Optimal control problems with constraints in $L^1(\Omega)$ or $L^\infty(0, T; L^1(\Omega))$.

Analysis, discretization and optimization methods.

Mariano Mateos
(Universidad de Oviedo, España)

Based on a joint work with Eduardo Casas (U. de Cantabria) and Karl Kunisch (U. of Graz)

Control en Tiempos de Crisis
December 9, 2021

Outline

Outline

1. Introduction
   - Sparse solutions of naïve least squares problems
   - The parabolic optimal control problem
2. A problem governed by an elliptic equation
   - Statement and first properties
   - Optimality conditions and sparsity properties
3. Numerical discretization
   - Finite element discretization
   - Optimality conditions and sparsity properties
   - Convergence and error estimates
4. Optimization methods
   - Projected gradiente method
   - Semismooth Newton method
   - Numerical results

The most simple least-squares problem

A naïve problem

Given $\boldsymbol{x}_d \in \mathbb{R}^2$, find

$$\bar{\boldsymbol{x}} = \arg\min_{\boldsymbol{x} \in \mathbb{R}^2} \frac{1}{2} \|\boldsymbol{x} - \boldsymbol{x}_d\|_2^2.$$

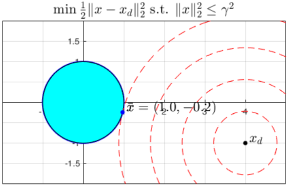
- Of course the solution is $\bar{\boldsymbol{x}} = \boldsymbol{x}_d$

Introducing constraints

A differentiable constraint $\|\mathbf{x}\|_2^2 \leq 1$

Given $\mathbf{x}_d \in \mathbb{R}^2$, find

$$\bar{\mathbf{x}} = \arg \min_{\|\mathbf{x}\|_2^2 \leq 1} \frac{1}{2}\|\mathbf{x} - \mathbf{x}_d\|_2^2.$$



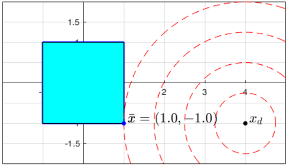
- It is quite typical that the solution has some “small” component. In this case $\bar{x}_2 = -0.2 \neq 0$.
- In practice, this “small” components, can be undesirable.

Bound constraints

Bound constraints. $\alpha \leq x_i \leq \beta$

Given $\mathbf{x}_d \in \mathbb{R}^2$, find

$$\bar{\mathbf{x}} = \arg \min_{\alpha \leq x_i \leq \beta} \frac{1}{2} \|\mathbf{x} - \mathbf{x}_d\|_2^2.$$



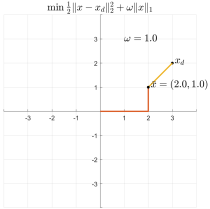
- Bound constraints do not help to get rid of “nonzero” components.

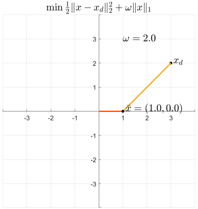
Looking for sparsity: ℓ^1 regularization.

ℓ^1 regularization

Given $\mathbf{x}_d \in \mathbb{R}^2$ and $\omega > 0$, find

$$\bar{\mathbf{x}} = \arg\min_{\mathbf{x}\in\mathbb{R}^2} \frac{1}{2}\|\mathbf{x} - \mathbf{x}_d\|_2^2 + \omega\|x\|_1.$$





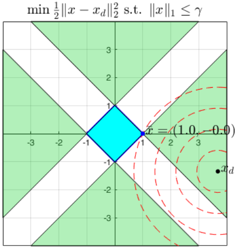
- “Tune” ω to get zero components.

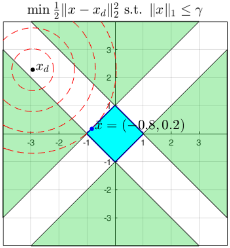
The approach of this talk: ℓ^1 constraints

A non-differentiable constraint: $\|\boldsymbol{x}\|_1 \leq 1$.

Given $\boldsymbol{x}_d \in \mathbb{R}^2$, find

$$\bar{\boldsymbol{x}} = \arg \min_{\|\boldsymbol{x}\|_1 \leq 1} \frac{1}{2}\|\boldsymbol{x} - \boldsymbol{x}_d\|_2^2.$$





- For data in the "green zone", the solution will have a "zero"component.

Outline

Model problem. Parabolic equation.

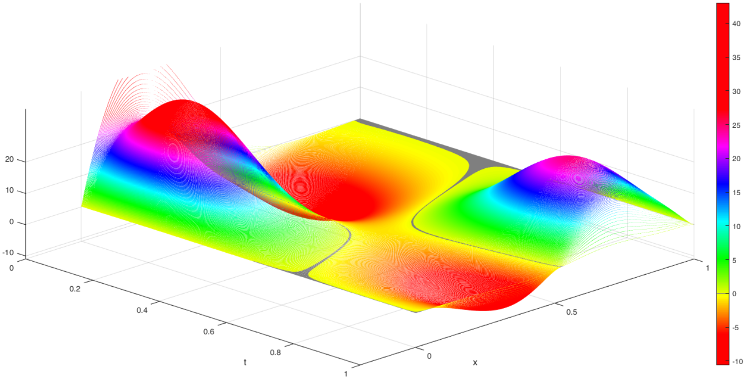
$$(P) \min_{u \in U_{\rm ad} \cap L^\infty(Q)} J(u) = \|y_u - y_d\|^2_{L^2(Q)} + \frac{\kappa}{2}\|u\|^2_{L^2(Q)}$$

$$\partial_t y_u + Ay_u + f(y_u) = u \text{ in } Q,\ \ y_u = 0 \text{ on } \Sigma, y_u(0) = y_0 \text{ in } \Omega$$

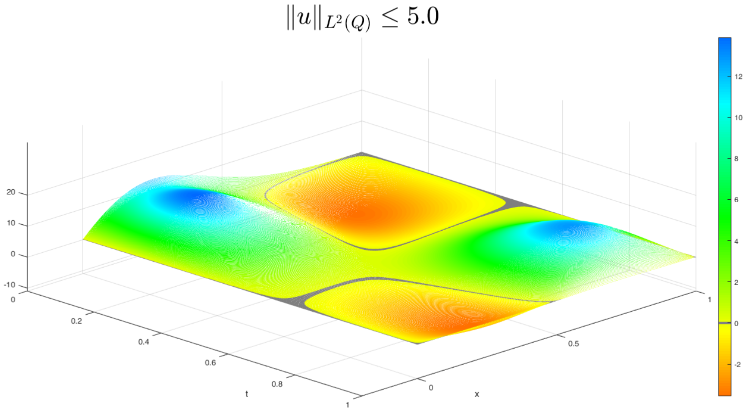
$$U_{\rm ad} = \{u \in L^\infty(0,T;L^1(\Omega)) :\ \|u(t)\|_{L^1(\Omega)} \le \gamma \text{ for a.a. } t \in (0,T)\}$$

- $\kappa > 0$,
- $\Omega \subset \mathbb{R}^d$, $d = 1, 2, 3$ bounded domain with boundary Γ, Lipschitz for $d \ge 2$,
- $T > 0$, $Q = \Omega \times (0,T)$, $\Sigma = \Gamma \times (0,T)$
- $y_d \in L^p(0,T;L^q(\Omega))$, for some p, q good enough,
- A an elliptic operator, f of class C^2 such that $f'(y) \ge C_f \in \mathbb{R}$ for all $y \in \mathbb{R}$.

