Parameter depending and turnpike control

Enrique Zuazua

DeustoTech (Bilbao) - UAM (Madrid) - UPMC (Paris)
www.cmc.deusto.es
Funded by ERC Advanced Grant DyCon and the French ANR

Zagreb, November 2017

Outline

Many control problems arising in engineering, biomedicine and social sciences, lead to natural questions of control in long time horizons.

1. Sustainable growth
2. Cronical diseases
3. New generation of supersonic aircrafts

Challenges:

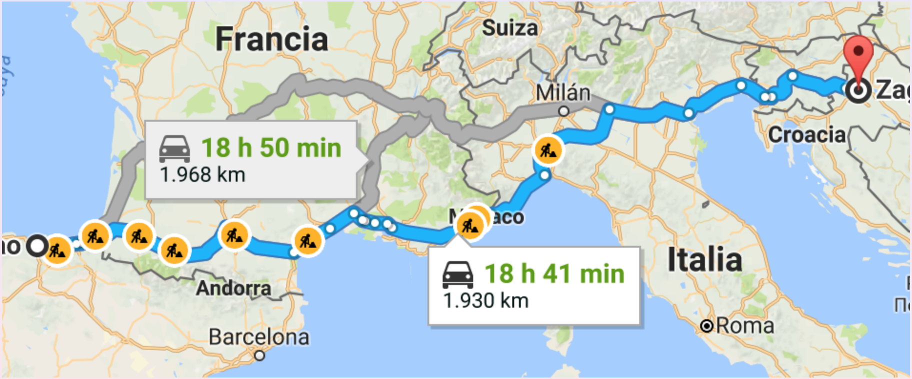
1. Develop specific tools for long time control horizons.
2. Build numerical schemes capable of reproducing accurately the control dynamics in long time intervals (geometric integration, asymptotic preserving schemes).

Outline

Origins

Although the idea goes back to John von Neumann in 1945, Lionel W. McKenzie traces the term to Robert Dorfman, Paul Samuelson, and Robert Solow's "Linear Programming and Economics Analysis" in 1958, referring to an American English word for a Highway:

> ... *There is a fastest route between any two points; and if the origin and destination are close together and far from the turnpike, the best route may not touch the turnpike. But if the origin and destination are far enough apart, it will always pay to get on to the turnpike and cover distance at the best rate of travel, even if this means adding a little mileage at either end.*

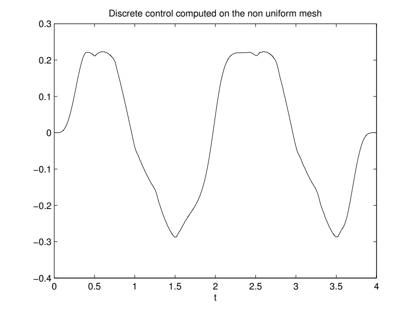

Francia
Suiza
Milán
Croacia
18 h 50 min
1.968 km
18 h 41 min
1.930 km
Italia
Andorra
Barcelona
Roma

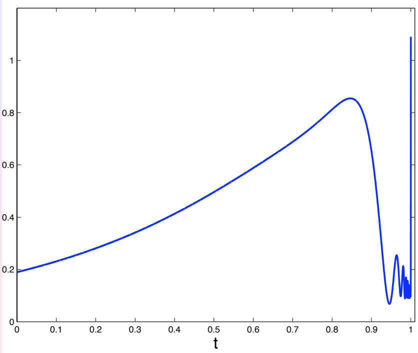
Tunrpike property ≡ Asymptotic simplification

The turnpike property...

1. ... ensures that optimal strategies for the steady-state problem lead to nearly optimal ones for the time-dependent dynamics.
2. ... is employed systematically much beyond the class of problems for which the principle can be proved to hold rigorously.
3. ... can be of use in many contexts such as mesh adaptivity, parameter-dependent problems, etc.
4. ... it yields a method to ensure robust control, independent of the initial datum under consideration.

Examples where controls seem to fail the turnpike property





Typical dynamics of controls for wave and heat like equations, as solutions of the corresponding adjoint systems.

