Journal of Computational Mathematics, Vol.9, No.4, 1991, 348–359.

A DUAL ALGORITHM FOR MINIMIZING A QUADRATIC FUNCTION WITH TWO QUADRATIC CONSTRAINTS*[1)]

Yuan Ya-xiang
(Department of Applied Mathematics and Theoretical Physics, University of Cambridge, England)

Abstract

In this paper, we present a dual algorithm for minimizing a convex quadratic function with two quadratic constraints. Such a minimization problem is a subproblem that appears in some trust region algorithms for general nonlinear programming. Some theoretical properties of the dual problem are given. Global convergence of the algorithm is proved and a local superlinear convergence result is presented. Numerical examples are also provided.

§1. The Problem

In this paper, we present a dual algorithm for minimizing a convex quadratic function with two special quadratic constraints. The problem has the form:

$$\min_{d\in\Re^n} \Phi(d) \equiv g^T d + \frac{1}{2}d^T Bd, \tag{1.1}$$

subject to

$$\|d\|_2 \le \Delta, \tag{1.2}$$

$$\|A^T d + c\|_2 \le \xi, \tag{1.3}$$

where $g \in \Re^n, B \in \Re^{n\times n}, A \in \Re^{n\times m}, c \in \Re^m, \Delta > 0, \xi \ge 0$ and B is a symmetric matrix. Problem (1.1)–(1.3) is a subproblem of some trust region algorithms for constrained optimization (for example, see Celis, Dennis and Tapia, 1985; and Powell and Yuan, 1986). Some theoretical properties of the problem are presented in Yuan (1987) for general B, but now we restrict attention to the case when B is positive definite, because we have not yet found a reliable method for computing the global solution in the general case.

The algorithm, given in Section 3, is based on Newton's method for the dual program of the following problem:

$$\min_{d\in\Re^n} \Phi(d) \equiv g^T d + \frac{1}{2}d^T Bd, \tag{1.4}$$

subject to

$$\|d\|_2^2 \le \Delta^2, \tag{1.5}$$

$$\|A^T d + c\|_2^2 \le \xi^2, \tag{1.6}$$

* Received August 16, 1988.

1) Research Supported by a Research Fellowship of Fitzwilliam College, Cambridge and by a research grant from the Chinese Academy of Sciences.

which is equivalent to (1.1)–(1.3).

In the next section, we give some theoretical properties of the dual problem. Then an algorithm is presented in Section 3, and convergence properties of the algorithm are given in Section 4. Numerical results are reported in Section 5, and finally a short discussion is given in Section 6.

§2. Dual Theory

For the dual variables $\lambda \geq 0, \mu \geq 0$ we define the matrix

$$H(\lambda,\mu) = B + \lambda I + \mu A A^T, \tag{2.1}$$

and the vector

$$d(\lambda,\mu) = -H(\lambda,\mu)^{-1}(g + \mu A c). \tag{2.2}$$

We also define the function

$$\Psi(\lambda,\mu) = \Phi(d(\lambda,\mu)) + \frac{1}{2}\lambda(\|d(\lambda,\mu)\|_2^2 - \Delta^2) + \frac{1}{2}\mu(\|A^T d(\lambda,\mu) + c\|_2^2 - \xi^2). \tag{2.3}$$

The dual problem for (1.4)–(1.6) is

$$\max_{(\lambda,\mu)\in\Re^2_+} \Psi(\lambda,\mu), \tag{2.4}$$

where we use the notation $\Re^2_+ = \{\lambda \geq 0, \mu \geq 0\}$. The relation of the dual problem to the primal problem is given in Lemma 2.2 below. One advantage of working with the dual problem (2.4) is that it has only two variables. Moreover, because gradients and second-order derivatives of $\Psi(\lambda,\mu)$ can be easily computed, (2.4) can be solved by applying Newton's method. Because the vector (2.2) is the value of $d(\lambda,\mu)$ that minimises the righthand side of expression (2.3), direct calculations show that

$$\nabla\Psi(\lambda,\mu) = \frac{1}{2}\begin{pmatrix} \|d(\lambda,\mu)\|_2^2 - \Delta^2 \\ \|A^T d(\lambda,\mu) + c\|_2^2 - \xi^2 \end{pmatrix}, \tag{2.5}$$

$$\nabla^2\Psi(\lambda,\mu) = -\begin{pmatrix} d(\lambda,\mu)^T H(\lambda,\mu)^{-1} d(\lambda,\mu) & d(\lambda,\mu)^T H(\lambda,\mu)^{-1} y(\lambda,\mu) \\ d(\lambda,\mu)^T H(\lambda,\mu)^{-1} y(\lambda,\mu) & y(\lambda,\mu)^T H(\lambda,\mu)^{-1} y(\lambda,\mu) \end{pmatrix}, \tag{2.6}$$

where $y(\lambda,\mu)$ is the vector

$$y(\lambda,\mu) = A(A^T d(\lambda,\mu) + c). \tag{2.7}$$

It is easy to see that $\Psi(\lambda,\mu)$ is a concave function. Another advantage of working with the dual problem (2.4) is that, as shown in the following lemma, the gradient and the Jacobian of $\Psi(\lambda,\mu)$ are both bounded above, even if the constraints (1.2) and (1.3) are inconsistent.

Lemma 2.1. *Let $d(\lambda,\mu)$ be defined by (2.2). Then*

$$\max_{(\lambda,\mu)\in\Re^2_+} \|d(\lambda,\mu)\|_2 \tag{2.8}$$

is finite. Consequently, $\nabla\Psi(\lambda,\mu)$ and $\nabla^2\Psi(\lambda,\mu)$ are bounded above in $\Re^2_+$.

Proof. The definition (2.2) shows that

$$\|d(\lambda,\mu)\|_2 \leq \|H(\lambda,\mu)^{-1} g\|_2 + \|H(\lambda,\mu)^{-1}\mu A c\|_2. \tag{2.9}$$