal of modeling: guaranteed prediction

Not only an approximate (numerical) solution, but also

1. error caused by the selection of a particular model from hierarchy of relevant mathematical models
2. error caused by the numerical method delivering the approximate solution
3. an assessment of uncertainty in model outputs caused by uncertain input data

Outline

Epistemic and aleatory uncertainty in modeling

Epistemic uncertainty

– the lack of knowledge.

In principle, it can (often) be reduced through improving measuring instruments as well as data collecting and mining.

Aleatory uncertainty

– the inherent variation associated with the modeled system.

Take, for example, the randomness of material parameters, or the variability of the weather.

Consequently, our mathematical models are burdened with uncertainty in input data.

Epistemic and aleatory uncertainty in modeling

Epistemic uncertainty

– the lack of knowledge.

In principle, it can (often) be reduced through improving measuring instruments as well as data collecting and mining.

Aleatory uncertainty

– the inherent variation associated with the modeled system.

Take, for example, the randomness of material parameters, or the variability of the weather.

Consequently, our mathematical models are burdened with uncertainty in input data.

Epistemic and aleatory uncertainty in modeling

Epistemic uncertainty

– the lack of knowledge.

In principle, it can (often) be reduced through improving measuring instruments as well as data collecting and mining.

Aleatory uncertainty

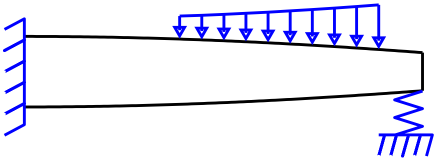
– the inherent variation associated with the modeled system.

Take, for example, the randomness of material parameters, or the variability of the weather.

Consequently, our mathematical models are burdened with uncertainty in input data.

Examples from engineering

A cantilever beam

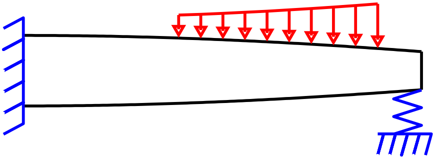


modeled by

$$(cEt^3(x)v''(x))'' = f(x) \text{ on } [a,b], \quad \text{and boundary conditions}$$

Examples from engineering

A cantilever beam



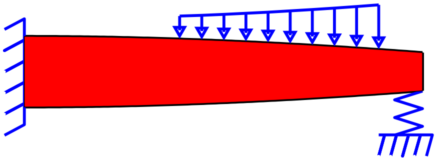
modeled by

$$(cEt^3(x)v''(x))'' = f(x) \text{ on } [a,b], \quad \text{and boundary conditions}$$

where the distributed load f is an uncertain function.

Examples from engineering

A cantilever beam



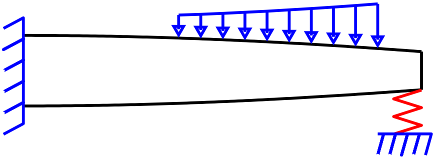
modeled by

$$(c{\color{red}E}t^3(x)v''(x))'' = f(x) \text{ on } [a,b], \quad \text{and boundary conditions}$$

where the Young modulus E is an uncertain scalar parameter (or a function).

Examples from engineering

A cantilever beam



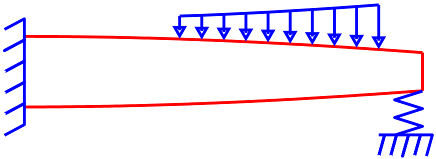
modeled by

$$(cEt^3(x)v''(x))'' = f(x) \text{ on } [a,b], \quad \text{and boundary conditions}$$

where the boundary condition (spring stiffness) is represented by an uncertain scalar parameter.

Examples from engineering

A cantilever beam



modeled by

$$(cEt^3(x)v''(x))'' = f(x) \text{ on } [a, b], \quad \text{and boundary conditions}$$

where the beam shape is determined by an uncertain function t defined on a possibly uncertain interval $[a, b]$.

Basic settings

$\mathcal{U}_{\text{ad}}$ . . . set of admissible parameters

Basic settings

$\mathcal{U}_{\mathrm{ad}}$ . . . set of admissible parameters

Examples

$\mathcal{U}_{\mathrm{ad}}$ is an interval if a scalar parameter is uncertain.

$\mathcal{U}_{\mathrm{ad}}$ is a convex subset of $\mathbb{R}^n$ if a vector is uncertain.

$\mathcal{U}_{\mathrm{ad}}$ is a convex set of functions if a function is uncertain.

