Greedy optimal control for elliptic equations

Applications to turnpike control

Martin Lazar

University of Dubrovnik

in collaboration with:

Víctor Hernández-Santamaría (*DeustoTech, University of Deusto*)

Enrique Zuazua *(University of Deusto, UAM)*

Workshop on Optimal Control of Dynamical Systems and Applications

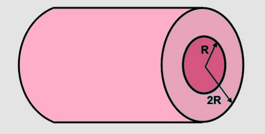
Osijek, May, 2018

Hrvatska zaklada za znanost

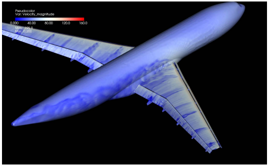
ConDyS

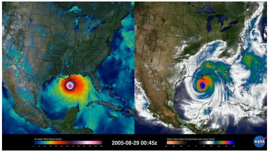
Parameter dependent problems

Real life applications (may) depend on a large number of parameters



examples: thickness, conductivity, density, length, humidity, pressure, curvature,. . .





Parameter dependent problems (Cont.)

- When dealing with applications and simulations, we would like to explore within different parameter configurations.
- From the control point of view, this implies solving a different problem for each configuration.
- Computationally expensive.

Our goal

Apply greedy theory to have a **robust and efficient** numerical solvers.

Parameter dependent control problem

$\Omega \subset \mathbb{R}^N, \quad \omega \subset \Omega.$

Consider the system

$$\begin{cases} -\text{div}(a(x,\nu)\nabla y) + c\,y = \chi_\omega u & \text{in } \Omega, \\ y = 0 & \text{on } \partial\Omega, \end{cases} \tag{1}$$

∘ ν is a parameter ∘ $u \in L^2(\omega)$ is a control ∘ $c = c(x) \in L^\infty(\Omega)$

Optimal control problem (OCP$_\nu$)

$$\min_{u \in L^2(\omega)} J_\nu(u) = \frac{1}{2}|u|^2_{L^2(\omega)} + \frac{\beta}{2}\|y - y_d\|^2_{L^2(\Omega)},$$

$\exists!$ optimal solution is well-known (Lions, Tröltzsch,...)

Parameter dependent control problem (cont.)

Characterization of the solution: optimal pair $(\bar{u}, \bar{y})$

$$\bar{u} = -\chi_\omega \bar{q}$$

where $(\bar{y}, \bar{q})$ solve the optimality system:

$$\begin{cases} -\text{div}(a(x,\nu)\nabla \bar{y}) + c\,\bar{y} = -\chi_\omega \bar{q}, & \text{in } \Omega, \\ -\text{div}(a(x,\nu)\nabla \bar{q}) + c\,\bar{q} = \beta\,(\bar{y} - y^d), & \text{in } \Omega, \\ \bar{y} = \bar{q} = 0, & \text{on } \partial\Omega. \end{cases} \quad (2)$$

As the state y depends on ν, also the control u depends on ν.

From the practical point of view,

- Measure parameter ν and determine $u_\nu = \arg\min_{u \in L^2(\omega)} J_\nu(u)$ using classical methods (iterative methods, ...)
- Repeat the process for each new value of ν.

Can we do it better?

Greedy control

Assume that ν ranges within a compact set $\mathcal{K} \subset \mathbb{R}^d$ and $a_\nu = a(x, \nu)$ are bounded functions satysfing

$$0 < a_1 \leq a_\nu \leq a_2, \qquad \nu \in K.$$

In this way, we ensure that each control can be uniquely determined by

$$\bar{u}_\nu = -\chi_\omega \bar{q}$$

where $(\bar{y}, \bar{q})$ solve the optimality system. Consider the set of controls $\bar{u}_\nu$ for each possible value $\nu \in \mathcal{K}$. That is,

$$\bar{\mathcal{U}} = \{\bar{u}_\nu : \nu \in \mathcal{K}\}$$

The idea

To determine a finite number of values of ν that yield the best possible approximation of the control manifold $\bar{\mathcal{U}}$

Description of the method

We look for a *small* number of parameters $\nu \in \mathcal{K}$ approximating the manifold $\bar{\mathcal{U}}$ in the sense of the Kolmogorov width. **Roughly**, the Kolmogorov width measures how well we can approximate $\bar{\mathcal{U}}$ by a finite dimensional space.