The control problem for the heat equation[a]

[a] Joint work with M. Gugat, V. Hernández-Santamaria, M. Lazar, A. Porretta, E. Trélat...

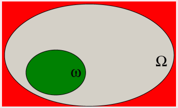
Let $n \geq 1$ and $T > 0$, Ω be a simply connected, bounded domain of $\mathbb{R}^n$ with smooth boundary Γ, $Q = (0, T) \times \Omega$ and $\Sigma = (0, T) \times \Gamma$:

$$\begin{cases} y_t - \Delta y = f 1_\omega & \text{in} \quad Q \\ y = 0 & \text{on} \quad \Sigma \\ y(x,0) = y^0(x) & \text{in} \quad \Omega. \end{cases} \tag{1}$$

$1_\omega =$ the characteristic function of ω of Ω where the control is active. We assume that $y^0 \in L^2(\Omega)$ and $f \in L^2(Q)$ so that (1) admits a unique solution

$$y \in C\left([0, T]; L^2(\Omega)\right) \cap L^2\left(0, T; H_0^1(\Omega)\right).$$

$$y = y(x,t) = solution = state, \; f = f(x,t) = control$$



We want to minimise the cost:

$$J(f) = \frac{1}{2}\int_0^T \int_\omega f^2 dxdt + \frac{1}{2}\int_\Omega |y(x,T) - y_d|^2 dx \qquad (2)$$

making a compromise between reaching the target u_d and energy consumption f.

The classical optimality system (Pontryaguin's principle) guarantees that the control is of the form

$$f = \varphi$$

where φ is the solution of the adjoint equation:

$$\begin{cases} -\varphi_t - \Delta\varphi = 0 & \text{in } Q \\ \varphi = 0 & \text{on } \Sigma \\ \varphi(T) = y(T) - y_d & \text{in } \Omega. \end{cases} \qquad (3)$$

Lack of turnpike?

Better balanced controls

Let us now consider the control f minimising a compromise between the norm of the state and the control among the class of admissible controls:

$$min\, \frac{1}{2}\Big[\int_0^T \int_\Omega |y|^2 dxdt + \int_0^T \int_\omega |f|^2 dxdt + \frac{1}{2}\int_\Omega |y(x,T) - y_d|^2 dx\Big].$$

Then the Optimality System reads

$$y_t - \Delta y = -\varphi 1_\omega \text{ in } Q$$

$$y = 0 \text{ on } \Sigma$$

$$y(x,0) = y^0(x) \text{ in } \Omega$$

$$-\varphi_t - \Delta\varphi = y \text{ in } Q$$

$$\varphi = 0 \text{ on } \Sigma.$$

$$\varphi(T) = y(T) - y_d \quad \text{ in } \Omega.$$

We now observe a coupling between φ and y on the adjoint state equation!

New Optimality System Dynamics

What is the dynamic behaviour of solutions of the new fully coupled OS?
For the sake of simplicity, assume $\omega = \Omega$.
The dynamical system now reads

$$y_t - \Delta y = -\varphi$$

$$\varphi_t + \Delta \varphi = -y$$

This is a forward-backward parabolic system.
A spectral decomposition exhibits the characteristic values

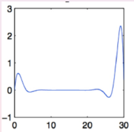
$$\mu_j^{\pm} = \pm\sqrt{1 + \lambda_j^2}$$

where $(\lambda_j)_{j \geq 1}$ are the (positive) eigenvalues of $-\Delta$.
Thus, the system is the superposition of growing + diminishing real exponentials.

The turnpike property for the heat equation

This new dynamic behaviour, combining exponentially stable and unstable branches, is compatible with the turnpike behavior.