Basic settings

$\mathcal{U}_{\text{ad}}$ ... set of admissible parameters

$D(a)u = f$ ... state problem dependent on $a \in \mathcal{U}_{\text{ad}}$
Consequently, its solution $u \equiv u(a)$ also depends on $a \in \mathcal{U}_{\text{ad}}$.

Basic settings

$\mathcal{U}_{\text{ad}}$ . . . set of admissible parameters

$D(a)u = f$ . . . state problem dependent on $a \in \mathcal{U}_{\text{ad}}$
Consequently, its solution $u \equiv u(a)$ also depends on $a \in \mathcal{U}_{\text{ad}}$.

Examples

A boundary value problem for an ordinary or a partial differential equation dependent on a.

An initial value problem dependent on a.

A variational inequality dependent on a (then "$=$" is inappropriate).

Basic settings

$\mathcal{U}_{\text{ad}}$ ... set of admissible parameters

$D(a)u = f$ ... state problem dependent on $a \in \mathcal{U}_{\text{ad}}$
Consequently, its solution $u \equiv u(a)$ also depends on $a \in \mathcal{U}_{\text{ad}}$.

$\Phi(a, u(a))$ or $\Psi(a) \equiv \Phi(a, u(a))$ ... quantity of interest, criterion-functional.

Basic settings

$\mathcal{U}_{\mathrm{ad}}$ ... set of admissible parameters

$D(a)u = f$ ... state problem dependent on $a \in \mathcal{U}_{\mathrm{ad}}$
Consequently, its solution $u \equiv u(a)$ also depends on $a \in \mathcal{U}_{\mathrm{ad}}$.

$\Phi(a, u(a))$ or $\Psi(a) \equiv \Phi(a, u(a))$ ... quantity of interest, criterion-functional.

Examples

Displacement, temperature, local mechanical stress or stress invariants, concentration of chemicals, etc.

Basic settings

$\mathcal{U}_{\text{ad}}$ . . . set of admissible parameters

$D(a)u = f$ . . . state problem dependent on $a \in \mathcal{U}_{\text{ad}}$
Consequently, its solution $u \equiv u(a)$ also depends on $a \in \mathcal{U}_{\text{ad}}$.

$\Phi(a, u(a))$ or $\Psi(a) \equiv \Phi(a, u(a))$ . . . quantity of interest, criterion-functional.

Assumptions:

- $D(a)u = f$ is uniquely solvable for each $a \in \mathcal{U}_{\text{ad}}$.
- Ψ is continuous and bounded on $\mathcal{U}_{\text{ad}}$.

Outline

From a stochastic DE to a deterministic DE

Let us consider a DE-based boundary value problem (BVP) in D dimensions whose parameter is a random function.

It is approximated by the truncated Karhunen-Loève expansion (TKLE) that is determined by N eigenpairs of a compact selfadjoint operator defined through the covariance function of the random function.

By replacing the random function by its TKLE and considering the weak formulation of the BVP, we can infer an (D+N)-dimensional deterministic BVP.

Remark: In practice, the determination of the covariance function of the random function is difficult due to insufficient experimental data. Assumptions and estimates are employed.

Solving multidimensional BVPs

Comments:

- Approximate solution by the FEM tailored to solving multidimensional BVPs.
    - It is said that $\approx 10 - 20$ dimensions are manageable.
- Error estimates and convergence rates estimates are possible.

Bibliography: Babuška, Nobile, Tempone, Webster, Zouraris; Schwab, Todor

Remark: DE burdened with white noise (Brownian motion, Wiener process) – stochastic differential equations and calculus (Itô).

Strength and weakness of stochastic methods

Stochastic methods can deliver extremely important and valuable assessment of uncertainty.

To perform well, they need input data whose probabilistic characteristics (probability distribution function, covariance function) can be and often are difficult to obtain in necessary quality and/or quantity.

Outline

The worst (case) scenario method (WSM)

In practice, the maximum of the criterion-functional over $\mathcal{U}_{\text{ad}}$ is often important: maximum temperature, mechanical stress, etc.