In order to achieve this goal we rely on greedy algorithms and reduced bases methods for parameter dependent PDEs or abstract equations in Banach spaces.

- A. Cohen, R. DeVore, Kolmogorov widths under holomorphic mappings, *IMA Journal on Numerical Analysis*, to appear
- A. Cohen, R. DeVore, Approximation of high-dimensional parametric PDEs, arXiv preprint, 2015.
- Y. Maday, O. Mula, A. T. Patera, M. Yano, The generalized Empirical Interpolation Method: stability theory on Hilbert spaces with an application to the Stokes equation, submitted

The pure greedy method

X – a Banach space $K \subset X$ – a compact subset.

The method approximates K by a a series of finite dimensional linear spaces V_n (a linear method).

The algorithm

The first step Choose $x_1 \in K$ such that

$$\|x_1\|_X = \max_{x \in K} \|x\|_X .$$

The general step Having found $x_1..x_n$, denote $V_n = \operatorname{span}\{x_1, \ldots, x_n\}$. Choose the next element

$$x_{n+1} := \arg\max_{x \in K} \operatorname{dist}(x, V_n) . \qquad (3)$$

The algorithm stops when $\sigma_n(K) := \max_{x \in K} \operatorname{dist}(x, V_n)$ becomes less than the given tolerance ε.

The greedy idea

The greedy idea

Which one you are going to choose?

Sometimes it is hard to solve the maximisation problem (3).

The weak greedy method

– a relaxed version of the pure one.

The algorithm

Fix a constant $\gamma \in \langle 0, 1]$.

The first step Choose $x_1 \in K$ such that

$$\|x_1\|_X \geq \gamma \max_{x \in K} \|x\|_X .$$

The general step

Having found $x_1..x_n$, denote $V_n = \operatorname{span}\{x_1, \ldots, x_n\}$.

Choose the next element

$$\operatorname{dist}(x_{n+1}, V_n) \geq \gamma \max_{x \in K} \operatorname{dist}(x, V_n) . \tag{4}$$

The algorithm stops when $\sigma_n(K) := \max_{x \in K} \operatorname{dist}(x, V_n)$ becomes less than the given tolerance ε.

Efficiency

In order to estimate **the efficiency of the (weak) greedy algorithm** we compare its approximation rates $\sigma_n(K)$ with the best possible one.

The Kolmogorov n width, $d_n(K)$

– measures how well K can be approximated by a subspace in X of a fixed dimension n.

$$d_n(K) := \inf_{\dim Y = n} \sup_{x \in K} \inf_{y \in Y} \|x - y\|_X .$$

Thus $d_n(K)$ represents optimal approximation performance that can be obtained by a n-dimensional linear space.

The greedy approximation rates have same decay as the Kolmogorov widths.

Performance obstacles

- The set K in general consists of infinitely many vectors.
- In practical implementations the set K is often unknown (e.g. it represents the family of solutions to parameter dependent problems).

Performance obstacles

- The set K in general consists of infinitely many vectors.
  Finite discretisation of K.
- In practical implementations the set K is often unknown (e.g. it represents the family of solutions to parameter dependent problems).

Performance obstacles

- The set K in general consists of infinitely many vectors.
  Finite discretisation of K.
- In practical implementations the set K is often unknown (e.g. it represents the family of solutions to parameter dependent problems).
  One uses some **surrogate** value replacing the exact distance appearing in (4) by some uniformly equivalent term.

Performance obstacles

- The set K in general consists of infinitely many vectors.
  Finite discretisation of K.
- In practical implementations the set K is often unknown (e.g. it represents the family of solutions to parameter dependent problems).
  One uses some **surrogate** value replacing the exact distance appearing in (4) by some uniformly equivalent term.

Practical realisation depends crucially on an existence of an appropriate surrogate .

The vectors chosen by the greedy procedure are the snapshots.

Their computation can be time consuming and computational expensive (offline part).

Los Alamos National Laboratory

The vectors chosen by the greedy procedure are the snapshots.

Their computation can be time consuming and computational expensive (offline part).

Los Alamos National Laboratory

Once having chosen the snapshots, one should easily approximate any value $x \in K$ (online part).

The vectors chosen by the greedy procedure are the snapshots.

Their computation can be time consuming and computational expensive (offline part).

Los Alamos National Laboratory

Once having chosen the snapshots, one should easily approximate any value $x \in K$ (online part).