Controls and trajectories exhibit the expected dynamics:



In this particular example turnpike means that optimal pairs $(f(t), y(t))$ are exponentially close to the steady-state optimal pair characterised as the minima for the functional

$$J_s(g) = \frac{1}{2}\int_\omega g^2 dxdt + \frac{1}{2}\int_\Omega \left[z^2 + |z(x) - z_d|^2\right] dx \tag{4}$$

where $z = z(x)$ solves

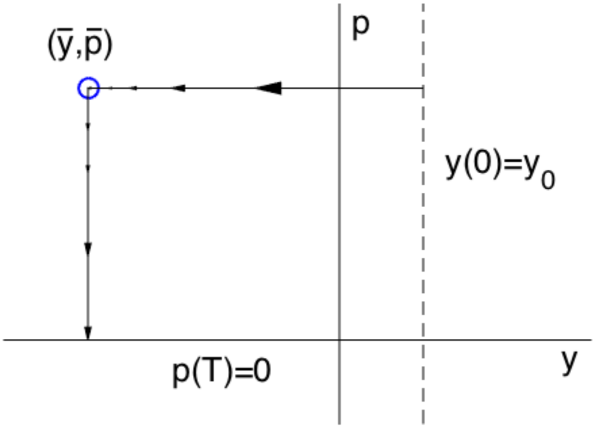
$$\begin{cases} -\Delta z = g 1_\omega & \text{in} \quad \Omega \\ z = 0 & \text{on} \quad \partial\Omega, \end{cases} \tag{5}$$

Namely

$$||y(t) - z|| + ||f(t) - g|| \leq C[\exp(-\mu t) + \exp(-\mu(T - t))]$$

if $T >> 1$.

The turnpike path



Consider the transport equation with unknown velocity v,

$$f_t + v f_x = 0,$$

and take averages with respect to v. Then

$$g(x,t) = \int f(x,t;v)\rho(v)dv$$

then, for the Gaussian density ρ:

$$\rho(v) = (4\pi)^{-1/2} \exp(-v^2/4)$$

$$g(x,t) = h(x,t^2); \quad h_t - h_{xx} = 0.$$

One can then employ parabolic techniques based on Carleman inequalities.[1][2]

[1] E. Z., Averaged controllability, Automatica, 50 (2014)

[2] Q. Lü & E. Z. Average Controllability for Random Evolution Equations, JMPA, 2016.

[3] Assume that the system depends on a parameter $\nu \in K \subset \mathbf{R}^d$, $d \geq 1$, K being a compact set, and controllability being fulfilled for all values of ν.

$$\begin{cases} x'(t) = A(\nu)x(t) + Bu(t), \ 0 < t < T, \\ x(0) = x^0. \end{cases} \tag{6}$$

Controls $u(t, \nu)$ are chosen to be of minimal norm satisfying the controllability condition:

$$x(T, \nu) = x^1, \tag{7}$$

and lead to a manifold of dimension d in $[L^2(0, T)]^M$:

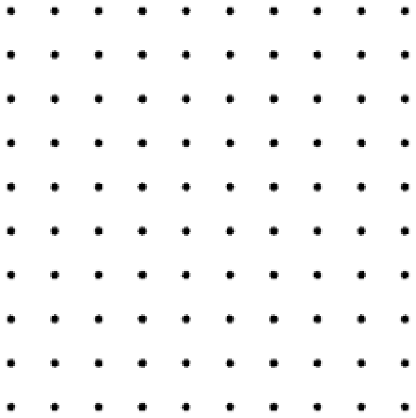
$$\nu \in K \subset \mathbf{R}^d \to u(t, \nu) \in [L^2(0, T)]^M.$$

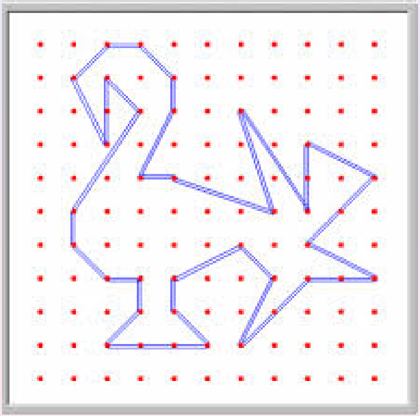
This manifold inherits the regularity of the mapping $\nu \to A(\nu)$.

To diminish the computational cost we look for the very distinguished values of ν that yield the best possible approximation of this manifold.

[3] M. Lazar & E. Zuazua, Greedy controllability of finite dimensional linear systems, Automatica, 2017.

