Generalized Penalty Methods
Solving the resulting subproblems

Stefan Takacs
Seminar Infinite Dimensional Optimization

Jan, 26, 2010

Doctoral Program
Computational Mathematics
Numerical Analysis and Symbolic Computation

supported by

LAND OBERÖSTERREICH
FWF Der Wissenschaftsfonds.
JOHANNES KEPLER UNIVERSITÄT LINZ | JKU

Outline

1. Introduction
   - Formulation of the (sub-)problem
   - Some properties
2. Newton's method with line search
   - Formulation of the algorithm
   - Convergence analysis
3. Newton's method with smoothed Newton step
   - Assumptions and formulation of the algorithm
   - Convergence analysis
   - Application to model problem

Outline

Optimization problem

In this talk we want to solve problems like

$$(P_q) \qquad \min_{z\in Z} J_q(z)$$

$$\text{subject to} \qquad Ez = 0$$

$$\text{where} \qquad J_q(z) := J(z) + \Psi_q(z)$$

$$:= J(z) + \underbrace{\sum_{i=1}^{m} \int_{\Omega_i} \psi_{i,q_i}(g_i(z)(x) - \varphi_i(x))\mathrm{d}x}_{\Psi_{i,q_i}(z):=}$$

- Z is a Hilbert space
- $E : Z \to V$ bounded and linear

$$Z_E := \{z \in Z \ : \ Ez = 0\}$$

- $g_i : Z \to L^{r_i}(\Omega_i)$ and $\varphi_i \in L^{r_i}(\Omega_i)$
- $\psi_{i,q} : \mathbb{R} \to (-\infty, \infty]$ is a penalty function
- Assume q to be fixed

Model problem

$$(MP_q) \qquad \min_{z=(y,u)} J_q(z) := J(z) + \Psi_q(z)$$

$$:= \frac{1}{2}\|y - y_d\|^2_{L^2(\Omega)} + \frac{\gamma}{2}\|u - u_d\|^2_{L^2(\Omega)} + \Psi_q(z)$$

subject to

$$-\Delta y = u \quad \text{in } \Omega$$

$$y = 0 \quad \text{on } \partial\Omega$$

where

$$\Psi_q(z) := \sum_{i=1}^{3} \int_\Omega \psi_{i,q_i}(g_i(z)(x) - \varphi_i(x))\mathrm{d}x$$

$$g_1(z) := u \qquad \varphi_1 := \varphi_u$$

$$g_2(z) := y \qquad \varphi_2 := \varphi_y$$

$$g_3(z) := |\nabla y(x)|_2 \qquad \varphi_3 := \varphi_g$$

Basic assumptions

(P1) There is a feasible point z_f

(P2) The cost functional J is **convex** and lower semi continuous **(l.s.c.)** on Z_E

(P3) $\exists \alpha > 0$ s.t. $\forall z_1, z_2 \in Z_E$:

$$J(z_2) \geq J(z_1) + J'(z_1, z_2 - z_1) + \frac{\alpha}{2} \|z_2 - z_1\|^2$$

(P4) For every closed convex set $U \subset L^{r_i}(\Omega_i)$ the pre-image $\{z \in Z_E \;:\; g_i(z) \in U\}$ is **closed**. Further all g_i are convex in z (for a.e. $x \in \Omega_i$)

(Q1) $\psi_{i,q}$ is **convex, l.s.c. and increasing** and

$$(-\infty, 0) \subset \text{dom}\psi_{i,q}$$

Idea

- We have to solve a general non-linear optimization problem with equality constraints
- (Within the model problem) J itself is smooth
- **Idea:** Use a Newton like method
- **But:** Ψ_q is not sufficiently smooth

Additional assumptions on the problem

(P5a) J is twice continuously differentiable on Z_E
$D^2 J(z)$ is bounded within bounded subsets of Z_E

(P5b) g_i has the representation $g_i(z)(x) = \eta_i((A_i\, z)(x))$, where
$A_i \in \mathcal{L}(Z, L^2(\Omega_i)^{d_i})$ is a continuous linear operator; $d_i \geq 1$
$\eta_i : \mathbb{R}^{d_i} \to \mathbb{R}$ is a convex function

Note that (P5b) does not imply that g_i is twice continuously differentiable, even if η_i is smooth.