The surrogate

In practical implementations, the set $\bar{\mathcal{U}}$ is unknown.

Given two parameters ν_1 and ν_2, how can we measure the distance between $\bar{u}_{\nu_1}$ and $\bar{u}_{\nu_2}$?

Recall that we want to avoid to compute $\bar{u}_\nu$.

The surrogate

In practical implementations, the set $\bar{\mathcal{U}}$ is unknown.

Given two parameters ν_1 and ν_2, how can we measure the distance between $\bar{u}_{\nu_1}$ and $\bar{u}_{\nu_2}$?

Recall that we want to avoid to compute $\bar{u}_\nu$.

Standard residual: Suppose that we have computed u_{ν_1}

$$|u_{\nu_1} - u_{\nu_2}| \sim |\nabla J_{\nu_2}(u_{\nu_1}) - \nabla J_{\nu_2}(u_{\nu_2})| = \nabla J_{\nu_2}(u_{\nu_1})$$

Compute $\nabla J_{\nu_2}(u_{\nu_1}) = u_{\nu_1} + \beta S^*_{\nu_2}(S_{\nu_2} u_{\nu_1} - y_d)$,
S_ν is to control-to-state operator. This means

$$S^*_{\nu_2}(S_{\nu_2} u_{\nu_1} - y_d) = -\chi_\omega q$$

where q is obtained by Solving a cascade system.

$$\begin{cases} -\text{div}(a_{\nu_2} \nabla y) + c\, y = \chi_\omega u_{\nu_1}, & \text{in } \Omega, \\ -\text{div}(a_{\nu_2} \nabla q) + c\, q = \beta\, (y - y^d), & \text{in } \Omega, \end{cases}$$

Numerical results

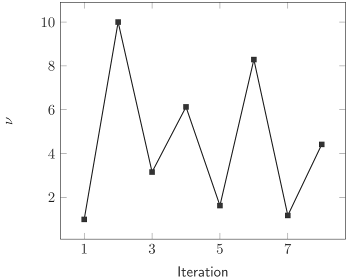
Numerical examples

- $\Omega = (0,1)^2$ in 2-D or $\Omega = (0,1)$ in 1-D.
- Uniform meshes, i.e., meshes with constant discretization steps in each direction, $N = 400$.
- We will approximate the operator $\mathcal{A} = -\text{div}(a(x,\nu)\nabla \cdot)$ by using the standard 5-point discretization.
- Discretize-then-optimize.
- $\nu \in \mathcal{K} = [1, 10]$.
- $\mathcal{K}$ sampled in 100 equidistant points.

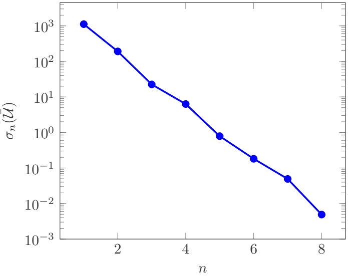
Greedy test # 1

$\circ\, a(x, \nu) = 1 + \nu(x_1^2 + x_2^2)$, $\quad \circ\, c(x) = \sin(2\pi x_1)\sin(2\pi x_2)$,
$\circ\, y_d = sin(\pi x_1)$, $\quad \circ\, \beta = 10^4$, $\quad \circ\, \varepsilon = 0.005$

