Generalized Penalty Methods

Stefan Takacs

Seminar Infinite Dimensional Optimization

Jan, 19, 2010

Doctoral Program
Computational Mathematics
Numerical Analysis and Symbolic Computation

supported by

LAND OBERÖSTERREICH

FWF Der Wissenschaftsfonds.

JOHANNES KEPLER UNIVERSITÄT LINZ | JKU

Outline

Outline

1. Introduction
   - General problem
   - Some properties
2. Generalized Penalty Methods
   - Barrier function
   - Convergence result
   - Construction of generalized penalty functions
3. Error estimates
   - A-priori error estimates
   - Application to examples of penalty functions
4. Duality
   - Lagrange Multipliers
   - Estimating Lagrange Multipliers

Optimization problem

$$(P) \qquad \min_{z \in Z} J(z)$$

subject to

$$Ez = 0$$

$$g_i(z)(x) \leq \varphi_i(x) \qquad \text{for a.e. } x \in \Omega_i,\ i = 1, \dots m$$

- 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 $\varphi_i \in L^{r_i}(\Omega_i)$

$$Z_F := \{z \in Z \ : \ Ez = 0 \text{ and } g_i(z)(x) \leq \varphi_i(x) \text{ for a.e. } x \in \Omega_i\}$$

Optimization problem

$$(P) \qquad \min_{z \in Z} J(z)$$

subject to

$$Ez = 0$$

$$g_i(z)(x) \leq \varphi_i(x) \qquad \text{for a.e. } x \in \Omega_i, \; i = 1, \dots m$$

- 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 $\varphi_i \in L^{r_i}(\Omega_i)$

$$Z_F := \{z \in Z \; : \; Ez = 0 \text{ and } g_i(z)(x) \leq \varphi_i(x) \text{ for a.e. } x \in \Omega_i\}$$

Optimization problem

$$(P) \qquad \min_{z \in Z} J(z)$$

subject to

$$Ez = 0$$

$$g_i(z)(x) \le \varphi_i(x) \qquad \text{for a.e. } x \in \Omega_i,\ i = 1, \dots m$$

- 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 $\varphi_i \in L^{r_i}(\Omega_i)$

$$Z_F := \{z \in Z \ : \ Ez = 0 \text{ and } g_i(z)(x) \le \varphi_i(x) \text{ for a.e. } x \in \Omega_i\}$$

Model problem

$$\text{(MP)} \qquad \min_{z=(y,u)} J(z) = \frac{1}{2}\|y - y_d\|^2_{L^2(\Omega)} + \frac{\gamma}{2}\|u - u_d\|^2_{L^2(\Omega)}$$

subject to

$$\begin{aligned} -\Delta y(x) &= u(x) && \text{in } \Omega \\ y(x) &= 0 && \text{on } \partial\Omega \\ u(x) &\leq \varphi_u(x) && \text{f.a.e } x \in \Omega \\ y(x) &\leq \varphi_y(x) && \text{f.a.e } x \in \Omega \\ |\nabla y(x)|_2 &\leq \varphi_g(x) && \text{f.a.e } x \in \Omega \end{aligned}$$

Model problem

$$(MP) \qquad \min_{z=(y,u)} J(z) = \frac{1}{2}\|y-y_d\|^2_{L^2(\Omega)} + \frac{\gamma}{2}\|u-u_d\|^2_{L^2(\Omega)}$$

subject to

$$-\Delta y(x) = u(x) \qquad \text{in } \Omega$$

$$y(x) = 0 \qquad \text{on } \partial\Omega$$

$$u(x) \leq \varphi_u(x) \qquad \text{f.a.e } x \in \Omega$$

$$y(x) \leq \varphi_y(x) \qquad \text{f.a.e } x \in \Omega$$

$$|\nabla y(x)|_2 \leq \varphi_g(x) \qquad \text{f.a.e } x \in \Omega$$

Model problem

$$
\text{(MP)} \qquad \min_{z=(y,u)} J(z) = \frac{1}{2}\|y - y_d\|^2_{L^2(\Omega)} + \frac{\gamma}{2}\|u - u_d\|^2_{L^2(\Omega)}
$$

subject to

$$
\begin{aligned}
-\Delta y(x) = u(x) &\qquad \text{in } \Omega \\
y(x) = 0 &\qquad \text{on } \partial\Omega \\
u(x) \leq \varphi_u(x) &\qquad \text{f.a.e } x \in \Omega \\
y(x) \leq \varphi_y(x) &\qquad \text{f.a.e } x \in \Omega \\
|\nabla y(x)|_2 \leq \varphi_g(x) &\qquad \text{f.a.e } x \in \Omega
\end{aligned}
$$

Assumptions on the Problem

(P1) There is a feasible point $z_f \in Z_F$

(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. $x \in \Omega_i$)

(MP) fulfills the properties (P2) – (P4)

(P1) – (P4) guarantee unique solution

Assumptions on the Problem

(P1) There is a feasible point $z_f \in Z_F$

(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. $x \in \Omega_i$)

(MP) fulfills the properties (P2) – (P4)
(P1) – (P4) guarantee unique solution

Assumptions on the Problem

(P1) There is a feasible point $z_f \in Z_F$

(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. $x \in \Omega_i$)

(MP) fulfills the properties (P2) – (P4)

(P1) – (P4) guarantee unique solution

Assumptions on the Problem

(P1) There is a feasible point $z_f \in Z_F$

(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. $x \in \Omega_i$)