Typical solution of an unconstrained problem (1D)

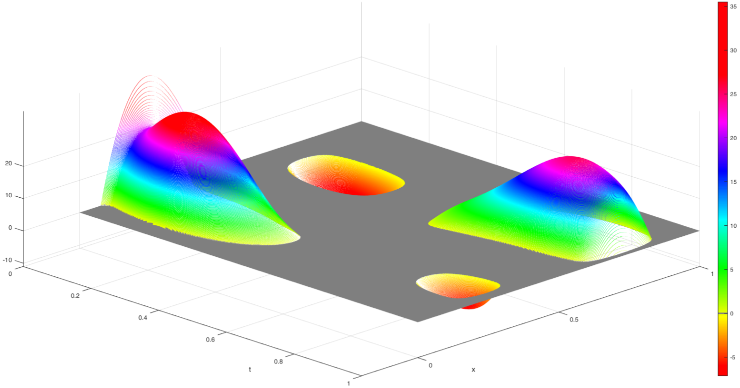


Typical solution of a problem (1D) with differentiable constraint

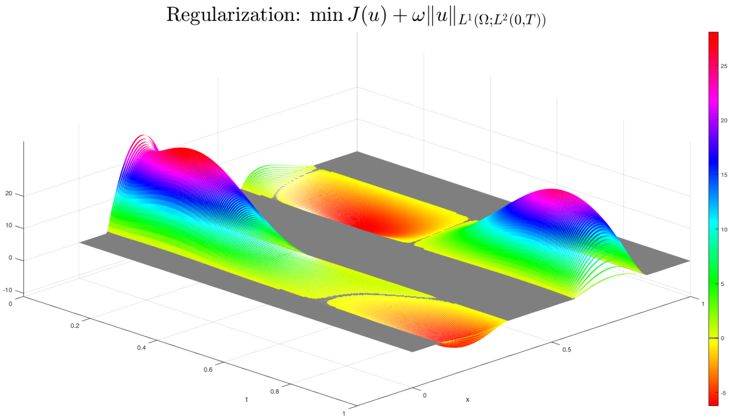


Typical solution of a problem (1D) with sparsity promoting term

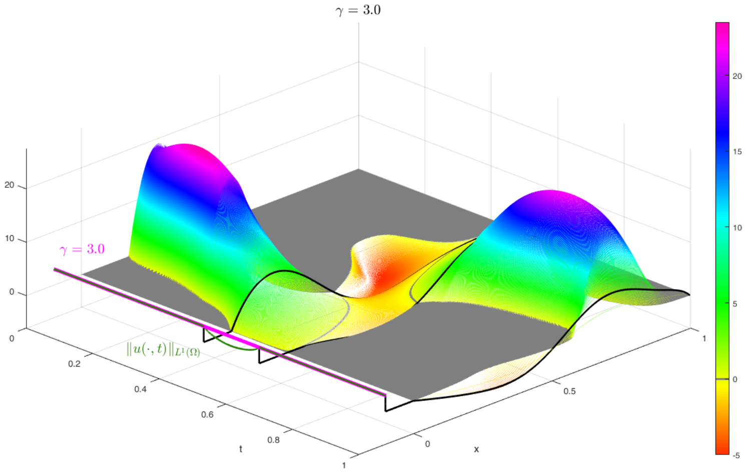
Regularization: $\min J(u) + \omega \|u\|_{L^1(Q)}$



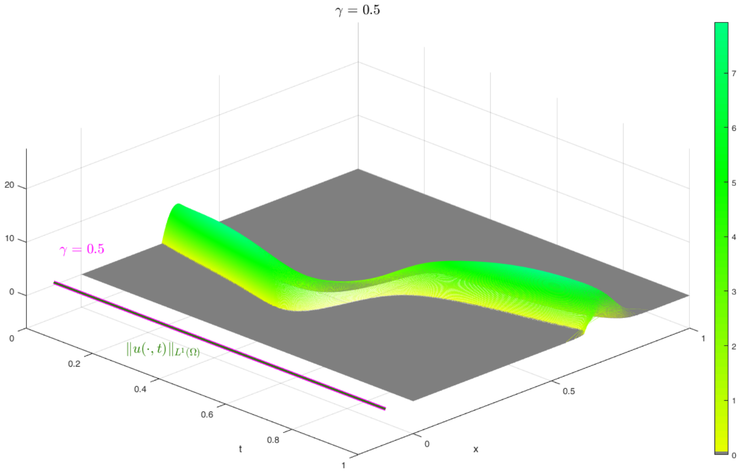
Typical solution of a problem (1D) with directional sparsity promoting term



Solution with the constraint $\|u(t)\|_1 \leq 3$ for all t (1D)



Solution with the constraint $\|u(t)\|_1 \leq 0.5$ for all t (1D)



Some comments and difficulties

- Allow strong and interesting non-linearities such as e^y, or y^{2n+1}, or $(y-a)(y-b)(y-c)$.
- Problem of existence of solution:
  - On one hand, J is not coercive in $L^\infty(Q)$
  - But on the other, $u \in L^2(Q)$ is not an appropriate datum for the nonlinear state equation.
  - A new truncation argument is used.
- The constraint is not differentiable
  - It will induce sparse solutions! $\mathrm{supp}(u(t))$ will be "small" when $\|u(t)\|_{L^1(\Omega)} = \gamma$

E. Casas and K. Kunisch (2021). "Optimal control of semilinear parabolic equations with non-smooth pointwise-integral control constraints in time-space". In: *Appl. Math. Optim.*, To appear

About the discretization

- Discretization in time: discontinuous Galerkin (dG0, implicit Euler).
  Continuous Galerkin (cG1, Crank-Nicholson) also possible.
- Discretization in space: finite elements.
- State and adjoint state, Lagrange P1 element (continuous piecewise linear functions)
- Control:
  - Piecewise constant functions works ok.
  - For Lagrange P1 element sparsity properties may be lost!!
  - We can solve this issue using the trapezoid rule to approximate $\|u(t)\|_{L^1(\Omega)}$ and $\|u(t)\|^2_{L^2(\Omega)}$.

E. Casas, K. Kunisch, and M. Mateos (2021). "Error estimates for the numerical approximation of optimal control problems with non-smooth pointwise-integral control constraints". In: *Submitted*

Outline

A more simple problem (drop time dependence)

Problem (E)

$$\min_{u \in U_{\rm ad} \cap L^\infty(\Omega)} J(u) = \|y_u - y_d\|^2_{L^2(\Omega)} + \frac{\kappa}{2}\|u\|^2_{L^2(\Omega)}$$

$$Ay_u + f(y_u) = u \text{ in } \Omega,\ \ y_u = 0 \text{ on } \Gamma$$

$$U_{\rm ad} = \{u \in L^1(\Omega) :\ \|u\|_{L^1(\Omega)} \leq \gamma\}$$

$y_d \in L^2(\Omega)$, f of class C^2, $f'(y) \geq 0$ for all $y \in \mathbb{R}$.

$$Ay(x) = -\sum_{i,j=1}^{n} \partial_{x_i}(a_{ij}\partial_{x_j} y(x)) + a_0(x)y(x),$$

where $a_0(x) \in L^\infty(\Omega)$, $a_0(x) \geq 0$, $a_{i,j} \in C^{0,1}(\bar{\Omega})$ and there exists $\Lambda > 0$ s.t.

$$\sum_{i,j=1}^{n} a_{ij}(x)\, \xi_i\xi_j \geq \Lambda|\xi|^2 \text{ for a.a. } x \in \Omega,\ \forall \xi \in \mathbb{R}^n$$

Study of the state equation.

