Mathematics of PDE constrained optimization
Discrete concepts
1. Basic approaches

Michael Hinze

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Oberwolfach, November 22-26, 2010

Mother Problem

$$
(\mathbb{P}) \quad \left\{ \begin{array}{l} \min_{(y,u)\in Y\times U} J(y,u) := \frac{1}{2}\|y-z\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\|u\|^2_U \\ \text{s.t.} \\ \quad \begin{array}{rclll} -\Delta y & = & Bu & & \text{in } \Omega, \\ y & = & 0 & & \text{on } \partial\Omega, \end{array} \\ \text{and} \\ u \in U_{\mathrm{ad}} \subseteq U. \end{array} \right. \tag{0.1}
$$

Here, $\Omega \subset \mathbb{R}^n$ denotes an open, bounded sufficiently smooth (polyhedral) domain, $Y := H^1_0(\Omega)$, the operator $B : U \to H^{-1}(\Omega) \equiv Y^*$ denotes the (linear, continuous) control operator, and U_{ad} is assumed to be a closed and convex subset of the Hilbert space U.

Example

1. $U := L^2(\Omega),\ B : L^2(\Omega) \to H^{-1}(\Omega)$ Injection, , $U_{\mathrm{ad}} := \{v \in L^2(\Omega); a \leq v(x) \leq b$ a.e. in $\Omega\}, a, b \in L^\infty(\Omega)$.
2. $U := H^1(\Omega),\ B : H^1(\Omega) \to H^{-1}(\Omega)$ Injection, , $U_{\mathrm{ad}} := \{v \in L^2(\Omega); a \leq v(x) \leq b$ a.e. in $\Omega\}, a, b \in L^\infty(\Omega)$.
3. $U := \mathbb{R}^m,\ B : \mathbb{R}^m \to H^{-1}(\Omega),\ Bu := \sum_{j=1}^m u_j F_j,\ F_j \in H^{-1}(\Omega)$ given ,

   $U_{\mathrm{ad}} := \{v \in \mathbb{R}^m; a_j \leq v_j \leq b_j\}, a < b.$

We already know that problem $(\mathbb{P})$ admits a unique solution $(y,u) \in H_0^1(\Omega) \times U$, and that $(\mathbb{P})$ equivalently can be rewritten as the optimization problem

$$\min_{u \in U_{\mathrm{ad}}} \hat{J}(u) \tag{0.2}$$

for the reduced functional

$$\hat{J}(u) := J(y(u),u) \equiv J(SBu,u)$$

over the set U_{ad}, where $S : Y^* \to Y$ denotes the solution operator associated with $-\Delta$. We further know that the first order necessary (and here also sufficient) optimality conditions take the form

$$\langle \hat{J}'(u), v-u \rangle_{U^*,U} \geq 0 \text{ for all } v \in U_{\mathrm{ad}} \tag{0.3}$$

Here

$$\hat{J}'(u) = \alpha(u,\cdot)_U + B^*S^*(SBu - z) \equiv \alpha(u,\cdot)_U + B^*p,$$

with $p := S^*(SBu - z) \in Y^{**}$ denoting the adjoint variable. The function p in our reflexive setting satisfies

$$\begin{aligned} -\Delta p &= y - z && \text{in } \Omega, \\ p &= 0 && \text{on } \partial\Omega. \end{aligned}$$

With the Riesz isomorphism $R : U^* \to U$ and the orthogonal projection $P_{U_{ad}} : U \to U_{ad}$ we have that (0.4) is equivalent to

$$u = P_{U_{ad}}\left(u - \sigma\nabla\hat{J}(u)\right) \text{ for all } \sigma > 0, \tag{0.4}$$

where

$$\nabla\hat{J}(u) = R\hat{J}'(u)$$

denotes the gradient of $\hat{J}(u)$.

To discretize $(\mathbb{P})$ we concentrate on Finite Element approaches and make the following assumptions.

