Introduction to PDE-Constrained Optimization

Part I

Stefan Volkwein

University of Konstanz, Department of Mathematics and Statistics, Chair for Numerical Optimization

Research School – Domain Decomposition for Optimal Control Problems, Luminy, Sept. 5-9, 2022

Our Motivation for the Research

PDE-constrained optimization: $\mathcal{X}_{\text{ad}}$ suitable set of admissible solutions

$$\min J(y,u) \quad \text{subject to (s.t.)} \quad (y,u) \text{ solves a PDE system and } (y,u) \in \mathcal{X}_{\text{ad}} \qquad \textbf{(P)}$$

Optimal Design (e.g., shape or topology optimization):
- What is an optimal configuration for a given application?

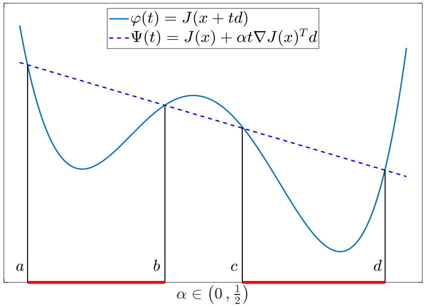
Optimal Control (e.g., drug treatment for patients, autonomous driving):
- How should/can we control a time-dependent process with partial/uncertain measurement information in order to guarantee stability of the process?

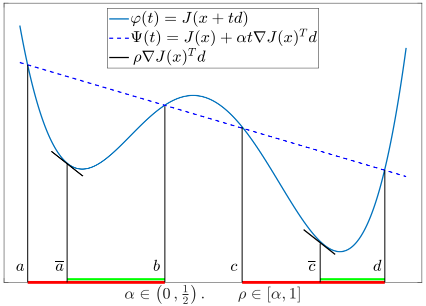
Parameter Estimation/Inverse Problems (e.g., lithium-ion battery models):
- Which unknown (and not measurable) PDE parameters should be chosen so that the PDE model describes the real phenomena sufficiently accurate?

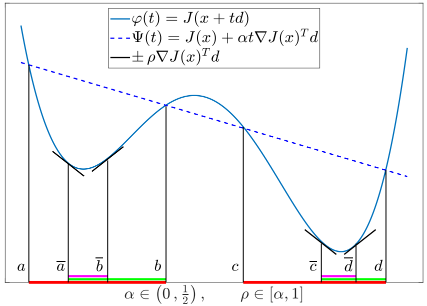
Optimal Experimental Design:
- How should we carry out experiments to get measurements so that the unknown PDE parameters can be optimally estimated?

Outline of the Talk

1 Constrained Nonlinear Optimization







References:

[1] Kelley: *Iterative Methods for Optimization*, SIAM, 1999
[2] Luenberger/Ye; *Linear and Nonlinear Programming*, Springer, 2016
[3] Nocedal/Wright: *Numerical Optimization*, Springer, 2006

Finite-Dimensional Optimization

Minimization Problem:

$$\min_x J(x) \quad \text{subject to (s.t.)} \quad x \in \mathbb{R}^n \text{ satisfies } e(x) = 0 \text{ in } \mathbb{R}^m \qquad (\mathbf{P}_1)$$

with cost functional $J : \mathbb{R}^n \to \mathbb{R}$ and equality constraint function $e : \mathbb{R}^n \to \mathbb{R}^m$ $(n \geq m)$

Set of feasible points: $\mathcal{F}(\mathbf{P}_1) = \{x \in \mathbb{R}^n \,|\, e(x) = 0\}$

Existence of solutions: proof by arguments from analysis

- assumptions on J and e (continuity)
- properties of the admissible set $\mathcal{F}(\mathbf{P}_1)$ (closedness, boundedness, ...)
- boundedness of minimizing sequences
- extraction of convergent subsequence (Bolzano-Weierstraß)

Bad News: no constructive procedure available for getting an optimal solution $\bar{x}$

First-Order Optimality Conditions – 1

Minimization Problem: J and e continuously differentiable

$$\min_x J(x) \quad \text{subject to (s.t.)} \quad x \in \mathbb{R}^n \text{ satisfies } e(x) = 0 \text{ in } \mathbb{R}^m \qquad (\mathbf{P}_1)$$

Locally Optimal Solution: $\bar{x} \in \mathcal{F}(\mathbf{P}_1) = \{x \in \mathbb{R}^n \,|\, e(x) = 0\}$

Lagrange Functional: $\mathscr{L}(x,p) = J(x) + \langle e(x), p\rangle_{\mathbb{R}^m} = J(x) + \sum_{i=1}^m e_i(x)p_i$ for $(x,p) \in \mathbb{R}^n \times \mathbb{R}^m$

Constrained Qualification Condition: Jacobian matrix $e'(\bar{x}) \in \mathbb{R}^{m\times n}$ is surjective
$\Rightarrow$ Tangent plane to the closed set $\mathcal{F}(\mathbf{P}_1)$ at $\bar{x}$ is characterized by $\ker e'(\bar{x})$.

KKT Conditions: If $e'(\bar{x})$ is surjective, there exists a unique $\bar{p} = (\bar{p}_i) \in \mathbb{R}^m$ such that

$$\nabla_x \mathscr{L}(\bar{x},\bar{p}) = \nabla J(\bar{x}) + \sum_{i=1}^m \bar{p}_i \nabla e_i(\bar{x}) \stackrel{!}{=} 0 \in \mathbb{R}^n, \quad \nabla_p \mathscr{L}(\bar{x},\bar{p}) = e(\bar{x}) \stackrel{!}{=} 0 \in \mathbb{R}^m$$

First-Order Optimality Conditions – 2

Notations: $\nabla J(x) \in \mathbb{R}^{n\times 1}$ (column vector), $e'(x) = \begin{pmatrix} \nabla e_1(x)^\top \\ \vdots \\ \nabla e_m(x)^\top \end{pmatrix} \in \mathbb{R}^{m\times n}$ with $\operatorname{rank} e'(x) \le m \le n$.