(MP) fulfills the properties (P2) – (P4)
(P1) – (P4) guarantee unique solution

Assumptions on the Problem

(P1) There is a feasible point $z_f \in Z_F$

(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. $x \in \Omega_i$)

(MP) fulfills the properties (P2) – (P4)

(P1) – (P4) guarantee unique solution

Assumptions on the Problem

(P1) There is a feasible point $z_f \in Z_F$

(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. $x \in \Omega_i$)

(MP) fulfills the properties (P2) – (P4)
(P1) – (P4) guarantee unique solution

Outline

Trivial barrier function

(P) is equivalent to

$$\min_{z \in Z_E} J(z) + \Psi(z) := J(z) + \sum_{i=1}^{m} \mathcal{I}_{(-\infty,0]}(g_i(z)(x) - \varphi_i(x))\mathrm{d}x,$$

where the indicator function $\mathcal{I}_{(-\infty,0]}$ is given by

$$\mathcal{I}_{(-\infty,0]}(t) := \begin{cases} 0 & \text{if } t \leq 0 \\ \infty & \text{if } t > 0. \end{cases}$$

Generalized penalty function

$$(P_q) \qquad \min_{z \in Z_E} J_q(z) := J(z) + \Psi_q(z) := J(z) + \sum_{i=1}^{m} \psi_{i,q_i}(g_i(z)(x) - \varphi_i(x))\mathrm{d}x$$

Consider $q \in Q$ and $q \to \overline{q}$

Idea: $\psi_{i,q_i} \to \mathcal{I}_{(-\infty,0]}$ for $q \to \overline{q}$

Generalized penalty function

$$(P_q) \qquad \min_{z \in Z_E} J_q(z) := J(z) + \Psi_q(z) := J(z) + \sum_{i=1}^{m} \psi_{i,q_i}(g_i(z)(x) - \varphi_i(x))\mathrm{d}x$$

Consider $q \in Q$ and $q \to \overline{q}$

Idea: $\psi_{i,q_i} \to \mathcal{I}_{(-\infty,0]}$ for $q \to \overline{q}$

Convergence result

Theorem (Gfrerer (2.1): Convergence)

Assume (P1) – (P4) and that the ***penalty function*** Ψ_q *fulfills for some* $q \in Q\backslash\{\overline{q}\}$*:*

(1) $\Psi_q : Z \to \mathbb{R} \cup \{\infty\}$ *is well defined and convex and l.s.c on* Z_E

(2) $\Psi_q(z_f) < \infty$ *(finite at some feasible point)*

(3) $\Psi_q(z) \geq \sum_{i=1}^m (b_{i,q}\langle \eta_i^*, z\rangle + a_{i,q})$ *for all* $z \in Z_E$

Then:

- (P_q) *has a unique solution (called* $\overline{z}_q$*)*

Convergence result (cont.)

Theorem (Gfrerer (2.1): Convergence)

*Assume (P1) – (P4) and that the **penalty function** fulfills:*

(1) – (3) *as before for all $q \in Q\backslash\{\overline{q}\}$ uniformly*

(η_i^* does not depend on q; $a_{i,q}$, $b_{i,q} \to 0$ for $q \to \overline{q}$)

(4) *for every $\hat{z} \in Z_E\backslash Z_F$ (**not feasible** for (P)):*
*$\forall R > 0$ there exist **neighborhoods** $U_{\hat{z}} \subset Z_E$ and $U_{\overline{q}} \subset Q$*

$$\inf_{q \in U_{\overline{q}}} \inf_{z \in U_{\hat{z}}} \Psi_q(z) \geq R$$

(5) *for every $\hat{z} \in Z_F$ (**feasible** for (P)):*
there is some family $(\hat{z}_q)_{q \in Q\backslash\{\overline{q}\}} \subset Z_E$ with $\lim_{q \to \overline{q}} \hat{z}_q = \hat{z}$
such that $\limsup_{q \to \overline{q}} \Psi_q(\hat{z}_q) \leq 0$

Then:

- $\lim_{q \to \overline{q}} \overline{z}_q = \overline{z}$ *(convergence to the exact solution of (P))*

Convergence result (cont.)

Theorem (Gfrerer (2.1): Convergence)

Assume (P1) – (P4) and that the ***penalty function*** *fulfills:*

(1) – (3) *as before for all* $q \in Q \backslash \{\overline{q}\}$ *uniformly*

(η_i^* *does not depend on* q; $a_{i,q}$, $b_{i,q} \to 0$ *for* $q \to \overline{q}$)

(4) *for every* $\hat{z} \in Z_E \backslash Z_F$ (***not feasible*** *for* (P))*:*
$\forall R > 0$ *there exist* ***neighborhoods*** $U_{\hat{z}} \subset Z_E$ *and* $U_{\overline{q}} \subset Q$

$$\inf_{q \in U_{\overline{q}}} \inf_{z \in U_{\hat{z}}} \Psi_q(z) \geq R$$

(5) *for every* $\hat{z} \in Z_F$ (***feasible*** *for* (P))*:*
there is some family $(\hat{z}_q)_{q \in Q \backslash \{\overline{q}\}} \subset Z_E$ *with* $\lim_{q \to \overline{q}} \hat{z}_q = \hat{z}$
such that $\limsup_{q \to \overline{q}} \Psi_q(\hat{z}_q) \leq 0$

Then:

- $\lim_{q \to \overline{q}} \overline{z}_q = \overline{z}$ *(convergence to the exact solution of* (P)*)*

Convergence result (cont.)