To determine the “worst” scenario (anti-optimization (Elishakoff)), we maximize Ψ by searching for

$$a^0 = \arg\max_{a \in \mathcal{U}_{\text{ad}}} \Psi(a).$$

If also the “best" scenario

$$a_0 = \arg\min_{a \in \mathcal{U}_{\text{ad}}} \Psi(a)$$

is found, then the range of $\Psi|_{\mathcal{U}_{\text{ad}}}$ is given by

$$I_\Psi = [\Psi(a_0), \Psi(a^0)].$$

Inverse problems

Observation

If a desirable ouput u_{given} is known on a domain Ω and if

$$\Psi(a) = \int_{\Omega} (u(a) - u_{\text{given}})^2 \, dx ,$$

then the search for the best scenario is, in fact, a parameter identification problem.

Pros and cons of the WSM

Pros

- The knowledge of $\mathcal{U}_{\text{ad}}$ is sufficient, no probabilistic features.
- Guaranteed range of $\Psi|_{\mathcal{U}_{\text{ad}}}$.
- The WSM can be combined with other methods, see later.

Cons

- Based on the search for global extremes.
- Probability or possibility is not considered though the occurrence of the extremal values of Ψ is rare in many practical problems.

Outline

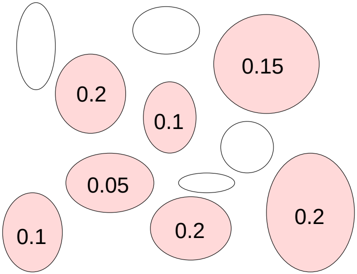
Dempster-Shafer evidence theory

X . . . universal set

P_X . . . power set of X

$m : P_X \to [0,1]$ . . . basic probability assignment
It must satisfy $m(\emptyset) = 0$ and $\sum_{\text{all } A \in P_X} m(A) = 1$.

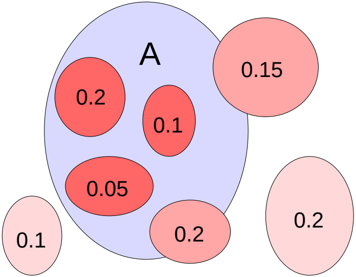
Let $m(A_i) > 0$ only for a finite number of sets $A_i \in P_X$;
A_i are called *focal elements*.



Definition (*Belief* and *Plausibility*)

The belief measure of $A \in P_X$: $Bel(A) = \sum_{\text{all } A_i \subseteq A} m(A_i)$.

The plausibility measure of $A \in P_X$: $Pl(A) = \sum_{\text{all } A_i \bigcap A \neq \emptyset} m(A_i)$.



$$Bel(A) = 0.2 + 0.1 + 0.05 = 0.35$$

$$Pl(A) = 0.2 + 0.1 + 0.05 + 0.15 + 0.2 = 0.7$$

Observation

$0 \leq Bel(A) \leq Pl(A) \leq 1 \quad A \in P_X$

Interpretations of *Pl* and *Bel* (sort of B/W scenarios)

$m(A_i)$ provides a measure of the amount of “likelihood" that is assigned to A_i.

$Pl(A)$ provides an upper bound on the likelihood of A.
$Bel(A)$ provides a lower bound on the likelihood of A.

$Pl(A)$ is the largest probability for A that is consistent with all available evidence.
$Bel(A)$ is the smallest probability for A that is consistent . . .

An upper limit (*Pl*) and a lower limit (*Bel*) on the strength of evidence at hand.

Observation: A sort of the best (*Pl*) and the worst (*Bel*) scenarios.

Dempster-Shafer theory and the WSM

Let A_i, where $i = 1, \ldots, N$, be both the focal elements of m and admissible sets.

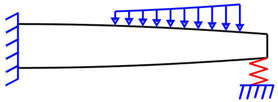
Let $\Psi(A_i)$, the range of $\Psi|_{A_i}$, be calculated via the worst and best scenarios.

I_k is a focal element of the probability assignment m_Ψ defined through m, A_i, and $\Psi(A_i)$ (extension principle):

$$m_\Psi(I_k) = \sum_{\{i\colon I_k = \Psi(A_i)\}} m(A_i), \quad k = 1, \ldots, M.$$

Thus m_Ψ, the basic probability assignment in the range of Ψ is established, and the relationship between the range of Ψ and various sets can be assessed through *Bel* and *Pl*.

Example (Uncertain spring stiffness)



$A_1, \ldots, A_5$, five intervals for the spring stiffness parameter.
Their respective basic probability assignment values are equal to 0.1, 0.4, 0.1, 0.25, and 0.15.

Ψ stands for the beam tip displacement.