Assumption

$\Omega \subset \mathbb{R}^n$ denotes a polyhedral domain, $\bar{\Omega} = \cup_{j=1}^{nt} \bar{T}_j$ with admissible quasi-uniform sequences of partitions $\{T_j\}_{j=1}^{nt}$ of Ω, i.e. with $h_{nt} := \max_j \operatorname{diam} T_j$ and $\sigma_{nt} := \min_j \{\sup \operatorname{diam} K; K \subseteq T_j\}$ there holds $c \le \frac{h_{nt}}{\sigma_{nt}} \le C$ uniformly in nt with positive constants $0 < c \le C < \infty$ independent of nt. We abbreviate $\tau_h := \{T_j\}_{j=1}^{nt}$.

In order to tackle $(\mathbb{P})$ numerically we shall distinguish two different approaches. The first is called

First discretize, then optimize,

the second

First optimize, then discretize.

It will turn out that both approaches under certain circumstances lead to the same numerical results. However, from a structural point of view they are completely different.

First discretize, then optimize

All quantities in $(\mathbb{P})$ are discretized a-priori:

- **replace Y and U by finite dimensional subspaces Y_h and U_d,**
- **the set U_{ad} by some discrete counterpart U_{ad}^d, and**
- **the functionals, integrals and dualities by appropriate discrete surrogates.**

Finite element space: For $k \in \mathbb{N}$

$$W_h := \{v \in C^0(\bar{\Omega}); v_{|T_j} \in \mathbb{P}_k(T_j) \text{ for all } 1 \leq j \leq nt\} =: \langle \phi_1, \ldots, \phi_{ng} \rangle, \text{ and}$$

$$Y_h := \{v \in W_h, v_{|\partial\Omega} = 0\} =: \langle \phi_1, \ldots, \phi_n \rangle \subseteq Y,$$

with some $0 < n < ng$.

Ansatz for discrete state: $y_h(x) = \sum_{i=1}^{n} y_i \phi_i$.

Discrete control space: with $u^1, \dots, u^m \in U$, we set

- $U_d := \langle u^1, \dots, u^m \rangle$, and
- $U_{\mathrm{ad}}^d := P_{U_{\mathrm{ad}}}^d(U_d)$, where
- $P_{U_{\mathrm{ad}}}^d : U \to U_{\mathrm{ad}}$ is a sufficiently smooth (nonlinear) mapping.

With $C \subset \mathbb{R}^m$ denoting a convex closed set we assume

$$U_{\mathrm{ad}}^d = \left\{ u \in U; u = \sum_{j=1}^m s_j u^j, s \in C \right\}.$$

Finally let $z_h := Q_h z = \sum_{i=1}^{ng} z_i \phi_i$, where $Q_h : L^2(\Omega) \to W_h$ denotes a continuous projection operator.

Now we replace problem $(\mathbb{P})$ by

$$(\mathbb{P}_{(h,d)}) \begin{cases} \min_{(y_h,u_d)\in Y_h\times U_d} J_{(h,d)}(y,u) := \frac{1}{2}\|y_h - z_h\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\|u_d\|^2_U \\ \text{s.t.} \\ \quad a(y_h, v_h) \quad = \quad \langle Bu_d, v_h\rangle_{Y^*,Y} \qquad \text{for all } v_h \in Y_h, \\ \text{and} \\ u_d \in U^d_{\text{ad}}. \end{cases} \tag{0.5}$$

Here, we have set $a(y,v) := \int_\Omega \nabla y \nabla v dx$.

Introduce Finite Element matrices:

- Stiffness matrix: $A := (a_{ij})_{i,j=1}^{n}$, $a_{ij} := a(\phi_i, \phi_j)$,
- mass matrix $M := (m_{ij})_{i,j=1}^{ng}$, $m_{ij} := \int\limits_{\Omega} \phi_i \phi_j dx$, the
- control matrix $E := (e_{ij})_{i,j=1}^{n,m}$, $e_{ij} = \langle Bu^j, \phi_i \rangle_{Y^*,Y}$, and the
- control mass matrix $F := (f_{ij})_{i,j=1}^{m}$, $f_{ij} := (u^i, u^j)_U$.

