Penalty and Barrier Methods for Optimal Control Problems with Control Constraints

Max Winkler

Universität der Bundeswehr München
Institut für Mathematik und Bauinformatik

11^{th} November 2011
5^{th} German-Polish Conference on Optimization
Cracow

Outline

Problem Formulation

- $\Omega \subset \mathbb{R}^d$ ($d \in \{2,3\}$) bounded domain with $C^{0,1}$-boundary,
- $\alpha > 0$ and b, $y_d \in L_\infty(\Omega)$ fixed.

Model Problem

$$J(y,u) := \frac{1}{2}\|y - y_d\|^2_{L_2(\Omega)} + \frac{\alpha}{2}\|u\|^2_{L_2(\Omega)} \to \min!$$

$$\text{s. t.} \quad \begin{aligned} -\Delta y &= u && \text{in } \Omega, \\ \frac{\partial y}{\partial n} + y &= 0 && \text{on } \Gamma, \\ u \in U_{ad} &:= \{u \in L_2(\Omega) : u \le b, \text{ a. e. in } \Omega\}. \end{aligned} \qquad \text{(P)}$$

Universität der Bundeswehr München

The State Equation

We define the *control-to-state-mapping* $S : U \to U$ by

$$y := S\,u \iff a(y,v) = (u,v) \quad \forall\, v \in V$$

$$\text{with} \quad a(v,w) := \int_\Omega \nabla v \cdot \nabla w + \int_\Gamma v\,w.$$

From Lax-Milgram-Lemma follows

Theorem

The images $S\,u$ belong to the space $V := H^1(\Omega)$ and S is well-defined. Further, there exists a $C > 0$ such that

$$\|S\,u\|_{H^1(\Omega)} \le C_a \|u\|.$$

hold. If, additionally, Ω is convex, the images $S\,u$ belong to $H^2(\Omega)$.

The Reduced Model Problem

The reduced model problem is defined by

$$J(u) := \frac{1}{2}\|S\,u - y_d\|^2 + \frac{\alpha}{2}\|u\|^2 \to \min! \quad \text{s.t.} \quad u \in U_{ad}. \qquad \text{(RP)}$$

Theorem (Existence of a solution of (RP))

There exists a unique $\bar{u} \in U_{ad}$ which solves (RP). Further, $\bar{u} \in U_{ad}$ is a solution of (RP), iff

$$(S^*(S\bar{u} - y_d) + \alpha\bar{u}, u - \bar{u}) \geq 0 \qquad \forall u \in U_{ad}. \qquad \text{(VI)}$$

Optimality Conditions

Theorem (Alternative Optimality Conditions (Hinze et. al. (2009))

The following statements are equivalent:

1. $\bar{u} \in U_{ad}$ *solves* (VI),
2. *For arbitrary* $\beta > 0$ *the optimal control* $\bar{u} \in U_{ad}$ *is a fixed-point of*

$$F_\beta \colon U \to U, \qquad F_\beta u := \Pi_{ad}\left(u - \beta\, J'(u)\right)$$

with the projection operator $[\Pi_{ad} u](x) := \min\{b(x), u(x)\}$.

3. *There exists a non-negative function* $\bar{\lambda} \in L_2(\Omega)$, *such that*

$$J'(\bar{u}) + \bar{\lambda} = 0 \quad \text{and} \quad \bar{\lambda}(\bar{u} - b) = 0$$

are satified, almost everywhere in Ω.

The Penalty Functional

We define a functional Φ with

$$\lim_{s\to 0} \Phi(u-b,s) = \begin{cases} \infty, & \text{for } u \notin U_{ad} \\ 0, & \text{for } u \in \operatorname{int} U_{ad}. \end{cases} \qquad (1)$$

Using a finite-dimensional penalty function $\phi : \mathbb{R} \times \mathbb{R}_+ \to \mathbb{R}$ condition (1) is fulfilled by

$$\Phi(u-b,s) := \int_\Omega \phi(u(x)-b(x),s)\,\mathrm{d}x.$$

The Penalty Functional

We define a functional Φ with

$$\lim_{s\to 0} \Phi(u-b,s) = \begin{cases} \infty, & \text{for } u \notin U_{ad} \\ 0, & \text{for } u \in \text{int } U_{ad}. \end{cases} \tag{1}$$

Using a finite-dimensional penalty function $\phi : \mathbb{R} \times \mathbb{R}_+ \to \mathbb{R}$ condition (1) is fulfilled by