KKT Conditions: If $e'(\bar{x})$ is surjective, there exists a unique $\bar{p} \in \mathbb{R}^m$ such that

$$\nabla_x \mathscr{L}(\bar{x},\bar{p}) = \nabla J(\bar{x}) + \sum_{i=1}^m \bar{p}_i \nabla e_i(\bar{x}) = \nabla J(\bar{x}) + e'(\bar{x})^\top \bar{p} \stackrel{!}{=} 0 \in \mathbb{R}^n, \quad \nabla_p \mathscr{L}(\bar{x},\bar{p}) = e(\bar{x}) \stackrel{!}{=} 0 \in \mathbb{R}^m$$

Comments:

- $\nabla J(\bar{x}) \in \operatorname{span}\{\nabla e_1(\bar{x}),\ldots,\nabla e_m(\bar{x})\} \subset \mathbb{R}^n$
- $e'(\bar{x})^\top \bar{p} = -\nabla J(\bar{x})$ and $e'(\bar{x})^\top$ is injective $\Rightarrow$ uniqueness of $\bar{p}$
- $n+m$ equations $\nabla_x \mathscr{L}(\bar{x},\bar{p}) = 0$ and $\nabla_p \mathscr{L}(\bar{x},\bar{p}) = 0$ for $n+m$ unknowns $(\bar{x},\bar{p})$
- If J and e are twice continuously differentiable, apply Newton's method

$$\begin{pmatrix} \nabla_{xx}\mathscr{L}(x^k,p^k) & e'(x^k)^\top \\ e'(x^k) & 0 \end{pmatrix} \begin{pmatrix} x_\delta \\ p_\delta \end{pmatrix} = - \begin{pmatrix} \nabla_x \mathscr{L}(x^k,p^k) \\ \nabla_p \mathscr{L}(x^k,p^k) \end{pmatrix}, \quad (x^{k+1},p^{k+1}) = (x^k,p^k) + (x_\delta,p_\delta)$$

Second-Order Optimality Conditions

Minimization Problem: J and e twice continuously differentiable

$$\min_x J(x) \quad \text{subject to (s.t.)} \quad x \in \mathbb{R}^n \text{ satisfies } e(x) = 0 \text{ in } \mathbb{R}^m \tag{$\mathbf{P}_1$}$$

Necessary KKT Conditions: If $e'(\bar{x})$ is surjective, there exists a unique $\bar{p} \in \mathbb{R}^n$ such that

$$\nabla J(\bar{x}) + e'(\bar{x})^\top \bar{p} \overset{!}{=} 0 \in \mathbb{R}^n, \quad e(\bar{x}) \overset{!}{=} 0 \in \mathbb{R}^m \tag{1}$$

Second-Order Sufficient Optimality Conditions: If $(\bar{x}, \bar{p})$ satisfy (1) and $e'(\bar{x})$ is surjective, there exists a unique $\bar{p} \in \mathbb{R}^n$ such that

$$v^\top \nabla_{xx} \mathscr{L}(\bar{x}, \bar{p}) v > 0 \quad \text{for all } v \in \ker e'(\bar{x}) \setminus \{0\}$$

where $\nabla_{xx} \mathscr{L}(\bar{x}, \bar{p}) = \nabla^2 J(\bar{x}) + \sum_{i=1}^m p_i \nabla^2 e_i(\bar{x}) \in \mathbb{R}^{n \times n}$ holds.

Quadratic Programming

Consider: $A \in \mathbb{R}^{m\times n}$ surjective, $Q \in \mathbb{R}^{n\times n}$ symmetric & positive definite on $\ker A$, $q \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}$

$$\min_x J(x) = \frac{1}{2}x^\top Qx - q^\top x + c \quad \text{s.t.} \quad x \in \mathcal{F}(\mathbf{P}_1) = \{x \in \mathbb{R}^n \,|\, e(x) = Ax - b = 0\} \qquad (\mathbf{P}_1)$$

Sufficient KKT Conditions: Since $e'(\bar{x}) = A$ is surjective, there is a unique $\bar{p} \in \mathbb{R}^n$ such that

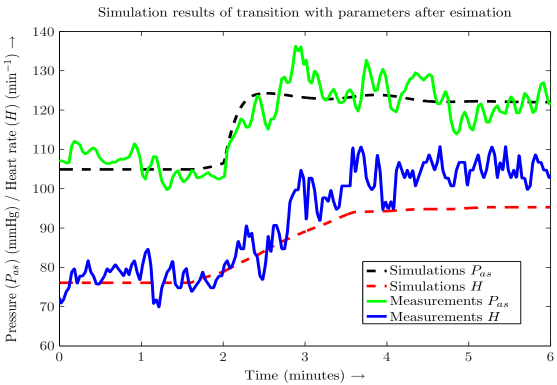
$$\nabla J(\bar{x}) + e'(\bar{x})^\top \bar{p} = Q\bar{x} - q + A^\top \bar{p} \stackrel{!}{=} 0 \in \mathbb{R}^n, \quad e(\bar{x}) = A\bar{x} - b \stackrel{!}{=} 0 \in \mathbb{R}^m$$

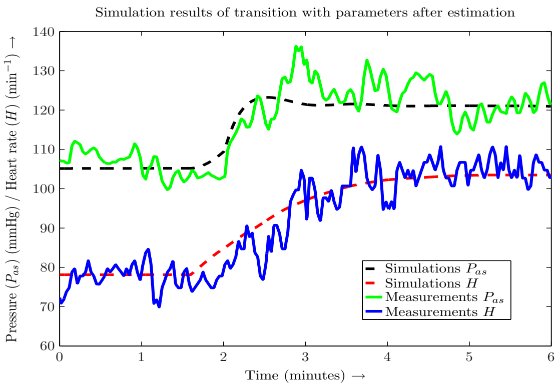
i.e., we obtain the uniquely solvable linear (saddlepoint) system

