CO 466/666: Continuous Optimization
Problem Set 1

H. Wolkowicz

Winter 2012.
Handed out: 2012-Jan-11.
Due: 2012-Jan-20, at the start of class.

Contents

1 Characterizations of Optimality

Define
$$f(x) := x_1^2 + x_2^2 + \gamma x_1 x_2 + x_1 + 2x_2,$$
where $\gamma \in \mathbb{R}$ is a parameter.

1. For each value of the parameter γ, find all the stationary points $\{x : \nabla f(x) = 0\}$.

2. Which of the stationary points are global minima? Which are local minima? (Why?)

2 Steepest Descent Method

Suppose that $f : \mathbb{R}^n \to \mathbb{R}$ is continuously differentiable. Suppose that the method of steepest descent with exact line search is applied and the search directions $d_k, k = 1, 2, \ldots$ are obtained.

1. Show that the search directions d_{k+1}, d_k are orthogonal for all k.

2. Apply the method to the simple quadratic $f(x) = 10x_1^2 + x_2^2$ from the initial point $\begin{pmatrix} 1/10 \\ 1 \end{pmatrix}$.

3. From [1], we know that the method has linear convergence (Q-linear) with a rate $R = \frac{\lambda_1 - \lambda_n}{\lambda_1 + \lambda_n}$, where λ_1, λ_n are the largest and smallest eigenvalues of the Hessian. Use MATLAB to verify numerically that this rate is achieved in this example. (Note that the rate can get arbitrarily close to 1 if the Hessian is sufficiently ill-conditioned.)

3 Newton's Method

Suppose that $f : \mathbb{R}^n \to \mathbb{R}$ is twice continuously differentiable. Let $x_c \in \mathbb{R}^n$ be the current estimate for a minimum of f.

1. Write down the quadratic model at x_c. Use this to derive Newton's method. (You can assume that the Hessian at x_c is positive definite.)

2. Describe what is meant by the statement *Newton's method is scale free*, and prove the statement. Is this still true for Newton's method with a line search? Why or why not?

References

[1] H. Akaike. On a successive transformation of probability distribution and its application to the analysis of the optimum gradient method. *Ann. Inst. Statist. Math. Tokyo*, 11:1–16, 1959. 2