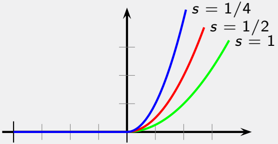
$$\Phi(u-b,s) := \int_\Omega \phi(u(x)-b(x),s)\,\mathrm{d}x.$$

Assumption: ϕ is differentiable w. r. t. the first argument and there exists a function $\psi : \mathbb{R} \to \mathbb{R}$ such that

$$\frac{\partial}{\partial t}\phi(t,s) = \psi\left(\frac{t}{s}\right) \qquad \forall\, t \in dom(\psi).$$

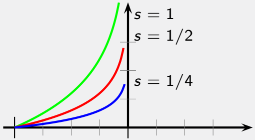
Quadratic Loss Penalty

$$\psi(t) := \max\{0, t\},$$
$$\phi(t,s) := \frac{1}{2\,s}\max^2\{0, t\}.$$



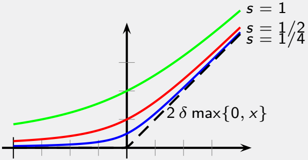
Logarithmic Barrier

$$\psi(t) := \frac{1}{-t}, \qquad \text{for } t < 0,$$
$$\phi(t,s) := -s\,\ln(-t), \quad \text{for } t < 0.$$



Smoothed Exact Penalties

$$\psi(t) := \delta\left(1 + \frac{t}{\sqrt{t^2+1}}\right),$$
$$\phi(t,s) := \delta\left(t + \sqrt{t^2+s^2}\right).$$



We replace the constrained $u \leq b$ by our penalty functional and define the auxiliary problems

Auxiliary Problem (AP)

$$J(u,s) := \frac{1}{2}\|S\,u - y_d\|^2 + \frac{\alpha}{2}\|u\|^2 + \int_\Omega \phi(u(x) - b(x), s)\mathrm{d}x \to \min! \qquad \text{(AP)}$$

s. t. $u \in U$.

- unconstrained optimization problem
- $s > 0$ is the penalty parameter

Existence of Solutions and Optimality Condition

By differentiation we get the optimality condition

$$S^*(S\,\bar{u} - y_d) + \alpha\,\bar{u} + \psi\left(\frac{\bar{u} - b}{s}\right) = 0.$$

Inserting the *state* $y := S\,u$ and *adjoint state* $p := -S^*(y - y_d)$ we get

Existence of Solutions and Optimality Condition

By differentiation we get the optimality condition

$$S^*(S\,\bar{u} - y_d) + \alpha\,\bar{u} + \psi\left(\frac{\bar{u} - b}{s}\right) = 0.$$

Inserting the *state* $y := S\,u$ and *adjoint state* $p := -S^*(y - y_d)$ we get

Theorem

For arbitrary $s > 0$ problem (AP) possesses a unique solution $u(s)$. An element $u(s) \in U$ is a solution of (AP), iff $y(s), p(s) \in V$ exist such that

$$\begin{aligned} a(y, p(s)) - (y_d - y(s), y) &= 0 && \forall\ y \in V, \\ a(y(s), v) - (u(s), v) &= 0 && \forall\ v \in V, \\ \alpha\,u(s) - p(s) + \psi\left(\frac{u(s) - b}{s}\right) &= 0 && \textit{f.ü. in } \Omega \end{aligned}$$

holds.

Convergence of $u(s)$

We consider the quadratic loss penalty

$$\psi(t) := \max\{0, t\}, \qquad \phi(t,s) := \frac{1}{2\,s}\max^2\{0, t\}.$$

Therefor holds

Theorem (Convergence of auxiliary solutions)

The solutions $u(s)$ of problem (AP) converge towards the optimal control $\bar{u}$ of (P) in $L_2(\Omega)$, i. e.

$$\lim_{s\to 0} \|u(s) - \bar{u}\|_{L_2(\Omega)} = 0.$$

Further, a constant $C > 0$ exists such that

$$\|u(s) - \bar{u}\|_{L_2(\Omega)} \leq C\,s. \tag{2}$$

So far, only $\|u(s) - \bar{u}\| \leq C\,s^{\frac{1}{2}}$ could be proven (Grossmann/Kunz/Meischner 2009).

