CONTROL AND OPTIMIZATION FOR NON-LOCAL AND FRACTIONAL DIFFERENTIAL EQUATIONS

Umberto Biccari and Enrique Zuazua

Chair of Computational Mathematics, Bilbao, Basque Country, Spain

Chair for Dynamics, Control and Numerics, Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany.

Universidad Autónoma de Madrid, Spain.

umberto.biccari@deusto.es enrique.zuazua@fau.de

cmc.deusto.es dcn.nat.fau.eu

PART IV: numerical approximation of differential equations and numerical control

LECTURE 12: numerical control of PDE

The gradient descent method

Numerical approximation of controls for PDE

We have seen that the controllability of a PDE is equivalent to the minimization of a convex functional

$$F(u) = \frac{1}{2}\int_0^T |Bu|^2\,dt + \frac{1}{2}|x(T)|^2 \qquad \textbf{direct functional}$$
$$F : L^2(0,T;U) \to \mathbb{R}$$

$$J(p_T) = \frac{1}{2}\int_0^T |B^*p|^2\,dt + \langle x_0, p(0)\rangle \qquad \textbf{adjoint functional}$$
$$J : H \to \mathbb{R}$$

We then look for **efficient computational algorithms** to find the minimizer of J and, from it, the control for our equation.

The gradient descent method

Consider a functional $J : X \to \mathbb{R}$ fulfilling two main assumptions

$|\nabla J(u) - \nabla J(v)| \leq L|u - v|$ **Lipschitz gradient**

$\langle \nabla J(u) - \nabla J(v), u - v\rangle \geq \alpha |u - v|^2$ **α-convexity**

Gradient descent scheme

$$u_{k+1} = u_k - \rho \nabla J(u_k)$$

Under the above assumptions, we have that, taking $0 < \rho < 2/(\alpha + L)$, the iterative method converges to the minimizer u^* of J

$$|u_k - u^*| \leq \left(1 - \frac{2\rho\alpha L}{\alpha + L}\right)^{\frac{k}{2}} |u_0 - u^*|.$$

If $\rho = 2/(\alpha + L)$, then

$$|u_k - u^*| \leq \left(\frac{\sigma - 1}{\sigma + 1}\right)^{\frac{k}{2}} |u_0 - u^*|,$$

where $\sigma = \alpha^{-1}L$ is the **conditioning number** of the problem.

The gradient descent method

We see that the convergence of the scheme is influenced by three factors.

- The **step-size** ρ, that shall be taken small enough. We stress that ρ does not need to be taken convex, but it can be $\rho = \rho_k$, i.e. changing at each iteration.
- The distance $|u_0 - u^*|$ of the initialization u_0 from the minimizer u^*.
- The conditioning of the problem, through the constant

$$\mathcal{C}_{GD} := \frac{\sigma - 1}{\sigma + 1}$$

Remark

Notice that $\lim_{\sigma \to +\infty} \mathcal{C}_{GD} = 1$, meaning that the convergence of the algorithm deteriorates for badly-conditioned problems.

ILLUSTRATIVE EXAMPLE:

$$\min_{x \in \mathbb{R}} \left(\frac{1}{2} x^\top Q_\tau x - b^\top x \right)$$

$$Q_\tau = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \tau & 0 \\ 0 & 0 & \tau^2 \end{pmatrix} \quad b = -\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

$$\sigma = \frac{\lambda_{max}}{\lambda_{min}} = \tau^2$$

τ	iterations	ρ
2	27	4
5	161	25
10	633	100
20	2511	400
50	15619	2500

Meza, Steepest descent, 2010

Computation of the gradient

How can we compute the gradient of the functional?

CASE 1: computation of ∇F

STEP 1: we first measure the rate of change of F in any direction $\zeta \in L^2(0,T;U)$ by calculating the directional derivative as follows:

$$D_\zeta F(u) = \frac{\mathrm{d}}{\mathrm{d}\epsilon} F(u+\epsilon\zeta)\Big|_{\epsilon=0} = \int_0^T \langle u, \zeta\rangle_U \, dt + \langle x(T), z(T)\rangle_H, \tag{1}$$

where $z \in L^2(0,T;H)$ is the solution of the following equation

$$\begin{cases} z'(t) = Az(t) + B\zeta, & 0 < t < T, \\ z(0) = 0. \end{cases} \tag{2}$$

STEP 2: let $p \in L^2(0,T;H)$ be the solution of the adjoint problem

$$\begin{cases} p'(t) = -A^* p(t), & 0 < t < T, \\ p(T) = -x(T). \end{cases} \tag{3}$$

Multiplying (2) by p_ν and integrating by parts we obtain

$$\langle x(T), z(T)\rangle_H = -\int_0^T \langle \zeta, B^* p\rangle_U \, dt.$$

Computation of the gradient

How can we compute the gradient of the functional?