Theorem

- *For every* $u \in L^2(\Omega)$ *there exists a unique* $y_u \in Y := H_0^1(\Omega) \cap C^{0,\nu}(\bar{\Omega})$*, for some* $\nu \in (0,1)$*, solution of the state equation.*
- $G : L^2(\Omega) \to Y$, $G(u) = y_u$ *is of class* C^2.
- *For all* $u, v \in L^2(\Omega)$, $G'(u)v = z_v$*, where* $z_v \in Y$ *is the unique solution of*

$$Az + f'(y_u)z = v \text{ in } \Omega, \;\; z = 0 \text{ on } \Gamma.$$

Parabolic problem

For the parabolic problem, we may lose differentiability for data $u, v \in L^2(Q)$ and a general f.

For conditions on f for the existence of solution, see e.g., H. Amann and P. Quittner (2006). "Optimal control problems governed by semilinear parabolic equations with low regularity data". In: *Adv. Differential Equations* 11.1, pp. 1–33.

Differentiability properties of the functional.

- $J : L^2(\Omega) \to \mathbb{R}$ is of class C^2

$$J'(u)v = \int_\Omega (\varphi_u + \kappa u)v dx$$

$$J''(u)v^2 = \int_\Omega [(1 - f''(y_u)\varphi_u)z_v^2 + kv^2]dx$$

where the adjoint state $\varphi_u \in Y$ is the unique solution of

$$A^*\varphi + f'(y_u)\varphi = y_u - y_d \text{ in } \Omega,\ \varphi = 0 \text{ on } \Gamma.$$

- Denote $j(u) = \|u\|_{L^1(\Omega)}$. j is convex and Lispchitz.

$$\lambda \in \partial j(u) \iff \lambda(x) \begin{cases} = 1 & \text{if } u(x) > 0 \\ \in [-1, 1] & \text{if } u(x) = 0 \\ = -1 & \text{if } u(x) < 0 \end{cases}$$

Existence of a global solution of (E)

The elliptic case is easier than the parabolic case. No need to truncate.

$$(E^2) \min_{u \in U_{\mathrm{ad}} \cap L^2(\Omega)} J(u)$$

- (E^2) has a solution $\bar{u}$ (direct method of the calculus of variations).
- Suppose $\bar{u}$ is a local solution of (E^2). Since $U_{\mathrm{ad}} \cap L^2(\Omega)$ is convex, then

  $$J'(\bar{u})(u-\bar{u}) = \int_\Omega (\bar{\varphi} + \kappa\bar{u})(u-\bar{u})dx \geq 0 \ \forall u \in U_{\mathrm{ad}} \cap L^2(\Omega),$$

  where $\bar{\varphi} = \varphi_{\bar{u}}$. Denote $\bar{\mu} = -\bar{\varphi} - \kappa\bar{u}$ and write

  $$\int_\Omega \bar{\mu}(u-\bar{u})dx \leq 0 \ \forall u \in U_{\mathrm{ad}} \cap L^2(\Omega).$$

- Test this inequality with $u(x) = \mathrm{sign}(\bar{\mu}(x))|\bar{u}(x)|$ to deduce $\bar{\mu}(x)$ and $\bar{u}(x)$ have the same sign.
- So for a.a $x \in \Omega$, we have

  $$\kappa|\bar{u}(x)| + |\bar{\mu}(x)| = |\kappa\bar{u}(x) + \bar{\mu}(x)| = |\bar{\varphi}(x)| \leq \|\bar{\varphi}\|_{L^\infty(\Omega)}$$

  And both $\bar{u}$ and $\bar{\mu}$ belong to $L^\infty(\Omega)$.

Outline

First order optimality conditions

Theorem

Suppose $\bar{u} \in U_{\mathrm{ad}} \cap L^\infty(\Omega)$ is a local solution of (E) (local in the sense of $L^2(\Omega)$). Then, there exist $\bar{y}, \bar{\varphi} \in Y$ and $\bar{\mu} \in L^\infty(\Omega)$ such that

$$A\bar{y} + f(\bar{y}) = \bar{u} \text{ in } \Omega, \;\; \bar{y} = 0 \text{ on } \Gamma$$

$$A^*\bar{\varphi} + f'(\bar{y})\bar{\varphi} = \bar{y} - y_d \text{ in } \Omega, \;\; \bar{\varphi} = 0 \text{ on } \Gamma.$$

$$\bar{\varphi} + \kappa\bar{u} + \bar{\mu} = 0$$

$$\int_\Omega \bar{\mu}(u - \bar{u})dx \leq 0 \;\forall u \in U_{\mathrm{ad}}$$

Proof

- $U_{\text{ad}} \cap L^\infty(\Omega) = \{u \in L^\infty(\Omega) : j(u) := \|u\|_{L^1(\Omega)} \leq \gamma\}$ is convex, so

$$J'(\bar{u})(u - \bar{u}) \geq 0 \ \forall u \in U_{\text{ad}} \cap L^\infty(\Omega).$$

  And using the expression for $J'(\bar{u})$, we can write

$$\int_\Omega (\bar{\varphi} + \kappa \bar{u})(u - \bar{u}) \geq 0 \ \forall u \in U_{\text{ad}} \cap L^\infty(\Omega). \qquad (^*)$$

- Using that $\bar{\varphi} + \kappa \bar{u} \in L^\infty(\Omega)$, we have

$$\int_\Omega (\bar{\varphi} + \kappa \bar{u})(u - \bar{u}) \geq 0 \ \forall u \in U_{\text{ad}}$$

  Proof: Given $u \in U_{\text{ad}}$, test $(^*)$ for $u_k = \text{proj}_{[-k,k]}(u(x)) \in U_{\text{ad}} \cap L^\infty(\Omega)$ and take limits as $k \to +\infty$.

- Define $\bar{\mu} = -\bar{\varphi} - \kappa \bar{u}$

Some consequences of the first order conditions

1. $\bar{u} = \mathrm{Proj}_{U_{\mathrm{ad}}}\left(-\frac{1}{\kappa}\bar{\varphi}\right)$ (projection in the sense of $L^2(\Omega)$).
2. $\bar{\mu}(x)\bar{u}(x) = |\bar{\mu}(x)\bar{u}(x)|$
3. $j(\bar{u}) < \gamma \Rightarrow \bar{\mu} \equiv 0$
4. Name $\bar{\omega} = \|\bar{\mu}\|_{L^\infty(\Omega)}$. We have a first **sparsity property:** If $j(\bar{u}) = \gamma$ and $\bar{\mu} \not\equiv 0$, then $\mathrm{supp}\bar{u} \subset \{x \in \Omega : \ |\bar{\mu}(x)| = \bar{\omega}\}$.
5. $-\kappa\bar{u}(x) = (\bar{\varphi}(x) - \bar{\omega})^+ + (\bar{\varphi}(x) + \bar{\omega})^-$ (*soft thresholding*)
6. 5, together with $\bar{\varphi} \in Y = H_0^1(\Omega) \cap C^{0,\nu}(\bar{\Omega})$, implies $\bar{u}, \bar{\mu} \in Y$.
7. 5 also implies a second **sparsity property:** $\bar{u}(x) = 0 \iff |\bar{\varphi}(x)| \le \bar{\omega}$.

