Journal of Computational Mathematics, Vol.21, No.1, 2003, 71–84.

LINEAR SYSTEMS ASSOCIATED WITH NUMERICAL METHODS FOR CONSTRAINED OPITMIZATION [*1)]

Y. Yuan [†]

(*State Key Laboratory of Scientific and Engineering Computing, Institute of Computational Mathematics and Scientific/Engineering Computing, Chinese Academy of Sciences, P. O. Box 2719, Beijing 100080, China*)

Dedicated to the 80th birthday of Professor Zhou Yulin

Abstract

Linear systems associated with numerical methods for constrained optimization are discussed in this paper. It is shown that the corresponding subproblems arise in most well-known methods, no matter line search methods or trust region methods for constrained optimization can be expressed as similar systems of linear equations. All these linear systems can be viewed as some kinds of approximation to the linear system derived by the Lagrange-Newton method. Some properties of these linear systems are analyzed.

Key words: Constrained optimization, Linear equations, Lagrange-Newton method, Trust region, Line search

1. Introduction

General nonlinear optimization problems have the form:

$$\min_{x\in\Re^n} f(x) \tag{1.1}$$

subject to

$$c_i(x) = 0, \qquad i = 1, 2, \cdots, m_e, \tag{1.2}$$
$$c_i(x) \geq 0, \qquad i = m_e + 1, \cdots, m, \tag{1.3}$$

where $m \geq m_e \geq 0$ are two non-negative integers. From the Kuhn-Tucker theory, at a local solution x^* of (1.1)-(1.3), there exist Lagrange multipliers $\lambda_i (i = 1, 2, \cdots, m)$ such that

$$\nabla f(x^*) - \sum_{i=1}^{m} \lambda_i \nabla c_i(x^*) = 0, \tag{1.4}$$
$$\lambda_i \geq 0, \quad \lambda_i c_i(x^*) = 0, \quad i = m_e + 1, \cdots, m. \tag{1.5}$$

Let $\mathcal{E} = \{1, 2, \cdots, m_e\}$, and $\mathcal{I}^* = \{i| \quad c_i(x^*) = 0, i = m_e + 1, \cdots, m\}$ be the index set of all active inequality constraints. The first order necessary condition (1.4)-(1.5) can be written as

$$\nabla f(x^*) - \sum_{i\in\mathcal{E}\cup\mathcal{I}^*} \lambda_i \nabla c_i(x^*) = 0. \tag{1.6}$$

[*] Received September 30, 2002.

[1)] Research partially supported by Chinese NSF grants 19731010 and CAS knowledge innovation program.

† Email address: yyx@lsec.cc.ac.cn.

Thus, when the iterates are close to a solution, inequality constraints can be treated as equality constraints by applying the active set strategy. Therfore, for simplicity, some of the methods we discussed in the paper are for equality constrained problem

$$\min_{x\in\Re^n} \quad f(x) \tag{1.7}$$
$$s.\quad t. \qquad c(x) = 0. \tag{1.8}$$

Some methods require the iterates staying in the interior of the feasible region, therefore only inequality constraints are considered. For these methods, we can only apply to inequality constrained problems:

$$\min_{x\in\Re^n} \quad f(x) \tag{1.9}$$
$$s.\ t. \qquad c(x) \geq 0. \tag{1.10}$$

Almost all numerical methods for nonlinear optimization are iterative. For a line search method, a search direction d_k will be generated and a suitable point $x_k + \alpha_k d_k$ is chosen so that a reduction in a merit function (which is a penalty function) will be obtained. For a trust region method, a trial step s_k is computed in a trust region, and some criterion will be used to decide whether the step s_k should be accepted.

For unconstrained problem ($m = m_e = 0$), the Newton's method is

$$x_{k+1} = x_k - (\nabla^2 f(x_k))^{-1} \nabla f(x_k), \tag{1.11}$$

which has a local quadratic convergence property if the Hessian matrix is positive definite at the solution. The Newton step $d = -(\nabla^2 f(x_k))^{-1}\nabla f(x_k)$ can be obtained by solving the following linear system

$$(\nabla^2 f(x_k))d = -\nabla f(x_k). \tag{1.12}$$

A very important class of methods for unconstrained optimization, quasi-Newton methods, define the search direction by solving

$$B_k d = -\nabla f(x_k), \tag{1.13}$$

where B_k is a quasi-Newton matrix. The linear system determines the next iterate, therefore plays the essential role for the convergence rate of the method. It is well known([3]) that the superlinear convergence of quasi-Newton methods is equavalent to

$$\lim_{k\to\infty} \frac{\|(B_k - \nabla^2 f(x_k))d_k\|}{\|d_k\|} = 0. \tag{1.14}$$

For constrained optimization problems, the search directions or the trial steps are computed by solving some subproblems. These subproblems are some kinds of approximation to the orginal optimization problem. Most of these subproblems are simple optimization problems. For example, the quadratic subproblem of the sequential quadratic programming method for (1.1)-(1.3) has the form

$$\min_{d\in\Re^n} d^T \nabla f(x_k) + \frac{1}{2} d^T B_k d \tag{1.15}$$

$$s.\quad t. \qquad c_i(x_k) + d^T \nabla c_i(x_k) = 0, \qquad i = 1, \cdots, m_e; \tag{1.16}$$
$$c_i(x_k) + d^T \nabla c_i(x_k) \geq 0, \qquad i = m_e + 1, \cdots, m, \tag{1.17}$$