Let $I_k = \Psi(A_k)$, $k = 1, 2, \ldots, 5$. It is calculated that
$I_1 = [77, 80]$, $m_\Psi(I_1) = 0.1$; $I_2 = [69, 74]$, $m_\Psi(I_2) = 0.4$;
$I_3 = [73, 79]$, $m_\Psi(I_3) = 0.1$; $I_4 = [71, 78]$, $m_\Psi(I_4) = 0.25$;
$I_5 = [76, 83]$, $m_\Psi(I_5) = 0.15$.

Example (Uncertain spring stiffness (cont.))

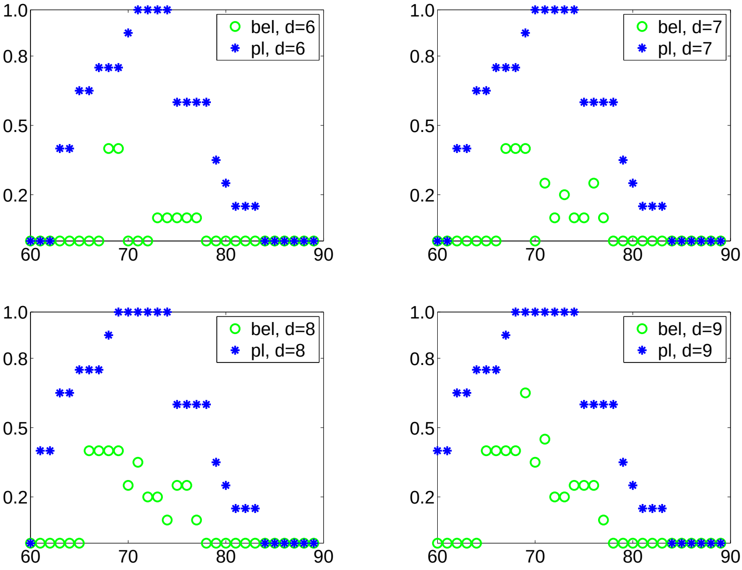
To analyze the (uncertain) behavior of the quantity of interest, let us graph

$$Bel([x, x+d]) \text{ and } Pl([x, x+d]),$$

where $d \in \{6, 7, 8, 9\}$ is fixed and $x \in [60, 90]$.

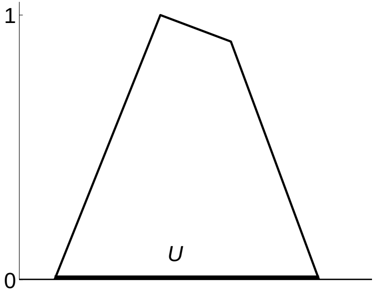
In other words, the chosen intervals $[x, x+d]$ will be weighted by the evidence we have about the behavior of Ψ.

This information can help the analyst to make a decision.


bel, d=6
pl, d=6
bel, d=7
pl, d=7
bel, d=8
pl, d=8
bel, d=9
pl, d=9
1.0
0.8
0.5
0.2
60
70
80
90

Outline

Fuzzy set theory (Zadeh)

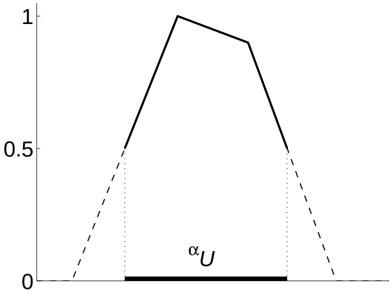


A fuzzy set U is identified with μ_U, the membership function

$$\mu_U : X \to [0, 1],$$

where the real value in $[0, 1]$ represents the degree to which $x \in X$ belongs to the set U. The higher the value, the stronger the membership.

Fuzzy set theory: α-cut



For $\alpha \in (0,1]$, a subset ${}^{\alpha}U$ comprising all $x \in X$ such that $\mu_U(x) \geq \alpha$ is called the α-cut.

It will be convenient to have $U \equiv {}^{0}U = \mathrm{supp}(\mu_U)$.

Fuzzy set theory and the WSM

Let $\mathcal{U}_{\text{ad}} \equiv U$ and $\mu_{\mathcal{U}_{\text{ad}}}$ be given.
Let ${}^{\alpha}\mathcal{U}_{\text{ad}}$ be the α-cut of $\mathcal{U}_{\text{ad}}$, $\alpha \in [0, 1]$.
For α, the best and worst scenarios determine

$$ {}^{\alpha}I_{\Psi} = [\Psi(a_{0,\alpha}), \Psi(a^{0,\alpha})]. $$

These ${}^{\alpha}I_{\Psi}$ are the α-cuts of $I_{\Psi} \equiv {}^{0}I_{\Psi} = \{\Psi(a) |\ a \in \mathcal{U}_{\text{ad}}\}$, the fuzzy range of Ψ.