Theorem (Gfrerer (2.1): Convergence)

Assume (P1) – (P4) and that the ***penalty function*** *fulfills:*

(1) – (3) *as before for all* $q \in Q \backslash \{\overline{q}\}$ *uniformly*

(η_i^* does not depend on q; $a_{i,q}$, $b_{i,q} \to 0$ for $q \to \overline{q}$)

(4) *for every* $\hat{z} \in Z_E \backslash Z_F$ *(**not feasible** for (P)):*
$\forall R > 0$ *there exist* ***neighborhoods*** $U_{\hat{z}} \subset Z_E$ *and* $U_{\overline{q}} \subset Q$

$$\inf_{q \in U_{\overline{q}}} \inf_{z \in U_{\hat{z}}} \Psi_q(z) \geq R$$

(5) *for every* $\hat{z} \in Z_F$ *(**feasible** for (P)):*
there is some family $(\hat{z}_q)_{q \in Q \backslash \{\overline{q}\}} \subset Z_E$ *with* $\lim_{q \to \overline{q}} \hat{z}_q = \hat{z}$
such that $\limsup_{q \to \overline{q}} \Psi_q(\hat{z}_q) \leq 0$

Then:

- $\lim_{q \to \overline{q}} \overline{z}_q = \overline{z}$ *(convergence to the exact solution of (P))*

Convergence result (cont.)

Theorem (Gfrerer (2.1): Convergence)

*Assume (P1) – (P4) and that the **penalty function** fulfills:*

(1) – (3) *as before for all* $q \in Q \backslash \{\overline{q}\}$ *uniformly*

(η_i^* *does not depend on* q; $a_{i,q}$, $b_{i,q} \to 0$ *for* $q \to \overline{q}$)

(4) *for every* $\hat{z} \in Z_E \backslash Z_F$ *(**not feasible** for (P)):*
$\forall R > 0$ *there exist **neighborhoods*** $U_{\hat{z}} \subset Z_E$ *and* $U_{\overline{q}} \subset Q$

$$\inf_{q \in U_{\overline{q}}} \inf_{z \in U_{\hat{z}}} \Psi_q(z) \geq R$$

(5) *for every* $\hat{z} \in Z_F$ *(**feasible** for (P)):*
there is some family $(\hat{z}_q)_{q \in Q \backslash \{\overline{q}\}} \subset Z_E$ *with* $\lim_{q \to \overline{q}} \hat{z}_q = \hat{z}$
such that $\limsup_{q \to \overline{q}} \Psi_q(\hat{z}_q) \leq 0$

Then:

- $\lim_{q \to \overline{q}} \overline{z}_q = \overline{z}$ *(convergence to the exact solution of (P))*

Construction of generalized penalty functions

- Know properties that should be fulfilled by Ψ_q
- Goal: construct functions $\psi_{i,q}$ such that these properties are fulfilled

Some examples for generalized penalty functions

- Quadratic penalty function:

$$\psi_\kappa(t) = \kappa \max\{0, t\}^2$$

- Logarithmic barrier function:

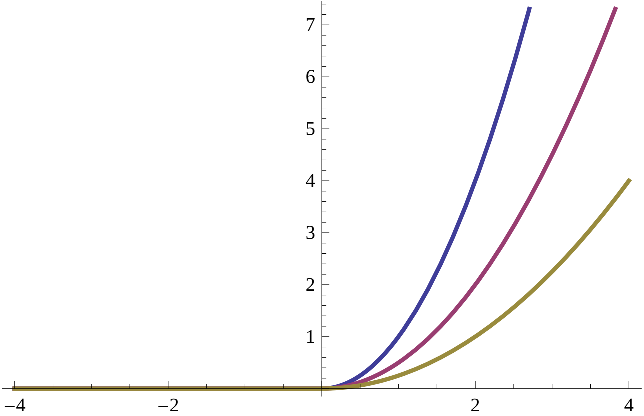
$$\psi_\kappa(t) = \begin{cases} -\kappa \ln(-t) & \text{if } t < 0 \\ \infty & \text{if } t \geq 0 \end{cases}$$

- Combined logarithmic-quadratic penalty function:

$$\psi_{\kappa,\epsilon}(t) = \begin{cases} -\kappa \ln(-t) & \text{if } t \leq -\epsilon \\ \kappa(-\ln(\epsilon) + \frac{t+\epsilon}{\epsilon} + \frac{(t+\epsilon)^2}{2\epsilon^2}) & \text{if } t > -\epsilon \end{cases}$$

where $Q := \{(\kappa, \epsilon) \in \mathbb{R}_+^\infty \ : \ \kappa \geq a\epsilon^{3/2} \text{ and } \kappa^{1/2} \ln \epsilon > -b\}$ $(a, b > 0 \text{ const})$

Some examples for generalized penalty functions



Some examples for generalized penalty functions

- Quadratic penalty function:

$$\psi_\kappa(t) = \kappa \max\{0, t\}^2$$

- Logarithmic barrier function:

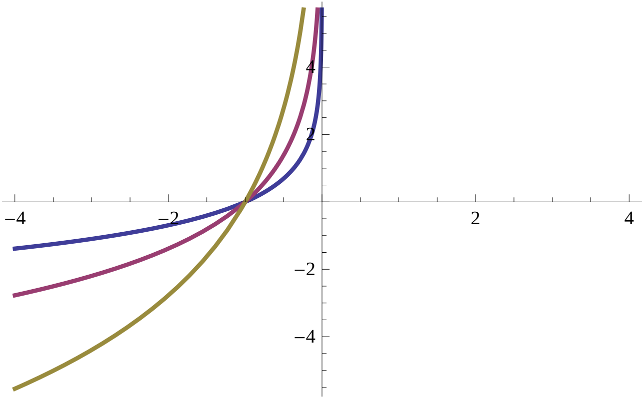
$$\psi_\kappa(t) = \begin{cases} -\kappa \ln(-t) & \text{if } t < 0 \\ \infty & \text{if } t \geq 0 \end{cases}$$

- Combined logarithmic-quadratic penalty function:

$$\psi_{\kappa,\epsilon}(t) = \begin{cases} -\kappa \ln(-t) & \text{if } t \leq -\epsilon \\ \kappa(-\ln(\epsilon) + \frac{t+\epsilon}{\epsilon} + \frac{(t+\epsilon)^2}{2\epsilon^2}) & \text{if } t > -\epsilon \end{cases}$$

where $Q := \{(\kappa, \epsilon) \in \mathbb{R}^\infty_+ \; : \; \kappa \geq a\epsilon^{3/2} \text{ and } \kappa^{1/2} \ln \epsilon > -b\}$ $(a, b > 0 \text{ const})$

Some examples for generalized penalty functions



Some examples for generalized penalty functions

- Quadratic penalty function:

$$\psi_\kappa(t) = \kappa \max\{0, t\}^2$$

- Logarithmic barrier function:

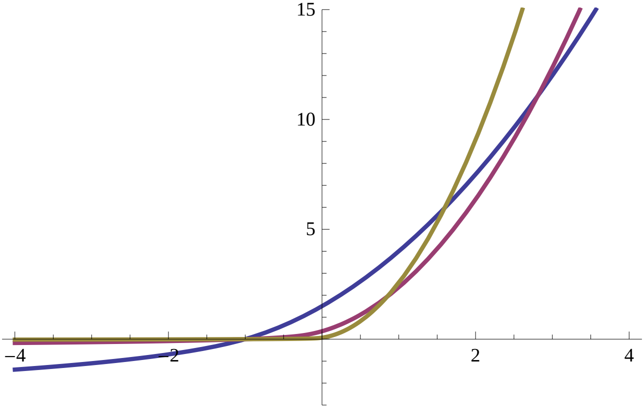
$$\psi_\kappa(t) = \begin{cases} -\kappa \ln(-t) & \text{if } t < 0 \\ \infty & \text{if } t \geq 0 \end{cases}$$

- Combined logarithmic-quadratic penalty function:

$$\psi_{\kappa,\epsilon}(t) = \begin{cases} -\kappa \ln(-t) & \text{if } t \leq -\epsilon \\ \kappa(-\ln(\epsilon) + \frac{t+\epsilon}{\epsilon} + \frac{(t+\epsilon)^2}{2\epsilon^2}) & \text{if } t > -\epsilon \end{cases}$$

where $Q := \{(\kappa, \epsilon) \in \mathbb{R}_+^\infty \; : \; \kappa \geq a\epsilon^{3/2} \text{ and } \kappa^{1/2} \ln \epsilon > -b\}$ $(a, b > 0 \text{ const})$

Some examples for generalized penalty functions



Basic assumptions on generalized penalty functions

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

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

(Q2) $\lim_{q\to\overline{q}} \psi_{i,q}(t) = \mathcal{I}_{(-\infty,0]}(t)$ pointwise for all $t \neq 0$

(PQ1) One of the following conditions hold:

- $g_i(z_f) \le \varphi_i - \delta$ for $\delta > 0$ (in the **interior**)
- $0 \in \mathrm{dom}\psi_{i,q}$ for all $q \in Q\backslash\{\overline{q}\}$ and $\lim_{q\to\overline{q}} \psi_{i,q}(0) = 0$ (*good* behavior for 0)

Basic assumptions on generalized penalty functions

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

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

(Q2) $\lim_{q\to\overline{q}} \psi_{i,q}(t) = \mathcal{I}_{(-\infty,0]}(t)$ pointwise for all $t \neq 0$

(PQ1) One of the following conditions hold:

- $g_i(z_f) \leq \varphi_i - \delta$ for $\delta > 0$ (in the **interior**)
- $0 \in \text{dom}\psi_{i,q}$ for all $q \in Q\backslash\{\overline{q}\}$ and $\lim_{q\to\overline{q}} \psi_{i,q}(0) = 0$ (*good* behavior for 0)

We notice that (Q1), (Q2) and (PQ1) are fulfilled if

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

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

- $\lim_{q\to\overline{q}} \psi_{i,q}(t) = \mathcal{I}_{(-\infty,0]}(t)$ pointwise for all t

These conditions hold for the **quadratic** penalty function and the **combined** penalty function.

Basic assumptions on generalized penalty functions

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

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

(Q2) $\lim_{q \to \overline{q}} \psi_{i,q}(t) = \mathcal{I}_{(-\infty,0]}(t)$ pointwise for all $t \neq 0$

(PQ1) One of the following conditions hold:

- $g_i(z_f) \leq \varphi_i - \delta$ for $\delta > 0$ (in the **interior**)
- $0 \in \mathrm{dom}\psi_{i,q}$ for all $q \in Q \backslash \{\overline{q}\}$ and $\lim_{q \to \overline{q}} \psi_{i,q}(0) = 0$ (*good* behavior for 0)

We notice that (Q1), (Q2) and (PQ1) are fulfilled for the **logarithmic barrier function** iff **there exists an interior point.**

Convergence result

Corollary

Assume that (P1)–(P4), (Q1), (Q2) and (PQ1) hold. Then for each $q \in Q\backslash\{\overline{q}\}$ the problems (P_q) each have a unique solution $\overline{z}_q$ and

$$\lim_{q \to \overline{q}} \overline{z}_q = \overline{z},$$

i.e., the solutions converge to the solution of (P).