CASE 1: computation of ∇F

STEP 3: replacing this last expression in (1), we obtain that, for any $\zeta \in L^2(0,T;U)$,

$$D_\zeta F_\nu(u) = \int_0^T \langle u - B^* p, \zeta \rangle_U \, dt.$$

the gradient of the functional F is then given by the expression $\nabla F(u) = u - B^* p$. Consequently, the GD scheme to minimize F becomes

$$u_{k+1} = u_k - \rho \Big(u_k - B^* p_k \Big).$$

Applying the GD scheme for minimizing the functional $F(u)$ requires to solve at each iteration the coupled system

$$\begin{cases} x'(t) = Ax(t) + Bu(t), & 0 < t < T, \\ p'(t) = -A^* p(t), & 0 < t < T, \\ x(0) = x_0, \quad p(T) = -x(T) & \end{cases}$$

Computation of the gradient

How can we compute the gradient of the functional?

CASE 2: computation of ∇J

With a similar procedure as the one to obtain ∇F, we can show that

$$\nabla J(p_T) = x(T),$$

with

$$\begin{cases} x'(t) = Ax(t) + BB^* p(t), & 0 < t < T, \\ p'(t) = -A^* p(t), & 0 < t < T, \\ x(0) = x_0, \quad p(T) = p_T \end{cases}$$

Consequently, the GD scheme to minimize J becomes

$$p_{T,k+1} = p_{T,k} - \rho x_k(T).$$

The conjugate gradient method

The **conjugate gradient** method is an efficient algorithm to solve **linear systems**, that can also be applied in optimization.

CASE 1: direct functional F

STEP 1: recall the gradient of F: $\nabla F(u) = u - B^* p$ with (x, p) solution of the coupled system

$$\begin{cases} x'(t)Ax(t) + Bu(t), & 0 < t < T \\ p'(t) = -A^* p(t), & 0 < t < T \\ x(0) = x_0, \; p(T) = -x(T) =: p_T \end{cases}$$

Moreover, notice that $x(t) = e^{tA}x_0 + z(t)$, with

$$\begin{cases} z'(t) = Az(t) + Bu(t), \quad 0 < t < T, \\ z(0) = 0. \end{cases} \tag{4}$$

The conjugate gradient method

CASE 1: direct functional F

STEP 2: we can readily check that ∇F can be rewritten in the form

$$\nabla F(u) = \underbrace{(I + \mathcal{L}_T^* \mathcal{L}_T)}_{\mathbb{A}} u + \underbrace{\mathcal{L}_T^*(e^{TA} x_0)}_{-b}, \tag{5}$$

where the operators $\mathcal{L}_T$ and $\mathcal{L}_T^*$ are defined as

$$\begin{array}{rccc} \mathcal{L}_T : & L^2(0,T;U) & \longrightarrow & H \\ & u & \longmapsto & z(T) \end{array} \quad \text{and} \quad \begin{array}{rccc} \mathcal{L}_T^* : & H & \longrightarrow & L^2(0,T;U) \\ & p_T & \longmapsto & B^* p, \end{array}$$

Since the minimizer u^* of F has to satisfy $\nabla F(u^*) = 0$, we see from (5) that computing u^* is equivalent to solve the linear system

$$\mathbb{A} u = b. \tag{6}$$

The conjugate gradient method

The CG methodology amounts to solve (6) through the following iterative procedure.