Getting ready for the optimization algorithm

First order optimality conditions may suggest what kind of optimization algorithm we can use

- $\bar{u} = \mathrm{Proj}_{U_{\mathrm{ad}}}\left(-\frac{1}{\kappa}\bar{\varphi}\right)$, so a fixed-point algorithm could be used *as a last resource!*
- $\nabla J(u) = \varphi_u + \kappa u$, so we can use a projected gradient algorithm.
  Barzilai-Borwein strategy for the choice of the step size results in an incredibly good performance, but ...
  - The convergence is linear.
  - The computation of the gradient requires the solve of the *non-linear* state equation.
  - At each step, we have to project onto the $L^1(\Omega)$ ball
  - For the parabolic, we have to do these projections for every instant of time.
  - Our problem maybe/is non convex.
- A question arises: Could we use semismooth Newton?
  - Superlinear convergence.
  - Only linear PDEs must be solved.
  - Well known globalization and continuation techniques.
  - Successful for problems with bound constraints or problems with $L^1(\Omega)$ regularization.

More consequences of the first order conditions

Lemma

$$\int_\Omega \bar{\mu}(u - \bar{u})dx \leq 0 \; \forall u \in U_{\mathrm{ad}} \iff \bar{\mu} \in \bar{\omega}\partial j(\bar{u}), \; \textit{where} \; \bar{\omega} = \|\bar{\mu}\|_{L^\infty(\Omega)}$$

Connection with $L^1(\Omega)$ penalization for some $\omega > 0$

$$(E_\omega) \; \min_{u \in L^\infty(\Omega)} J(u) + \omega \|u\|_{L^1(\Omega)}$$

If $u^* \in L^\infty(\Omega)$ is a local solution of (E_ω), then there exist $y^* = y_{\bar{u}} \in Y$, $\varphi^* = \varphi_{\bar{u}} \in Y$ and $\lambda^* \in \partial j(u^*)$ such that

$$\varphi^* + \kappa u^* + \omega \lambda^* = 0.$$

Remark

- The solution $\bar{u}$ of (E), satisfies first order necessary conditions of problem $(E_{\bar{\omega}})$.
- The difficulty is that we do not know $\bar{\omega}$ beforehand.

Towards semismooth Newton

- Now it is standard to write the subdifferential condition using max and min

$$\bar{u}(x) - \max\{0, \bar{u}(x) + C(\bar{\mu}(x) - \bar{\omega})\} - \min\{0, \bar{u}(x) + C(\bar{\mu}(x) + \bar{\omega})\} = 0$$

  for a.a. $x \in \Omega$ and all $C > 0$
- On the other hand, $j(\bar{u}) < \gamma \Rightarrow \bar{\mu} \equiv 0$ implies

$$\bar{\omega} \geq 0, \ j(\bar{u}) - \gamma \leq 0, \ \bar{\omega} \cdot (j(\bar{u}) - \gamma) = 0$$

  This complementarity system can be written as

$$\bar{\omega} - \max\{0, \bar{\omega} + D \cdot (j(\bar{u}) - \gamma)\} = 0 \ \forall D > 0$$

First order optimality conditions revisited

- Suppose $\bar{u} \in U_{\text{ad}} \cap L^\infty(\Omega)$ is a local solution of (E). Then, there exist $\bar{y}, \bar{\varphi} \in Y$, $\bar{\mu} \in L^\infty(\Omega)$ and $\bar{\omega} \in \mathbb{R}$ such that

$$A\bar{y} + f(\bar{y}) = \bar{u} \text{ in } \Omega, \;\; \bar{y} = 0 \text{ on } \Gamma$$

$$A^*\bar{\varphi} + f'(\bar{y})\bar{\varphi} = \bar{y} - y_d \text{ in } \Omega, \;\; \bar{\varphi} = 0 \text{ on } \Gamma.$$

$$\bar{\varphi} + \kappa\bar{u} + \bar{\mu} = 0$$

$$\bar{u} - \max\{0, \bar{u} + C \cdot (\bar{\mu} - \bar{\omega})\} - \min\{0, \bar{u} + C \cdot (\bar{\mu} + \bar{\omega})\} = 0 \text{ for a.a. } x \in \Omega, \;\; \forall C > 0$$

$$\bar{\omega} - \max\{0, \bar{\omega} + D \cdot (j(\bar{u}) - \gamma)\} = 0 \; \forall D > 0$$

- We can try semismooth Newton on this system.

Second order conditions

$$C_{\bar{u}} = \left\{ v \in L^2(\Omega) : J'(\bar{u})v = 0 \text{ and } j'(\bar{u}; v) \begin{cases} = 0 & \text{if } j(\bar{u}) = \gamma \text{ and } \bar{\mu} \not\equiv 0 \\ \leq 0 & \text{if } j(\bar{u}) = \gamma \text{ and } \bar{\mu} \equiv 0 \end{cases} \right\}$$

- Necessary conditions: if $\bar{u}$ is a local solution of (E), then

$$J''(\bar{u})v^2 \geq 0 \, \forall v \in C_{\bar{u}}.$$

- Sufficient conditions: if $\bar{u}$ satisfies first order optimality conditions and

$$J''(\bar{u})v^2 > 0 \, \forall v \in C_{\bar{u}} \setminus \{0\}$$

then $\bar{u}$ is a local quadratic solution in the sense of $L^2(\Omega)$.

Why do we work so hard to get second order conditions

About the importance of second order conditions, see

E. Casas and F. Tröltzsch (2015). “Second order optimality conditions and their role in PDE control”. In: *Jahresber. Dtsch. Math.-Ver.* 117.1, pp. 3–44

Further uses of second order conditions

- For $\tau > 0$ define the extended cone.

$$C_{\bar{u}}^{\tau} = \Bigg\{ v \in L^2(\Omega) : |J'(\bar{u})v| \le \tau \|v\|_{L^2(\Omega)} \text{ and}$$

$$\left\{ \begin{array}{ll} |j'(\bar{u}; v)| \le \tau \|v\|_{L^2(\Omega)} & \text{if } j(\bar{u}) = \gamma \text{ and } \bar{\mu} \not\equiv 0 \\ j'(\bar{u}; v) \le \tau \|v\|_{L^2(\Omega)} & \text{if } j(\bar{u}) = \gamma \text{ and } \bar{\mu} \equiv 0 \end{array} \right. \Bigg\}$$

- If $\bar{u}$ satisfies first order optimality conditions and $J''(\bar{u})v^2 > 0 \ \forall v \in C_{\bar{u}} \setminus \{0\}$, then there exist $\varepsilon > 0$ and $\tau > 0$ such that

$$J''(u)v^2 \ge \tau \|v\|^2_{L^2(\Omega)} \ \forall u \in B_{L^2(\Omega)}(\bar{u}, \varepsilon), \ \forall v \in C_{\bar{u}}^{\tau}$$

- This will be very useful to obtain error estimates and convergence of Newton-like methods.