Additional assumptions on the problem

(P5a) J is twice continuously differentiable on Z_E
$D^2 J(z)$ is bounded within bounded subsets of Z_E

(P5b) g_i has the representation $g_i(z)(x) = \eta_i((A_i\, z)(x))$, where
$A_i \in \mathcal{L}(Z, L^2(\Omega_i)^{d_i})$ is a continuous linear operator; $d_i \geq 1$
$\eta_i : \mathbb{R}^{d_i} \rightarrow \mathbb{R}$ is a convex function

Note that (P5b) does not imply that g_i is twice continuously differentiable, even if η_i is smooth.

Additional assumptions on the problem

(PQ2a) $\pi_i : \mathbb{R}^{d_i} \times \mathbb{R} \to \mathbb{R}$, $\pi_i(s,t) := \psi_{i,q_i}(\eta_i(s) - t)$ is twice continuously differentiable w.r.t. s for all $t \in \mathbb{R}$
$\exists\ L > 0$ const. s.t.

$$|D^2_{ss}\pi_i(s,t)|_2 \leq L$$

$$|D^2_{ss}\pi_i(s,t) - D^2_{ss}\pi_i(\bar{s},t)|_2 \leq L|s-\bar{s}|_2$$

for all $(s,\bar{s},t) \in \mathbb{R}^{d_i} \times \mathbb{R}^{d_i} \times \mathbb{R}$ where

$$|D^2_{ss}\pi(s,t)|_2 := \sup\{\langle D^2_{22}\pi_i(s,t)h,k\rangle \ :\ |h|_2 = 1, |k|_2 = 1\}$$

(PQ2b) $\Psi_{i,q}$ is continuous on Z

Note that in this setting $\Psi_{i,q}(z) = \int_{\Omega_i} \pi_i((A_i\, z)(x), \varphi_i(x))dx$

Additional assumptions on the problem

(PQ2a) $\pi_i : \mathbb{R}^{d_i} \times \mathbb{R} \to \mathbb{R}$, $\pi_i(s,t) := \psi_{i,q_i}(\eta_i(s) - t)$ is twice continuously differentiable w.r.t. s for all $t \in \mathbb{R}$
$\exists\ L > 0$ const. s.t.

$$|D^2_{ss}\pi_i(s,t)|_2 \leq L$$

$$|D^2_{ss}\pi_i(s,t) - D^2_{ss}\pi_i(\bar{s},t)|_2 \leq L|s - \bar{s}|_2$$

for all $(s, \bar{s}, t) \in \mathbb{R}^{d_i} \times \mathbb{R}^{d_i} \times \mathbb{R}$ where

$$|D^2_{ss}\pi(s,t)|_2 := \sup\{\langle D^2_{22}\pi_i(s,t)h, k\rangle \ : \ |h|_2 = 1, |k|_2 = 1\}$$

(PQ2b) $\Psi_{i,q}$ is continuous on Z

Note that in this setting $\Psi_{i,q}(z) = \int_{\Omega_i} \pi_i((A_i\, z)(x), \varphi_i(x))\mathrm{d}x$

Discussion for the model problem

The conditions (P5) and (PQ2) are

- not fulfilled for combined logarithmic-quadratic penalty function for original setting and model problem
- **Replace** constraint 3 by:

$$\hat{g}_3(z) := \sqrt{1 + |\nabla y|_2^2} \leq \sqrt{1 + \varphi_g^2} =: \hat{\varphi}_g$$

Then: The conditions are fulfilled for combined penalty function with

$$\begin{aligned} \pi_1(s,t) &= \psi_{1,q}(s-t) & A_1(y,u) &= u \\ \pi_2(s,t) &= \psi_{2,q}(s-t) & A_2(y,u) &= y \\ \pi_3((s_1,s_2),t) &= \psi_{3,q}\left(\sqrt{1+|(s_1,s_2)|_2^2} - t\right) & A_3(y,u) &= \nabla y \end{aligned}$$

Differentiability

Lemma (Gfrerer (5.1) - Differentiability)