Using these quantities allows us to rewrite $(\mathbb{P}_{(h,d)})$ as finite-dimensional optimization problem:

$$(\mathbb{P}_{(n,m)}) \begin{cases} \min_{(y,s)\in\mathbb{R}^n\times\mathbb{R}^m} Q(y,s) := \frac{1}{2}(y-z)^t M(y-z) + \frac{\alpha}{2} s^t F s \\ \text{s.t.} \\ \quad Ay \quad = \quad Es \\ \text{and} \\ s \in C. \end{cases} \tag{0.6}$$

Admissibility is characterized by the closed, convex set $C \subset \mathbb{R}^m$.

Since the matrix A is spd, problem $(\mathbb{P}_{(n,m)})$ is equivalent to minimizing the reduced functional

$$\hat{Q}(s) := Q(A^{-1}Es, s)$$

over the set C.

Problem $(\mathbb{P}_{(n,m)})$ admits a unique solution $(y(s), s) \in \mathbb{R}^n \times C$ which is characterized by the finite dimensional variational inequality

$$(\nabla \hat{Q}(s), t - s)_{\mathbb{R}^m} \geq 0 \text{ for all } t \in C, \tag{0.7}$$

with

$$\nabla \hat{Q}(s) = \alpha Fs + E^t A^{-t} M(A^{-1}Es - z) \equiv \alpha Fs + E^t p,$$

where

$$p := A^{-t} M(A^{-1}Es - z).$$

Comparing

$$\nabla \hat{Q}(s) = \alpha F s + E^t A^{-t} M (A^{-1} E s - z) \equiv \alpha F s + E^t p$$

with

$$\nabla \hat{J}(u) = \alpha u + R B^* S^* (S B u - z) \equiv \alpha u + R B^* p$$

from the infinite-dimensional problem, we note that transposition takes the role of the Riesz isomorphism R,

- **the matrix F takes the role of the identity in U,**
- **the matrix M takes the role of the identity in $L^2(\Omega)$,**
- **the matrix E takes the role the control operator B, and**
- **the matrix A^{-1} that of the solution operator S.**

Problem ($\mathbb{P}_{(n,m)}$) now can be solved numerically with the help of appropriate solution algorithms, which should exploit the structure of the problem. We fix the following

Remark

In the First discretize, then optimize approach the discretization of the adjoint variable $\boldsymbol{p}$ is determined by the test space for the discrete state $\boldsymbol{y_h}$.

In the First optimize, then discretize approach discussed next, this is different.

First optimize, the discretize

Starting point: the first order necessary optimality conditions for problem ($\mathbb{P}$);

$$
(\mathbb{OS}) \quad \left\{ \begin{array}{rcll} -\Delta y & = & Bu & \text{in } \Omega, \\ y & = & 0 & \text{on } \partial\Omega, \\ -\Delta p & = & y - z & \text{in } \Omega, \\ p & = & 0 & \text{on } \partial\Omega, \\ (\alpha u + RB^* p, v - u)_U & \geq & 0 & \text{for all } v \in U_{\text{ad}}. \end{array} \right. \tag{0.8}
$$

- **Discretize everything related to the state y, the control u, and to functionals, integrals, and dualities as in the First discretize, then optimize approach.**
- **In addition, we have the freedom to also select an appropriate discretization of the adjoint variable p.**

For p we choose continuous Finite Elements of order l on τ, which leads to the Ansatz

$$p_h(x) = \sum_{i=1}^{q} p_i \chi_i(x),$$

where

$$\langle \chi_1, \dots, \chi_q \rangle \subset Y$$

denotes the Ansatz space for the adjoint variable.