One can show that the combination of (Q1), (Q2) and (PQ1) implies the assumptions of the main theorem.

Outline

A-priori error estimates

Theorem (Gfrerer (3.3): Error estimate)

Assume *(P1) – (P4), (Q1), (Q2), (PQ1) and*

(Q3') *For each* $q \in Q \backslash \{\overline{q}\}$ *there are* $c_{i,q} \geq 0$ *and* $d_{i,q} \geq 0$ *s.t.*

$$\psi_{i,q}(t/2) \leq c_{i,q}\psi_{i,q}(t) + d_{i,q} \qquad \forall t \in [-1, 0)$$

A-priori error estimates

Theorem (Gfrerer (3.3): Error estimate)

Then: *there is a neighborhood $U_{\overline{q}} \subset Q$ and a constant L such that for all $q \in U_{\overline{q}}$ we have*

$$\|\overline{z}_q - \overline{z}\|_Z^2 \leq \frac{2}{\alpha}\left(L\,\mathrm{dist}(\overline{z}_q, Z_F) + \sum_{i=1}^{m} \int_{\Omega_i} \psi_{i,q}^{\#}(g_i(\overline{z}_q)(x) - \varphi_i(x), -(g_i(\overline{z}_q)(x) - \varphi_i(x)))\mathrm{d}x \right),$$

where

$$\psi_{i,q}^{\#}(u,h) := \begin{cases} \psi_{i,q}'(u,h) & \text{if } u \in \mathrm{int\ dom}\ \psi_{i,q} \\ \lim_{t\downarrow 0} \frac{\psi_{i,q}(u+th)-\psi_{i,q}(u)}{t} \in \mathbb{R} \cup \{-\infty\} & \text{if } u \in \mathrm{bd\ dom}\ \psi_{i,q}, h < 0 \\ 0 & \text{otherwise} \end{cases}$$

Application to combined penalty function

- Combined logarithmic-quadratic penalty function:

$$\psi_{\kappa,\epsilon}(t) = \begin{cases} -\kappa \ln(-t) & \text{if } t \leq -\epsilon \\ \kappa(-\ln(\epsilon) + \frac{t+\epsilon}{\epsilon} + \frac{(t+\epsilon)^2}{2\epsilon^2}) & \text{if } 0 > t > -\epsilon \end{cases}$$

- We obtain the following derivatives:

$$\psi^{\#}_{\kappa,\epsilon}(t,-t) := \begin{cases} \kappa & \text{if } t \leq -\epsilon \\ -\kappa(\frac{2t}{\epsilon} + \frac{t^2}{\epsilon^2}) \leq \kappa & \text{if } t > -\epsilon \end{cases}$$

- Plugging into the last theorem shows:

$$\|\overline{z}_{(\kappa,\epsilon)} - \overline{z}\|_Z^2 \leq \frac{2}{\alpha}\left(L \operatorname{dist}(\overline{z}_{(\kappa,\epsilon)}, Z_F) + \sum_{i=1}^{m} \kappa \, |\Omega_i| \right)$$

Application to combined penalty function

- Combined logarithmic-quadratic penalty function:

$$\psi_{\kappa,\epsilon}(t) = \begin{cases} -\kappa \ln(-t) & \text{if } t \le -\epsilon \\ \kappa(-\ln(\epsilon) + \frac{t+\epsilon}{\epsilon} + \frac{(t+\epsilon)^2}{2\epsilon^2}) & \text{if } 0 > t > -\epsilon \end{cases}$$

- We obtain the following derivatives:

$$\psi^{\#}_{\kappa,\epsilon}(t,-t) := \begin{cases} \kappa & \text{if } t \le -\epsilon \\ -\kappa(\frac{2t}{\epsilon} + \frac{t^2}{\epsilon^2}) \le \kappa & \text{if } t > -\epsilon \end{cases}$$

- Plugging into the last theorem shows:

$$\|\overline{z}_{(\kappa,\epsilon)} - \overline{z}\|_Z^2 \le \frac{2}{\alpha}\left(L \operatorname{dist}(\overline{z}_{(\kappa,\epsilon)}, Z_F) + \sum_{i=1}^{m} \kappa \, |\Omega_i| \right)$$

Application to combined penalty function

- Combined logarithmic-quadratic penalty function:

$$\psi_{\kappa,\epsilon}(t) = \begin{cases} -\kappa \ln(-t) & \text{if } t \leq -\epsilon \\ \kappa(-\ln(\epsilon) + \frac{t+\epsilon}{\epsilon} + \frac{(t+\epsilon)^2}{2\epsilon^2}) & \text{if } 0 > t > -\epsilon \end{cases}$$

- We obtain the following derivatives:

$$\psi^{\#}_{\kappa,\epsilon}(t,-t) := \begin{cases} \kappa & \text{if } t \leq -\epsilon \\ -\kappa(\frac{2t}{\epsilon} + \frac{t^2}{\epsilon^2}) \leq \kappa & \text{if } t > -\epsilon \end{cases}$$