Outline

Finite element discretization

- Suppose Ω is convex and polygonal or polyhedral.
- Consider a quasi-uniform family of triangulations $\{\mathcal{T}_h\}_{h>0}$
- Define

$$a(y,\eta) = \int_\Omega \left(\sum_{i,j=1}^{n} a_{i,j} \partial_{x_i} y \partial_{x_j} \eta + a_0 y \eta \right) dx \ \forall y, \eta \in H^1(\Omega),$$

$$Y_h = \{ y_h \in C(\bar{\Omega}) : y_{h|T} \in P^1(T) \ \forall T \in \mathcal{T}_h \text{ and } y_h = 0 \text{ on } \Gamma \},$$

$$U_h = \{ u_h \in C(\bar{\Omega}) : u_{h|T} \in P^1(T) \ \forall T \in \mathcal{T}_h \}.$$

- Denote $\{e_i\}_{i=1}^N$, $N = N(h)$, the nodal basis associated to the nodes $\{x_i\}_{i=1}^N$ of the triangulation and suppose $\{x_i\}_{i=1}^{N_I}$, $N_I = N_I(h)$ are the interior nodes.

$$y_h(x) = \sum_{i=1}^{N_I} y_i e_i(x), \qquad u_h(x) = \sum_{i=1}^{N} u_i e_i(x)$$

$$\boldsymbol{y} = (y_1, \ldots, y_{N_I}, 0, \ldots, 0)^T, \qquad \boldsymbol{u} = (u_1, \ldots, u_N)^T$$

State discretization

- For every $u \in L^2(\Omega)$ there exists a unique $y_h(u) \in Y_h$ such that

$$a(y_h, \eta_h) + \int_\Omega f(y_h)\eta_h dx = \int_\Omega u\eta_h dx \; \forall \eta_h \in Y_h.$$

- Since Ω is convex, $y_u \in Y_2 := Y \cap H^2(\Omega)$ for every control $u \in L^2(\Omega)$, and

$$\|y_u - y_h(u)\|_{L^2(\Omega)} \leq C(\|u\|_{L^2(\Omega)} + 1)h^2.$$

Parabolic problem

For the parabolic problem, we also need for later use an error estimate in $L^\infty(Q)$.

- $G_h : U_h \to Y_h$, $G_h(u_h) = y_h(u_h)$ is of class C^2. For all $u_h, v_h \in U_h$, $z_h(v_h) = G'(u_h)v_h \in Y_h$ is the unique solution of the linear equation

$$a(z_h, \eta_h) + \int_\Omega f'(y_h(u_h))z_h\eta_h dx = \int_\Omega v_h\eta_h dx \; \forall \eta_h \in Y_h$$

Control discretization

- Piecewise constant discretization can be done, but let us focus on *continuous piecewise linear controls.*
- Notice that although we could try to use

$$\|u_h\|_{L^1(\Omega)} = \int_\Omega \left| \sum_{i=1}^N u_i e_i \right| dx$$

$$\|u_h\|^2_{L^2(\Omega)} = \sum_{i,j=1}^N u_i u_j \int_\Omega e_i e_j dx = \boldsymbol{u}^T \mathcal{M} \boldsymbol{u},$$

  where $\mathcal{M}$ is the mass matrix of the fem basis, $\boldsymbol{u} = (u_1, \ldots, u_N)^T$,
- the solutions of the discrete problem

$$\min_{u_h \in U_{\mathrm{ad}} \cap U_h} \frac{1}{2} \|y_h(u) - y_d\|^2_{L^2(\Omega)} + \frac{\kappa}{2} \|u_h\|^2_{L^2(\Omega)}$$

  need not satisfy the same sparsity properties as the solutions of the continuous problem.
- And also, the numerical computation of $\|u_h\|_{L^1(\Omega)}$ can be computationally expensive.

The "correct" way to discretize the control

- Instead, use the *diagonal* lump matrix $\mathcal{L}$, where $\ell_{ii} = \int_\Omega e_i dx$ and define

$$(u_h, v_h)_h = \boldsymbol{u}^T \mathcal{L} \boldsymbol{u} = \sum_{i=1}^N u_i v_i \ell_{ii}, \; \|u_h\|_h^2 = (u_h, u_h)_h, \qquad j_h(u_h) = \sum_{i=1}^N |u_i| \ell_{ii}$$

  These are approximations using the trapezoid rule. Let I_h be the nodal interpolator:

$$\|u_h\|_h^2 = \int_\Omega I_h(u_h^2) dx, \qquad j_h(u_h) = \int_\Omega I_h |u_h| dx$$

- We consider problem

$$(E_h) \min_{u_h \in U_{h,\mathrm{ad}}} J_h(u_h) := \frac{1}{2} \|y_h(u_h) - y_d\|_{L^2(\Omega)}^2 + \frac{\kappa}{2} \|u_h\|_h^2$$

  where

$$U_{h,\mathrm{ad}} = \{u_h \in U_h : \; j_h(u_h) \leq \gamma\}$$

- Notice that $j(u_h) \leq j_h(u_h)$, and hence $U_{h,\mathrm{ad}} \subset U_{\mathrm{ad}}$.

Existence of solution and differentiability properties

- $\|u_h\|_h$ is an equivalent norm in U_h to $\|u_h\|_{L^2(\Omega)}$, with equivalence constants independent of h.
  So J_h is coercive. Since $U_{h,\mathrm{ad}}$ is closed, there exists a solution $\bar{u}_h \in U_{h,\mathrm{ad}}$.
- $J_h : U_h \to \mathbb{R}$ is of class C^2. For all $u_h, v_h \in U_h$

$$J_h'(u_h)v_h = \int_\Omega \varphi_h(u_h)v_h dx + \kappa(u_h, v_h)_h,$$

  where $\varphi_h(u_h) \in Y_h$ is the unique solution of the linear equation

$$a(\eta_h, \varphi_h) + \int_\Omega f'(y_h(u_h))\varphi_h\eta_h dx = \int_\Omega (y_h(u_h) - y_d)\eta_h dx \; \forall \eta_h \in Y_h.$$

- $j_h : U_h \to \mathbb{R}$ is convex and Lipschitz. $\lambda_h \in \partial j_h(u_h)$, the convex subdifferential of u_h w.r.t. the scalar product $(\cdot, \cdot)_h$, iff

$$\lambda_i \begin{cases} = 1 & \text{if } u_i > 0 \\ \in [-1, 1] & \text{if } u_i = 0 \\ = -1 & \text{if } u_i < 0 \end{cases}$$

Dealing with the discrete adjoint state

- The expression $J_h'(u_h)v_h = \int_\Omega \varphi_h(u_h)v_h dx + \kappa(u_h, v_h)_h$ mixes two different scalar products, and can be difficult to handle.
- For $\varphi \in L^1(\Omega)$ we use Carstensen's quasi-interpolation operator $\Pi_h : L^1(\Omega) \to U_h$ by

$$\Pi_h\varphi = \sum_{i=1}^{N} \frac{\int_\Omega \varphi e_i dx}{\int_\Omega e_i dx}\, e_i$$

- Among many other properties, we have that

$$\int_\Omega \varphi v_h dx = (\Pi_h\varphi, v_h)_h \ \forall \varphi \in L^1(\Omega)\ \forall v_h \in U_h.$$

- Given $u_h \in U_h$, I will name $\phi_h(u_h) = \Pi_h\varphi_h(u_h) \in U_h$.
  We have $\boldsymbol{\phi} = \mathcal{L}^{-1}\mathcal{M}\boldsymbol{\varphi}$, i.e., $\phi_i = \frac{\int_\Omega \varphi_h(u_h)e_i dx}{\int_\Omega e_i dx}$, $i = 1, \ldots, N$.
- Now

$$J_h'(u_h)v_h = (\phi_h(u_h) + \kappa u_h, v_h)_h = \sum_{i=1}^{N}(\phi_i + \kappa u_i)v_i\ell_{ii}$$

