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

6th June 2011

Optimization
Exercises 5

✓ Exercise 17 (5 Points)

Let $f : \mathbb{R} \to \mathbb{R}$ a four times differentiable, approximative function – which means that the function values $f(x)$ are not known exactly, but with some error tolerance ϵ_{tol}:

$$f_{\text{known}}(x) = f_{\text{exact}}(x) + \epsilon_{\text{tol}}(x) \qquad \text{where } |\epsilon_{\text{tol}}(x)| \leq \epsilon \text{ for some known } \epsilon > 0.$$

Determine the derivative of f numerically by central differences and prove that the error ϵ_H arising in the numerical approximation of the second derivative of f is of order $\epsilon_H = \mathrm{O}(\epsilon^{\frac{4}{9}})$.

Exercise 18

Let $f \in \mathcal{C}^2(\mathbb{R}^n, \mathbb{R})$. Verify the formula

$$\frac{\partial}{\partial x_i}\frac{\partial}{\partial x_j} f(x) = \frac{f(x + \epsilon e_i + \epsilon e_j) - f(x + \epsilon e_i) - f(x + \epsilon e_j) + f(x)}{\epsilon^2} + \mathrm{O}(\epsilon)$$

for the approximation of the Hessian matrix by evaluations of f.

Exercise 19

The angle η_k between the search direction d_k and the steepest descent direction $-\nabla f(x_k)$ is defined by

$$\cos \eta_k := \frac{-\langle \nabla f(x_k), d_k \rangle}{\|\nabla f(x_k)\|_2 \; \|d_k\|_2}. \tag{$*$}$$

Consider now the Newton-like method

$$x_{k+1} = x_k + \alpha_k d_k \qquad \text{with} \qquad d_k = -B_k^{-1} \nabla f(x_k)$$

where $B_k \in \mathbb{R}^{n\times n}$ are symmetric, positive definite matrices with uniformly bounded condition numbers, i.e.

$$\mathrm{cond}_2(B_k) = \|B_k\|_2 \; \|B_k^{-1}\|_2 \le M \qquad \text{for all } k \ge 0.$$

Show that

$$\cos \eta_k \ge \frac{1}{M}.$$

Hint: First prove that $\|Bx\|_2 \ge \frac{\|x\|_2}{\|B^{-1}\|_2}$ for any non-singular matrix B, then use $(*)$ to prove the claim.

Exercise 20

Let the function ϕ be given by

$$\phi(\alpha) = f(x_k + \alpha d_k)$$

where d_k is a descent direction, i.e., $\langle \nabla f(x_k), d_k \rangle < 0$.

Derive that the quadratic function interpolating $\phi(0)$, $\phi'(0)$ and $\phi(\alpha_0)$ is given by

$$\phi_q(\alpha) = \left(\frac{\phi(\alpha_0) - \phi(0) - \phi'(0)\alpha_0}{\alpha_0^2} \right) \alpha^2 + \phi'(0)\alpha + \phi(0)$$

by making the ansatz

$$\phi_q(\alpha) = a_0 + a_1\alpha + a_2\alpha^2$$

and using the interpolaters to calculate the coefficients a_0, a_1, a_2 of ϕ_q.

Assume now that the sufficient decrease condition

$$f(x_k + \alpha d_k) \le f(x_k) + c_1\alpha\langle \nabla f(x_k), d_k \rangle$$

is not satisfied at α_0.

Show that ϕ_q has positive curvature and that the minimizer α^* of ϕ_q satisfies

$$\alpha^* < \frac{\alpha_0}{2(1-c_1)}.$$

Remark: Since c_1 is chosen to be quite small in practise, this indicates that α^* cannot be much greater than $\frac{\alpha_0}{2}$ (and may be smaller), which gives us an idea of the new step length.

Deadline: Monday, 13$^{\text{th}}$ June, 8:30 am