- Plugging into the last theorem shows:

$$\|\overline{z}_{(\kappa,\epsilon)} - \overline{z}\|_Z^2 \leq \frac{2}{\alpha}\left(L \operatorname{dist}(\overline{z}_{(\kappa,\epsilon)}, Z_F) + \sum_{i=1}^{m} \kappa \, |\Omega_i| \right)$$

Application to logarithmic barrier function

- Logarithmic barrier function:

$$\psi_\kappa(t) = \begin{cases} -\kappa \ln(-t) & \text{if } t < 0 \\ \infty & \text{if } t \geq 0 \end{cases}$$

- We obtain the following derivatives:

$$\psi_\kappa^{\#}(t, -t) := \begin{cases} \kappa & \text{if } t < 0 \\ 0 & \text{if } t \geq 0 \end{cases}$$

- Plugging into the last theorem shows:

$$\|\overline{z}_{(\kappa,\epsilon)} - \overline{z}\|_Z^2 \leq \frac{2}{\alpha} \left(0 + \sum_{i=1}^{m} \kappa \, |\Omega_i| \right)$$

Application to logarithmic barrier function

- Logarithmic barrier function:

$$\psi_\kappa(t) = \begin{cases} -\kappa \ln(-t) & \text{if } t < 0 \\ \infty & \text{if } t \geq 0 \end{cases}$$

- We obtain the following derivatives:

$$\psi_\kappa^\#(t, -t) := \begin{cases} \kappa & \text{if } t < 0 \\ 0 & \text{if } t \geq 0 \end{cases}$$

- Plugging into the last theorem shows:

$$\|\bar{z}_{(\kappa,\epsilon)} - \bar{z}\|_Z^2 \leq \frac{2}{\alpha}\left(0 + \sum_{i=1}^{m} \kappa\, |\Omega_i|\right)$$

Application to logarithmic barrier function

- Logarithmic barrier function:

$$\psi_\kappa(t) = \begin{cases} -\kappa \ln(-t) & \text{if } t < 0 \\ \infty & \text{if } t \geq 0 \end{cases}$$

- We obtain the following derivatives:

$$\psi_\kappa^\#(t, -t) := \begin{cases} \kappa & \text{if } t < 0 \\ 0 & \text{if } t \geq 0 \end{cases}$$

- Plugging into the last theorem shows:

$$\|\overline{z}_{(\kappa,\epsilon)} - \overline{z}\|_Z^2 \leq \frac{2}{\alpha} \left(0 + \sum_{i=1}^{m} \kappa \, |\Omega_i| \right)$$

Summary

- For all three cases some error estimates are possible
- The combined logarithmic-quadratic penalty function can be bounded by essential the same bounds as the logarithmic barrier functions but it works without having an interior point
- For the combined logarithmic-quadratic penalty function κ and ϵ can be adjusted to ensure both contributions of the error to be small

Outline

Lagrange Multipliers

- Sometimes Lagrange multipliers are of our interest (e.g. Stokes problem: multiplier p is the pressure)
- Methods based on the optimality system deliver the solution and the Lagrange multipliers
- For penalty methods the existence of Lagrange multipliers is not required and the multipliers are not computed during the iteration
- If we assume a **constraint qualification condition** ensuring the **existence** of multipliers, the multipliers **can be approximated** using the generalized penalty methods

Lagrange Multipliers

- Sometimes Lagrange multipliers are of our interest (e.g. Stokes problem: multiplier p is the pressure)
- Methods based on the optimality system deliver the solution and the Lagrange multipliers
- For penalty methods the existence of Lagrange multipliers is not required and the multipliers are not computed during the iteration
- If we assume a **constraint qualification condition** ensuring the **existence** of multipliers, the multipliers **can be approximated** using the generalized penalty methods

Lagrange Multipliers

- Sometimes Lagrange multipliers are of our interest (e.g. Stokes problem: multiplier p is the pressure)
- Methods based on the optimality system deliver the solution and the Lagrange multipliers
- For penalty methods the existence of Lagrange multipliers is not required and the multipliers are not computed during the iteration
- If we assume a **constraint qualification condition** ensuring the **existence** of multipliers, the multipliers **can be approximated** using the generalized penalty methods

Lagrange Multipliers

- Sometimes Lagrange multipliers are of our interest (e.g. Stokes problem: multiplier p is the pressure)
- Methods based on the optimality system deliver the solution and the Lagrange multipliers
- For penalty methods the existence of Lagrange multipliers is not required and the multipliers are not computed during the iteration
- If we assume a **constraint qualification condition** ensuring the **existence** of multipliers, the multipliers **can be approximated** using the generalized penalty methods

Additional assumption

(Q3) For each $q \in Q \backslash \{\overline{q}\}$ we assume:

- $\psi_{i,q}$ to be differentiable on $\mathbb{R}$
- $\limsup_{t \to \infty} t^{-r_i} \psi_{i,q}(t) < \infty$
- $\lim_{q \to \overline{q}} \psi_{i,q}(0) = 0$

Notice:

- (Q3) is fulfilled by the **quadratic** penalty method, by the **combined quadratic-logarithmic** penalty method but not by the **logarithmic barrier** method
- (Q3) implies the (PQ1) and (Q3')
- (Q3) and (Q1) imply $|\psi_{i,q}(t)| \leq C_{i,q}(1 + |t|^{r_i})$ for all t. This implies $\int_{\Omega_i} \psi_{i,q}(u(x))dx$ to be l.s.c., convex, real-valued on $L^{r_i}(\Omega_i)$ and **differentiable**

