Norwegian University of Science and Technology
Department of Mathematical Sciences

TMA4180 Optimization Theory
Spring 2013

Exercise set 4

Tutorial: No tutorial on February 14. If you have questions, contact Markus directly.

1 **a)** Find the global minima (in $\mathbb{R}^2$) of the function

$$f(x,y) = 2x^2 + y^2 - 2xy - 2x^3 + x^4. \tag{1}$$

List the general results you are using.

Solution:

We compute ∇f *and* $\nabla^2 f$*:*

$$\nabla f(x,y) = \begin{bmatrix} 4x - 2y - 6x^2 + 4x^3 \\ 2y - 2x \end{bmatrix},$$
$$\nabla^2 f(x,y) = \begin{bmatrix} 4 - 12x + 12x^2 & -2 \\ -2 & 2 \end{bmatrix}.$$

The candidate points will be solutions of

$$y = x, \qquad 4x - 6x^2 + 4x^3 = 2y,$$

which are easily seen to be

$$x_{(1)} = (0,0)^{\mathrm{T}},$$
$$x_{(2)} = (\tfrac{1}{2}, \tfrac{1}{2})^{\mathrm{T}},$$
$$x_{(3)} = (1,1)^{\mathrm{T}}.$$

Now,

$$\nabla^2 f(0,0) = \begin{bmatrix} 4 & -2 \\ -2 & 2 \end{bmatrix} > 0,$$
$$\nabla^2 f(\tfrac{1}{2}, \tfrac{1}{2}) = \begin{bmatrix} 1 & -2 \\ -2 & 2 \end{bmatrix}, \textit{indefinite},$$
$$\nabla^2 f(1,1) = \begin{bmatrix} 4 & -2 \\ -2 & 2 \end{bmatrix} > 0.$$

Thus, only $(0,0)^{\mathrm{T}}$ *and* $(1,1)^{\mathrm{T}}$ *are minima, and both are strict since the Hessian is positive definite. The point in the middle is a saddle point. The function values in both minima are equal to* 0*, so they are both global.*

The problem may also be solved by the following trick*:*

$$f(x,y) = 2x^2 + y^2 - 2xy - 2x^3 + x^4 = (x-y)^2 + (x-x^2)^2.$$

Thus, the global minimum is 0*, which is obtained for* $x = x^2$*,* $y = x$*, i.e.* $(0,0)^{\mathrm{T}}$ *and* $(1,1)^{\mathrm{T}}$*.*

b) Estimate the drop in the error per iteration (expressed in terms of the appropriate norm) of the steepest descent method near the global minima in **a)**.

Solution: If we denote the Hessian near the global minima by A*, the error estimate is given by*

$$\frac{\|x_{k+1} - x^*\|_A}{\|x_k - x^*\|_A} \leq \frac{\kappa(A) - 1}{\kappa(A) + 1}.$$

The Hessians at $(0,0)^{\mathrm{T}}$ *and* $(1,1)^{\mathrm{T}}$ *are the same,*

$$A = \begin{bmatrix} 4 & -2 \\ -2 & 2 \end{bmatrix},$$

with eigenvalues $\lambda_1 = 3 + \sqrt{5}$ *and* $\lambda_2 = 3 - \sqrt{5}$*. Hence* $\kappa(A) = (3+\sqrt{5})/(3-\sqrt{5}) \approx 6.8541$*, and*

$$\frac{\|x_{k+1} - x^*\|_A}{\|x_k - x^*\|_A} \leq \frac{\kappa(A) - 1}{\kappa(A) + 1} \approx 0.745,$$

that is, about a 25% decrease per iteration.

2 When it is easy to compute first and second derivatives of a one-dimensional function (that is, $x \in \mathbb{R}$ and $f(x) \in \mathbb{R}$), it is possible to combine a trust region algorithm with Newton's method for finding the minimum. Outline an algorithm for this.
Hint: First derive a quadratic approximation to the function. Show that minimizing this function corresponds to the Newton step, plus an investigation involving the endpoints of the domain.
Solution: We start at a point x_k *and have now* an interval on the real line *as the trust region,*

$$D_k = [x_k - \Delta_k, x_k + \Delta_k].$$

The quadratic approximation is the simple parabola

$$m_k(x) = f(x_k) + f'(x_k)(x - x_k) + \tfrac{1}{2} f''(x_k)(x - x_k)^2,$$

and solving

$$x_{k+1} = \arg\min_{x \in D_k} m_k(x)$$

means finding the minimum of the parabola within the interval D_k*. Clearly, the global minimum of* $m_k(x)$ *is given by*

$$f'(x_k) + f''(x_k)(x_g - x_k) = 0,$$

corresponding to the Newton step,

$$x_g = x_k - \frac{1}{f''(x_k)} f'(x_k).$$

Thus, x_{k+1} may be equal to x_g or being one of the endpoints.

The rest of the algorithm is as before, starting by considering the actual vs. the predicted decrease

$$\rho = \frac{f(x_k) - f(x_{k+1})}{f(x_k) - m_k(x_{k+1})}.$$

We increase Δ if $\rho > \beta$ and shrink Δ if $\rho < \alpha$, $0 < \alpha < \beta < 1$. It may be reasonable to let $x_{k+1} = x_k$ when ρ is very small, at least when $\rho < 0$.

3 Solve the following problems by use of the trust region method, using the file `trustdemo.m` which can be found on the lecture plan.

Try using the Cauchy point, the dogleg method as well as the exact solution of the TR problem.

$$f(x) = x_1^2 - 5x_1x_2 + x_2^4 - 25x_1 - 8x_2 \tag{2}$$

$$f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2 \tag{3}$$

$$f(x) = e^{x_1}(4x_1^2 + 2x_2^2 + 4x_1x_2 + 2x_2 + 1) \tag{4}$$

Compare these methods with the ones you tried in Exercise 2.

Solution: *You just have to modify the file. I think this is self-explanatory.*