Naive versus smart sampling of K





Our work relies on recent ones on greedy algorithms and reduced bases methods:

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

M. A. Grepl, M Kärche, *Reduced basis a posteriori error bounds for parametrized linear-quadratic elliptic optimal control problems*, CRAS Paris, 2011.

S. Volkwein, PDE-Constrained Multiobjective Optimal Control by Reduced-Order Modeling, IFAC CPDE2016, Bertinoro.

Description of the Method

We look for the realisations of the parameter ν ensuring the best possible approximation of the manifold of controls

$$\nu \in K \subset \mathbf{R}^d \to u(t,\nu) \in [L^2(0,T)]^M$$

(of dimension d in $[L^2(0,T)]^M$) in the sense of the Kolmogorov width.[4]

Greedy algorithms search for the values of ν leading to the most distinguished controls $u(t,\nu)$, those that are farther away one from each other.

Given an error ε, the goal is to find $\nu_1, ..., \nu_{n(\epsilon)}$, so that for all parameter values ν the corresponding control $u(t,\nu)$ can be approximated by a linear combination of $u(t,\nu_1), ..., u(t,\nu_{n(\epsilon)})$ with an error $\leq \epsilon$.

An of course to do it with a minimum number $n(\epsilon)$.

[4]Ensure the optimal rate of approximation by means of all possible finite-dimensional subspaces.

Step 1. *Characterization of minimal norm controls by adjoints*

The adjoint system depends also on the parameter ν:

$$-\varphi'(t) = A^*(\nu)\varphi(t), t \in (0,\ T);\ \varphi(T) = \varphi^0. \tag{8}$$

The control is

$$u(t,\nu) = B^*\varphi(t,\nu),$$

where $\varphi(t,\nu)$ is the solution of the adjoint system associated to the minimizer of the following quadratic functional in $\mathbf{R^N}$:

$$J_\nu\left(\varphi^0(\nu)\right) = \frac{1}{2}\int_0^T \left|B^*\varphi(t,\nu)\right|^2 dt - < x^1, \varphi^0 > + < x^0, \varphi(0,\nu) > .$$

The functional is continuous and convex, and its coercivity is guaranteed by the Kalman rank condition that we assume is satisfied for all value of ν.

Step 2. *Controllability distance*

Given two parameter values ν_1 and ν_2, how can we measure the distance between $u(t, \nu_1)$ and $u(t, \nu_2)$?

Roughly: Compute the residual

$$||x(T, \nu_2) - x^1||$$

for the solution of the state equation ν_2 achieved by the control $u(t, \nu_1)$.

More precisely: Solve the Optimality System (OS):

$$-\varphi'(t) = A^*(\nu_2)\varphi(t) \ t \in (0,\ T);\ \varphi(T) = \varphi_1^0.$$
$$x'(t) = A(\nu_2)x(t) + BB^*\varphi(t, \nu_2),\ 0 < t < T,\ x(0) = x^0.$$

Then

$$\left|\nabla J_{\nu_2}(\varphi_1^0)\right| = ||x(T, \nu_2) - x^1|| \sim ||\varphi_1^0 - \varphi_2^0||.$$

Within the class of controls of minimal L^2-norm, given by the adjoint, $u = B^*\varphi$, the residual $||x(T, \nu) - x^1||$ is a measure of the distance to the exact control, and also to the true minimiser.

Offline algorithm

Step 3. *Initialisation of the weak-greedy algorithm.* Choose any ν in K, $\nu = \nu_1$, and compute the minimizer of J_{ν_1}. This leads to φ_1^0.

Step 4. *Recursive choice of ν's.*
Assuming we have $\nu_1, ..., \nu_p$, we choose ν_{p+1} as the maximiser of

$$\max_{\nu \in K} \min_{\phi \in span[\varphi_j^0, j=1,...,p]} |\nabla J_\nu(\phi)|$$

We take ν_{p+1} as the one realizing this maximum.
Note that

$$|\nabla J_\nu(\phi)| = ||x(T,\nu) - x^1||.$$

$x(T,\nu)$ being the solution obtained by means of the control $u = B^*\phi(t,\nu)$, ϕ being the solution of the adjoint problem associated to the initial datum ϕ^0 in $span[\varphi_j^0, j = 1, ..., p]$.