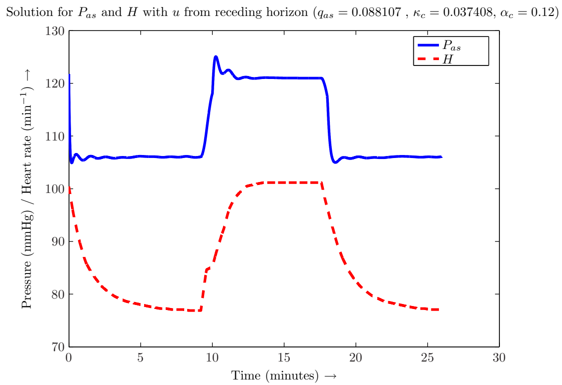
$$\begin{pmatrix} Q & A^\top \\ A & 0 \end{pmatrix} \begin{pmatrix} \bar{x} \\ \bar{p} \end{pmatrix} = \begin{pmatrix} q \\ b \end{pmatrix}$$

Second-Order Sufficient Optimality Condition: $\mathcal{L}_{xx}(x,p) = Q$ positive definite on $\ker A$

2 Optimization of Linear Dynamical Systems







References:

[4] Gerdts: *Optimal Control of ODEs and DAEs*, De Gruyter, 2011

Linear-Quadratic Problem

Consider:

$$\min_{(y,u)} J(y,u) = \frac{1}{2}\int_0^T y(t)^\top Q y(t) + u(t)^\top R u(t)\,\mathrm{d}t \quad \text{s.t.} \quad \begin{cases} \dot y(t) = Ay(t) + Bu(t) \text{ for } t \in (0,T] \\ y(0) = y_\circ \end{cases} \qquad (\mathbf{P}_2)$$

with infinite-dimensional state $y \in \mathcal{Y} = H^1(0,T;\mathbb{R}^n)$ and control $u \in \mathcal{U} = L^2(0,T;\mathbb{R}^m)$

Equality Constraints: For $\mathcal{X} = \mathcal{Y} \times \mathcal{U}$ and $\mathcal{Z} = L^2(0,T;\mathbb{R}^n) \times \mathbb{R}^n$ let $e : \mathcal{X} \to \mathcal{Z}$ be given as

$$e(x) = \begin{pmatrix} e_1(x) \\ e_2(x) \end{pmatrix} = \begin{pmatrix} \dot y - Ay - Bu \\ y(0) - y_\circ \end{pmatrix} \in \mathcal{Z} \quad \text{for } x = (y,u) \in \mathcal{X}$$

Constrained Optimization Problem:

$$\min_x J(x) \quad \text{s.t.} \quad x \in \mathcal{X} \text{ satisfies } e(x) = 0 \qquad (\mathbf{P}_2)$$

Problem Setting: $T > 0$, $Q \in \mathbb{R}^{n\times n}$ symmetric and positive semidefinite, $R \in \mathbb{R}^{m\times m}$ symmetric positive definite, $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $y_\circ \in \mathbb{R}^n$

Lagrange Functional

Consider:

$$\min_{(y,u)} J(y,u) = \frac{1}{2}\int_0^T y(t)^\top Q y(t) + u(t)^\top R u(t)\,\mathrm{d}t \quad \text{s.t.} \quad \begin{cases} \dot y(t) = Ay(t) + Bu(t) \text{ for } t \in (0,T] \\ y(0) = y_\circ \end{cases} \qquad (\mathbf{P}_2)$$

Equality Constraint: For $\mathcal{X} = \mathcal{Y} \times \mathcal{U}$ and $\mathcal{Z} = L^2(0,T;\mathbb{R}^n) \times \mathbb{R}^n$ let $e : \mathcal{X} \to \mathcal{Z}$ be given as

$$e(x) = \begin{pmatrix} e_1(x) \\ e_2(x) \end{pmatrix} = \begin{pmatrix} \dot y - Ay - Bu \\ y(0) - y_\circ \end{pmatrix} \in \mathcal{Z} \quad \text{for } x = (y,u) \in \mathcal{X}$$

Lagrange Functional: For $x = (y,u) \in \mathcal{X}$ and $p = (p_1, p_2) \in \mathcal{Z}$ we set

$$\begin{aligned} \mathscr{L}(x,p) &= J(x) + \langle e(x), p\rangle_{\mathcal{Z}} = J(x) + \langle e_1(x), p_1\rangle_{L^2(0,T;\mathbb{R}^n)} + \langle e_2(x), p_2\rangle_{\mathbb{R}^n} \\ &= \frac{1}{2}\int_0^T y(t)^\top Q y(t) + u(t)^\top R u(t)\,\mathrm{d}t + \int_0^T \big(\dot y(t) - Ay(t) - Bu(t)\big)^\top p_1(t)\,\mathrm{d}t + \big(y(0) - y_\circ\big)^\top p_2 \end{aligned}$$

Constraint Qualification

Equality Constraint: For $\mathcal{X} = \mathcal{Y} \times \mathcal{U}$ and $\mathcal{Z} = L^2(0,T;\mathbb{R}^n) \times \mathbb{R}^n$ let $e : \mathcal{X} \to \mathcal{Z}$ be given as

$$e(x) = \begin{pmatrix} e_1(x) \\ e_2(x) \end{pmatrix} = \begin{pmatrix} \dot{y} - Ay - Bu \\ y(0) - y_\circ \end{pmatrix} \in \mathcal{Z} \quad \text{for } x = (y,u) \in \mathcal{X}$$

Jacobian: For $x = (y,u) \in \mathcal{X}$ and direction $x^\delta = (y^\delta, u^\delta) \in \mathcal{X}$ we obtain

$$e'(x)x^\delta = \begin{pmatrix} e_1'(x)x^\delta \\ e_2'(x)x^\delta \end{pmatrix} = \begin{pmatrix} \dot{y}^\delta - Ay^\delta - Bu^\delta \\ y^\delta(0) \end{pmatrix} \in \mathcal{Z}$$