Then the membership function $\mu_{I_{\Psi}}$ can be constructed via

$$ \mu_{I_{\Psi}}(y) = \max\{\alpha |\ y \in {}^{\alpha}I_{\Psi}\}, \quad y \in I_{\Psi}, $$

to asses the fuzziness of Ψ, the quantity of interest.
(A numerical example will be given later.)

Outline

1. A broader context
   - Ultimate goal of modeling
   - Omnipresent uncertainty
2. Stochastic approaches
   - Transformation of a stochastic DE to a deterministic DE
3. Non-stochastic methods
   - The worst (case) scenario method (WSM)
   - Dempster-Shafer theory
   - Fuzzy set theory
4. DE-driven problems: uncertain functions
   - The worst scenario method
   - The WSM and fuzzy sets

DE-driven problems with uncertain data

We will focus on worst/best scenario problems where $D(a)u = f$, the state problem, stems from a differential equation and a is an uncertain function.

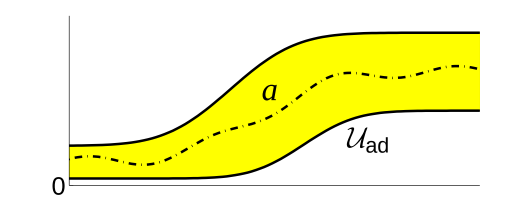
Boundary value problems with uncertain data

Motivation: $D(a)u = f$ stands for quasilinear heat conduction

$$-\operatorname{div}(a(u)\nabla u) = f \text{ in } \Omega \subset \mathbb{R}^2$$

Dirichlet, Neumann, Newton (Robin) or mixed BC,

where a is uncertain, $a \in \mathcal{U}_{\text{ad}}$.



A typical $\mathcal{U}_{\text{ad}}$ comprises bounded, positive ($a \geq c > 0$) functions with uniformly bounded derivative (no wild oscillations).
$\mathcal{U}_{\text{ad}}$ is *compact* in the space of continuous functions.

Boundary value problems with uncertain data

Quasilinear heat conduction, criterion-functional

$$-\operatorname{div}(a(u)\nabla u) = f \text{ in } \Omega \subset \mathbb{R}^2 \quad \& \quad \text{BC}$$

Temperature monitored in a fixed subdomain $G \subset \Omega$:

$$\Psi(a) = (\operatorname{meas} G)^{-1} \int_G u(a)\, dx .$$

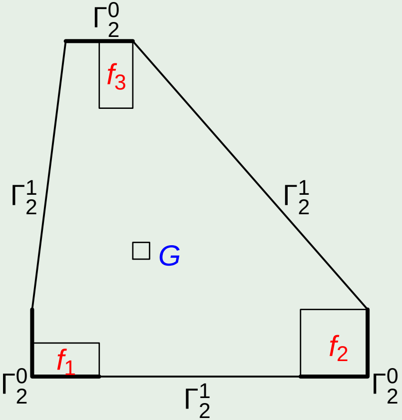
Worst scenario problem: Find

$$a^0 = \arg\max_{a \in \mathcal{U}_{\text{ad}}} \Psi(a).$$

Example (Quasilinear heat conduction)

$-\operatorname{div}(a(u)\nabla u) = f$ in $\Omega \subset \mathbb{R}^2$ & BC, a is a diag. matrix

$$\Psi(a) = (\operatorname{meas} G)^{-1} \int_G u(a)\,dx.$$



Γ_2^0 insulation (Dirichlet); Γ_2^1 Newton (Robin), f_i heat sources

Approximate WS problem

To approximate $u(a)$ by $u_h(a)$, we resort to finite elements, finite differences, boundary elements, ...

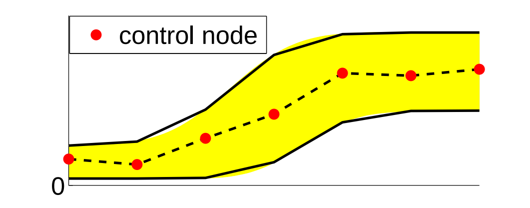
$\mathcal{U}_{\text{ad}}$ is approximated by $\mathcal{U}_{\text{ad}}^M$ that is identifiable with a compact subset of $\mathbb{R}^M$.