Step 5. *Stopping criterion.* Stop if the max $\leq \epsilon$.

Online part

Step 6. For a specific realisation of ν solve the finite-dimensional reduced minimisation problem:

$$\min_{\phi \in span[\varphi_j^0, j=1,\dots,p]} J_\nu(\phi)$$

and choose the control $u(t,\nu)$. This minimises yields:

$$u(t,\nu) = B^* \varphi(t,\nu),$$

$\varphi(t,\nu)$ being the solution of the adjoint problem with datum ϕ at $t = T$.

The same applies when K is infinite-dimensional provided its Kolmogorov width decays polynomially.

Theorem

The weak-greedy algorithm above leads to an optimal approximation method.
More precisely, if the set of parametres K is finite-dimensional, and the map $\nu \to A(\nu)$ is analytic, for all $\alpha > 0$ there exists $C_\alpha > 0$ such that for all other values of ν the control $u(\cdot, \nu)$ can be approximated by linear combinations of the weak-greedy ones as follows:

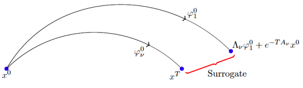
$$dist(u(\cdot, \nu); span[u(\cdot; \nu_j) : j = 1, ..., k]) \leq C_\alpha k^{-\alpha}.$$

5

[5] The approximation of the controls has to be understood in the sense above: Taking the control given by the corresponding adjoint solution, achieved by minimising the functional J over the finite-dimensional subspace generated by the adjoints for the distinguished parameter-values.

Potential improvements

1. **Find cheaper surrogates.** Is there a reduced model leading to lower bounds on controllability distances without solving the full Optimality System?



2. All this depends on the initial and final data: x_0, x_1.
   Can the search of the most relevant parameter-values ν be done independent of x_0, x_1?
   In other words, get **lower bounds on the controllability distances** between (A_1, B_1) and (A_2, B_2).

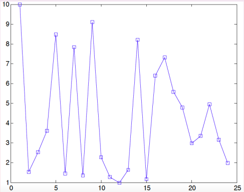
Semi-discrete wave equation

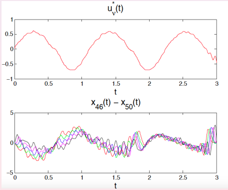
1. Finite difference approximation of the $1-d$ wave equation with 50 nodes in the space-mesh.
2. Unknown velocity v ranging within $[1, \sqrt{10}]$.
3. Discrete parameters taken over an equi-distributed set of 100 values
4. Boundary control
5. Sinusoidal initial data given: $y_0 = \sin(\pi x); y_1 \equiv 0$. Null final target.
6. Time of control $T = 3$.
7. Weak-greedy requires 20 snapshots.
8. Approximate control with error 0.5 in each component.
9. The algo stops after 24 iterations.
10. Offline time: 2 312 seconds (personal notebook with a 2.7 GHz processor and DDR3 RAM with 8 GB and 1,6 GHz).
11. Online time for one realisation ν: 7 seconds
12. Computational time for one single parameter value with standard methods: 51 seconds.

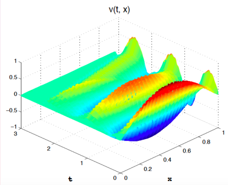
Choose a number at random within $[1, 10]$

But, please, choose π

The greedy algo leads to:

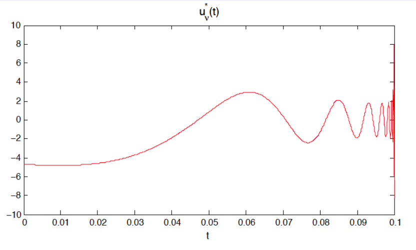


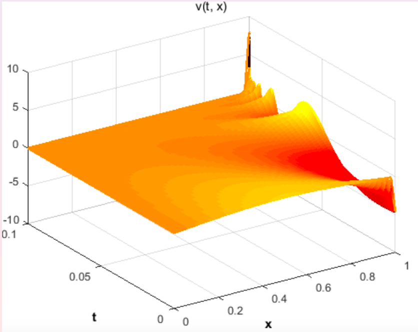




Semi-discrete heat equation

1. Finite difference approximation of the $1-d$ heat equation with 50 nodes in the space-mesh.
2. Unknown diffusivity v ranging within $[1,2]$.
3. Discrete parameters taken over an equi-distributed set of 100 values
4. Boundary control
5. Sinusoidal initial data given: $y_0 = \sin(\pi x)$. Null final target.
6. Time of control $T = 0.1$.
7. Weak-greedy requires 20 snapshots.
8. Approximate control with error 10^{-4} in each component.
9. The algo stops after 3 iterations: $\nu = 1.00, 1.18, 1.45$.
10. Offline time: 213 seconds.
11. Online time for one realisation $\nu = \sqrt{2}$: 1.5 seconds
12. Computational time for one single parameter value with standard methods: 37 seconds.


$u_v^*(t)$
10
8
6
4
2
0
−2
−4
−6
−8
−10
0
0.01
0.02
0.03
0.04
0.05
0.06
0.07
0.08
0.09
0.1
t


v(t, x)
10
5
0
-5
-10
0.1
0.05
t
0
0
0.2
0.4
0.6
0.8
1
x

Open problems and perspectives

- The method be extended to PDE. But analyticity of controls with respect to parameters has to be ensured to guarantee optimal Kolmogorov widths. This typically holds for elliptic and parabolic equations. But not for wave-like equations.
  Indeed, solutions of
  $$y_{tt} - v^2 y_{xx} = 0$$
  do not depend analytically on the coefficient v.
  One expects this to be true for heat equations in the context of null-controllability. Still this needs to be rigorously proved.
- Cheaper surrogates need to be found so to make the recursive choice process of the various $\nu's$ faster.
- Find surrogates for the controllability distance between two PDEs (of the same type).
  1. For wave equations in terms of distances between the dynamics of the Hamiltonian systems of bicharacteristic rays?
  2. For $1-d$ wave equations in terms of spectral distances?
  3. For heat equations?

Problem formulation

M. Choulli & E. Z. CRAS Paris, 2016

To better understand the complexity of the problem of applying the greedy methodology for control systems, independently of the initial and final data under consideration, it is natural to consider the following **diffusive equation as a model problem**.

Let Ω be a bounded domain of $\mathbb{R}^n$, $n \geq 1$. Fix $0 < \sigma_0 < \sigma_1$ and consider the class of scalar diffusivity coefficients

$$\Sigma = \{\sigma \in L^\infty(\Omega);\ \sigma_0 \leq \sigma \leq \sigma_1 \text{ a.e. in } \Omega\}.$$

For $\sigma \in \Sigma$, let $A_\sigma : H_0^1(\Omega) \to H^{-1}(\Omega)$ be the bounded operator given by

$$A_\sigma u = -\text{div}(\sigma \nabla u).$$

The inverse or resolvent operator $R_\sigma : H^{-1}(\Omega) \to H_0^1(\Omega)$.

The goal is to implement the greedy algo in the class of resolvent operators.

The existing theory gives the answer for a given right hand side term:

$$-\mathrm{div}(\sigma \nabla u) = f.$$

But we are interested on searching the most representative realisations of the resolvents as operators, independently of the value of f.
The analog at the control theoretical level would be to do it for the inverse of the Gramian operators rather than proceeding as above, for each specific data to be controlled.

The question under consideration is. How to find a surrogate (lower bound) for

$$dist(R_\sigma, span[R_{\sigma_1}, ..., R_{\sigma_k}])$$

?

The question is easy to solve when dealing with two resolvents R_1 and R_2.
But seems to become non-trivial in the general case.