Constraint Qualification: Let $\mathscr{F} = (f, f_\circ) \in \mathcal{Z}$ be arbitrary. Then we consider

$$\begin{aligned} e'(x)x^\delta = \mathscr{F} \text{ in } \mathcal{Z} \quad &\Leftrightarrow \quad e_1'(x)x^\delta = f \text{ in } L^2(0,T;\mathbb{R}^n) \text{ and } e_2'(x)x^\delta = f_\circ \text{ in } \mathbb{R}^n \\ &\Leftrightarrow \quad \dot{y}^\delta - Ay^\delta - Bu^\delta = f \text{ in } L^2(0,T;\mathbb{R}^n) \text{ and } y^\delta(0) = f_\circ \text{ in } \mathbb{R}^n \end{aligned}$$

For any $u^\delta \in \mathcal{U}$ there exists a (unique) solution $y^\delta \in \mathcal{Y} \Rightarrow e'(x) : \mathcal{X} \to \mathcal{Z}$ is surjective

First-Order Optimality Conditions

Consider:

$$\min_{(y,u)} J(y,u) = \frac{1}{2}\int_0^T y(t)^\top Q y(t) + u(t)^\top R u(t)\,\mathrm{d}t \quad \text{s.t.} \quad \begin{cases} \dot{y}(t) = Ay(t) + Bu(t) \text{ for } t \in (0,T] \\ y(0) = y_\circ \end{cases} \qquad (\mathbf{P}_2)$$

Unique Optimal Solution: $\bar{x} = (\bar{y},\bar{u}) \in \mathcal{X}$

Lagrange Functional: For $x = (y,u) \in \mathcal{X}$ and $p = (p_1, p_2) \in \mathcal{Z}$ we set

$$\mathscr{L}(x,p) = \frac{1}{2}\int_0^T y(t)^\top Q y(t) + u(t)^\top R u(t)\,\mathrm{d}t + \int_0^T \big(\dot{y}(t) - Ay(t) - Bu(t)\big)^\top p_1(t)\,\mathrm{d}t + \big(y(0) - y_\circ\big)^\top p_2$$

Optimality Conditions: There exists a unique $\bar{p} = (\bar{p}_1, \bar{p}_2) \in \mathcal{Z}$ satisfying

$$\mathscr{L}_y(\bar{y},\bar{u},\bar{p}) \text{ in } \mathcal{Y}', \quad \mathscr{L}_u(\bar{y},\bar{u},\bar{p}) \text{ in } \mathcal{U}', \quad \mathscr{L}_p(\bar{y},\bar{u},\bar{p}) = 0 \text{ in } \mathcal{Z}'$$

State Equation

Lagrange Functional: For $x = (y, u) \in \mathcal{X}$ and $p = (p_1, p_2) \in \mathcal{Z}$ we set

$$\mathscr{L}(x,p) = \frac{1}{2}\int_0^T y(t)^\top Q y(t) + u(t)^\top R u(t)\,\mathrm{d}t + \int_0^T \big(\dot{y}(t) - Ay(t) - Bu(t)\big)^\top p_1(t)\,\mathrm{d}t + \big(y(0) - y_\circ\big)^\top p_2$$

Optimality Conditions: There exists a unique $\bar{p} = (\bar{p}_1, \bar{p}_2) \in \mathcal{Z}$ satisfying

$$\mathscr{L}_y(\bar{y}, \bar{u}, \bar{p}) \text{ in } \mathcal{Y}', \quad \mathscr{L}_u(\bar{y}, \bar{u}, \bar{p}) \text{ in } \mathcal{U}', \quad \mathscr{L}_p(\bar{y}, \bar{u}, \bar{p}) = 0 \text{ in } \mathcal{Z}'$$

State/Primal Equation: $\mathscr{L}_p(\bar{y}, \bar{u}, \bar{p}) = 0$ in $\mathcal{Z}' \Leftrightarrow \langle \mathscr{L}_p(\bar{y}, \bar{u}, \bar{p}), p\rangle_{\mathcal{Z}',\mathcal{Z}} = 0$ for all $p = (p_1, p_2) \in \mathcal{Z}$

$$0 = \langle \mathscr{L}_p(\bar{y}, \bar{u}, \bar{p}), p\rangle_{\mathcal{Z}',\mathcal{Z}} = \int_0^T \big(\dot{\bar{y}}(t) - A\bar{y}(t) - B\bar{u}(t)\big)^\top p_1(t)\,\mathrm{d}t + \big(\bar{y}(0) - y_\circ\big)^\top p_2$$

$$\Leftrightarrow \dot{\bar{y}} = A\bar{y} + B\bar{u} \text{ in } L^2(0,T;\mathbb{R}^n) \text{ and } \bar{y}(0) = y_\circ \text{ in } \mathbb{R}^n$$

Adjoint/Dual Equation

Lagrange Functional: For $x = (y,u) \in \mathcal{X}$ and $p = (p_1, p_2) \in \mathcal{Z}$ we set

$$\mathscr{L}(x,p) = \frac{1}{2}\int_0^T y(t)^\top Qy(t) + u(t)^\top Ru(t)\,\mathrm{d}t + \int_0^T \big(\dot{y}(t) - Ay(t) - Bu(t)\big)^\top p_1(t)\,\mathrm{d}t + \big(y(0) - y_\circ\big)^\top p_2$$

Adjoint/Dual Equation: $\mathscr{L}_y(\bar{y},\bar{u},\bar{p}) = 0$ in $\mathcal{Y}' \Leftrightarrow \langle \mathscr{L}_y(\bar{y},\bar{u},\bar{p}), y\rangle_{\mathcal{Y}',\mathcal{Y}} = 0$ for all $y \in \mathcal{Y}$

$$0 = \langle \mathscr{L}_y(\bar{y},\bar{u},\bar{p}), y\rangle_{\mathcal{Y}',\mathcal{Y}} = \int_0^T \bar{y}(t)^\top Qy(t)\,\mathrm{d}t + \int_0^T \big(\dot{y}(t) - Ay(t)\big)^\top \bar{p}_1(t)\,\mathrm{d}t + y(0)^\top \bar{p}_2$$

Variational arguments:

$\int_0^T \dot{y}(t)^\top \bar{p}_1(t)\,\mathrm{d}t = -\int_0^T \dot{\bar{p}}_1(t)^\top y(t)\,\mathrm{d}t$ for all $y \in C_0^\infty(0,T;\mathbb{R}^n) \hookrightarrow \mathcal{Y} \hookrightarrow L^2(0,T;\mathbb{R}^n)$

$\Rightarrow \quad 0 = \int_0^T \big(-\dot{\bar{p}}_1(t) - A^\top \bar{p}_1(t) + Q\bar{y}(t)\big)^\top y(t)\,\mathrm{d}t$ for all $y \in C_0^\infty(0,T;\mathbb{R}^n) \hookrightarrow L^2(0,T;\mathbb{R}^n)$

$\Rightarrow \quad 0 = -\dot{\bar{p}}_1 - A^\top \bar{p}_1 + Q\bar{y}$ in $L^2(0,T;\mathbb{R}^n)$

$\Rightarrow \quad 0 = \langle \mathscr{L}_y(\bar{y},\bar{u},\bar{p}), y\rangle_{\mathcal{Y}',\mathcal{Y}} = \big(y(t)\bar{p}_1(t)\big)\Big|_{t=0}^{t=T} + y(0)^\top \bar{p}_2 \quad \Rightarrow \quad \bar{p}_1(T) = 0$ and $\bar{p}_2 = \bar{p}_1(0)$

Optimality Condition

Lagrange Functional: For $x = (y, u) \in \mathcal{X}$ and $p = (p_1, p_2) \in \mathcal{Z}$ we set

$$\mathscr{L}(x,p) = \frac{1}{2}\int_0^T y(t)^\top Q y(t) + u(t)^\top R u(t)\,\mathrm{d}t + \int_0^T \big(\dot{y}(t) - Ay(t) - Bu(t)\big)^\top p_1(t)\,\mathrm{d}t + \big(y(0) - y_\circ\big)^\top p_2$$

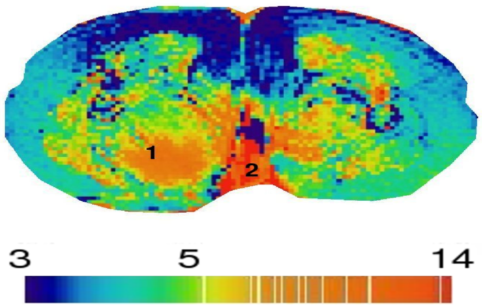
Optimality Condition: $\mathscr{L}_u(\bar{y},\bar{u},\bar{p}) = 0$ in $\mathcal{U}' \Leftrightarrow \langle \mathscr{L}_u(\bar{y},\bar{u},\bar{p}), u\rangle_{\mathcal{U}',\mathcal{U}} = 0$ for all $u \in \mathcal{U}$

$$\begin{aligned} 0 = \langle \mathscr{L}_u(\bar{y},\bar{u},\bar{p}), u\rangle_{\mathcal{U}',\mathcal{U}} &= \int_0^T \bar{u}(t)^\top R u(t)\,\mathrm{d}t + \int_0^T \big(-Bu(t)\big)^\top \bar{p}_1(t)\,\mathrm{d}t \\ &= \int_0^T \big(R\bar{u}(t) - B^\top \bar{p}_1(t)\big)^\top u(t)\,\mathrm{d}t \quad \Rightarrow \quad \bar{u} = R^{-1}B^\top \bar{p}_1 \text{ in } \mathcal{U} = L^2(0,T;\mathbb{R}^m) \end{aligned}$$

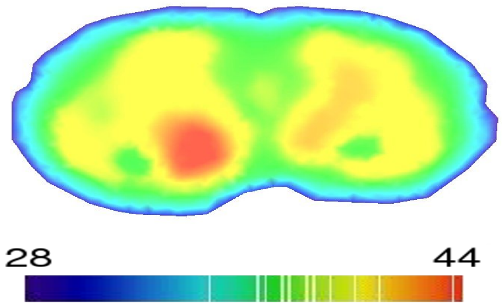
KKT System: coupled ODE system

$$\begin{aligned} \dot{\bar{y}}(t) &= A\bar{y}(t) + B\bar{u}(t) && \text{for } t \in (0,T], \quad \bar{y}(0) = y_\circ && \text{(state equation)} \\ -\dot{\bar{p}}_1(t) &= A^\top \bar{p}_1(t) - Q\bar{y}(t) && \text{for } t \in [0,T), \quad \bar{p}_1(T) = 0, \quad \bar{p}_2 = \bar{p}_1(0) && \text{(adjoint equation)} \\ \bar{u}(t) &= R^{-1}B^\top \bar{p}_1(t) && \text{for } t \in (0,T) && \text{(optimality condition)} \end{aligned}$$

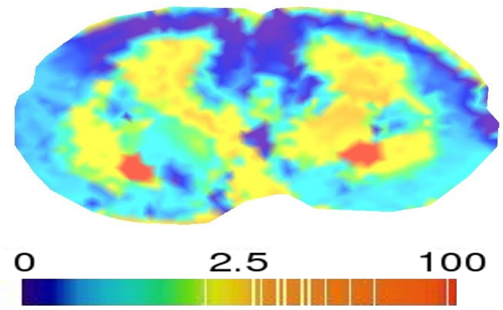
3 Optimization of Linear Elliptic PDEs



Messwerte



identifizierte Temperatur



berechnete Perfusion

References:

[05] Hinze/Pinnau/Ulbrich/Ulbrich: *Optimization with PDE Constraints*, Springer, 2009
[06] Lions: *Optimal Control of Systems Governed by Partial Differential Equations*, Springer-Verlag, 1971
[07] Tröltzsch: *Optimal Control pf Partial Differential Equations*, AMS, 2010