Assume: *(P4), (P5b), (PQ2)*

Then: *$\Psi_{i,q}$ is twice Gâteaux-differentiable and $\Psi_{i,q} \in C^{1,1}(Z)$ where*

$$\langle D\Psi_{i,q}(z), h\rangle = \int_{\Omega_i} \langle D_s\pi_i(A_i z(x), \varphi_i(x)), A_i h(x)\rangle \mathrm{d}x$$

$$\langle D^2\Psi_{i,q}(z)h, k\rangle = \int_{\Omega_i} \langle D^2_{ss}\pi_i(A_i z(x), \varphi_i(x))A_i h(x), A_i k(x)\rangle \mathrm{d}x$$

and some continuity result on the second derivative holds:
for all $z \in Z$:

$$\lim_{z'\to z} \sup_{h\in\mathcal{B},\tilde{k}\in\tilde{\mathcal{K}}} \int_{\Omega_i} |\langle (D^2_{ss}\pi_i(A_i z'(x), \varphi(x)) - D^2_{ss}\pi_i(A_i z(x), \varphi(x)))h(x), \tilde{k}(x)\rangle|\mathrm{d}x = 0$$

Differentiability

Lemma (Gfrerer (5.1) - Differentiability)

Assume: *(P4), (P5b), (PQ2)*
Then: $\Psi_{i,q}$ *is twice Gâteaux-differentiable and* $\Psi_{i,q} \in C^{1,1}(Z)$ *and some continuity result on the second derivative holds:*
for all $z \in Z$*:*

$$\lim_{z' \to z} \sup_{h \in \mathcal{B}, \tilde{k} \in \tilde{\mathcal{K}}} \int_{\Omega_i} |\langle (D^2_{ss}\pi_i(A_i z'(x), \varphi(x)) - D^2_{ss}\pi_i(A_i z(x), \varphi(x))) h(x), \tilde{k}(x) \rangle| \mathrm{d}x = 0$$

for every bounded subset $\mathcal{B} \subset L^2(\Omega_i)^{d_i}$ *and every* $\tilde{\mathcal{K}}$ *with* ***either****:*

- $\tilde{\mathcal{K}}$ *is bounded in* $L^{\hat{r}_i}(\Omega_i)^{d_i}$ *with* $\hat{r}_i > 2$ ***or***
- $\tilde{\mathcal{K}} = \{R(x)k(x) \ : \ k(x)$ *belongs to* $\mathcal{K} \subset L^2(\Omega_i)^{d_i}$ *compact and* $R \in \mathcal{R} \subset L^\infty(\Omega_i)^{d_i \times d_i}$ *bounded*$\}$.

Outline

Newton's method with line search

If we have a twice differentiable convex function, it is reasonable to apply Newtons' method:

1. Choose $0 < \gamma < 1$, $z^0 \in Z_E$; Set $n := 0$
2. Compute $h^n \in Z_E$ such that it minimizes
$$\frac{1}{2}\langle D^2 J_q(z^n)h, h\rangle + \langle DJ_q(z^n), h\rangle$$
3. Line search: Choose $\sigma_n \in \{1, \frac{1}{2}, \frac{1}{4}, \ldots\}$ s.t.
$$J_q(z^n + \sigma_n h^n) \leq J_q(z^n) + \gamma\sigma_n \langle DJ_q(z^n), h^n\rangle$$
4. Set $z^{n+1} := z^n + \sigma_n h^n$; Set $n := n + 1$ and goto 2 if stop. crit. is not fulfilled

Newton's method with line search

If we have a twice differentiable convex function, it is reasonable to apply Newtons' method:

1. Choose $0 < \gamma < 1$, $z^0 \in Z_E$; Set $n := 0$
2. Compute $h^n \in Z_E$ such that it minimizes

$$\frac{1}{2}\langle D^2 J_q(z^n)h, h\rangle + \langle DJ_q(z^n), h\rangle$$

3. Line search: Choose $\sigma_n \in \{1, \frac{1}{2}, \frac{1}{4}, \ldots\}$ s.t.

$$J_q(z^n + \sigma_n h^n) \leq J_q(z^n) + \gamma\sigma_n\langle DJ_q(z^n), h^n\rangle$$

4. Set $z^{n+1} := z^n + \sigma_n h^n$; Set $n := n+1$ and goto 2 if stop. crit. is not fulfilled

Some remarks

- In (MP) the PDE is part of the constraints of the quadratic subproblems.
- The quadratic subproblems

$$\min_{h \in Z_E} \frac{1}{2} \langle D^2 J_q(z^n) h, h \rangle + \langle D J_q(z^n), h \rangle$$

can be solved e.g. using the optimality system:
Find a stationary point $(h, p) \in Z \times V$ of

$$\frac{1}{2} \langle D^2 J_q(z^n) h, h \rangle + \langle D J_q(z^n), h \rangle + \langle E h, p \rangle$$

Leads to KKT-system:

$$\begin{array}{rcccl} D^2 J_q(z^n) h & + & E^* p & = & -D J_q(z^n) \\ E h & & & = & 0 \end{array}$$

Convergence of Newton's method

Theorem (Gfrerer (5.3); Convergence)

Assume *(P1) – (P5), (Q1) and (PQ2) and let z^n be generated by Newton's method with line search.*
Then: $\lim_{n\to\infty} z^n = \overline{z}_q$.