Approximate WS problem

To approximate $u(a)$ by $u_h(a)$, we resort to finite elements, finite differences, boundary elements, ...

$\mathcal{U}_{\text{ad}}$ is approximated by $\mathcal{U}_{\text{ad}}^M$ that is identifiable with a compact subset of $\mathbb{R}^M$.

Example ($\mathcal{U}_{\text{ad}}$ approximated by $\mathcal{U}_{\text{ad}}^M$)



a^M controlled by the vertical position of seven nodes constrained by the bounds and through the Lipschitz constant.

Approximate WS problem

To approximate $u(a)$ by $u_h(a)$, we resort to finite elements, finite differences, boundary elements, ...

$\mathcal{U}_{\text{ad}}$ is approximated by $\mathcal{U}^M_{\text{ad}}$ that is identifiable with a compact subset of $\mathbb{R}^M$.

The approximate worst scenario problem: Find

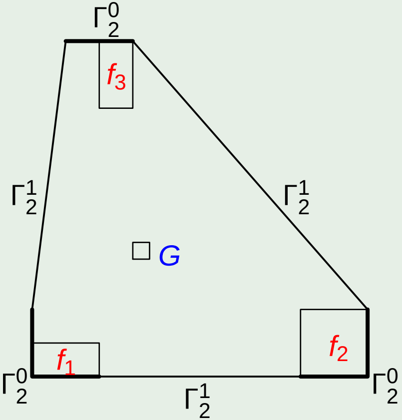
$$a^{M0} = \operatorname*{arg\,max}_{a^M \in \mathcal{U}^M_{\text{ad}}} \Phi(a^M, u_h(a^M)).$$

The approximate best scenario problem is analogous.

Example (Quasilinear heat conduction)

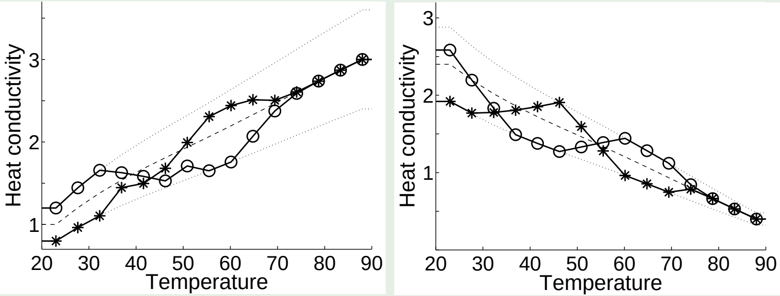
$-\operatorname{div}(a(u)\nabla u) = f$ in $\Omega \subset \mathbb{R}^2$ & BC, a is a diag. matrix

$$\Psi(a) = (\text{meas}\, G)^{-1} \int_G u(a)\, dx\,.$$



Γ_2^0 insulation (Dirichlet); Γ_2^1 Newton (Robin), f_i heat sources

Example (Quasilinear heat conduction)



Conductivities in horizontal (left) and vertical (right) directions.

$*$ max. problem, $\Phi(a^{M0}, u_h(a^{M0})) = 32.1$

$\circ$ min. problem, $\Phi(a_0^M, u_h(a_0^M)) = 30.9$

Ingredients of the WSM algorithm

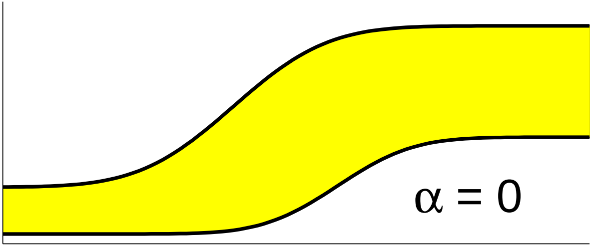
- $\mathcal{U}_{\text{ad}}^{M}$ (p.-w. linear functions, splines)
- state problem solver
- constrained global optimization
  - sensitivity analysis: differentiation of $\Phi(a, u(a))$ w.r.t. a, or $\Phi(a^M, u_h(a^M))$ w.r.t. a^M
    (also automatic differentiation)
  - gradient methods (can be trapped in local optima)
  - optimality conditions (can lead to local optima)
  - genetic algorithms (tend to be computationally expensive)

These tools are also used in shape optimization, PDE constrained optimization, or methods for solving inverse problems.