Infinite-Dimensional Optimization

Consider: $\Omega \subset \mathbb{R}^d$ bounded ($d \in \{1,2,3\}$), $y_\mathrm{d} \in L^2(\Omega)$, $\sigma > 0$, $V = H^1(\Omega)$, $\mathcal{U} = L^2(\Omega)$

$$\min_{(y,u)} J(y,u) = \frac{1}{2}\int_\Omega |y - y_\mathrm{d}|^2 + \sigma\,|u|^2 \,\mathrm{d}\mathbf{x} \quad \text{s.t.} \quad -\Delta y + y = u \text{ in } \Omega, \quad \frac{\partial y}{\partial \mathbf{n}} = 0 \text{ on } \Gamma = \partial\Omega \qquad (\mathbf{P}_3)$$

Setting: $\mathcal{X} = V \times \mathcal{U}$, cost functional $J : \mathcal{X} \to \mathbb{R}$, constraint operator $e : \mathcal{X} \to V'$ with

$$\langle e(x), \varphi \rangle_{V',V} = \int_\Omega \nabla y \cdot \nabla \varphi + y\varphi - u\varphi \,\mathrm{d}\mathbf{x} \quad \text{for } x = (y,u) \in \mathcal{X},\ \varphi \in V$$

i.e., $e(x) = 0$ in $V' \Leftrightarrow x = (y,u)$ satisfies $\int_\Omega \nabla y \cdot \nabla \varphi + y\varphi \,\mathrm{d}\mathbf{x} = \int_\Omega u\varphi \,\mathrm{d}\mathbf{x}$ for all $\varphi \in V$.

Optimization Problem with Equality Constraints:

$$\min J(x) \quad \text{s.t.} \quad x \in \mathcal{X} \text{ satisfies } e(x) = 0 \text{ in } V' \qquad (\mathbf{P}_3)$$

Unconstrained Reduced Problem

Neumann Problem: $V = H^1(\Omega)$, $\mathcal{U} = L^2(\Omega)$

$$\min_{(y,u)} J(y,u) = \frac{1}{2}\int_\Omega |y - y_\mathsf{d}|^2 + \sigma\,|u|^2\,\mathrm{d}\mathbf{x} \quad \text{s.t.} \quad \int_\Omega \nabla y\cdot\nabla\varphi + y\varphi\,\mathrm{d}\mathbf{x} = \int_\Omega u\varphi\,\mathrm{d}\mathbf{x} \text{ for all } \varphi\in V \qquad (\mathbf{P}_3)$$

Control-to-State Mapping: For $u\in\mathcal{U}$ let $y(u) = \mathscr{S}u \in V$ be the unique solution to

$$\langle e(y(u),u),\varphi\rangle_{V',V} = \int_\Omega \nabla y(u)\cdot\nabla\varphi + y(u)\varphi - u\varphi\,\mathrm{d}\mathbf{x} \quad \text{for all } \varphi\in V$$

Then, $e(y(u),u) = 0$ in V' holds. Moreover, $\mathscr{S}$ is linear and bounded ($\|y(u)\|_V \le \|u\|_\mathcal{U}$)

Reduced Cost Functional: $\hat{J}(u) = J(y(u),u)$ for $u\in\mathcal{U}$

Reduced Problem:

$$\min_u \hat{J}(u) = \frac{1}{2}\int_\Omega |\mathscr{S}u - y_\mathsf{d}|^2 + \sigma\,|u|^2\,\mathrm{d}\mathbf{x} \qquad (\hat{\mathbf{P}}_3)$$

Existence of Optimal Controls

Constrained Optimization: $\min_{(y,u)} J(y,u) = \frac{1}{2}\int_\Omega |y-y_\mathrm{d}|^2 + \sigma\,|u|^2\,\mathrm{d}\mathbf{x}$ s.t. $e(y,u)=0$ $\quad(\mathbf{P}_3)$

Unconstrained Reduced Problem: $\min_u \hat{J}(u) = \frac{1}{2}\int_\Omega |\mathscr{S}u-y_\mathrm{d}|^2 + \sigma\,|u|^2\,\mathrm{d}\mathbf{x}$ $\quad(\hat{\mathbf{P}}_3)$

Existence of a unique Optimal Control: $\mathcal{U} = L^2(\Omega)$

- $\{u^k\}_{k\in\mathbb{N}}$ minimizing sequence for $(\hat{\mathbf{P}}_3)$, i.e., $\lim_{k\to\infty}\hat{J}(u^k) = \inf\{\hat{J}(u)\,|\,u\in\mathcal{U}\} \geq 0$
- $\hat{J}$ radially unbounded, i.e., $\hat{J}(u)\to\infty$ for $\|u\|_\mathcal{U}\to\infty \Rightarrow \|u^k\|_\mathcal{U}$ bounded
- $\bar{u}$ solution to $(\hat{\mathbf{P}}_3)$, because $u^k \rightharpoonup \bar{u}$ in $\mathcal{U}$ and weakly lower semicontinuity of $\hat{J}$ imply

$$\infty > \inf\{\hat{J}(u)\,|\,u\in\mathcal{U}\} = \lim_{k\to\infty}\hat{J}(u^k) = \liminf_{k\to\infty}\hat{J}(u^k) \geq \hat{J}(\bar{u})$$

- $\hat{J}$ quadratic $\Rightarrow$ $(\hat{\mathbf{P}}_3)$ convex $\Rightarrow$ unique optimal solution $\bar{u}$
- $\bar{x} = (\bar{y},\bar{u})$ unique solution to $(\mathbf{P}_3)$ with $\bar{y} = \mathscr{S}\bar{u}$
- Utilized facts: linear, bounded control-to-state mapping $\mathscr{S}$, reflexivity of $\mathcal{U}$, $\sigma > 0$

First-Order Sufficient Optimality Conditions

Neumann Problem: $V = H^1(\Omega)$, $\mathcal{U} = L^2(\Omega)$, $\mathcal{X} = V \times \mathcal{U}$

$$\min_{(y,u)} J(y,u) = \frac{1}{2}\int_\Omega |y - y_\mathrm{d}|^2 + \sigma\,|u|^2\,\mathrm{d}\mathbf{x} \quad \text{s.t.} \quad \int_\Omega \nabla y \cdot \nabla \varphi + y\varphi\,\mathrm{d}\mathbf{x} = \int_\Omega u\varphi\,\mathrm{d}\mathbf{x} \text{ for all } \varphi \in V \qquad (\mathbf{P}_3)$$