If Ψ_q (and therefore J_q) is twice continuously differentiable on Z_E, then the algorithm converges q-superlinear.

Continuous Differentiability

Corollary (Gfrerer (5.2) - Continuous Differentiability)

Assume: *(P4), (P5b), (PQ2)*

and moreover ***either***

- A_i *is compact from* Z_E *into* $L^2(\Omega_i)^{d_i}$ ***or***
- $A_i \in \mathcal{L}(Z_E, L^{\hat{r}_i}(\Omega_i)^{d_i})$ *with* $r_i > 2$.

Then: $\Psi_{i,q}$ *is twice* ***continuously*** *differentiable on* Z_E

Model problem

- H_0^1 is compactly embedded in $L^2 \Rightarrow A_2(y,u) = y$ is compact $\Rightarrow_{\text{Corr. 5.2}} \Psi_{2,q}$ is twice continuously differentiable
- $-\Delta : H_0^1(\Omega) \to H^{-1}(\Omega)$ is homeomorphism $\Rightarrow$ $\nabla \circ (-\Delta)^{-1} \in \mathcal{L}(H^{-1}(\Omega), L^2(\Omega)^d)$. Since L^2 is compactly embedded in $H^{-1}(\Omega) \Rightarrow \nabla \circ (-\Delta)^{-1}$ is compact on $L^2 \Rightarrow$ $A_3(y,u) = \nabla y$ is compact on Z_E
  $\Rightarrow_{\text{Corr. 5.2}} \Psi_{3,q}$ is twice continuously differentiable
- **But:** for $A_1(y,u) = u$ the assumptions of Corr. 5.2 are not fulfilled.
  One can show: $\Psi_{1,q}$ is nowhere twice Fréchet differentiable

So we cannot show q-superlinear convergence.

Outline

Smoothed Newton step

- In this section we modify Newton's algorithm such that the method converges q-superlinear also in the case that Ψ_q is not twice continuously differentiable

Additional assumptions

Recall: (P5b) tells us: $g_i(z) = \eta_i(A_i z)$

(P6a) We can split the constraints ($\exists m'$):

- for $m'+1, \dots, m$ ("good" constraints):
  we have $\Psi_{i,q} \in C^2(Z_E)$
  Remark (Corr. 5.3) sufficient:
  - $A_i \in \mathcal{L}(Z, L^2(\Omega_i)^{d_i})$ is compact **or**
  - $A_i \in \mathcal{L}(Z, L^{\tilde{r}_i}(\Omega_i)^{d_i})$ where $\tilde{r}_i > 2$
- for $i = 1, \dots, m'$ (the others):
  $A_i = \gamma_i B + C_i$ with

  $B \in \mathcal{L}(Z, L^2(\tilde{\Omega})^{\tilde{d}})$ is **common** lin. operator

  $C_i \in \mathcal{L}(Z, L^2(\tilde{\Omega})^{\tilde{d}})$ are **compact** lin. operators

  $\gamma_i \in \mathbb{R}$

  Notice: all (these) A_i live in the same spaces (e.g. all $d_i = \tilde{d}$)

(P6b) The mapping $\mathcal{H} : Z \to V \times L^2(\tilde{\Omega})^{\tilde{d}}$, $\mathcal{H}(z) := (Ez, Bz)$ is surjective.