Outline

First order optimality conditions

Let $\bar{u}_h \in U_{h,\mathrm{ad}}$ be a local solution of (P_h). Then there exist $\bar{y}_h, \bar{\varphi}_h \in Y_h$ and $\bar{\mu}_h \in U_h$ such that

$$a(\bar{y}_h, \eta_h) + \int_\Omega f(\bar{y}_h)\eta_h dx = \int_\Omega \bar{u}_h \eta_h dx \; \forall \eta_h \in Y_h,$$

$$a(\eta_h, \bar{\varphi}_h) + \int_\Omega f'(\bar{y}_h)\bar{\varphi}_h \eta_h dx = \int_\Omega (\bar{y}_h - y_d)\eta_h dx \; \forall \eta_h \in Y_h$$

$$\bar{\phi}_i + \kappa \bar{u}_i + \bar{\mu}_i = 0 \; \forall i = 1, \ldots, N$$

$$(\bar{\mu}_h, u_h - \bar{u}_h)_h \leq 0 \; \forall u_h \in U_{h,\mathrm{ad}}$$

where

$$\bar{\phi}_i = \frac{\int_\Omega \bar{\varphi}_h e_i dx}{\int_\Omega e_i dx}, \; i = 1, \ldots, N$$

Sparsity properties of the discrete solution

For a fixed index $i \in \{1, \ldots, N\}$, using test controls $u_h \in U_{h,\mathrm{ad}}$ satisfying $u_j = \bar{u}_j$ if $i \neq j$, we can deduce element-wise analogous properties to those of local solutions of (P):

1. For every $i = 1, \ldots, N$, $\bar{u}_i$ and $\bar{\mu}_i$ have the same sign.
2. $j_h(\bar{u}_h) < \gamma \Rightarrow \bar{\mu}_h \equiv 0$.
3. Name $\bar{\omega}_h = \|\bar{\mu}_h\|_{L^\infty(\Omega)} = \|\bar{\boldsymbol{\mu}}\|_\infty = \max\{|\bar{\mu}_i|,\ i = 1, \ldots, N\}$.
   If $j_h(\bar{u}_h) = \gamma$ and $\bar{\mu}_h \not\equiv 0$, then $\bar{u}_i \neq 0 \Rightarrow |\bar{\mu}_i| = \bar{\omega}_h$.
4. $-\kappa \bar{u}_i = (\bar{\phi}_i - \bar{\omega}_h)^+ + (\bar{\phi}_i + \bar{\omega}_h)^-$ *(Soft thresholding)*
5. ⑤ implies a second **sparsity property:** $\bar{u}_i = 0 \iff |\bar{\phi}_i| \leq \bar{\omega}_h$.
6. $\bar{\mu}_h \in \bar{\omega}_h \partial j_h(\bar{u}_h)$

Outline

Convergence

Theorem

For every $h > 0$, let $\bar{u}_h \in U_{h,\mathrm{ad}}$ be a local solution of (E_h). Then $\{\bar{u}_h\}_{h>0}$ is bounded in $L^\infty(\Omega)$ and if a sequence $\bar{u}_h \overset{*}{\rightharpoonup} \bar{u}$ in $L^\infty(\Omega)$ as $h \to 0$, then $\bar{u}$ is a local solution of (E) in the sense of $L^2(\Omega)$,

$$\lim_{h\to 0} \|\bar{u}_h - \bar{u}\|_{L^2(\Omega)} = 0, \qquad \lim_{h\to 0} J_h(\bar{u}_h) = J(\bar{u}) \tag{*}$$

Theorem

Let $\bar{u} \in U_{\mathrm{ad}}$ be a strict local minimum of (E) in the sense of $L^2(\Omega)$. Then there exist $h_0 > 0$ and $\varepsilon > 0$ and a sequence $\{\bar{u}_h\}_{h<h_0}$ of local minima of (E_h) such that (*) holds and

$$J_h(\bar{u}_h) = \min\{J_h(u_h) : \; u_h \in U_{h,\mathrm{ad}} \cap \bar{B}_{L^2(\Omega)}(\bar{u}, \varepsilon)\}.$$

Parabolic problem

Same results for the parabolic problem, but proofs require more steps and error estimates in $L^\infty(Q)$ for the state equation.

Error estimates

Theorem

Let $\bar{u}$ be a local solution of (E) in the $L^2(\Omega)$ sense such that $J''(\bar{u})v^2 > 0$ for all $v \in C_{\bar{u}} \setminus \{0\}$, and let $\{\bar{u}_h\}$ be a sequence of local solutions of problems (E_h) such that $\bar{u}_h \to \bar{u}$ in $L^2(\Omega)$. Then, there exist $h_0 > 0$ and $C > 0$ such that

$$\|\bar{u}_h - \bar{u}\|_{L^2(\Omega)} \leq Ch \, \forall h < h_0.$$

- Remark: In general, we only know that $\bar{u} \in H_0^1(\Omega)$, so $\|\bar{u} - \Pi_h \bar{u}\|_{L^2(\Omega)} \leq Ch$, and in this sense, the above estimate can be seen as “optimal”.

Parabolic problem

For the parabolic problem, we obtain

$$\|\bar{u}_h - \bar{u}\|_{L^2(Q)} \leq C(h + \tau) \, \forall h < h_0, \, \forall \tau < \tau_0,$$

where τ is the time step size.

Outline

Projected gradient method

Algorithm 1: Projected gradient with Barzilai-Borwein stepsize

Initialize $k = 0$, u_0, s_0, $g_0 = \nabla J(u_0)$
while *not convergence* **do**
  $\tilde{u}_{k+1} = u_k - s_k g_k$ `// Move in a descent direction`
  $u_{k+1} = \mathrm{Proj}_{U_{\mathrm{ad}}}(\tilde{u}_{k+1})$ `// Project onto the` L^1 `ball`
  $g_{k+1} = \nabla J(u_{k+1})$ `// Compute the new gradient`
  $\delta_u = u_{k+1} - u_k$, $\delta_g = g_{k+1} - g_k$ `// Compute new stepsize`
  Compute either $s_{k+1} = \frac{(\delta_g, \delta_u)}{(\delta_g, \delta_g)}$ or $s_{k+1} = \frac{(\delta_u, \delta_u)}{(\delta_u, \delta_g)}$
  $k := k + 1$
end

J. Barzilai and J. M. Borwein (1988). "Two-point step size gradient methods". In: *IMA J. Numer. Anal.* 8.1, pp. 141–148

L. Condat (2016). "Fast projection onto the simplex and the l_1 ball". In: *Math. Program.* 158.1-2, Ser. A, pp. 575–585

A. Ang (n.d.). *Notes.* URL: https://angms.science/notes.html

Comments

- The computation of the gradient requires solving the **nonlinear** state equation and the linear adjoint state equation.
- In finite dimension $\nabla J(u_h) = \phi_h(u_h) + \kappa u_h$, and $(\cdot,\cdot) = (\cdot,\cdot)_h$.
- The stepsize for steepest descent is $s_k = \arg\min_s J(u_k - s g_k)$. Barzilai-Borwein strategies satisfy respectively for each of the choices

$$s_{k+1} = \arg\min_s \|\delta_u - s\delta_g\|_h^2 \text{ or } s_{k+1} = \arg\min_s \|s\delta_u - \delta_g\|_h^2$$

- The projection onto the L^1-ball is computationally expensive.

Parabolic problem

For the parabolic problem, we have to project N_τ times at each step, where N_τ is the number of time steps of the discretization.

