A COMPOSITE NEWTON-RAPHSON GRADIENT METHOD FOR THE SOLUTION OF SYSTEMS OF EQUATIONS

WILLIAM L. HART AND THEODORE S. MOTZKIN

1. Introduction. This article was motivated by the desire to obtain an iterative method for solving a system of equations, linear or not, into which all equations would enter symmetrically, and which would be suitable for numerical application, particularly on a high speed digital computing machine.

The general problem considered is the solution of a system of k equations $\{f_j(x)=0\}$ in n unknowns $(x_1, \cdots, x_n)=x$ where, as throughout the paper, all variables and function values are real. Each step of our method consists in obtaining, from one approximation x to a solution of the system, the next approximation by adding to x the vector sum of corrections parallel to the gradients of the k functions $f_j(x)$. The lengths of the corrections are regulated by individual weights and by use of a factor $\rho \neq 0$. The component gradient correction for a single equation $f_j(x)=0$ is of the Newton-Raphson type because the correction, if applied to an initial approximation $x^{(0)}$, gives a point annihilating the usual linear approximation to $f_j(x)$ for x near $x^{(0)}$.

After considering in § 2 the well known formula for a gradient correction to an approximate solution $x^{(0)}$ of a single equation $f(x)=0$, the method of composite gradient corrections for a general system is described in § 3. In § 4, we apply the method to a system of k linear equations in n unknowns, and prove that, for an arbitrary approximation $x^{(0)}$ to a solution of the system, we obtain a sequence $\{x^{(m)}\}$ which tends with a geometric rate of convergence to a point $\hat{x}$, nearest to $x^{(0)}$, of the set which satisfies the system in a sense of weighted least squares. Section 5 treats a fairly general system with an isolated solution $\hat{x}$. The sequence $\{x^{(m)}\}$ of § 3 is proved to converge to $\hat{x}$ if the initial approximation $x^{(0)}$ is sufficiently near $\hat{x}$. Section 6 considers the implicit function $x=x(t)$ defined by a related system of n equations $f(x;\tau)=0$, where $\tau=(\tau_1, \cdots, \tau_k)$, and $\tau=\tau(t)$, $0 \leq t \leq 1$. It is proved that, if $0=t_0<t_1<\cdots<t_l=1$ is a fine enough partition of the t-interval, then the sequence $\{x^{(m)}\}$ of § 3 tends to $x(t_i)$ if $x^{(0)}=x(t_{i-1})$. This result yields a small arc method for computing the points $x(t_i)$ in sequence.

There is an extensive literature on the solution of *linear* systems $\{f_j(x)=0\}$ by iterative processes where each iteration involves a correction related to a specified direction, in particular that of some gradient;

Received October 7, 1955. Presented to the American Mathematical Society, April 30, 1955. This work was sponsored (in part) by the Office of Ordnance Research, U.S. Army.