Additional assumptions

Recall: (P5b) tells us: $g_i(z) = \eta_i(A_i z)$

(P6a) We can split the constraints ($\exists m'$):

- for $m'+1, \ldots, m$ ("good" constraints):
  we have $\Psi_{i,q} \in C^2(Z_E)$
  Remark (Corr. 5.3) sufficient:
  - $A_i \in \mathcal{L}(Z, L^2(\Omega_i)^{d_i})$ is compact **or**
  - $A_i \in \mathcal{L}(Z, L^{\tilde{r}_i}(\Omega_i)^{d_i})$ where $\tilde{r}_i > 2$
- for $i = 1, \ldots, m'$ (the others):
  $A_i = \gamma_i B + C_i$ with

  $B \in \mathcal{L}(Z, L^2(\tilde{\Omega})^{\tilde{d}})$ is **common** lin. operator

  $C_i \in \mathcal{L}(Z, L^2(\tilde{\Omega})^{\tilde{d}})$ are **compact** lin. operators

  $\gamma_i \in \mathbb{R}$

  Notice: all (these) A_i live in the same spaces (e.g. all $d_i = \tilde{d}$)

(P6b) The mapping $\mathcal{H} : Z \rightarrow V \times L^2(\tilde{\Omega})^{\tilde{d}}$, $\mathcal{H}(z) := (Ez, Bz)$ is surjective.

Newton's method with smoothed Newton step

1. Choose $0 < \gamma < 1$, $z^0 \in Z_E$; Set $n := 0$
2. Compute $\zeta_1 \in Z$ such that it minimizes

$$\begin{aligned} G_{z^n}(\zeta) &:= \frac{1}{2}\langle D^2 J_q(z^n)\zeta, \zeta\rangle + \langle DJ_q(z^n), \zeta\rangle \\ &= \frac{1}{2}\langle D^2 J(z^n)\zeta, \zeta\rangle + \langle DJ(z^n), \zeta\rangle \\ &\quad + \sum_{i=1}^{m}\left(\frac{1}{2}\langle D^2\Psi_{i,q}(z^n)\zeta, \zeta\rangle + \langle D\Psi_{i,q}(z^n), \zeta\rangle\right) \end{aligned}$$

   subject to $E\zeta = 0$ and $B\zeta = 0$

   Derive a multiplier $(v_1^*, \nu^*) \in V^* \times (L^2(\tilde{\Omega})^{\tilde{d}})^*$ such that

$$DG_{z^n}(\zeta_1) + E^* v_1^* + B^* \nu^* = 0$$

3. Compute $\zeta_2 \in \mathcal{U}$
4. Compute $\zeta_3 \in Z$
5. Set $z^{n+1} := z^n + \zeta_1 + \zeta_2 + \zeta_3$; Set $n := n + 1$ and goto 2 if

Newton's method with smoothed Newton step

1. Choose $0 < \gamma < 1$, $z^0 \in Z_E$; Set $n := 0$
2. Compute $\zeta_1 \in Z$
3. Compute $\zeta_2 \in \mathcal{U}$ ($\approx (\text{Ker } B)^\perp$) such that it minimizes

$$
\begin{aligned}
T_n(\zeta) \quad := \quad & \langle E^* v_1^*, \zeta \rangle \\
& + \frac{1}{2} \langle D^2 J(z^n)(\zeta_1 + \zeta), \zeta_1 + \zeta \rangle + \langle DJ(z^n), \zeta_1 + \zeta \rangle \\
& + \sum_{i=1}^{m'} \int_{\tilde{\Omega}} (\pi_i((A_i(z^n + \zeta_1) + \gamma_i B \zeta)(x), \varphi_i(x)) + \\
& \qquad \langle D_s \pi_i (A_i(z^n + \zeta_1)), \varphi_i(x)), C_i \zeta(x) \rangle) \mathrm{d}x \\
& + \sum_{i=m'+1}^{m} \left(\frac{1}{2} \langle D^2 \Psi_{i,q}(z^n)(\zeta_1 + \zeta), \zeta_1 + \zeta \rangle \right. \\
& \qquad \left. + \langle D\Psi_{i,q}(z^n), \zeta_1 + \zeta \rangle \right)
\end{aligned}
$$

4. Compute $\zeta_3 \in Z$

Newton's method with smoothed Newton step

1. Choose $0 < \gamma < 1$, $z^0 \in Z_E$; Set $n := 0$
2. Compute $\zeta_1 \in Z$
3. Compute $\zeta_2 \in \mathcal{U}$
4. Compute $\zeta_3 \in Z$ such that it minimizes