Outline

Semismooth Newton method (ongoing work)

$X^\infty = Y_2 \times Y_2 \times L^\infty(\Omega) \times L^\infty(\Omega) \times \mathbb{R}$, $X^2 = L^2(\Omega) \times L^2(\Omega) \times L^2(\Omega) \times L^2(\Omega) \times \mathbb{R}$.
For fixed positive C and D, define $F : X^\infty \to X^2$ as

$$F\begin{pmatrix} y \\ \varphi \\ u \\ \mu \\ \omega \end{pmatrix} = \begin{pmatrix} Ay + f(y) - u \\ A^*\varphi + f'(y)\varphi - y + y_d \\ \varphi + \kappa u + \mu \\ u - \max\{0, u + C(\mu - \omega)\} - \min\{0, u + C(\mu + \omega)\} \\ \omega - \max\{0, \omega + D(j(u) - \gamma)\} \end{pmatrix}$$

For given $\boldsymbol{x} = (y, \varphi, u, \mu, \omega) \in X^\infty$, we define the active and inactive sets as

$$\begin{aligned} \mathbb{A}^+ &= \{x \in \Omega : u(x) + C(\mu(x) - \omega) > 0\}, \\ \mathbb{A}^- &= \{x \in \Omega : u(x) + C(\mu(x) + \omega) < 0\}, \\ \mathbb{A} &= \mathbb{A}^+ \cup \mathbb{A}^-, \\ \mathbb{I} &= \Omega \setminus \mathbb{A}. \end{aligned}$$

Slant derivative of F

There exists $\rho > 0$ such that the function F is slantly differentiable in $B(\boldsymbol{x}, \rho)$. Moreover, for every $\boldsymbol{x} \in X^\infty$, define $G(\boldsymbol{x}) \in \mathcal{L}(X^\infty, X^2)$ as

$$G \begin{pmatrix} y \\ \varphi \\ u \\ \mu \\ \omega \end{pmatrix} \begin{pmatrix} \delta_y \\ \delta_\varphi \\ \delta_u \\ \delta_\mu \\ \delta_\omega \end{pmatrix} = \begin{pmatrix} A\delta_y + f'(y)\delta_y - \delta_u \\ A^*\delta_\varphi + f'(y)\delta_\varphi - (1 - f''(y)\varphi)\delta_y \\ \delta_\varphi + \kappa\delta_u + \delta_\mu \\ \delta_u \chi_{\mathbb{I}} - C\delta_\mu \chi_{\mathbb{A}} + C\delta_\omega(\chi_{\mathbb{A}^+} - \chi_{\mathbb{A}^-}) \\ \begin{cases} \delta_\omega & \text{if } \omega + D(j(u) - \gamma) \le 0 \\ -D\int_\Omega \lambda\delta_u dx & \text{if } \omega + D(j(u) - \gamma) > 0 \end{cases} \end{pmatrix}$$

where $\lambda \in \partial j(u)$. Then G is a slanting function for F.

Semismooth Newton's method

Algorithm 2: Semismooth Newton's method

Initialize $k = 0$, $x_0 = (y_0, \varphi_0, u_0, \mu_0, \omega_0)$
while *not convergence* **do**
  Compute the active and inactive sets $\mathbb{A}_k^+, \mathbb{A}_k^-, \mathbb{A}_k, \mathbb{I}_k$
  if $\omega_k + D \cdot (j(u_k) - \gamma) > 0$ **then**
    Choose $\lambda_k \in \partial j(u_k)$
  end
  Solve $G(x_k)\delta_x = -F(x_k)$
  Choose a stepsize $s_k \leq 1$
  $x_{k+1} = x_k + s_k\delta_x$
end

Solving the linear system $G(x_k)\delta_x = -F(x_k)$

$$A\delta_y + f'(y)\delta_y = -Ay - f(y) + u + \delta_u \text{ in } \Omega,\ \delta_y = 0 \text{ on } \Gamma \tag{1}$$

$$A^*\delta_\varphi + f'(y)\delta_\varphi = \delta_y(1 - f''(y)\varphi) - A^*\varphi - f'(y)\varphi + y - y_d \text{ in } \Omega,\ \delta_\varphi^{\mathbb{I}} = 0 \text{ on } \Gamma \tag{2}$$

$$\delta_\varphi + \kappa\delta_u + \delta_\mu = -\varphi - \kappa u - \mu \tag{3}$$

$$\delta_u\chi_{\mathbb{I}} - C\delta_\mu\chi_{\mathbb{A}} + C\delta_\omega(\chi_{\mathbb{A}^+} - \chi_{\mathbb{A}^-}) = -u + \max\{0, u + C(\mu - \omega)\} + \min\{0, u + C(\mu + \omega)\} \tag{4}$$

$$\delta_\omega = -\omega \text{ if } \omega + D(j(u) - \gamma) \le 0 \tag{5_1}$$

$$-\int_\Omega \lambda\delta_u dx = j(u) - \gamma \text{ if } \omega + D(j(u) - \gamma) > 0 \tag{5_2}$$

Solving the linear system $G(x_k)\delta_x = -F(x_k)$

- Notice that $\delta_u \chi_{\mathbb{I}} = -u\chi_{\mathbb{I}}$, so we split $\delta_u = -u\chi_{\mathbb{I}} + \delta_u \chi_{\mathbb{A}}$.
- Using this we split $\delta_y = \delta_y^{\mathbb{I}} + \delta_y^{\mathbb{A}}$ and $\delta_\varphi = \delta_\varphi^{\mathbb{I}} + \delta_\varphi^{\mathbb{A}}$
  - Fixed part:

$$\begin{aligned} A\delta_y^{\mathbb{I}} + f'(y)\delta_y^{\mathbb{I}} =& - Ay - f(y) + u - u\chi_{\mathbb{I}} \text{ in } \Omega,\ \delta_y^{\mathbb{I}} = 0 \text{ on } \Gamma, \\ A^*\delta_\varphi^{\mathbb{I}} + f'(y)\delta_\varphi^{\mathbb{I}} =& \delta_y^{\mathbb{I}}(1 - f''(y)\varphi) - A^*\varphi - f'(y)\varphi + y - y_d \text{ in } \Omega,\ \delta_\varphi^{\mathbb{I}} = 0 \text{ on } \Gamma, \end{aligned}$$

  - Part linear in $\delta_u \chi_{\mathbb{A}}$

$$\begin{aligned} A\delta_y^{\mathbb{A}} + f'(y)\delta_y^{\mathbb{A}} =& \delta_u \chi_{\mathbb{A}} \text{ in } \Omega,\ \delta_y^{\mathbb{A}} = 0 \text{ on } \Gamma \\ A^*\delta_\varphi^{\mathbb{A}} + f'(y)\delta_\varphi^{\mathbb{A}} =& \delta_y^{\mathbb{A}}(1 - f''(y)\varphi) \text{ in } \Omega,\ \delta_\varphi^{\mathbb{A}} = 0 \text{ on } \Gamma \end{aligned}$$

- We can use the fourth equation to deduce $\delta_\mu = -\mu + (\chi_{\mathbb{A}^+} - \chi_{\mathbb{A}^-})(\omega + \delta_\omega)$ in $\mathbb{A}$.
- We will use the third equation to compute $\delta_\mu = -\delta_\varphi - \varphi - \mu$ in $\mathbb{I}$ once δ_φ has been computed.

