Universität Konstanz
Fachbereich Mathematik und Statistik
Prof. Dr. Stefan Volkwein
Martin Gubisch, Roberta Mancini, Stefan Trenz

23rd May 2011

Optimization
Exercises 4

✓ Exercise 13 (5 Points)

Let $f \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R})$ and $d_k \in \mathbb{R}^n$ a direction of descent in the point $x_k \in \mathbb{R}^n$. Further, assume that f is limited from below on the ray $\{x_k + td_k \mid t > 0\}$

Show that for any given parameters $0 < \alpha < \rho < 1$ there are is a step-size t such that the Wolfe-Powell conditions

$$f(x_k + td_k) \leq f(x_k) + \alpha t \nabla f(x_k)^t d_k$$
$$\langle \nabla f(x_k + td_k), d_k \rangle \geq \rho \langle \nabla f(x_k), d_k \rangle$$

or the strict Wolfe-Powell conditions

$$f(x_k + td_k) \leq f(x_k) + \alpha t \nabla f(x_k)^t d_k$$
$$|\langle \nabla f(x_k + td_k), d_k \rangle| \leq \rho |\langle \nabla f(x_k), d_k \rangle|,$$

respectively, hold in an open neighbourhood of t.

Optimization under boundary constraints.

Until now, we looked for local minimal points x^* of a sufficiently smooth, real-valued function f in an open set $\Omega \subseteq \mathbb{R}^n$:

$$x^* = \arg\min_{x \in \Omega} f(x).$$

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

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

If Ω is closed, the situation is slightly more complicated: Local minimizers on the boundary are possible, but here the gradient condition is not a necessary criterion.

Let $\Omega \subseteq \mathbb{R}^n$ a closed interval, i.e. there are $L_i, R_i \in \mathbb{R}$ $(i = 1, ..., n)$ such that

$$\Omega = \prod_{i=1}^{n} [L_i, R_i] = \{x \in \mathbb{R}^n \mid \forall i = 1, ..., n : L_i \leq x \leq R_i\},$$

and $f \in \mathcal{C}^2(\Omega, \mathbb{R})$. Notice that $\nabla f : \Omega^\circ \to \mathbb{R}^n$ can be expanded on the boundary of Ω since $f \in \mathcal{C}^2$ implies that ∇f is Lipschitz continous on Ω°.

Exercise 14

(a) Let $x^* \in \Omega$ a local minimizer of f, i.e.

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

Prove 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.

(b) Let $P : \mathbb{R}^n \to \Omega$ the canonical projection

$$(Px)_i := \begin{cases} L_i & \text{if } x_i \leq L_i \\ x_i & \text{if } x_i \in [L_i, R_i] \\ R_i & \text{if } x_i \geq R_i \end{cases}$$

and

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

Prove that

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

Exercise 15

Let L the Lipschitz constant for ∇f. Prove that

$$\forall \lambda \in \left(0, \frac{2(1-\alpha)}{L}\right] : f(x(\lambda)) - f(x) \leq -\frac{\alpha}{\lambda} ||x - x(\lambda)||^2.$$

This condition coincides with the Armijo condition for the classical line-search case.

The Gradient Projection Algorithm.

We modify the general descent algorithm `gradmethod` with modified Armijo step-size choice such that the algorithm can be applied for the situation above:

```
function X = gradproj(x,f,grad(f),N,epsilon,t0,alpha,beta)

while termination criterion (1) is not fulfilled
  find stepsize lambda such that (2) holds
  set x = x(lambda)
end
```

where the termination criteria are

```
(1.1)   ||x-x(1)|| < epsilon             (success)   or
(1.2)   ||grad(f)(x)|| < epsilon         (success)   or
(1.3)   number of iteration points > N   (failure)
```

and the step-size choice is provided by

```
      while modified armijo condition not fulfilled
(2)     reduce lambda
      end
```

Our objective is to prove that the generated iteration sequence has a convergent subsequence which converges towards a stationary point of f, cp. Satz 3.8 in the lecture notes.

Exercise 16

Let $(x_n)_{n\in\mathbb{N}}$ an iteration sequence created by `gradproj`.

(a) Show that $(f(x_n))_{n\in\mathbb{N}}$ converges.

(b) Show that $(x_n)_{n\in\mathbb{N}}$ has at least one convergent subsequence and that all accumulation points of $(x_n)_{n\in\mathbb{N}}$ are stationary points of f.

(c) Show that x^* is a stationary point of f if and only if $x^* = P(x^* - \lambda\nabla f(x^*))$ holds.

Deadline: Monday, 30$^{\text{th}}$ May, 8:30 am