SUPERLINEARLY CONVERGENT QUASI-NEWTON ALGORITHMS FOR NONLINEARLY CONSTRAINED OPTIMIZATION PROBLEMS*

U.M. GARCIA PALOMARES**
O.L. MANGASARIAN
University of Wisconsin, Madison, Wis., U.S.A.

Received 29 April 1974
Revised manuscript received 27 October 1975

A class of algorithms for nonlinearly constrained optimization problems is proposed. The subproblems of the algorithms are linearly constrained quadratic minimization problems which contain an updated estimate of the Hessian of the Lagrangian. Under suitable conditions and updating schemes local convergence and a superlinear rate of convergence are established. The convergence proofs require among other things twice differentiable objective and constraint functions, while the calculations use only first derivative data. Rapid convergence has been obtained in a number of test problems by using a program based on the algorithms proposed here.

1. Introduction

We develop in this paper a class of algorithms for finding Kuhn–Tucker points of the following nonlinear programming problem:

\[
\text{minimize } \{f(x) | g(x) \leq 0\} \quad (1.1)
\]

where \(f\) and \(g\) are differentiable functions from \(\mathbb{R}^n\) into \(\mathbb{R}\) and \(\mathbb{R}^m\) respectively. Our algorithms require calculations with gradients, hence \(f\) and \(g\) should at least be differentiable. However our convergence proofs require, among other things, twice differentiability of \(f\) and \(g\). Starting with an initial and possibly infeasible guess \(x_0\), of a solution, and an initial guess \(u_0\) of the Kuhn–Tucker multipliers, the algorithms construct a sequence \(\{(x_i, u_i)\}\) which under suitable conditions converges to a Kuhn–Tucker point \((\bar{x}, \bar{u})\) of 1.1. The convergence we establish is a local one and the main conditions required for it are closeness of the starting point to a point \((\bar{x}, \bar{u})\) satisfying the second order sufficiency conditions, the linear independence of the gradients of the active constraints and positivity of the multipliers associated with them. The subproblems solved are linearly constrained.

* Research supported by NSF Grant GJ-35292 at the University of Wisconsin.
** Now at Instituto Venezolano de Investigaciones Científicas (IVIC), Departamento De Matematicas, Apartado 1827, Caracas, Venezuela.
quadratic programming problems for which efficient and finite algorithms exist [3, 12]. These quadratic subproblems contain an \( n \times n \) matrix which is an estimate of the Hessian of the Lagrangian of 1.1. If this estimate is close enough to the Hessian a linear convergence rate is obtained, and if it converges to the Hessian a superlinear rate of convergence is obtained. If the exact Hessian is used the algorithm becomes that of Wilson [14] for which Robinson [11] has established a quadratic convergence rate.

In order to achieve superlinear convergence it is sufficient but not necessary that the updated estimate of the Hessian of the Lagrangian converge to the Hessian. In fact any updating scheme which satisfies a less stringent condition, condition (b) of the algorithm below will achieve superlinear convergence. Thus a "finite difference" approximation of the Hessian similar to that employed by Goldstein and Price [6], in which only gradients are computed, can be used. Numerical experiments on all of Colville's test problems [2] were carried out with a specific updating scheme [5], in which under suitable conditions the updated matrices converge to the Hessian. The test results are quite encouraging and are discussed in more detail in section 4.

The subproblems generated here by matrix updating schemes can be considered either as quadratic approximations of the subproblems of either Robinson's [10] or Wilson's [14] algorithms, both of which converge quadratically [11]. However in Robinson's algorithm the subproblems to be solved have nonlinear objective functions which in general are not quadratic. In Wilson's algorithm explicit evaluation of second derivatives is required which need not be the case in our algorithms. In addition the Lagrangians of our subproblems here can also be considered as estimates of a quadratic approximation of the Lagrangians of the original problem 1.1 [5]. Topkis [13] has proposed a different quadratically convergent algorithm. However his subproblems are nonlinearly constrained.

Throughout this paper all vectors will be column vectors. The transpose will be denoted by the superscript \( T \). The gradient or Jacobian with respect to \( x \) will be denoted by \( \nabla \), whereas the gradient or Jacobian with respect to any other variable will be denoted by the same symbol \( \nabla \) subscripted by that variable. Similarly \( \nabla^2 \) will denote the Hessian with respect to \( x \), whereas the Hessian with respect to another variable will be denoted by \( \nabla^2 \) subscripted by that variable. Subscripts to vectors will denote iteration number, while superscripts will denote a vector component. The symbol \( \|\cdot\| \) will denote an arbitrary but fixed norm.

2. Algorithm

The algorithm is described by the following steps. Let \( z = (x, u) \), and let \( \nabla^2 L(z) \) denote the \( n \times n \) Hessian with respect to \( x \) of the Lagrangian \( L(z) = f(x) + u^T g(x) \).

\[ \text{Step 1: Set } i = 0. \]

\[ \text{Step 2: Having } z_i = (x_i, u_i) \text{ find a Kuhn–Tucker point } z_{i+1} = (x_{i+1}, u_{i+1}) \text{ of the} \]