Lagrange Functional for $(\mathbf{P}_3)$:

$$\mathscr{L}(y,u,p) = J(y,u) + \langle e(y,u), p\rangle_{V',V} = \int_\Omega \Big(\frac{1}{2}|y - y_\mathrm{d}|^2 + \frac{\sigma}{}2|u|^2 + \nabla y \cdot \nabla p + yp - up\Big)\,\mathrm{d}\mathbf{x}$$

Optimality System: $\bar{x} = (\bar{y},\bar{u}) \in \mathcal{X}$ optimal solution. Then there exists $\bar{p} \in V$ satisfying

$$\mathscr{L}_y(\bar{y},\bar{u},\bar{p}) = 0 \text{ in } V', \quad \mathscr{L}_u(\bar{y},\bar{u},p) = 0 \text{ in } \mathcal{U}', \quad \mathscr{L}_p(\bar{y},\bar{u},\bar{p}) = 0 \text{ in } V'$$

since a constrained qualification condition holds ($e'(\bar{x}) : \mathcal{X} \to V'$ is surjective)

Note: derivatives are functionals!

State Equation

Lagrange Functional for $(\mathbf{P}_3)$: $\mathscr{L}(y,u,p) = \int_\Omega \Big(|y-y_\mathrm{d}|^2 + \frac{\sigma}{2}|u|^2 + \nabla y \cdot \nabla p + yp - up \Big)\,\mathrm{d}\mathbf{x}$

State/Primal Equation: $\mathscr{L}_p(\bar y,\bar u,\bar p) = 0$ in V' $\Leftrightarrow$ $\langle \mathscr{L}_p(\bar y,\bar u,\bar p), p\rangle_{V',V} = 0$ for all $p \in V$

$$0 = \langle \mathscr{L}_p(\bar y,\bar u,\bar p), p\rangle_{V',V} = \int_\Omega \big(\nabla \bar y \cdot \nabla p + \bar y p - \bar u p\big)\,\mathrm{d}\mathbf{x} \quad \text{for all } p \in V$$

i.e., $\bar y = y(\bar u) \in V$ is the weak solution to $-\Delta \bar y + \bar y = \bar u$ in Ω, $\frac{\partial \bar y}{\partial \mathbf{n}} = 0$ on Γ (state equation)

Stationarity: From

$$\int_\Omega \nabla y(u) \cdot \nabla \varphi + y(u)\varphi\,\mathrm{d}\mathbf{x} = \int_\Omega u\varphi\,\mathrm{d}\mathbf{x} \text{ for all } \varphi \in V'$$

we infer $\mathscr{L}_p(y(u),u,p) = 0$ in V' for any $(u,p) \in \mathfrak{U} \times V$, because

$$\langle \mathscr{L}_p(y(u),u,p), \varphi\rangle_{V',V} = \int_\Omega \big(\nabla y(u) \cdot \nabla \varphi + y(u)\varphi - u\varphi\big)\,\mathrm{d}\mathbf{x} = 0 \quad \text{for all } \varphi \in V$$

Adjoint Equation

Lagrange Functional for $(\mathbf{P}_3)$: $\mathscr{L}(y,u,p) = \int_\Omega \Big(|y-y_\mathsf{d}|^2 + \frac{\sigma}{2}|u|^2 + \nabla y \cdot \nabla p + yp - up \Big)\,\mathrm{d}\mathbf{x}$

Adjoint/Dual Equation: $\mathscr{L}_y(\bar y,\bar u,\bar p) = 0$ in $V' \Leftrightarrow \langle \mathscr{L}_y(\bar y,\bar u,\bar p), y\rangle_{V',V} = 0$ for all $y \in V$

$$
\begin{aligned}
0 = \langle \mathscr{L}_y(\bar y,\bar u,\bar p), y\rangle_{V',V} &= \int_\Omega \Big((\bar y - y_\mathsf{d})y + \nabla y \cdot \nabla \bar p + y\bar p \Big)\,\mathrm{d}\mathbf{x} \\
&= \int_\Omega \big(\nabla \bar p \cdot \nabla y + \bar p y - (y_\mathsf{d} - \bar y)y \big)\,\mathrm{d}\mathbf{x} \quad \text{for all } y \in V
\end{aligned}
$$

i.e., $\bar p \in V$ is the weak solution to $-\Delta \bar p + \bar p = y_\mathsf{d} - \bar y$ in Ω, $\frac{\partial \bar p}{\partial \mathbf{n}} = 0$ on Γ (adjoint equation)

Stationarity: From

$$
\int_\Omega \nabla p(u) \cdot \nabla \varphi + p(u)\varphi\,\mathrm{d}\mathbf{x} = \int_\Omega (y_\mathsf{d} - y(u))\varphi\,\mathrm{d}\mathbf{x} \text{ for all } \varphi \in V'
$$

we infer $\mathscr{L}_y(y(u),u,p(u)) = 0$ in V' for any $u \in \mathfrak{U}$

Optimality Condition

Lagrange Functional for $(\mathbf{P}_3)$: $\mathscr{L}(y,u,p) = \int_\Omega \Big(|y - y_\mathrm{d}|^2 + \frac{\sigma}{2}|u|^2 + \nabla y \cdot \nabla p + yp - up \Big)\,\mathrm{d}\mathbf{x}$

Optimality: $\mathscr{L}_u(\bar y, \bar u, \bar p) = 0$ in $\mathfrak{U}' \Leftrightarrow \langle \mathscr{L}_u(\bar y, \bar u, \bar p), u\rangle_{\mathfrak{U}',\mathfrak{U}} = 0$ for all $u \in \mathfrak{U}$

$$0 = \langle \mathscr{L}_u(\bar y, \bar u, \bar p), u\rangle_{\mathfrak{U}',\mathfrak{U}} = \int_\Omega \big(\sigma \bar u u - u \bar p\big)\,\mathrm{d}\mathbf{x} = \int_\Omega \big(\sigma \bar u - \bar p\big) u\,\mathrm{d}\mathbf{x} = \langle \sigma \bar u - \bar p, v\rangle_{\mathfrak{U}} \quad \text{for all } v \in \mathfrak{U}$$