This leads to a new class of Inverse Problems

Distance between two resolvents

It is easy to get a surrogate for the distance between two resolvents R_1 and R_2 corresponding to two different diffusivities σ_1 and σ_2:

$$A_1 - A_2 = A_1(R_2 - R_1)A_2,$$

$$\Big|A_1 - A_2\Big| \leq \sigma_1^2\Big|R_1 - R_2\Big|.$$

$$\langle (A_1 - A_2)u, u\rangle_{-1,1} = \int_\Omega (\sigma_1 - \sigma_2)|\nabla u|^2 dx,$$

$$\int_\Omega (\sigma_1 - \sigma_2)|\nabla u|^2 dx \leq \Big|A_1 - A_2\Big|\Big|u\Big|^2_{H_0^1(\Omega)} \leq \sigma_1^2\Big|R_1 - R_2\Big|\Big|u\Big|^2_{H_0^1(\Omega)}.$$

Now taking $u = u_\epsilon$ so that $|\nabla u_\epsilon|^2$ **constitutes an approximation of the identity** (for each $x_0 \in \Omega$) we get

$$||\sigma_1 - \sigma_2||_\infty \leq \sigma_1^2\Big|R_1 - R_2\Big|.$$

This can be understoof in the context of Inverse Problems: The resolvent determines the diffusivity with Lipschitz continuous dependence

1d

Unfortunately, this argument does not seem to apply for estimating the distance to a subspace

$$R_1 - \sum_{j=1}^{k} \alpha_j R_j.$$

This is a non-standard inverse problems. We are dealing with linear combinations of $k+1$ resolvents and not only 2 as in classical identification problems

In $1-d$ the problem can be solved, thanks to the explicit representation of solutions[6]

$$-(\sigma(x)u_x)_x = f \text{ in } (0,1), \quad u_x(0) = 0 \text{ and } u(1) = 0. \tag{9}$$

$$u_x(x) = -\frac{1}{\sigma(x)} \int_0^x f(t)dt = -T_\sigma f \text{ a.e. } (0,1). \tag{10}$$

[6] Very much as in the context of homogenisation

$$||R_\sigma - R_{\widetilde{\sigma}}||_* = \left| \frac{1}{\widetilde{\sigma}(x)} - \frac{1}{\sigma(x)} \right|_{L^\infty((0,1))}.$$

$$\left(R_\tau f - \sum_{i=1}^{N} a_i R_i f \right)_x = \left(\sum_{i=1}^{N} \frac{a_i}{\sigma_i(x)} - \frac{1}{\tau(x)} \right) \int_0^x f(t)dt \ \text{ a.e. } (0,1) \quad (11)$$

$$\left| R_\tau - \sum_{i=1}^{N} a_i R_i \right|_* = \left| \sum_{i=1}^{N} \frac{a_i}{\sigma_i(x)} - \frac{1}{\tau(x)} \right|_{L^\infty((0,1))}. \quad (12)$$

This means that, in this $1d$ context, it suffices to apply the greedy algo in L^∞ within the class of coefficients $1/\sigma(x)$.

Multi-dimensional extension?

Consider control systems of the form

$$\begin{cases} x'(t) = A_j x(t) + Bu(t), \ 0 < t < T, \\ x(0) = x^0, \end{cases} \tag{13}$$

$j = 1, ..., k$.
Control operators:

$$P_j(x^0) = u_j(t), j = 1, ..., K.$$

Find a surrogate for

$$dist(P_j, span[P_\ell; \ell \neq j]) = sup_{||x^0||=1} dist(u_j(t), span[u_\ell(t) : \ell \neq j]).$$

We want an equivalent measure, but easier to be computed.

Our references

1. **Turnpike**
   1. A. Porretta & E. Z., SIAM J. Control. Optim., 51 (6) (2013), 4242 – 4273.
   2. E. Trélat & E. Z., JDE, 218 (2015) , 81-114.
   3. E. Z., Annual Reviews in Control, 44 (2017) 199 – 210.
2. **Averaged control**
   1. E. Z., Automatica, 50 (2014) 3077 – 3087
   2. Q. Lü & E. Z., J. Math. Pures Appl. 105 (2016) 367 – 414.
3. **Greedy control**
   1. M. Lazar & E. Z., Automatica 74 (2016) 327 – 340.
4. **Greedy + turnpike**
   1. M. Lazar, V. Hernández-Santamaria & E. Z., preprint