Optimization Theory

Applied Mathematics

Problem set 2

1. Show that the function $f(x) = -x^3 + 4x^2 + x$ has exactly one local minimum on the interval $[-1, 2]$ and no local maxima (hence, it is unimodal there).

   (a) Do first two steps of the golden-section search method for f on $[-1, 2]$. What is the new search interval after step 2 of the procedure? What is its length?

   (b) Do first two steps of the line search method based on quadratic interpolation for f, $a = -1$, $b = 2$ and $c = \frac{1}{2}$. What is the new search interval after step 2 of the procedure? What is its length in comparison to the one given by the golden-section search? Does it not contradict the fact that the second method is faster than the first one?

2. Apply the quadratic interpolation search method to the function from excercise 1. with different starting points: $a = -1$, $b = 2$ and $c = 1$. Show that the procedure stops after step 2, because two consecutive approximations are the same (even though they are far from the actual local minimum). Can this problem be fixed somehow by incorporating some special rule for such a situation in the method? How can this rule look like?

3. Show that the 2-dimensional function $f(x, y) = x^2 + 4xy - 2y^2$ has exactly one stationary point which is neither a local minimizer nor a local maximizer.

   (a) Show that for the starting point $(x_0, y_0) = a(2, 1)$ the steepest descent method with a constant stepsize α converges iff $\alpha \in \left(0, \frac{1}{2}\right)$.

   (b) Next show that for $(x_0, y_0) = b(1, -2)$ the steepest descent method with a constant stepsize α converges iff $\alpha \in \left(-\frac{1}{3}, 0\right)$. Of course $\alpha > 0$ by definition. How can you interpret the result then?

   (c) Show that for $(x_0, y_0) = a(2, 1) + b(1, -2)$, $a, b \neq 0$ the steepest descent method does not converge for any value $\alpha > 0$.

   (d) Compute the eigenvalues and the corresponding eigenvectors of the matrix defining the quadratic form $f(x, y)$. Explain the behaviour of the steepest descent method in parts (a) and (b) using the obtained results.

4. ([AnLu07]) The problem

$$\text{minimize } f(x, y) = x^2 + 2y^2 + 4x + 4y$$

   is solved using the steepest descent method with optimal stepsize and initial point $(x_0, y_0) = (0, 0)$. By means of induction, show that $(x_k, y_k) = \left(\frac{2}{3^k} - 2, \left(-\frac{1}{3}\right)^k - 1\right)$. Deduce the minimizer of $f(x, y)$.
   Hint: The function f is quadratic, so the expressions for the almost optimal stepsize obtained through quadratic approximation of f from the lecture give its optimal value.

5. Consider the function

$$f(x, y) = 7x^2 - 8xy + 13y^2.$$

   Show that it has exactly one stationary point x^*, which is the global minimizer of f. Show that the steepest descent method with stepsize chosen according to the formula $\alpha_k = \frac{1}{k+1}$ converges to x^* from any starting point x^0. To do that, compute the eigenvalues and the eigenvectors of the matrix defining the quadratic form $f(x, y)$ and then mimic what has been done in part (c) of excercise 3. What if $\alpha_k = \frac{\pi}{k}$? Does it change anything?

6. Show that the descent directions chosen by the steepest descent method with optimal stepsize in its subsequent iterations, $\nabla f(x_k)$ and $\nabla f(x_{k+1})$, $k = 1, 2, \ldots$, are orthogonal.

7. ([AnLu07]) Consider the two following minimization problems:

$$\text{minimize } f(x, y) = x^2 + y^2 - 0.2xy - 2.2x + 2.2y + 2.2,$$

$$\text{minimize } g(x, y) = 5x^2 + 4.075y^2 - 9xy + x.$$

(a) Show that both f and g are strictly convex, thus the steepest descent method converges to their unique minima from any starting point.

(b) It can be checked that for the first problem the steepest descent method with $(x_0, y_0) = (0, 0)$ converges very fast, while the convergence for the second problem when $(x_0, y_0) = (1, 1)$ is very slow. Try to justify these two behaviours using the result about the rate of convergence of the steepest descent method given during the lecture. How many iterations will it take at most before the value of the objective function is reduced to at most 10^{-10} more than its minimal value in each of these problems?

8. Solve the problem from excercise 4. using Newton's method with optimal stepsize. How can you explain its behaviour?

9. Consider the function $f(x) = x^3 - 3x$. Check that it has two stationary points, one of which is its minimizer.

   Suppose Newton's method with constant stepsize $\alpha \equiv 1$ is used to find this minimizer. Show that for any $x_0 > 0$ it converges to the minimizer that you have found.

10. ([BV04]) Show that the function $f(x) = \ln(e^x + e^{-x})$ has a unique minimizer $x^* = 0$. Show that Newton's method with constant stepsize α does not converge to x^* for any $x_0 \neq x^*$ if the value of α is taken too big.

11. Suppose we want to find the minimizer (x^*, y^*) of the function $f(x, y) = x^{2k} + y^{2l}$ with $k, l \in \mathbb{Z}^+$, $k \neq l$. What is this minimizer? Show that Newton's method with optimal stepsize converges to (x^*, y^*) in a finite number of iterations iff $x_0 = 0$ or $y_0 = 0$.

12. Show that for any convex function $f : \mathbb{R}^n \to \mathbb{R}$ with exactly one minimum x^* any gradient method satisfying the assumptions of the convergence theorem from the lecture converges to x^* from any starting point x_0.
    Hint: Show that the level set $\{x \in \mathbb{R}^n : f(x) \leq f(x_0)\}$ is compact, then apply the definition of compactness to the sequence (x_k).

References:

[BV04] S. Boyd, L. Vanderberghe, Convex Optimization, Cambridge University Press, 2004
[AnLu07] A. Antoniou, W.-S. Lu, Practical Optimization, Springer Science+Business Media, LLC, 2007