CG algorithm

input $d^0 = r^0 = b - \mathbb{A}u^0$
for $k \geq 1$ **do**

$$\alpha_k = \frac{(r^k)^\top r^k}{(d^k)^\top \mathbb{A} d^k}$$

$$x^{k+1} = x^k + \alpha_k d^k$$

$$r^{k+1} = r^k - \alpha_k d^k$$

$$\gamma_{k+1} = \frac{(r^{k+1})^\top r^{k+1}}{(r^k)^\top r^k}$$

$$d^{k+1} = r^{k+1} + \gamma_{k+1} d^k$$

end for

About convergence, we know that

$$|u_k - u^*| \leq 2 \left(\frac{\sqrt{\sigma} - 1}{\sqrt{\sigma} + 1} \right)^k |u_0 - u^*|.$$

Y. Saad, *Iterative methods for sparse linear systems*, SIAM, 2003.

The conjugate gradient method

CASE 2: adjoint functional F

With a similar procedure as for F, we can readily check that ∇J can be rewritten in the form

$$\nabla J(p_T) = \underbrace{\mathcal{L}_T \mathcal{L}_T^*}_{\mathbb{A}} p_T + \underbrace{e^{TA} x_0}_{-b} . \tag{7}$$

Since the minimizer p_T^* of F has to satisfy $\nabla J(p_T^*) = 0$, we see from (7) that computing p_T^* is equivalent to solve the linear system

$$\mathbb{A} p_T = b.$$

Numerical approximation of controls

The penalized Hilbert uniqueness method

The **penalized Hilbert Uniqueness Method** is a largely employed methodology to compute the controls for a general dynamical system

$$\begin{cases} \dot{y}(t) + Ay(t) = Bu(t), \quad t \in (0, T) \\ y(0) = y_0 \end{cases}$$

F. Boyer, *On the penalised HUM approach and its applications to the numerical approximation of null-controls for parabolic problems*, ESAIM: Proc., 2013.

R. Glowinski and J.-L. Lions, *Exact and approximate controllability for distributed parameter systems*, Acta Num., 1994.

R. Glowinski, J.-L. Lions and J. He, *Exact and approximate controllability for distributed parameter systems: a numerical approach*, Cambridge University Press, 2008.

The penalized Hilbert uniqueness method

Let H be a Hilbert space and $A : \mathcal{D}(A) \subset H \to H$ be an unbounded operator such that $-A$ generates an analytic semi-group.

Let U be another Hilbert space and $B : U \to \mathcal{D}(A)^*$ be a bounded operator.

Let $T > 0$ be given and, for any $y_0 \in H$ and $u \in L^2(0,T;U)$, let us consider the Cauchy problem

$$y_t + Ay = Bu \text{ in } (0,T), \quad y(0) = y_0. \tag{8}$$

The **penalized HUM approach** consists in finding the control of minimal $L^2(0,T;U)$ norm for (8) by means of the following optimization problem:

$$\begin{aligned} u_\beta &= \underset{u \in L^2(0,T;U)}{argmin} \; F_\beta(u) \\ F_\beta(u) &:= \frac{1}{2}\int_0^T \|u(t)\|_U^2 \, dt + \frac{1}{2\beta}\|y(T)\|_H^2 \, . \end{aligned} \tag{9}$$

Remark

Recall that, if (8) is controllable (either null, exactly or approximately), then for any $\beta > 0$, the functional F_β is strictly convex, continuous and coercive. Hence, it has a unique minimizer $u_\beta \in L^2(0,T;U)$.

Theorem

The following controllability properties hold.

1. Problem (8) is approximately controllable at time $T > 0$ from $y_0 \in H$ if and only if $\|y_\beta(T)\|_H \to 0$ as $\beta \to 0$, where y_β denotes the solution corresponding to u_β.
2. Problem (8) is null-controllable at time $T > 0$ from $y_0 \in H$ if and only if

$$\mathcal{E}_{y_0} := 2 \sup_{\beta>0} \left(\inf_{u \in L^2(0,T;U)} F_\beta(u) \right) < +\infty. \tag{10}$$

In this case, we have

$$\|u_\beta\|_{L^2(0,T;U)} \leq \sqrt{\mathcal{E}_{y_0}} \tag{11a}$$

$$\|y_\beta(T)\|_H \leq \sqrt{\mathcal{E}_{y_0}\beta}. \tag{11b}$$

Moreover, as $\beta \to 0$, $u_\beta \to \bar{u}$ strongly in $L^2(0, T; U)$, $\bar{u}$ being the optimal control obtained from the functional (9) without the second penalization term.

According to the previous theorem, there is then an essential difference between approximate and null controllability in this penalization context.

In both cases, the solution at time T of (8) corresponding to u_β **converges to zero in H as $\beta \to 0$**.

Nevertheless, for null-controllability, **this convergence has a precise rate $\sqrt{\beta}$** which is typically violated when only approximate controllability holds.