Outline

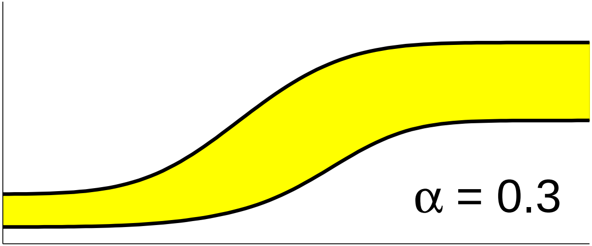
Fuzzy $\mathcal{U}_{\text{ad}}$, $\mathcal{U}_{\text{ad}}^{M}$: direct approach to α-cuts

The parameter α directly controls the lower and upper bound of the cut.



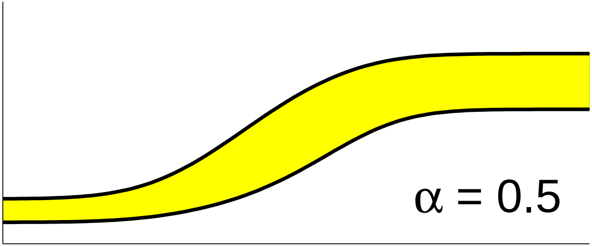
Fuzzy $\mathcal{U}_{\text{ad}}$, $\mathcal{U}_{\text{ad}}^{M}$: direct approach to α-cuts

The parameter α directly controls the lower and upper bound of the cut.



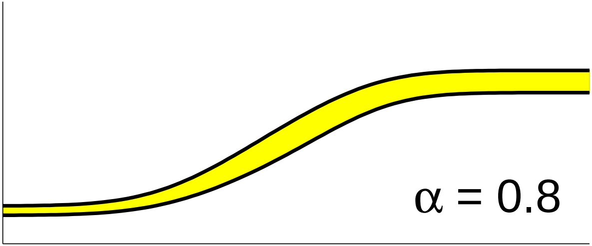
Fuzzy $\mathcal{U}_{\text{ad}}$, $\mathcal{U}_{\text{ad}}^{M}$: direct approach to α-cuts

The parameter α directly controls the lower and upper bound of the cut.



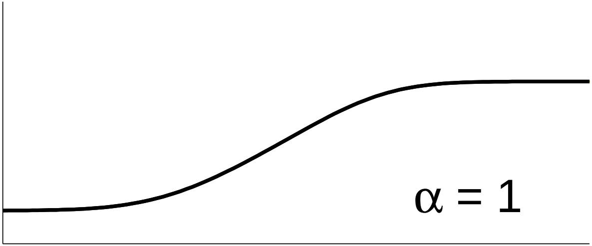
Fuzzy $\mathcal{U}_{\text{ad}}$, $\mathcal{U}_{\text{ad}}^{M}$: direct approach to α-cuts

The parameter α directly controls the lower and upper bound of the cut.



Fuzzy $\mathcal{U}_{\text{ad}}$, $\mathcal{U}_{\text{ad}}^M$: direct approach to α-cuts

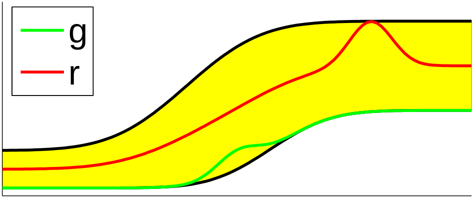
The parameter α directly controls the lower and upper bound of the cut.



Fuzzy $\mathcal{U}_{\text{ad}}$, $\mathcal{U}_{\text{ad}}^{M}$: direct approach to α-cuts

Pros: ${}^{\alpha}\mathcal{U}_{\text{ad}}$ easy to define; ${}^{\alpha}\mathcal{U}_{\text{ad}}^{M}$ leads to optimization problems with simple bounds;

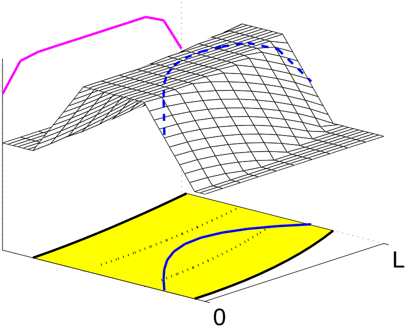
Contra: ${}^{\alpha}\mathcal{U}_{\text{ad}}$ has partly unnatural features:

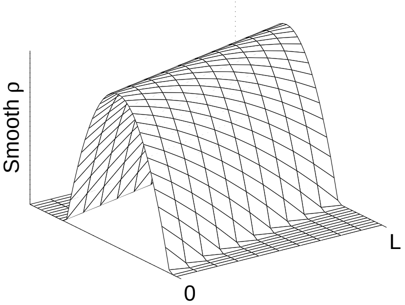


Both g and r *belong* or *do not belong* to the same ${}^{\alpha}\mathcal{U}_{\text{ad}}$, though g "lives its life on the outskirts of ${}^{\alpha}\mathcal{U}_{\text{ad}}$", but r "lives in the center except for a short deviation."

