Universität Konstanz Sommersemester 2014
Fachbereich Mathematik und Statistik
Prof. Dr. Stefan Volkwein
Roberta Mancini, Carmen Gräßle, Laura Lippmann, Kevin Sieg

Optimierung

http://www.math.uni-konstanz.de/numerik/personen/volkwein/teaching/

Sheet 4

Deadline for hand-in: 2014/06/18 at lecture

Optimization with boundary constraints.
So far we looked for (local) minimizer $x^* \in \mathbb{R}^n$ of a sufficiently smooth and real valued function $f : \mathbb{R}^n \to \mathbb{R}$ in an <u>open</u> set $\Omega \subseteq \mathbb{R}^n$:

$$x^* = \underset{x \in \Omega}{\operatorname{argmin}} f(x).$$

By differential calculus, we immediately received as a first order necessary condition:

$$f(x^*) \leq f(x) \text{ for all } x \in B_\epsilon(x^*) \qquad \Longrightarrow \qquad \forall x \in \Omega : \langle \nabla f(x^*), x \rangle = 0.$$

If Ω is <u>closed</u>, e.g.,

$$\Omega = \prod_{i=1}^{n} [a_i, b_i] = \{x \in \mathbb{R}^n \mid \forall i = 1, ..., n : \; a_i \leq x_i \leq b_i \, , \; a_i, b_i \in \mathbb{R}\},$$

the situation turns out to be slightly more complicated: if a (local) minimizer is located on the boundary, the gradient condition is not longer a necessary criterion. We will focus on that in the next exercise.

Exercise 12
Let $f : \Omega \to \mathbb{R}$, $f \in \mathcal{C}^0(\bar{\Omega}) \cap \mathcal{C}^1(\Omega^\circ)$, ∇f continuously expandable on $\bar{\Omega}$ and Ω as defined above. Further, let $x^* \in \Omega$ be a local minimizer of f, i.e.

$$\exists \, \epsilon > 0 : \; \forall x \in B_\epsilon(x^*) \cap \Omega : \; f(x^*) \leq f(x).$$

Show that the following modified first order condition holds:

$$\forall x \in \Omega : \langle \nabla f(x^*), x - x^* \rangle \geq 0.$$

Any x^* that fulfills this condition is called *stationary point* of f.

Exercise 13
For $f : \mathbb{R}^n \to \mathbb{R}$ let L be the Lipschitz constant of the gradient ∇f. The *canonical projection* of $x \in \mathbb{R}^n$ on the closed set $\Omega = \prod_{i=1}^n [a_i, b_i]$ is given by $P : \mathbb{R}^n \to \Omega$,

$$\big(P(x)\big)_i := \begin{cases} a_i & \text{if } x_i \leq a_i \\ x_i & \text{if } x_i \in (a_i, b_i) \\ b_i & \text{if } x_i \geq b_i \end{cases} \; .$$

Further we define

$$x(\lambda) := P(x - \lambda \nabla f(x)).$$

Prove that the following ***modified Armijo condition*** holds for all $\lambda \in \left(0, \frac{2(1-\alpha)}{L}\right]$:

$$f(x(\lambda)) - f(x) \;\leq\; -\frac{\alpha}{\lambda} \|x - x(\lambda)\|^2_{\mathbb{R}^n}.$$

Hints: The following ansatz with the fundamental theorem of calculus may be helpful:

$$f(x(\lambda)) - f(x) = \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t} f\Big(x - t\big(x - x(\lambda)\big)\Big) \mathrm{d}t.$$

You can use the following formula without proof:

$$\forall x, y \in \Omega : \; \big\langle y - x(\lambda),\; x(\lambda) - x + \lambda \nabla f(x) \big\rangle \;\geq\; 0.$$

Exercise 14 (4 Points)

Let $f : \mathbb{R}^n \to \mathbb{R}$ convex and differentiable and $C \subseteq \mathbb{R}^n$ a closed and non-empty convex set. Show in the following order that:

1. $x^* \in C$ minimizes f over $C \quad \Leftrightarrow \quad \langle \nabla f(x^*), c - x^* \rangle \geq 0 \; \forall c \in C$.
   Hint: it might be helpful to use (without proof) that, for a convex function f, it holds that $f(b) \geq f(a) + \langle \nabla f(a), b - a \rangle$, $a, b \in C$.

2. $\langle c - P(x), P(x) - x \rangle \geq 0 \; \forall c \in C$ and $x \in \mathbb{R}^n$ considering that one can determine the projection $P(x)$ of x in C solving the minimization problem

$$\min_{c \in C} f(c) = \min_{c \in C} \frac{1}{2} \|c - x\|^2.$$

3. $x^* \in C$ minimizes f over $C \quad \Leftrightarrow \quad x^* = P(x^* - \gamma \nabla f(x^*))$
   Hint: it might be helpful to use (without proof) that $\langle x - P(x - \gamma \nabla f(x)), P(x - \gamma \nabla f(x)) - x + \gamma \nabla f(x) \rangle \geq 0$, with $\gamma \in \mathbb{R}^+$.