Optimal control of parabolic equations using spectral calculus

Ivica Nakić

Faculty of Science, University of Zagreb

Joint work (in progress) with L. Grubišić, M. Lazar and M. Tautenhahn

Operator Theory and Krein Spaces, TU Wien, 2019

Financed by the HRZZ project IP-2016-06-2468

The problem

Inital condition or starting optimal control: solve the following problem

$$\min_{u \in \mathcal{H}} \{J(u) \colon \|y(T) - y^*\| \leq \varepsilon\},$$

where (in the weak sense)

$$\begin{cases} y'(t) + Ay(t) = f(t) & \text{for } 0 \leq t \leq T, \\ y(0) = u, \end{cases}$$

$$J(u) = \frac{\alpha}{2}\|u\|^2 + \frac{1}{2}\int_0^T \beta(t)\|y(t) - w(t)\|^2 \, \mathrm{d}t.$$

Here we assume: A selfadjoint, lower semi-bounded operator on a Hilbert space $\mathcal{H}$, $f \in L^2((0,T);\mathcal{H})$. Parameters: y^* is the target state, $\varepsilon > 0$ is the tolerance, $\alpha > 0$, $\beta \in L^\infty((0,T);[0,\infty))$ are weights, and $w \in L^2((0,T);\mathcal{H})$ is the desired trajectory of the system.

The solution

The solution $\hat{u}$ of the problem is given by

$$S_T\hat{u} = (\hat{\mu} + B)^{-1}(\hat{\mu} y_h^* + b) - y^* - y_h^*,$$

where

$$B = \alpha I + \int_0^T \beta(t) S_{2t}\, \mathrm{d}t, \quad b = \int_0^T \beta(t) S_{T+t} w_h(t)\, \mathrm{d}t,$$

$$y_h^* = y^* - \int_0^T S_\tau f(\tau)\, \mathrm{d}\tau, \quad w_h = w - \int_0^{\cdot} S_\tau f(\tau)\, \mathrm{d}\tau,$$

$\hat{\mu}$ is the unique solution of

$$G(\mu) := \|y_h^* - (\mu + B)^{-1}(\mu y_h^* + b)\| = \varepsilon,$$

and $\{S_t\}$ is the semigroup generated by $-A$.

Remarks

→ The optimal final state $\hat{y}$ is given by
$\hat{y} = (\hat{\mu} + B)^{-1}(\hat{\mu} y_h^* + b)$.

Remarks

→ The optimal final state $\hat{y}$ is given by $\hat{y} = (\hat{\mu} + B)^{-1}(\hat{\mu} y_h^* + b)$.

→ The solution of the unconstrained problem ($\varepsilon = \infty$) is given by the same formula with $\hat{\mu} = 0$.

Remarks

→ The optimal final state $\hat{y}$ is given by
$\hat{y} = (\hat{\mu} + B)^{-1}(\hat{\mu} y_h^* + b)$.

→ The solution of the unconstrained problem ($\varepsilon = \infty$) is given by the same formula with $\hat{\mu} = 0$.

→ The function G is decreasing.

Remarks

- → The optimal final state $\hat{y}$ is given by $\hat{y} = (\hat{\mu} + B)^{-1}(\hat{\mu} y_h^* + b)$.
- → The solution of the unconstrained problem ($\varepsilon = \infty$) is given by the same formula with $\hat{\mu} = 0$.
- → The function G is decreasing.
- → Let $g(\mu) = y_h^* - (\mu + B)^{-1}(\mu y_h^* + b)$, so $G(\mu) = \|g(\mu)\|$. Then $g(\mu) = y_h^* - x$, where x is the solution of the equation

$$(\mu + B)\, x = \mu y_h^* + b,$$

hence the calculation of $G(\mu)$ reduces to solving a linear equation.

Remarks

- The optimal final state $\hat{y}$ is given by $\hat{y} = (\hat{\mu} + B)^{-1}(\hat{\mu} y_h^* + b)$.
- The solution of the unconstrained problem ($\varepsilon = \infty$) is given by the same formula with $\hat{\mu} = 0$.
- The function G is decreasing.
- Let $g(\mu) = y_h^* - (\mu + B)^{-1}(\mu y_h^* + b)$, so $G(\mu) = \|g(\mu)\|$. Then $g(\mu) = y_h^* - x$, where x is the solution of the equation

$$(\mu + B)\, x = \mu y_h^* + b,$$

  hence the calculation of $G(\mu)$ reduces to solving a linear equation.
- The optimal final state is the solution of the equation

$$(\hat{\mu} + B)\, x = \hat{\mu} y_h^* + b,$$

  hence we obtain it for free.

Remarks

→ In applications, it is enough to find good μ ($\mu \geq \hat{\mu}$, μ close to $\hat{\mu}$), not the optimal one. One choice is to take

$$\mu = \frac{\|By_h^* - b\|}{\varepsilon} + \alpha + \int_0^T \frac{1}{2t}\beta(t)\mathrm{e}^{-2t\kappa}\,\mathrm{d}t,$$

where $A \geq \kappa$.