$$\frac{1}{2}\langle D^2 J(z^n)(\zeta_1 + \zeta_2 + \zeta), \zeta_1 + \zeta_2 + \zeta\rangle + \langle DJ(z^n), \zeta_1 + \zeta_2 + \zeta\rangle$$
$$+ \sum_{i=1}^{m'} \left(\frac{1}{2}\langle D^2\Psi_{i,q}(z^n + \zeta_1 + \zeta_2)\zeta, \zeta\rangle + \langle D\Psi_{i,q}(z^n + \zeta_1 + \zeta_2), \zeta\rangle\right)$$
$$+ \sum_{i=m'+1}^{m} \Big(\frac{1}{2}\langle D^2\Psi_{i,q}(z^n)(\zeta_1 + \zeta_2 + \zeta), \zeta_1 + \zeta_2 + \zeta\rangle$$
$$+ \langle D\Psi_{i,q}(z^n), \zeta_1 + \zeta_2 + \zeta\rangle\Big)$$

subject to $E(\zeta_2 + \zeta) = 0$

5. Set $z^{n+1} := z^n + \zeta_1 + \zeta_2 + \zeta_3$; Set $n := n + 1$ and goto 2 if

Newton's method with smoothed Newton step

1. Choose $0 < \gamma < 1$, $z^0 \in Z_E$; Set $n := 0$
2. Compute $\zeta_1 \in Z$
3. Compute $\zeta_2 \in \mathcal{U}$
4. Compute $\zeta_3 \in Z$
5. Set $z^{n+1} := z^n + \zeta_1 + \zeta_2 + \zeta_3$; Set $n := n + 1$ and goto 2 if stop. crit. is not fulfilled

Convergence rate

Lemma (Gfrerer (5.4); convergence rate)

Assume *(P1) – (P6), (Q1) and (PQ2).*
Then:

$$\|\zeta_1\|_Z + \|\zeta_2\|_Z + \|\zeta_3\|_Z + \|\nu^*\|_{L^2(\tilde{\Omega}^{\tilde{d}})} + \|v_3^* - v_1^*\|_{V^*} = \mathcal{O}(\|z^n - \bar{z}_q\|_{Z_E})$$

Convergence rate

Theorem (Gfrerer (5.5); convergence rate)

Assume *(P1) – (P6), (Q1) and (PQ2) and moreover*

- *that there is some bounded set* $\tilde{\mathcal{K}} \subset L^2(\tilde{\Omega})^{\tilde{d}}$ *which is either*
  - *bounded in* $L^{\hat{r}}(\tilde{\Omega})^{\tilde{d}}$ *with* $\hat{r} > 2$ *or*
  - *the elements have the form* $\tilde{k}(x) = R(x)k(y)$, *where* k *belongs to a compact subset* $\mathcal{K} \subset L^2(\tilde{\Omega})^{\tilde{d}}$ *and* R *or from a bounded subset* $\mathcal{R} \subset L^\infty(\tilde{\Omega})^{\hat{d}\times\hat{d}}$

  such that for the smoothed Newton steps we have $\operatorname{dist}(B\zeta_3, \|B\zeta_3\|\tilde{\mathcal{K}}) = o(\|z^n - \overline{z}_q\|_Z)$ *for all* $z^n \in Z_E$ *in some neighborhood of* $\overline{z}_q$.

Then*: there exists a increasing function* $\omega : \mathbb{R}_+ \to \mathbb{R}_+$ *with* $\lim_{t\to 0_+} \omega(t) = 0$ *such that*

$$\|z^{n+1} - \overline{z}_q\|_Z \leq \omega(\|z^n - \overline{z}_q\|_Z)\|z^n - \overline{z}_q\|_Z$$

(q-superlinear convergence)

Model problem

We can apply Newton's method with smoothed Newton step to (MP)

- **Recall:** $\Psi_{2,q}$ and $\Psi_{3,q}$ are twice continuously differentiable on $Z_E \Rightarrow m' = 1$
- $\Psi_{1,q}(y,u) = u$ is not twice continuously differentiable
  But a decomposition as in (P6) is possible:
  $B(y,u) := A_1(y,u) = u$, $\gamma_1 := 1$, $C_1(y,u) := 0$ and
  $\mathcal{U} := \{0\} \times L^2(\Omega)$ $(\approx (\mathrm{Ker} B)^\perp)$

Model problem: the algorithm

1. Let $z^n = (y^n, u^n)$ be some iterate
2. Find $\zeta_1 := (\zeta_{1,y}, \zeta_{1,u})$ such that it minimizes something **subject to**

   $$E(\zeta_y, \zeta_y) = 0, \text{ i.e., } -\Delta\zeta_y = \zeta_u \text{ in } \Omega \text{ with } \zeta_y = 0 \text{ on } \partial\Omega$$