Additional assumption

(Q3) For each $q \in Q \backslash \{\overline{q}\}$ we assume:

- $\psi_{i,q}$ to be differentiable on $\mathbb{R}$
- $\limsup_{t \to \infty} t^{-r_i} \psi_{i,q}(t) < \infty$
- $\lim_{q \to \overline{q}} \psi_{i,q}(0) = 0$

Notice:

- (Q3) is fulfilled by the **quadratic** penalty method, by the **combined quadratic-logarithmic** penalty method but not by the **logarithmic barrier** method
- (Q3) implies the (PQ1) and (Q3')
- (Q3) and (Q1) imply $|\psi_{i,q}(t)| \leq C_{i,q}(1 + |t|^{r_i})$ for all t. This implies $\int_{\Omega_i} \psi_{i,q}(u(x))dx$ to be l.s.c., convex, real-valued on $L^{r_i}(\Omega_i)$ and **differentiable**

Linearize the problem

Therefore we can linearize the problem (P_q).
We know that $0 \leq J_q'(\overline{z}_q, z - \overline{z}_q)$ holds.
So we can derive that $\overline{z}_q$ is the unique solution of

$$\min_{z \in Z_E} J(z) + \langle y_q^*, G_i(z) \rangle,$$

where

- $\langle y_q^*, y \rangle := \sum_{i=1}^m \langle l_{i,q}^*, y_i \rangle$
- $G(z) := (g_1(z), \dots, g_n(z))$
- $l_{i,q}^* \in L^{r_i}(\Omega_i)^*$ such that

$$\langle l_{i,q}^*, y_i \rangle := \int_{\Omega_i} \psi_{i,q}'(g_i(\overline{z}_q)(x) - \varphi_i(x), y_i(x)) \mathrm{d}x$$

Property of the limit of the multipliers

Theorem (Gfrerer (4.2): Property of limit point)

Assume:

- *(P1) – (P4), (Q1) – (Q3)*
- $\hat{Z}$ *to be a Banach space continuously embedded in* Z
- $\hat{Y}$ *to be a Banach space continuously embedded in* $Y := \prod_{j=1}^{m} L^{r_j}(\Omega_j)$ *s.t.* $G(\hat{Z}) \subset \hat{Y}$.
- $(q^n) \subset Q \backslash \{\overline{q}\}$ *with* $q^n \to \overline{q}$
- $\langle y^*_{q^n}, y \rangle := \sum_{j=0}^{m} \langle l^*_{j,q^n}, y_j \rangle$
- $y^*_{q^n} \to_{w^*} y^* \in \hat{Y}^*$ *(convergence)*

Then:

Property of the limit of the multipliers

Theorem (Gfrerer (4.2): Property of limit point)

Assume:

- ...

Then:

- *y^* belongs to the normal cone of C at $G(\bar{z})$*
  (i.e., $\langle y^*, c - G(\bar{z})\rangle \leq 0 \; \forall c \in C$)
  where $C := \{c = (c_1, \dots, c_n) \in \hat{Y} \; : \; c_j \leq \varphi_j \text{ a.e.i. } \Omega_j \; \forall j\}$
  $G(z) := (g_1(z), \dots, g_n(z))$
- *$\bar{z}$ is the unique solution of*

$$\min_{z \in \hat{Z}_E} J(z) + \langle y^*, G(z)\rangle$$

That means: *y^* is a multiplier.*

Convergence of the multipliers

Theorem (Gfrerer (4.3): Convergence)

Assume:

- *(P1) – (P4), (Q1) – (Q3)*

(Q4) $0_Y \in \text{int}(G(\hat{Z}_E) - C)$

Then:

- $\limsup_{q \to \overline{q}} ||y_q^*||_{\hat{Y}^*} < \infty$

This implies the existence of a weak--convergent subsequence.*

Notice that condition (Q4) is a constraint qualification condition

Convergence of the multipliers

Theorem (Gfrerer (4.3): Convergence)

Assume:

- *(P1) – (P4), (Q1) – (Q3)*

(Q4) $0_Y \in \operatorname{int}(G(\hat{Z}_E) - C)$

Then:

- $\limsup_{q \to \overline{q}} ||y_q^*||_{\hat{Y}^*} < \infty$

This implies the existence of a weak--convergent subsequence.*

Notice that condition (Q4) is a constraint qualification condition

Convergence of the multipliers (cont.)

Theorem (Gfrerer (4.4): Convergence)

Assume:

- *the constraints can be partitioned* $C = C_1 \times C_2 \subset \hat{Y}_1 \times \hat{Y}_2 = \hat{Y}$ $G = (G_1, G_2)$ *where* $G_i : \hat{Z} \to \hat{Y}_i$ $(i = 1, 2)$ *such that*
  - $c_1 \in \operatorname{int} G_1(\hat{Z}_E)$ *for some* $c_1 \in C_1$ *(or more generally,* $0 \in \operatorname{int}_{\hat{Y}_1}(G_1(\hat{Z}_E) - C_1)$*) and*
  - $G_1(\tilde{z}) \in C_1,\ G_2(\tilde{z}) \in \operatorname{int}_{\hat{Y}_2} C_2$ *for some* $\tilde{z} \in \hat{Z}_E$