Remarks

ConDyS

- → In applications, it is enough to find good μ ($\mu \geq \hat{\mu}$, μ close to $\hat{\mu}$), not the optimal one. One choice is to take

$$\mu = \frac{\|By_h^* - b\|}{\varepsilon} + \alpha + \int_0^T \frac{1}{2t}\beta(t)\mathrm{e}^{-2t\kappa}\,\mathrm{d}t,$$

  where $A \geq \kappa$.
- → In applications, we can use $B = \tilde{\beta}_0(A)$, where $\tilde{\beta}_0(\lambda) = \alpha + \int_0^T \beta(t)\exp(-2t\lambda)\,\mathrm{d}t$ and we can approximate b by ($w_h(t) \approx \sum_{i=1}^N w_i \chi_{[t_{i-1},t_i]}$)

$$\sum_{i=1}^N \tilde{\beta}_i(A) w_i, \text{ where } \tilde{\beta}_i(\lambda) = \int_{t_{i-1}}^{t_i} \beta(t)\exp(-(T+t)\lambda)\,\mathrm{d}t,$$

An example

Let A be positive definite operator and let $f = 0$. We take $\beta = \chi_{[T/3, 2T/3]}$ and assume that w does not depend on time.
Then

$$B = \alpha + \frac{1}{2} A^{-1} S_{2T/3}(I - S_{2T/3}),$$
$$b = A^{-1} S_{4T/3}(I - S_{T/3}) w.$$

Sensitivity of the problem

Let us perturbe all the parameters of the problem with perturbations $< \nu$, in respective norms, such that that the perturbed problem has the same structure (*A* still selfadjoint etc.). Then

Theorem

For small enough $\nu > 0$ the optimal solutions of the original and the perturbed problem differ by $< C\nu$ with explicit C.

Idea of the proof

Proof is mostly geometry.
We can assume $f = 0$. We work in $\tilde{\mathcal{H}} = \operatorname{Ran} S_T$, with the scalar product $\langle S_T^{-1}\cdot, S_T^{-1}\cdot\rangle$, and define

$$\omega(\cdot) = J(S_T^{-1}\cdot).$$

Then $\hat{y}$ is the unique solution of

$$\min_{y\in\tilde{\mathcal{H}}} \{\omega(y) \colon \|y - y^*\| \leq \varepsilon\}\,.$$

We define

$$W_c = \{y \in \tilde{\mathcal{H}} \colon \omega(y) \leq c\}.$$

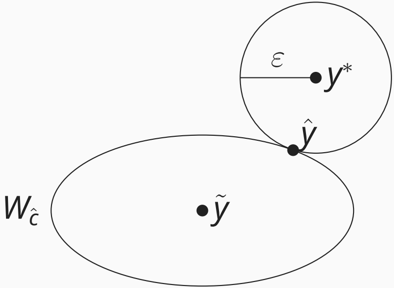
Let $\Pi_c(x)$ be a projection of x to W_c. Then

$$\Pi_{\hat{c}}(y^*) = \hat{y}, \text{ where } \hat{c} = \omega(\hat{y}).$$

Idea of the proof

The second geometric ingredient is the following result:

there exists $\hat{\gamma} > 0$ such that $y^* - \hat{y} = \hat{\gamma} \nabla \omega(\hat{y})$.



Non-homogeneous boundary condition

Suppose we have

$$\begin{cases} y'(t) + Ly(t) = 0 & \text{for } 0 \leq t \leq T, \\ Gy(t) = g(t), \\ y(0) = u, \end{cases}$$

where G is a boundary trace operator.
We assume that (L, G) forms so-called boundary control system.

Definition

A boundary control system is a pair of operators (L, G) where $L \in L(Z, X)$ and $G \in L(Z, U)$, if there exists $\beta \in \mathbb{C}$ such that:
G is surjective, Ker G is dense in X, $\beta - L$ restricted to Ker G is surjective, and $\mathrm{Ker}(\beta - L) \cap \mathrm{Ker}\, G = \{0\}$.

Non-homogeneous boundary condition

We define the operator A on X by $Au = Lu$ and $D(A) = \operatorname{Ker} G$. Let X_{-1} be the extrapolation space corresponding to A and let $\hat{A}$ be the extension of A to X_{-1}. There exists a unique $T \in L(U, X_{-1})$ such that the problem can be written as

$$\dot{y}(t) + \hat{A}y(t) = Tg(t).$$

So we are back in business if A is a lower semi-bounded selfadjoint operator ($\hat{A}$ inherits the properties of A) but X_{-1} is not a nice space to work with.

Non-homogeneous boundary condition

We define the operator A on X by $Au = Lu$ and $D(A) = \operatorname{Ker} G$. Let X_{-1} be the extrapolation space corresponding to A and let $\hat{A}$ be the extension of A to X_{-1}. There exists a unique $T \in L(U, X_{-1})$ such that the problem can be written as

$$\dot{y}(t) + \hat{A}y(t) = Tg(t).$$

So we are back in business if A is a lower semi-bounded selfadjoint operator ($\hat{A}$ inherits the properties of A) but X_{-1} is not a nice space to work with.

Fear not, B and b can still be seen as operator/element in X using the fact that $Tg(\cdot) = (\beta - \hat{A})h(\cdot)$, for h function with values in X.
If g is constant in time, B and b have nice formulas.

What's the point?

Standard (and much more general) solution is based on Lagrange multipliers, and to find the solution one needs to solve two coupled time-dependent problems, the original and the adjoint one.
We need to solve just one stationary problem, but with a more complicated operator and vector.

What's the point?

Standard (and much more general) solution is based on Lagrange multipliers, and to find the solution one needs to solve two coupled time-dependent problems, the original and the adjoint one.
We need to solve just one stationary problem, but with a more complicated operator and vector.

Should have some advantages. Constructing efficient numerical procedure is a work in progress. To learn about one possible approach go to Luka’s talk.

Look ahead

→ proper numerics

Look ahead

→ proper numerics
→ distributed control

Look ahead

→ proper numerics
→ distributed control
→ boundary control

Look ahead

- → proper numerics
- → distributed control
- → boundary control
- → non-selfadjoint case

ConDyS

Thanks for the attention!