   $B(\zeta_y, \zeta_y) = 0$, i.e., $\zeta_u = 0$ on Ω.
   Obviously this PDE has one unique solution: $\zeta_y \equiv 0$.
   The multiplier $v_1^* \in H_0^1(\Omega)$ is given by the variational problem

   $$\int_\Omega (\langle \nabla v_1^*, \nabla v \rangle + (y^n - y_d)v + D_s\pi_2(y^n, \varphi_y)v + \langle D_s\pi_3(\nabla y^n, \hat{\varphi}_g), \nabla v \rangle) = 0 \quad \forall v \in H_0^1(\Omega)$$

3. Find $\zeta_2 := (\zeta_{2,y}, \zeta_{2,u})$, where $\zeta_{2,y} = 0$ and $\zeta_{2,u} \in L^2(\Omega)$, where for each $x \in \Omega$ the value $\zeta_{2,u}(x)$ minimizes ($\zeta \in \mathbb{R}$)

   $$-v_1^*(x)\zeta + \beta((u^n(x) - u_d(x))\zeta + \frac{1}{2}\zeta^2) + \pi_1(u^n(x) + \zeta, \varphi_u)$$

4. Find $\zeta_3 := (\zeta_{3,y}, \zeta_{3,u})$ such that it minimizes

Model problem: the algorithm

1. Let $z^n = (y^n, u^n)$ be some iterate
2. Find $\zeta_1 := (\zeta_{1,y}, \zeta_{1,u})$ such that it minimizes something **subject to**

   $$E(\zeta_y, \zeta_y) = 0, \text{ i.e., } -\Delta\zeta_y = \zeta_u \text{ in } \Omega \text{ with } \zeta_y = 0 \text{ on } \partial\Omega$$

   $$B(\zeta_y, \zeta_y) = 0, \text{ i.e., } \zeta_u = 0 \text{ on } \Omega.$$

   Obviously this PDE has one unique solution: $\zeta_y \equiv 0$.
   The multiplier $v_1^* \in H_0^1(\Omega)$ is given by the variational problem

   $$\int_\Omega (\langle \nabla v_1^*, \nabla v\rangle + (y^n - y_d)v + D_s\pi_2(y^n, \varphi_y)v + \langle D_s\pi_3(\nabla y^n, \hat{\varphi}_g), \nabla v\rangle) = 0 \quad \forall v \in H_0^1(\Omega)$$

3. Find $\zeta_2 := (\zeta_{2,y}, \zeta_{2,u})$, where $\zeta_{2,y} = 0$ and $\zeta_{2,u} \in L^2(\Omega)$, where for each $x \in \Omega$ the value $\zeta_{2,u}(x)$ minimizes ($\zeta \in \mathbb{R}$)

   $$-v_1^*(x)\zeta + \beta((u^n(x) - u_d(x))\zeta + \frac{1}{2}\zeta^2) + \pi_1(u^n(x) + \zeta, \varphi_u)$$

4. Find $\zeta_3 := (\zeta_{3,y}, \zeta_{3,u})$ such that it minimizes

Model problem: the algorithm

1. Let $z^n = (y^n, u^n)$ be some iterate
2. The multiplier $v_1^* \in H_0^1(\Omega)$ is given by ...
3. Find $\zeta_2 := (\zeta_{2,y}, \zeta_{2,u})$, where $\zeta_{2,y} = 0$ and $\zeta_{2,u} \in L^2(\Omega)$, where for each $x \in \Omega$ the value $\zeta_{2,u}(x)$ minimizes ($\zeta \in \mathbb{R}$)

$$-v_1^*(x)\zeta + \beta((u^n(x) - u_d(x))\zeta + \frac{1}{2}\zeta^2) + \pi_1(u^n(x) + \zeta, \varphi_u)$$