Furthermore, when the problem is null controllable, we also have that **the control cost $\|u_\beta\|_{L^2(0,T;U)}$ remains uniformly bounded** which, together with (11b), yields the convergence of u_β to the solution $\bar{u}$ of the non-penalized problem.

The Fenchel-Rockafellar duality

One of the funding pillars of optimal control theory is that a convex optimization problem can be solved by applying duality in the sense of Fenchel and Rockafellar.

I. Ekeland and R. Temam, *Convex analysis and variational problems*, SIAM, 1999.

STEP 1: since the problem (8) is linear, we can write its solution as $y = e^{-tA}y_0 + z$, with

$$\begin{cases} z' + Az = Bu & \text{in } (0,T) \\ z(0) = 0 \end{cases} \tag{12}$$

STEP 2: let $\mathcal{L}_T : L^2(0,T;U) \to H$ be the linear continuous operator defined as $\mathcal{L}_T(u) = z(T)$. Then, the adjoint operator $\mathcal{L}_T^\star : H \to L^2(0,T;U)$ is given by $\mathcal{L}_T^\star(p_T) = B^*p$ where, for all $p_T \in H$, p solves

$$\begin{cases} -p' + A^*p = 0 & \text{in } (0,T), \\ p(T) = p_T \end{cases} \tag{13}$$

STEP 3: with the notation introduced, we have $F_\beta(u) = \widehat{F}(u) + G_\beta(\mathcal{L}_T u)$, where

$$\widehat{F}(u) := \frac{1}{2}\int_0^T \|u(t)\|_U^2 \, dt \quad \text{and} \quad G_\beta(\mathcal{L}_T u) := \frac{1}{2\beta}\left\|\mathcal{L}_T u + e^{-TA}y_0\right\|_H^2.$$

Since both $\widehat{F}$ and G_β are convex functionals, Fenchel-Rockafellar theory yields that

$$u_\beta = B^* p_\beta, \tag{14}$$

with p_β solution of (13) corresponding to the initial datum

$$p_{T,\beta} = \underset{p_T \in H}{argmin} J(p_T),$$

and $J(p_T) := \widehat{F}^\star(\mathcal{L}_T^\star p_T) + G_\beta^\star(-p_T)$, $\widehat{F}^\star$ and $G_\beta^\star$ being the **convex conjugates**

$$\begin{aligned} \widehat{F}^\star(u) &= \sup_{v \in L^2(0,T;U)} \left\{ \langle u, v\rangle_{L^2(0,T;U)} - \widehat{F}(v) \right\}, \quad u \in L^2(0,T;U) \\ G_\beta^\star(-p_T) &= \sup_{q_T \in H} \left\{ -\langle p_T, q_T\rangle_H - G_\beta(q_T) \right\}, \quad q_T \in H. \end{aligned} \tag{15}$$

The Fenchel-Rockafellar duality

STEP 4: it can be readily checked using (15) that

$$\widehat{F}^\star(\mathcal{L}_T^\star p_T) = \frac{1}{2}\int_0^T \|p(t)\|_H^2 \, dt$$

$$G_\beta^\star(-p_T) = \langle p_T, e^{-TA} y_0 \rangle_H + \frac{\beta}{2}\|p_T\|_H^2 \, .$$

Collecting everything, we then obtain that $J_\beta(p_T)$ is given by

$$J_\beta(p_T) = \frac{1}{2}\int_0^T \|p(t)\|_H^2 \, dt + \frac{\beta}{2}\|p_T\|_H^2 + \langle p_T, e^{-TA} y_0 \rangle_H.$$

THANK YOU FOR YOUR ATTENTION!

Funding

- European Research Council (ERC): grant agreements NO: 694126-DyCon and No.765579-ConFlex.
- MINECO (Spain): Grant PID2020-112617GB-C22 KILEARN
- Alexander von Humboldt-Professorship program
- DFG (Germany): Transregio 154 Project "Mathematical Modelling,Simulation and Optimization Using the Example of Gas Networks"
- COST Action grant CA18232, "Mathematical models for interacting dynamics on networks".

Fundación Deusto Fundazioa

Friedrich-Alexander-Universität Department Data Science | DDS

European Research Council
Established by the European Commission

European Commission

DyCon DYNAMIC CONTROL

Chair DYNAMICS, CONTROL AND NUMERICS FAU

Unterstützt von / Supported by
Alexander von Humboldt Stiftung / Foundation

UAM Universidad Autónoma de Madrid