$\Rightarrow \bar u = \bar p/\sigma$ in $\mathfrak{U}$; moreover, $\bar u \in V$ follows from $\bar p \in V$

Gradient: Riesz representant $\nabla_u \mathscr{L}(\bar y, \bar u, \bar p) \in \mathfrak{U}$ of $\mathscr{L}_u(\bar y, \bar u, \bar p) \in \mathfrak{U}'$ is given as

$$\nabla_u \mathscr{L}(\bar y, \bar u, \bar p) = \sigma \bar u - \bar p \in \mathfrak{U}$$

Relation to the Reduced Cost Functional

Reduced Cost: For a chosen $u \in \mathcal{U}$ we infer from $\hat{J}(u) = J(y(u),u)$ and $e(y(u),u) = 0$ in V'

$$\hat{J}(u) = J(y(u),u) = J(y(u),u) + \underbrace{\langle e(y(u),u),p\rangle_{V',V}}_{=0} = \mathscr{L}(y(u),u,p) \text{ for any } p \in V$$

Chain Rule: For every direction $v \in \mathcal{U}$ we have

$$\langle \hat{J}'(u),v\rangle_{\mathcal{U}',\mathcal{U}} = \langle \mathscr{L}_y(y(u),u,p),y'(u)v\rangle_{V',V} + \langle \mathscr{L}_u(y(u),u,p),v\rangle_{\mathcal{U}',\mathcal{U}} \quad \text{for any } p \in V \tag{2}$$

Recall: $\mathscr{L}_y(y(u),u,p(u)) = 0$ in V'

Gradient of the Reduced Cost: Choosing $p = p(u) \in V$ in (2) imply for all $v \in \mathcal{U}$

$$\langle \hat{J}'(u),v\rangle_{\mathcal{U}',\mathcal{U}} = \langle \mathscr{L}_u(y(u),u,p(u)),v\rangle_{\mathcal{U}',\mathcal{U}} = \int_\Omega \big(\sigma u - p(u)\big)v\,\mathrm{d}\mathbf{x} = \langle \sigma u - p(u),v\rangle_{\mathcal{U}} = \langle \nabla\hat{J}(u),v\rangle_{\mathcal{U}}$$

First-Order Optimality System / KKT-System

Optimality system:

$$-\Delta\bar{y}+\bar{y}=\bar{u} \quad \text{in } \Omega, \quad \frac{\partial\bar{y}}{\partial\mathbf{n}}=0 \quad \text{on } \Gamma \quad \text{(state equation)}$$

$$\sigma\bar{u}-\bar{p}=0 \quad \text{in } \Omega, \quad \text{(optimality conditon)}$$

$$-\Delta\bar{p}+\bar{p}=y_{\mathrm{d}}-\bar{y} \quad \text{in } \Omega, \quad \frac{\partial\bar{p}}{\partial\mathbf{n}}=0 \quad \text{on } \Gamma \quad \text{(adjoint equation)}$$

Reduction:

$$-\Delta\bar{y}+\bar{y}+\frac{\bar{p}}{\sigma}=0 \quad \text{in } \Omega, \quad \frac{\partial\bar{y}}{\partial\mathbf{n}}=0 \quad \text{on } \Gamma \quad \text{(state equation)}$$

$$-\Delta\bar{p}+\bar{p}+\bar{y}=y_{\mathrm{d}} \quad \text{in } \Omega, \quad \frac{\partial\bar{p}}{\partial\mathbf{n}}=0 \quad \text{on } \Gamma \quad \text{(adjoint equation)}$$

Numerical Solution Methods / Optimize-then-Discretize

Reduced Problem: $\min_u \hat{J}(u) = \frac{1}{2}\|\mathscr{S}u - y_\mathrm{d}\|^2_{L^2(\Omega)} + \sigma\|u\|^2_{L^2(\Omega)}\,\mathrm{d}\mathbf{x}$ $(\hat{\mathbf{P}}_1)$

Iteration of the Gradient Method:

1: Compute state $y^k = y(u^k)$ and adjoint $p^k = p(u^k)$;
2: Set $\nabla J(u^k) = \sigma u^k - p^k$;
3: Determine appropriate step size $t_k > 0$ (e.g., satisfying Armijo or Powell-Wolfe rule);
4: $u^{k+1} = u^k - t_k \nabla \hat{J}(u^k)$;

Note: $\|\mathscr{S}u - y_\mathrm{d}\|^2_{L^2(\Omega)} = \underbrace{\langle \mathscr{S}u, \mathscr{S}u\rangle_{L^2(\Omega)}}_{=\langle \mathscr{S}^*\mathscr{S}u,u\rangle_{L^2(\Omega)}} + 2\underbrace{\langle \mathscr{S}u, y_\mathrm{d}\rangle_{L^2(\Omega)}}_{=\langle \mathscr{S}^* y_\mathrm{d},u\rangle_{L^2(\Omega)}} + \|y_\mathrm{d}\|^2_{L^2(\Omega)}$, $\mathscr{S}^*: V \to \mathfrak{U}$ adjoint of $\mathscr{S}$

Conjugate Gradient Method: Let $g^0 = \sigma u^0 - p(u^0)$ and $d^0 = -g^0$ we compute for each k

1: Set $p^k = (\mathscr{S}^*\mathscr{S})d^k$, $t_k = \|g^k\|^2_{L^2(\Omega)} / \langle p^k, d^k\rangle_{L^2(\Omega)}$ and $u^{k+1} = u^k + t_k d^k$;
2: Determine $g^{k+1} = g^k + t_k d^k$;
3: Define $\beta_k = \|g^{k+1}\|^2_{L^2(\Omega)} / \|g^k\|^2_{L^2(\Omega)}$ and $d^{k+1} = -g^{k+1} + \beta_k d^k$;