ON THE USE OF DIRECTIONS OF NEGATIVE CURVATURE IN A MODIFIED NEWTON METHOD*

Jorge J. MORÉ
Argonne National Laboratory Argonne, IL, U.S.A.

Danny C. SORENSEN
University of Kentucky Lexington, KY, U.S.A.

Received 17 October 1977
Revised manuscript received 6 June 1978

We present a modified Newton method for the unconstrained minimization problem. The modification occurs in non-convex regions where the information contained in the negative eigenvalues of the Hessian is taken into account by performing a line search along a path which is initially tangent to a direction of negative curvature. We give termination criteria for the line search and prove that the resulting iterates are guaranteed to converge, under reasonable conditions, to a critical point at which the Hessian is positive semidefinite. We also show how the Bunch and Parlett decomposition of a symmetric indefinite matrix can be used to give entirely adequate directions of negative curvature.

Key words: Unconstrained Optimization, Modified Newton's Method, Descent Pairs, Directions of Negative Curvature, Symmetric Indefinite Factorization, Steplength Algorithm.

1. Introduction

Let \( f : \mathbb{R}^n \rightarrow \mathbb{R} \) be a continuously differentiable function in the open set \( \mathcal{D} \), and consider the problem of producing a sequence \( \{x_k\} \) that converges to a local minimizer \( x^* \) of \( f \). That is, we seek \( x^* \) such that

\[
    f(x^*) \leq f(x), \quad x \in N \cap \mathcal{D}
\]

with \( N \) some neighborhood of \( x^* \).

Algorithms for the solution of (1.1) are usually descent methods. A descent method determines a direction \( s_k \) at the iterate \( x_k \) such that \( \nabla f(x_k)^T s_k < 0 \). A line search then yields a step-length \( \alpha_k > 0 \) such that

\[
    f(x_k + \alpha_k s_k) < f(x_k),
\]

and thus it is sensible to let \( x_{k+1} = x_k + \alpha_k s_k \). Under some additional restrictions on the choice of \( \alpha_k \) one can show that

\[
    \lim_{k \to \infty} \frac{\nabla f(x_k)^T s_k}{\|s_k\|} = 0. \tag{1.2}
\]

Moreover, the vector \( s_k \) is usually related to \( \nabla f(x_k) \) in such a way that (1.2)

* Work performed under the auspices of the U.S. Department of Energy.
implies that \( \| \nabla f(x_k) \| \) converges to zero. Thus, every limit point \( x^* \) of \( \{ x_k \} \) satisfies \( \nabla f(x^*) = 0 \).

It is desirable to produce a sequence which converges to a point \( x^* \) with \( \nabla f(x^*) \) positive definite. This would imply that \( x^* \) is an isolated local minimizer of \( f \), and in particular, that \( x^* \) satisfies (1.1). In general, it is not possible to produce such a sequence. However, through the use of directions of negative curvature we shall be able to produce a sequence which converges to a point \( x^* \) with \( \nabla^2 f(x^*) \) positive semidefinite. For practical purposes, this is a very strong assertion. For instance, if the Hessian were known to be nonsingular at all critical points, then \( x^* \) would have to be a local minimizer. To see that theoretically it is not, in general, possible for \( \nabla^2 f(x^*) \) to be positive definite, consider an example of Wolfe [13]. In that example, steepest descent (and actually, any reasonable descent method) converges to a saddle point at which the Hessian is singular. Since the algorithm approaches the saddle point through a region in which \( f \) is strictly convex, there is no possible way to avoid this saddle point.

The idea of using directions of negative curvature appeared as early as 1968 [5, pp. 165–169], but recently there has been renewed interest [6, 7, 9, 10]. We are particularly indebted to the paper of McCormick [9]. In that paper, McCormick showed how a modification of the Armijo line search could be used with directions of negative curvature. Our Theorem 3.1 is a slight modification of McCormick's main result; its purpose is to isolate the main ingredients of McCormick's paper. In this paper we first extend McCormick's work by considering the practical generation of directions of negative curvature. We discuss two methods in Section 4. One is based on Gill and Murray's [7] modified Cholesky factorization, and the other based on Bunch and Parlett's [3] factorization of symmetric indefinite matrices. We show that Gill and Murray's method does not satisfy the requirements of our convergence theorems, but that the Bunch and Parlett factorization can be used to give entirely adequate directions of negative curvature.

In Section 5 we again extend McCormick's work by replacing the Armijo line search by a general line search with a satisfactory termination criteria. The results of this section provide a theoretically justified alternative to Fletcher and Freeman's [6] ad-hoc line search. Finally, in Section 6 we present our convergence results. In particular, we show how the line search can be used together with our directions of negative curvature to provide a very effective modified Newton method.

**Assumption 1.1.** Let \( f: \mathbb{R}^n \rightarrow \mathbb{R} \) have two continuous derivatives on the open set \( \mathcal{D} \), and assume that for some \( x_0 \) in \( \mathcal{D} \), the level set

\[
L(x_0) = \{ x \in \mathcal{D} : f(x) \leq f(x_0) \}
\]

is a compact subset of \( \mathcal{D} \).