Journal of Computational Mathematics, Vol.8, No.3, 1990, 195–201.

A NEW METHOD FOR EQUALITY CONSTRAINED OPTIMIZATION*

Jia Li-xing
(*Department of Mathematics, Shaanxi Normal University, Xi'an, Shaanxi, China*)

Abstract

This paper prensents a detailed derivation and description of a new method for solving equality constrained optimization problem. The new method is based upon the quadratic penalty function, but uses orthogonal transformations, derived from the Jacobian matrix of the constraints, to deal with the numerical ill-conditioning that affects the methods of this type.

At each iteration of the new algorithm, the orthogonal search direction is determined by a quasi-Newton method which can avoid the necessity of solving a set of equations and the step-length is chosen by a Armijo line search. The matrix which approaches the inverse of the projected Hessian of composite function is updated by means of the BFGS formula from iteration to iteration. As the penalty parameter approaches zero, the projected inverse Hessian has special structure which can guarantee us to obtain the search direction accurately even if the Hessian of composite function is ill-conditioned in the former penalty function methods.

§1. Introduction

We consider the problem

$$\text{minimize} \quad F(x), \quad x \in R^n, \tag{1.1a}$$

$$\text{subject to} \quad c(x) = 0, \quad c \in R^m, \quad m \le n \tag{1.1b}$$

where $F(x)$ and $c_i(x)(i = 1(1)m)$ are all twice continuously differentiable functions of x.

The former penalty function method for solving (1.1) is minimizing the composite function

$$\Phi(x, r) = F(x) + c^T c/2r \tag{1.2}$$

where $c^T c/2r$ is the quadratic penalty term and r is the penalty parameter. It is known that if x^* is the solution of (1.1) and $x^*(r)$ is the unconstrained minimum of (1.2), then under mild conditions [4],

$$\lim_{r \to 0} x^*(r) = x^*.$$

Thus, we will deal with the unconstrained problem

$$\text{minimize} \Phi(x, r) \tag{1.3}$$

in stead of (1.1). The method we suggest for solving (1.3) will generate a sequence that converges (as $r \to 0$) to solution x^* of the original problem (1.1).

* Received July 15, 1987.

§2. Basic Philosophy

In order to simplify our discussion, we introduce the following notations:

g : gradient of the objective functions F,
J : Jacobian of the constraints, $J = [\partial c_i / \partial x_j]$, $\text{rank}(J) = m$,
f : gradient of $\Phi (= g + J^T c/r)$,
Q : orthogonal matrix satisfying $JQ^T = [U^T, O]$, where U is upper triangular,
h : projected gradient, $h = Qg$,
H_0 : Hessian of F,
H_i : Hessian of c_i,
H_Φ : $H_\Phi = H_0 + (1/r) \sum_{i=1}^{m} c_i H_i$,
G : projected Hessian, $G = QH_\Phi Q^T$,
$\overline{B}$: an approximation to the inverse of projected Hessian of Φ.

If we solve (1.2) by applying Newton's method to the equation

$$f = g + J^T c/r = 0,$$

to perform one step of Newton's iteration, we must solve the equation

$$(H_\Phi + J^T J/r)p = -f \tag{2.1}$$

to determine the search direction p. The matrix H_Φ is generally well-behaved but for the standard penalty function method, $J^T J$ is of rank m and $\|J^T J/r\| \to \infty$ as $r \to 0$, with singularity occuring in the limit [5], which causes difficulty for solving (2.1). This is a fatal problem if we try to solve equation (2.1) directly. The second difficulty is that, while using Newton's method, we must supply a formula with which the second derivative matrices can be evaluated, and this can be a major disincentive if $F(x)$ and/or $c(x)$ are complicated functions of x. The third and final problem is the necessity of solving a set of equations at each iteration. Therefore, we may need to investigate a new method for solving (2.1).

§3. The New Method

In order to devise a new method, we rewrite equation (2.1) as

$$(H_\Phi + J^T J/r)p = -(g + J^T c/r), \tag{3.1}$$

and impose an orthogonal transformation on J^T, the transpose of the Jacobian of $c(x)$.. Then we can obtain an orthogonal matrix Q :

$$QJ^T = \begin{bmatrix} U \\ O \end{bmatrix}$$

where U is upper triangular and nonsingular since J has full row rank. Then we transform equation (3.1) using the orthogonal matrix Q :

$$Q(H_\Phi + J^T J/r)Q^T Qp = -Q(g + J^T c/r),$$