Case 1

If $\omega + D \cdot (j(u) - \gamma) \leq 0$, then $\delta_\omega = -\omega$. Using (4) we have $\delta_\mu = -\mu$ in $\mathbb{A}$. The third equation can be written now as

$$\delta_\varphi^{\mathbb{A}} + \kappa \delta_u = -\delta_\varphi^{\mathbb{I}} - \varphi - \kappa u \text{ in } \mathbb{A}$$

This equation, together with the two linear PDEs for $\delta_y^{\mathbb{A}}$ and $\delta_\varphi^{\mathbb{A}}$ are the optimality system of the **unconstrained linear quadratic control problem**

$$\min_{\delta_u \in L^2(\mathbb{A})} \frac{1}{2} \int_\Omega (1 - f''(y)\varphi) {\delta_y^{\mathbb{A}}}^2 dx + \frac{\kappa}{2} \int_{\mathbb{A}} \delta_u^2 dx - \int_{\mathbb{A}} \delta_u \Big(- \delta_\varphi^{\mathbb{I}} - \varphi - \kappa u \Big) dx$$

$$A\delta_y^{\mathbb{A}} + f'(y)\delta_y^{\mathbb{A}} = \delta_u \chi_{\mathbb{A}} \text{ in } \Omega,\ \delta_y^{\mathbb{A}} = 0 \text{ on } \Gamma$$

Use Conjugate Gradient to solve $\mathcal{A}\delta_u^{\mathbb{A}} = b$, where $\mathcal{A}\delta_u^{\mathbb{A}} = (\delta_\varphi^{\mathbb{A}} + \kappa \delta_u^{\mathbb{A}})\chi_{\mathbb{A}}$ and $b = (-\delta_\varphi^{\mathbb{I}} - \varphi - \kappa u)\chi_{\mathbb{A}}$.

Compare with §14 in E. Casas and M. Mateos (2017). "Optimal control of partial differential equations". In: *Comp. Math., Num. An. and Apps.* Vol. 13. SEMA SIMAI Springer Ser. Pp. 3–59

Case 2

Suppose now that $\omega + D \cdot (j(u) - \gamma) > 0$. Now equation (3) reads like

$$\delta_\varphi^{\mathbb{A}} + \kappa \delta_u = -\delta_\varphi^{\mathbb{I}} - \varphi - \kappa u - (\chi_{\mathbb{A}^+} - \chi_{\mathbb{A}^-})(\omega + \delta_\omega) \text{ in } \mathbb{A}$$

For the simplest choice of $\lambda \in \partial j(u)$, eq. (5) is

$$\int_\Omega (\chi_{\mathbb{A}^+} - \chi_{\mathbb{A}^-})\delta_u dx = \gamma - j(u)$$

These two equations, together with the two linear PDEs for $\delta_y^{\mathbb{A}}$ and $\delta_\varphi^{\mathbb{A}}$ are the optimality system of the **linearly constrained linear quadratic control problem**

$$\min_{\delta_u \in L^2(\mathbb{A})} \frac{1}{2} \int_\Omega (1 - f''(y)\varphi) {\delta_y^{\mathbb{A}}}^2 dx + \frac{\kappa}{2} \int_{\mathbb{A}} \delta_u^2 dx - \int_{\mathbb{A}} \delta_u \Big(-\delta_\varphi^{\mathbb{I}} - \varphi - \kappa u - (\chi_{\mathbb{A}^+} - \chi_{\mathbb{A}^-})\omega \Big) dx$$

$$\text{s.t.} \quad \int_{\mathbb{A}} (\chi_{\mathbb{A}^+} - \chi_{\mathbb{A}^-})\delta_u dx = \gamma - j(u)$$

Here, δ_ω plays the role of the Lagrange multiplier related to the linear constraint.

Outline

Numerical comparison of the optimization methods.

Consider the problem (E) with data $\Omega = (0,1)^n$, $n = 1, 2, 3$, $A = -\Delta$, $\kappa = 1$, $y_d(x) = \prod_{i=1}^{n} \sin(2\pi x_i)$ and $f(y) = e^y$ and $\gamma = 0.005 \cdot 2^{-n}$.

- We can solve the problem with both methods.
- Projected gradient with Barzilai-Borwein is faster in almost all cases for this example.
- But in dimension 3 with $h = 1/2^6$ (274 625 nodes) Newton's method is faster.
- The main reason is that in this case, the computation of y_u is very expensive.

Parabolic problem

Ongoing work. Since in this case the solution of the nonlinear equation is much more "expensive" in the parabolic case, we expect Newton's method to be faster.

Observed error estimates (Parabolic problem)

$n = 1$, $\Omega = (0, 1)$, $A = -\Delta$, $f \equiv 0$ (linear equation) $y_0 = 0$, $T = 1$, $\kappa = 10^{-4}$ and

$$y_d(x, t) = \exp(-20[(x - 0.2)^2 + (t - 0.2)^2]) + \exp(-20[(x - 0.7)^2 + (t - 0.9)^2])$$

Data from Ed. Casas, R. Herzog, and G. Wachsmuth (2017). "Analysis of spatio-temporally sparse optimal control problems of semilinear parabolic equations". In: *ESAIM Control Optim. Calc. Var.* 23.1, pp. 263–295

$h_i = \tau_i$	$\|\bar{u}_{\sigma_{I,I}} - \bar{u}_{\sigma_{i,i}}\|_{L^2(Q)}$	EOC
2^{-8}	1.76E − 1	−
2^{-9}	8.93E − 2	0.98
2^{-10}	4.49E − 2	0.99

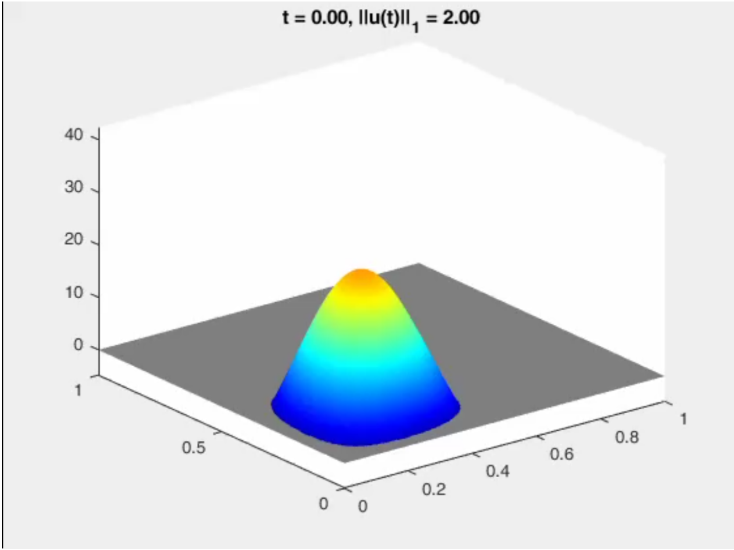
Table: Experimental order of convergence. Simultaneous refinement in space and time.

h_i	$\|\bar{u}_{\sigma_{I,I}} - \bar{u}_{\sigma_{i,I}}\|_{L^2(Q)}$	EOC
2^{-8}	1.10E − 02	−
2^{-9}	3.87E − 03	1.51
2^{-10}	1.34E − 03	1.53

τ_i	$\bar{u}_{\sigma_{I,I}} - \bar{u}_{\sigma_{I,i}}$	EOC
2^{-8}	1.76E − 1	−
2^{-9}	8.93E − 2	0.98
2^{-10}	4.49E − 2	0.99

Table: Left, refinement only in space. Right, refinement only in time.

Sparsity properties ($n = 2$, parabolic problem).



Σας ευχαρστω για προσοχη σας
Obrigado pela sua atenção
Grazie per la vostra attenzione
Merci beaucoup pour votre attention
Eskerrik asko zuen arretagatik
Grazas pola túa atención
Vielen Dank für Ihre Aufmerksamkeit
Gràcies per la seva atenció
Thank you for your attention
Gracias por su atención