Optimization Theory

Applied Mathematics

Problem set 2

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

   (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 $\alpha \in \left(0, \frac{1}{4}\right]$.

   (b) Next show that for $(x_0, y_0) = b(1,-2)$ the steepest descent method with a constant stepsize α converges iff $\alpha \in \left[-\frac{1}{6}, 0\right)$. 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 (b) and (c) using the obtained results.

2. ([AnLu07]) The steepest descent method with optimal stepsize is applied to solve the problem
$$\text{minimize } f(x,y) = 2x^2 - 2xy + y^2 + 2x - 2y.$$
Show by induction that for $(x_0, y_0) = (0,0)$, $(x_{2k+1}, y_{2k+1}) = \left(0, 1 - \frac{1}{5^k}\right)$. Find the minimizer of $f(x,y)$ using the obtained sequence of approximations.
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.

3. ([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 $(x_{k+1}, y_{k+1}) = \left(\frac{2}{3^k} - 2, \left(-\frac{1}{3}\right)^k - 1\right)$. Deduce the minimizer of $f(x,y)$.

4. 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.

5. 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).

6. ([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?

7. Solve the problem from excercise 2. using Newton's method with optimal stepsize. How can you explain its behaviour?

8. ([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.

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