$\circ\, t = 304s$

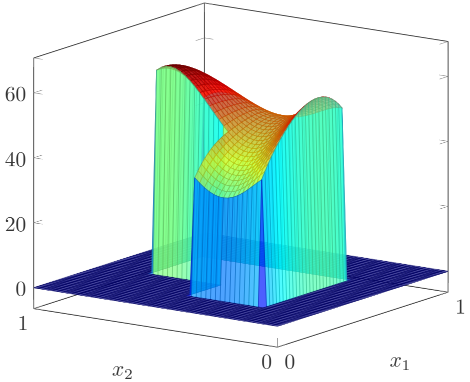


(a) Selected ν

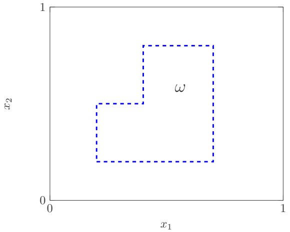


(b) Approximation error

Approximation for $\nu = \pi/2$



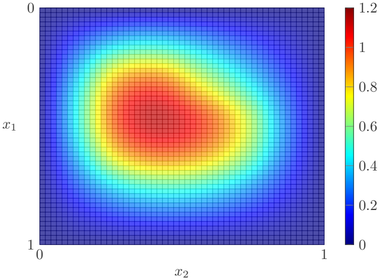
(c) The approximated control



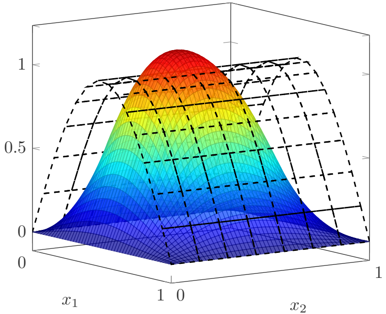
(d) The control set ω

$\circ \, |u^{\star}_{\pi/2} - \bar{u}_{\pi/2}|_{L^2(\omega)} \approx 1.45 \times 10^{-5}, \quad t_{\text{online}} = 0.45s, \quad t_{\text{iterative}} = 6.01s.$

Approximation for $\nu = \pi/2$ (cont.)



(e) The controlled state



(f) The state $y^\star_{\pi/2}$ and the target function y^d (dashed)

$\circ\ |y^\star_{\pi/2} - \bar{y}_{\pi/2}|_{L^2(\Omega)} \approx 1.15 \times 10^{-7}$

Connection with the turnpike problems

Time dependent control problem

Consider

$$\begin{cases} \partial_t y - \text{div}(a(x,\nu)\nabla y) + c\,y = \chi_\omega u & \text{in } Q = \Omega \times (0,T), \\ y = 0 & \text{on } \Sigma = \partial\Omega \times (0,T), \\ y(x,0) = y^0(x) & \text{in } \Omega. \end{cases} \quad (5)$$

and the control problem

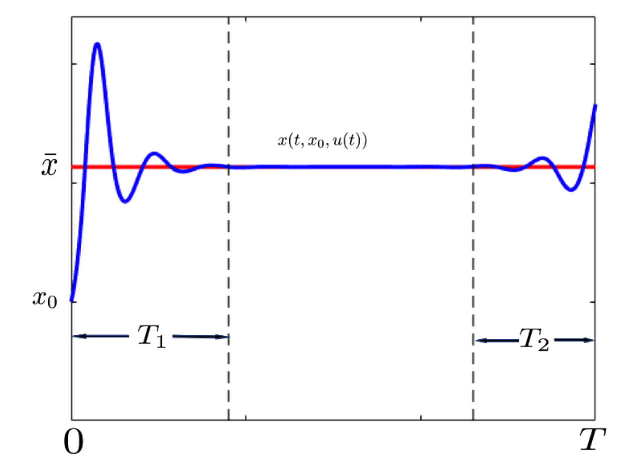
$$\min_u J_\nu{}^T(u) = \frac{1}{2}\int_0^T |u(t)|^2_{L^2(\omega)} dt + \frac{\beta}{2}\int_0^T \|y(t) - y^d\|^2_{L^2(\Omega)} dt.$$

The optimal solution (u^T, y^T) satisfies

$$\|y^T(t) - \bar{y}\|_{L^2(\Omega)} + \|u^T(t) - \bar{u}\|_{L^2(\Omega)} \leq K\left(e^{-\mu t} + e^{-\mu(T-t)}\right), \quad \forall t \in [0,T]$$

- Exponential convergence of the finite-time horizon control problem to the steady one as $T \to \infty$.

Turnpike behaviour



- optimal, time dependent control
- steady control

Greedy test # 2

We consider time-dependent version of the last example.

$\circ\, a(x,\nu) = 1 + \nu(x_1^2 + x_2^2)$, $\quad \circ\, c(x) = \sin(2\pi x_1)\sin(2\pi x_2)$,
$\circ\, y_d = sin(\pi x_1)$, $\quad \circ\, \beta = 10^4$, $\quad \circ\, \varepsilon = 0.005$
$\circ\Omega = (0,1)^2$ in 2-D or $\Omega = (0,1)$ in 1-D.

We take initial datum
$y_0(x) = \sin(3\pi x_1)\sin(2\pi x_2)$

The case $c(x) \geq 0$ (greedy test #2)

$$u(x,t) = u^{\star}_{\pi/2}(x), \qquad y_0(x) = \sin(3\pi x_1)\sin(2\pi x_2)$$