Matrices:

- adjoint stiffness matrix $\tilde{A} := (\tilde{a}_{ij})_{i,j=1}^{q}$, $\tilde{a}_{ij} := a(\chi_i, \chi_j)$,
- the matrix $\tilde{E} := (\tilde{e}_{ij})_{i,j=1}^{q,m}$, $\tilde{e}_{ij} = \langle Bu^j, \chi_i \rangle_{Y^*,Y}$,
- and the matrix $T := (t_{ij})_{i,j=1}^{n,q}$, $t_{ij} := \int\limits_{\Omega} \phi_i \chi_j dx$.

The discrete analogon to $(\mathbb{OS})$ reads

$$(\mathbb{OS})_{(n,q,m)} \quad \begin{cases} Ay &= Es, \\ \tilde{A}p &= T(y-z), \\ (\alpha Fs + \tilde{E}^t p, t-s)_{\mathbb{R}^m} &\geq 0 \text{ for all } t \in C. \end{cases} \tag{0.9}$$

Since the matrices A and $\tilde{A}$ are spd, this system is equivalent to the variational inequality

$$(\alpha Fs + \tilde{E}^t \tilde{A}^{-1} T(A^{-1}Es - z), t-s)_{\mathbb{R}^m} \geq 0 \text{ for all } t \in C. \tag{0.10}$$

Examples

1. $U := L^2(\Omega)$, $B : L^2(\Omega) \to H^{-1}(\Omega)$ Injection, , $U_{\text{ad}} := \{v \in L^2(\Omega); a \leq v(x) \leq b$ a.e. in $\Omega\}$, $a, b \in L^\infty(\Omega)$. Further let $k = l = 1$ (linear Finite Elements for y and p), $U_d := \langle u^1, \ldots, u^{nt}\rangle$, where $u^k_{|T_i} = \delta_{ki}$ $(k, i = 1, \ldots, nt)$ are piecewise constant functions (i.e. $m = nt$), $C := \prod_{i=1}^{nt} [a_i, b_i]$, where $a_i := a(barycenter(T_i))$, $b_i := b(barycenter(T_i))$.

2. As in 1., but $U_d := \langle \phi_1, \ldots, \phi_{ng}\rangle$ (i.e. $m = ng$), $C := \prod_{i=1}^{ng} [a_i, b_i]$, where $a_i := a(P_i)$, $b_i := b(P_i)$, with P_i $(i = 1, \ldots, ng)$ denoting the vertices of the triangulation τ.

3. (Compare Example 1): As in 1., but $U := \mathbb{R}^m$, $B : \mathbb{R}^m \to H^{-1}(\Omega)$, $Bu := \sum_{j=1}^{m} u_j F_j$, $F_j \in H^{-1}(\Omega)$ given , $U_{\text{ad}} := \{v \in \mathbb{R}^m; a_j \leq v_j \leq b_j\}$, $a < b$, $U_d := \langle e_1, \ldots, e_m\rangle$ with $e_i \in \mathbb{R}^m$ $(i = 1, \ldots, m)$ denoting the $i-$th unitvector, $C := \prod_{i=1}^{ng} [a_i, b_i] \equiv U_d$.

Discussion and implications

- **Choosing the same Ansatz spaces for the state y and the adjoint variable p in the First optimize, then discretize approach leads to an optimality condition which is identical to that of the First discretize, then optimize approach, since then $T \equiv M$.**
- **Choosing a different approach for p in general leads to a non-symmetric matrix T, with the consequence that the matrix $\alpha F + \tilde{E}^t \tilde{A}^{-1} T A^{-1} E$ no longer represents a symmetric matrix (and thus no Hessian), and**
- **the expression $\alpha F s + \tilde{E}^t \tilde{A}^{-1} T (A^{-1} E s - z)$ in general does not represent a gradient.**
- **There is up to now no general recipe which approach has to be preferred, and it should depend on the application and computational resources which approach to take for tackling the numerical solution of the optimization problem.**
- **However, the numerical approach taken should to some extent reflect and preserve the structure which is inherent in the infinite dimensional optimization problem ($\mathbb{P}$).**