4. Find $\zeta_3 := (\zeta_{3,y}, \zeta_{3,u})$ such that it minimizes

$$\begin{aligned}\int_\Omega \Big(\quad & (y^n - y_d)\zeta_y + \frac{1}{2}\zeta_y^2 \\ + \quad & \beta((u^n - u_d)(\zeta_u + \zeta_{2,u}) + \frac{1}{2}(\zeta_u + \zeta_{2,u})^2) \\ + \quad & D_s\pi_1(u^n + \zeta_{2,u}, \varphi_u)\zeta_u + \frac{1}{2}D_{ss}^2\pi_1(u^n + \zeta_{2,u}, \varphi_u)\zeta_u^2 \\ + \quad & D_s\pi_2(y^n, \varphi_y)\zeta_y + \frac{1}{2}D_{ss}^2\pi_2(y^n, \varphi_y)\zeta_y^2 \\ + \quad & \langle D_s\pi_3(\nabla y^n, \hat{\varphi}_g\rangle + \frac{1}{2}\langle D_{ss}^2\pi_3(\nabla y^n, \hat{\varphi}_g)\nabla\zeta_y, \zeta_y\rangle\Big)\end{aligned}$$

Model problem: the algorithm

1. Let $z^n = (y^n, u^n)$ be some iterate
2. The multiplier $v_1^* \in H_0^1(\Omega)$ is given by ...
3. Find $\zeta_2 := (\zeta_{2,y}, \zeta_{2,u})$
4. Find $\zeta_3 := (\zeta_{3,y}, \zeta_{3,u})$ such that it minimizes

$$
\begin{aligned}
\int_\Omega \Big(\quad & (y^n - y_d)\zeta_y + \frac{1}{2}{\zeta_y}^2 \\
+ \quad & \beta((u^n - u_d)(\zeta_u + \zeta_{2,u}) + \frac{1}{2}(\zeta_u + \zeta_{2,u})^2) \\
+ \quad & D_s\pi_1(u^n + \zeta_{2,u}, \varphi_u)\zeta_u + \frac{1}{2}D_{ss}^2\pi_1(u^n + \zeta_{2,u}, \varphi_u){\zeta_u}^2 \\
+ \quad & D_s\pi_2(y^n, \varphi_y)\zeta_y + \frac{1}{2}D_{ss}^2\pi_2(y^n, \varphi_y){\zeta_y}^2 \\
+ \quad & \langle D_s\pi_3(\nabla y^n, \hat{\varphi}_g\rangle + \frac{1}{2}\langle D_{ss}^2\pi_3(\nabla y^n, \hat{\varphi}_g)\nabla\zeta_y, \zeta_y\rangle\Big)
\end{aligned}
$$

Remarks

The multiplier $v_3^* \in H_0^1(\Omega)$ in step 3 fulfills

$$\begin{aligned} 0 &= \int_\Omega (\langle \nabla v_3^*, \nabla v\rangle + (y^n - y_d + \zeta_{3,y} + D_s\pi_2(y^n,\varphi_y) \\ &\quad + D_{ss}^2\pi_2(y^n,\varphi_y)\zeta_{3,y})v + \langle D_s\pi_3(\nabla y^n,\hat\varphi_g) \\ &\quad + D_{ss}^2\pi_3(\nabla y^n,\hat\varphi_g)\nabla\zeta_{3,y},\nabla v\rangle) \qquad \forall v \in H_0^1(\Omega) \\ 0 &= \beta(u^n - u_d + \zeta_{2,u} + \zeta_{3,u}) + D_s\pi_1(u^n + \zeta_{2,u},\varphi_u) \\ &\quad + D_{ss}^2\pi_1(u^n + \zeta_{2,u},\varphi_u)\zeta_{3,u} - v_3^* \end{aligned}$$

Deduce (using construction of ζ_2) for all $x \in \Omega$

$$(\beta + D_{ss}^2\pi_1(u^n(x) + \zeta_{2,u}(x),\varphi_u(x)))\zeta_{3,u}(x) = v_3^*(x) - v_1^*(x)$$

By convexity of $\pi_1(s,t)$ w.r.t. s have $D_{ss}^2\pi_1(u^n(x) + \zeta_{2,u}(x),\varphi_u(x)) \geq 0$ and since $H_0^1(\Omega)$ is compactly embedded in $L^2(\Omega)$, together with Lemma 5.4, the assumptions of the convergence theorem are fulfilled.

$\Rightarrow$ **algorithm converges for (MP) superlinearly.**

Remarks and Conclusions

- To ensure global convergence:
  E.g.: Apply alternating: smoothed Newton step and Newton step with line search
  Accept smoothed Newton step only if decrease in objective is achieved
- Numerical results show good results (if the approximation close enough to the exact solution)

Thanks for your attention!

Literature

- H. Gfrerer: Generalized Penalty Methods for a Class of Convex Optimization Problems with Pointwise Inequality Constraints