Universität der Bundeswehr München

The optimality condition is equivalent to the fixed-point equation

$$F_\beta\, \bar{u} := \Pi_{ad}\left(\bar{u} - \beta\, J'(\bar{u})\right) = \bar{u}. \tag{3}$$

Since F_β is contracting for sufficiently small β one obtains

$$\|u_n - \bar{u}\| \leq \frac{k^n}{1-k}\|u_1 - u_0\| \quad \text{with} \quad u_{n+1} = F_\beta u_n,\ u_0 \in U \text{ arbitrary.}$$

The optimality condition is equivalent to the fixed-point equation

$$F_\beta\,\bar{u} := \Pi_{ad}\left(\bar{u} - \beta\,J'(\bar{u})\right) = \bar{u}. \qquad (3)$$

Since F_β is contracting for sufficiently small β one obtains

$$\|u_n - \bar{u}\| \le \frac{k^n}{1-k}\|u_1 - u_0\| \quad \text{with} \quad u_{n+1} = F_\beta u_n,\ u_0 \in U \text{ arbitrary.}$$

With the choice $n = 0$ and $u_0 = u(s)$ follows

$$\|u(s) - \bar{u}\| \le \frac{1}{1-k}\|F_\beta\, u(s) - u(s)\|.$$

The optimality condition is equivalent to the fixed-point equation

$$F_\beta\,\bar{u} := \Pi_{ad}\left(\bar{u} - \beta\,J'(\bar{u})\right) = \bar{u}. \qquad (3)$$

Since F_β is contracting for sufficiently small β one obtains

$$\|u_n - \bar{u}\| \le \frac{k^n}{1-k}\|u_1 - u_0\| \quad \text{with} \quad u_{n+1} = F_\beta u_n,\; u_0 \in U \text{ arbitrary.}$$

With the choice $n = 0$ and $u_0 = u(s)$ follows

$$\|u(s) - \bar{u}\| \le \frac{1}{1-k}\|F_\beta\,u(s) - u(s)\|.$$

By the optimality condition

$$0 = J'(u(s)) + \psi\left(\frac{u(s) - b}{s}\right)$$

we have a representation of $J'(u(s))$. Inserting this into (3) yields

$$F_\beta\,u(s) = \Pi_{ad}\,u(s).$$

Universität der Bundeswehr München

The definition of Π_{ad} yields

$$u(s) - \Pi_{ad}\, u(s) = \max\{0, u(s) - b\}. \tag{4}$$

The definition of Π_{ad} yields

$$u(s) - \Pi_{ad}\, u(s) = \max\{0, u(s) - b\}. \tag{4}$$

Using the definition of $J(u, s)$, optimality of $u(s)$, saddle-point inequality, non-negativity of $\bar\lambda$ and Cauchy-Schwarz's inequality we have

$$\begin{aligned} J(u(s), s) \;&:=\; J(u(s)) + \frac{1}{2s}\|\max\{0, u(s) - b\}\|_{L_2}^2 \\ &\le\; J(\bar u, s) \;=\; J(\bar u) \\ &\le\; J(u(s)) + \big(\bar\lambda, u(s) - b\big) \\ &\le\; J(u(s)) + \int_\Omega \bar\lambda \max\{0, u(s) - b\} \\ &\le\; J(u(s)) + \|\bar\lambda\|_{L_2}\|\max\{0, u(s) - b\}\|_{L_2}. \end{aligned} \tag{5}$$

The definition of Π_{ad} yields

$$u(s) - \Pi_{ad}\, u(s) = \max\{0, u(s) - b\}. \tag{4}$$

Using the definition of $J(u,s)$, optimality of $u(s)$, saddle-point inequality, non-negativity of $\bar{\lambda}$ and Cauchy-Schwarz's inequality we have

$$\begin{aligned} J(u(s),s) \;&:=\; J(u(s)) + \frac{1}{2s}\|\max\{0, u(s)-b\}\|_{L_2}^2 \\ &\leq\; J(\bar{u}, s) \;=\; J(\bar{u}) \\ &\leq\; J(u(s)) + \left(\bar{\lambda}, u(s) - b\right) \\ &\leq\; J(u(s)) + \int_\Omega \bar{\lambda} \max\{0, u(s) - b\} \\ &\leq\; J(u(s)) + \|\bar{\lambda}\|_{L_2} \|\max\{0, u(s) - b\}\|_{L_2}. \end{aligned} \tag{5}$$

(4) and (5) together yield

$$\|\Pi_{ad}\, u(s) - u(s)\| \leq 2\|\bar{\lambda}\|\, s.$$

This proves (2).

Idea

Original Problem:

- The optimality condition reads

$$(-\bar{p} + \alpha\bar{u}, u - \bar{u}) \geq 0 \qquad \forall u \in U_{ad}. \tag{6}$$

- The projection formula

$$\bar{u} = \Pi_{ad}\left(\frac{1}{\alpha}\bar{p}\right) \tag{7}$$

is equivalent to (6).