Then: *condition (Q4) is fulfilled.*

Application to model problem

- $\Omega \subset \mathbb{R}^d$ with $C^{1,1}$ boundary
- Assume $\overline{u} \in L^r(\Omega)$ with $r > \max\{d, 2\}$
  (can be ensured using box constraints)
- Choose $\hat{Z} := W^{2,r}(\Omega) \cap W_0^{1,r}(\Omega) \times L^r(\Omega)$
- Choose $\hat{Y} := L^r(\Omega) \times C(\text{cl}\Omega) \times C(\text{cl}\Omega)$
  (since y and ∇y are continuous on $\text{cl}\Omega$ for $y \in W^{2,r}(\Omega)$)
- $G(y, u) := (u, y, |\nabla y|_2)$
- $C := C_u \times C_y \times C_g = \{u \in L^r(\Omega) \ : \ u \le \varphi_u\} \times \{y \in C(\text{cl}\Omega) \ : \ y \le \varphi_y\} \times \{g \in C(\text{cl}\Omega) \ : \ g \le \varphi_g\}$
- (Q4) is fulfilled if there exists some feasible $(y, u) \in \hat{Z}$ with $y \in \text{int}\, C_y$, $|\nabla y|_2 \in \text{int}\, C_g$
  (i.e., $\exists \delta > 0$ such that $y \le \varphi_y - \delta$ and $|\nabla y|_2 \le \varphi_g - \delta$)

Application to model problem

- $\Omega \subset \mathbb{R}^d$ with $C^{1,1}$ boundary
- Assume $\overline{u} \in L^r(\Omega)$ with $r > \max\{d, 2\}$
  (can be ensured using box constraints)
- Choose $\hat{Z} := W^{2,r}(\Omega) \cap W_0^{1,r}(\Omega) \times L^r(\Omega)$
- Choose $\hat{Y} := L^r(\Omega) \times C(\mathrm{cl}\Omega) \times C(\mathrm{cl}\Omega)$
  (since y and ∇y are continuous on $\mathrm{cl}\Omega$ for $y \in W^{2,r}(\Omega)$)
- $G(y, u) := (u, y, |\nabla y|_2)$
- $C := C_u \times C_y \times C_g = \{u \in L^r(\Omega) \ : \ u \leq \varphi_u\} \times \{y \in C(\mathrm{cl}\Omega) \ : \ y \leq \varphi_y\} \times \{g \in C(\mathrm{cl}\Omega) \ : \ g \leq \varphi_g\}$
- (Q4) is fulfilled if there exists some feasible $(y, u) \in \hat{Z}$ with $y \in \mathrm{int}\, C_y$, $|\nabla y|_2 \in \mathrm{int}\, C_g$
  (i.e., $\exists \delta > 0$ such that $y \leq \varphi_y - \delta$ and $|\nabla y|_2 \leq \varphi_g - \delta$)

Application to model problem

- $\Omega \subset \mathbb{R}^d$ with $C^{1,1}$ boundary
- Assume $\overline{u} \in L^r(\Omega)$ with $r > \max\{d, 2\}$
  (can be ensured using box constraints)
- Choose $\hat{Z} := W^{2,r}(\Omega) \cap W_0^{1,r}(\Omega) \times L^r(\Omega)$
- Choose $\hat{Y} := L^r(\Omega) \times C(\mathrm{cl}\Omega) \times C(\mathrm{cl}\Omega)$
  (since y and ∇y are continuous on $\mathrm{cl}\Omega$ for $y \in W^{2,r}(\Omega)$)
- $G(y,u) := (u, y, |\nabla y|_2)$
- $C := C_u \times C_y \times C_g = \{u \in L^r(\Omega) \ : \ u \leq \varphi_u\} \times$
  $\{y \in C(\mathrm{cl}\Omega) \ : \ y \leq \varphi_y\} \times \{g \in C(\mathrm{cl}\Omega) \ : \ g \leq \varphi_g\}$
- (Q4) is fulfilled if there exists some feasible $(y,u) \in \hat{Z}$ with $y \in \mathrm{int}\, C_y$, $|\nabla y|_2 \in \mathrm{int}\, C_g$
  (i.e., $\exists \delta > 0$ such that $y \leq \varphi_y - \delta$ and $|\nabla y|_2 \leq \varphi_g - \delta$)

Application to model problem

- $\Omega \subset \mathbb{R}^d$ with $C^{1,1}$ boundary
- Assume $\overline{u} \in L^r(\Omega)$ with $r > \max\{d, 2\}$

  (can be ensured using box constraints)
- Choose $\hat{Z} := W^{2,r}(\Omega) \cap W_0^{1,r}(\Omega) \times L^r(\Omega)$
- Choose $\hat{Y} := L^r(\Omega) \times C(\text{cl}\Omega) \times C(\text{cl}\Omega)$

  (since y and ∇y are continuous on clΩ for $y \in W^{2,r}(\Omega)$)
- $G(y, u) := (u, y, |\nabla y|_2)$
- $C := C_u \times C_y \times C_g = \{u \in L^r(\Omega) \; : \; u \leq \varphi_u\} \times$
  $\{y \in C(\text{cl}\Omega) \; : \; y \leq \varphi_y\} \times \{g \in C(\text{cl}\Omega) \; : \; g \leq \varphi_g\}$
- (Q4) is fulfilled if there exists some feasible $(y, u) \in \hat{Z}$ with $y \in \text{int}\, C_y$, $|\nabla y|_2 \in \text{int}\, C_g$

  (i.e., $\exists \delta > 0$ such that $y \leq \varphi_y - \delta$ and $|\nabla y|_2 \leq \varphi_g - \delta$)

Thanks for your attention!

Literature

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