Fuzzy $\mathcal{U}_{\text{ad}}$, $\mathcal{U}_{\text{ad}}^{M}$: integral approach to α-cuts

$\mu_{\mathcal{U}_{\text{ad}}}(a) = L^{-1}\int_0^L \varrho(x, a(x))\,dx$ is defined through $\varrho(x, y)$ with values in $[0, 1]$:

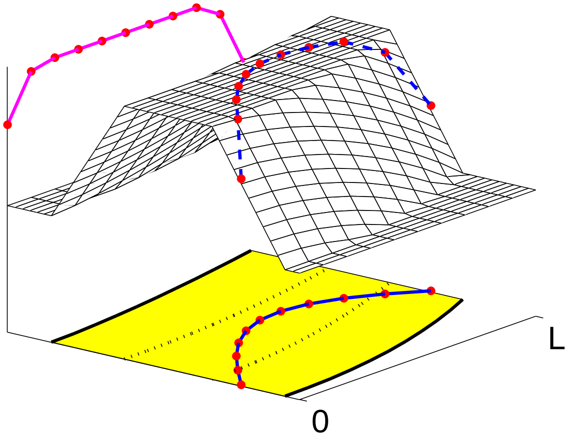




Pro: separation of g and r

Con: more demanding optimization, even nonsmooth; however:

Idea: If the control nodes do not cross the lines of nondifferentiability, the constraint is differentiable.



Example (ODE constrained problem with fuzzy input data)

$-(a(x)u'(x))' = f$ on $(0,1)$, $u(0) = 0 = u(1)$

$\Psi(a) = \int_0^1 (u(x) - \sin(2\pi x))^2 \, dx$,
f such that $\Psi(1 + x) = 0$.

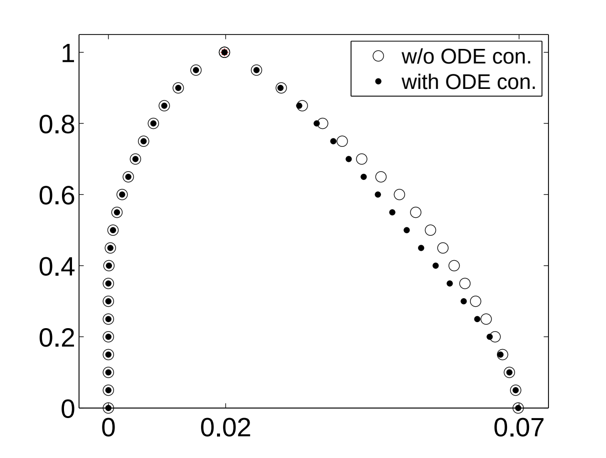
$\mathcal{U}_{\text{ad}} = \{a : \ |a - q| \leq 0.5$ and $|a' - q'| \leq 0.8\}$, where $q(x) = 1.5 + x^2$.

ϱ "triangular"

ODE-based constraint $|u(1/2)| \leq 0.06$, its gradient as well as $\nabla\Psi$ are calculated via adjoint equations by standard techniques.

The gradient of $\mu_{\mathcal{U}_{\text{ad}}}$ exists in all smooth subproblems and can be calculated from an analytical formula.

Example (Results (MATLAB®, NAG®))



Vertical axis: α-levels

Horizontal axis: the range of Ψ

Concluding remarks

- The WSM is useful by itself in analyzing the propagation of uncertainty through models.
- The WSM is used in extending the fuzzy or Dempster-Schafer theory characteristics of input data to output data (from $\mathcal{U}_{\text{ad}}$ to the range of $\Psi|_{\mathcal{U}_{\text{ad}}}$).
- The WSM shares many features with optimal shape design, PDE constrained optimization, and inverse problems.
- Various well-tried tools for theoretical analysis are at our disposal. First steps have been done.
- Various well-tried tools for computational analysis are at our disposal. However, their tailoring for uncertain input data problems is desirable.

Thank you for your attention.