Idea

Original Problem:

- The optimality condition reads

$$(-\bar{p} + \alpha\bar{u}, u - \bar{u}) \geq 0 \qquad \forall u \in U_{ad}. \tag{6}$$

- The projection formula

$$\bar{u} = \Pi_{ad}\left(\frac{1}{\alpha}\bar{p}\right) \tag{7}$$

  is equivalent to (6).

Auxiliary Problem:

- The optimality condition for the auxiliary problem is

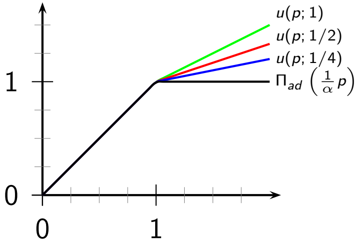
$$-p(s) + \alpha u(s) + \psi\left(\frac{u(s) - b}{s}\right) = 0$$

- Is it possible to find a formula similar to (7)?

In case of quadratic loss penalty we have a pointwise resolution

$$0 = \alpha\, u(x) - p(x) + \frac{1}{s}\max\{0, u(x) - b(x)\} \qquad \Longleftrightarrow$$

$$\boxed{u(p;s)(x) := \frac{1}{\alpha} p(x) - \max\left\{0, \frac{\frac{1}{\alpha} p(x) - b(x)}{1 + \alpha\, s}\right\}}\,.$$



Obviously, there holds

$$\lim_{s\to 0} u(p;s) = \Pi_{ad}\left(\frac{1}{\alpha} p\right).$$

We introduce the Operator

$$D\colon H^1(\Omega) \to H^1(\Omega)^* \qquad \text{with} \qquad \langle D\,y, v\rangle := a(y, v).$$

Theorem (Optimality Condition after Control Reduction)

Let $(y(s), p(s)) \in V \times V$ be the root of the operator equation

$$F : H^1(\Omega) \times H^1(\Omega) \times \mathbb{R}_+ \to H^1(\Omega)^* \times H^1(\Omega)^*$$

$$0 = F(y, p; s) := \begin{pmatrix} Dp + y - y_d \\ Dy - u(p; s) \end{pmatrix}. \tag{8}$$

Then $u(s) = u(p(s); s)$ solves the auxiliary problem (AP).

- Optimality condition depends only on smooth $H^1(\Omega)$-functions
- The elements $v(s) := (y(s), p(s))$ characterize a central path

Newton's Method for (8)

Problem: F is not Fréchet-Differentiable

(We remember: $u(p;s) = \frac{1}{\alpha}p - \max\left\{0, \frac{\frac{1}{\alpha}p - b}{1+\alpha\, s}\right\}$)

Idea: Convergence analysis of Newton's method using a weaker differentiability concept (Semi-Smooth Newton Methods)

Newton's Method for (8)

Problem: F is not Fréchet-Differentiable

(We remember: $u(p;s) = \frac{1}{\alpha}p - \max\left\{0, \frac{\frac{1}{\alpha}p-b}{1+\alpha\, s}\right\}$)

Idea: Convergence analysis of Newton's method using a weaker differentiability concept (Semi-Smooth Newton Methods)

Theorem (Convergence of Semi-Smooth Newton Method)

Assume, $S\,u \in H^2(\Omega)$ for all $u \in L_2(\Omega)$ and

$$meas(\Omega_\varepsilon) \le C\,\varepsilon \quad \text{where} \quad \Omega_\varepsilon := \left\{x \in \Omega \colon 0 < \left|\frac{1}{\alpha}p(s)(x) - b(x)\right| < \varepsilon\right\}$$

Then, constants $c_\rho, c_\delta > 0$ exist such that

$$\|v - v(s)\|_{L_\infty(\Omega)} \le c_\rho \Longrightarrow \|\tilde{v} - v(s)\|_{L_\infty(\Omega)} \le c_\delta\,\|v - v(s)\|^2_{L_\infty(\Omega)},$$

whereas $\tilde{v} := v + d$ with the Newton correction d.

A Path-Following Algorithm

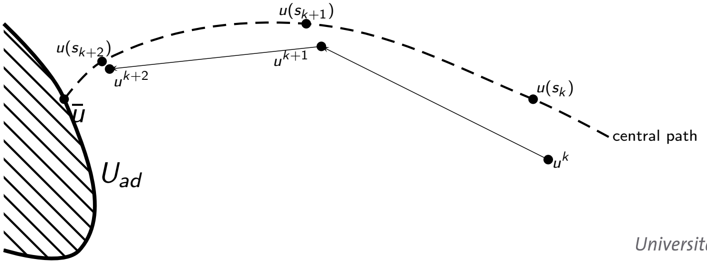
S1 Choose $y^0, p^0 \in V \times V$ and $\sigma \in (0,1)$, $s_0 > 0$, set $k = 0$

