12. A Modification of the Gradient Method and Function Extremization

By Tatsuo NODA
Department of Applied Mathematics,
Toyama Prefectural College of Technology

(Communicated by Kôsaku YOSIDA, M. J. A., Feb. 13, 1989)

1. Introduction. Let $x=(x_1, x_2, \cdots, x_n)$ be a vector in R^n and D a region contained in R^n. Let $f(x)$ be a real-valued nonlinear function defined on D. Define an n-dimensional vector $\nabla f(x)$ and an $n\times n$ matrix $H(x)$ by

$$\nabla f(x)=(\partial f(x)/\partial x_i) \qquad (1\leqq i\leqq n)$$

and

$$H(x)=(\partial^2 f(x)/\partial x_j \partial x_k) \qquad (1\leqq j, k\leqq n).$$

For a vector x, we shall use the norm defined by

$$\|x\|=\left(\sum_{i=1}^n x_i^2\right)^{1/2}.$$

The Euclidean norm and the spectral norm of an $n\times n$ matrix $A=(a_{ij})$, denoted by $\|A\|$ and $\|A\|_s$, are defined as

$$\|A\|=\left(\sum_{i=1}^n \sum_{j=1}^n a_{ij}^2\right)^{1/2}$$

and

$$\|A\|_s=\lambda^{1/2},$$

respectively, where λ is the maximum eigenvalue of A^*A and A^* is the transposed matrix of A.

Throughout this paper, we shall assume the following three conditions.

(A.1) $f(x)$ is two times continuously differentiable on D.

(A.2) There exists a point $\bar{x}\in D$ satisfying $\nabla f(x)=0$.

(A.3) The $n\times n$ symmetric matrix $H(\bar{x})$ is positive definite.

Let $U(\bar{x}\,;\delta)=\{x\,;\|x-\bar{x}\|<\delta\}$ be a neighbourhood of $\bar{x}$.

The following well-known theorem gives a sufficient condition for finding a local minimum of $f(x)$.

Theorem 1 ([3, Theorem 8.3]). *In addition to conditions* (A.1)–(A.3), *suppose that the following condition* (A.4) *holds.*

(A.4) *α is a constant satisfying* $0<\alpha<\dfrac{2}{\|H(\bar{x})\|_s}$.

Under conditions (A.1)–(A.4), *there exists a neighbourhood* $U(\bar{x}\,;\delta_0)\subset D$ *such that, for arbitrary* $x^{(0)}\in U(\bar{x}\,;\delta_0)$,

$$x^{(k)}\longrightarrow \bar{x} \quad as \quad k\longrightarrow\infty,$$

where the $x^{(k)}$ are generated by the gradient method

$$x^{(k+1)}=x^{(k)}-\alpha\nabla f(x^{(k)}). \tag{1.1}$$

The purpose of this paper is to show Theorem 2 by considering an