S2 Solve

$$\begin{pmatrix} I & D \\ D & -u_p(p^k; s_k) \end{pmatrix} \begin{pmatrix} \delta y \\ \delta p \end{pmatrix} = - \begin{pmatrix} D\,p^k + y^k - y_d \\ D\,y^k - u(p^k; s_k) \end{pmatrix}$$

S3 Set $\quad y^{k+1} = y^k + \delta y, \qquad p^{k+1} = p^k + \delta p$

S4 If no stopping criteria holds, set $s_{k+1} = \sigma \cdot s_k$ and $k = k + 1$ go to $S2$



- $\{\Omega_i\}_{i=0}^M$ quasi-uniform triangulation of Ω
- *conforming finite-element discretization* with

$$V_h := \big\{v \in C(\bar{\Omega})\colon\ v|_{\Omega_i} \in P_1(\Omega_i),\ i = 0, \ldots, M\big\} \subset V.$$

The *discrete control-to-state mapping* is defined by

$$\begin{aligned} S_h\colon U &\to V_h \subset U \\ y_h &= S_h\, u \qquad :\Longleftrightarrow\ a(y_h, v_h) = (u, v_h) \quad \forall\, v_h \in V_h. \end{aligned}$$

Discretized optimization problems

$$J_h(u) := \frac{1}{2}\|S_h\, u - y_d\|^2 + \frac{\alpha}{2}\|u\|^2 \to \min! \qquad \text{s.t.} \quad u \in U_{ad},$$

$$J_h(u,s) := \frac{1}{2}\|S_h\, u - y_d\|^2 + \frac{\alpha}{2}\|u\|^2 + \frac{1}{2\,s}\int_\Omega \max{}^2\{0, u - b\} \to \min!$$

$$\text{s.t.} \quad u \in U.$$

By applying the same theory as in the continuous case, we can derive:

Convergence of Newton's Method

For given p_h^0 and y_h^0, the sequence $\{(y_h^k, p_h^k)\}$, generated by

$$\begin{pmatrix} I_h & D_h \\ D_h & -u_p(p_h^k; s_k) \end{pmatrix} \begin{pmatrix} \delta y_h \\ \delta p_h \end{pmatrix} = - \begin{pmatrix} D_h\, p_h^k + y_h^k - y_d \\ D_h\, y_h^k - u(p_h^k; s_k) \end{pmatrix} \qquad (9)$$

$$(y_h^{k+1}, p_h^{k+1}) := (y_h^k, p_h^k) + (\delta y_h, \delta p_h)$$

converges towards the discrete auxiliary solution $(y_h(s), p_h(s))$.

- System (9) is finite-dimensional
- Analogous to the continuous case, we can implement a complete path-following algorithm

Mesh Independence

Remember: For the continuous solutions $\bar{u}$ and $u(s)$ holds

$$\|u(s) - \bar{u}\|_{L_2(\Omega)} \leq \frac{2}{1-k}\|\bar{\lambda}\|_{L_2(\Omega)}\, s.$$

Analogously, one can derive the same estimate for the discrete solutions $\bar{u}_h$ and $u_h(s)$. One can show boundedness of $\|\bar{\lambda}_h\|_{L_2(\Omega)}$, this leads to

Theorem (Grossmann / Winkler 2011)

For the discrete auxiliary solutions $u_h(s)$ holds

$$\lim_{s\to 0} \|u_h(s) - \bar{u}_h\| = 0$$

and there exists a constant $C > 0$ independent of h such that

$$\|u_h(s) - \bar{u}_h\| \leq C\, s$$

for all $s > 0$.

Conclusion and Outlook

- Conclusion
  - Quadratic Penalty Method for solving control-constrained optimal control problems
  - Linear convergence of auxiliary solutions $\|u(s) - \bar{u}\| \leq C\, s$
    - For continuous problem
    - For discrete problem with $C \neq C(h)$
  - Semi-Smooth Newton-Method for solving auxiliary problems
    - Quadratic convergence under additional assumptions
- Outlook
  - Convergence of Path-Following Algorithm
    - requires estimate for radius of convergence like $c_\rho \geq \bar{c}_\rho\, s$
    - derive bounds for reduction parameter σ $(s_{k+1} = \sigma s_k)$
  - Extension to other penalty-barrier functions
    - for logarithmic barrier and smoothed exact penalty only $\|u(s) - \bar{u}\| \leq C\, s^{1/2}$ could be proven
    - experiments show that locally $\mathcal{O}(s)$ is reached