Mathematical Programming 6 (1974) 62–88. North-Holland Publishing Company

VALIDATION OF SUBGRADIENT OPTIMIZATION

Michael HELD
IBM Systems Research Institute, New York, U.S.A.

Philip WOLFE, Harlan P. CROWDER
IBM Watson Research Center, Yorktown Heights, New York, U.S.A.

Received 6 August 1973
Revised manuscript received 26 November 1973

The "relaxation" procedure introduced by Held and Karp for approximately solving a large linear programming problem related to the traveling-salesman problem is refined and studied experimentally on several classes of specially structured large-scale linear programming problems, and results on the use of the procedure for obtaining exact solutions are given. It is concluded that the method shows promise for large-scale linear programming

1. Introduction

A wide variety of mathematical programming problems can be given the form

$$\text{maximize } w(\pi), \ \pi \in S\,, \tag{1.1}$$

where

$$w(\pi) = \min\{c_k + \pi\cdot v_k \colon k = 1, ..., K\}\,, \tag{1.2}$$

c_k is a scalar, $v_k = (v_{k1}, ..., v_{kn})$ is a real n-vector for all k, and S is a closed convex subset of E^n. ($\pi\cdot v_k$ above denotes the inner product.) If K is finite and S is a polyhedron, then the problem is alinear programming problem; otherwise it is the general problem of maximizing a concave function over S, for any concave function has such a representation [22, p. 102]. This paper reports some experiments testing an unusual and surprisingly successful procedure for the numerical solution

of certain problems of this kind. We call this procedure a *subgradient method.*

If the constraint set S can be specified by a simple set of linear inequalities (as it is in the problems we consider) and K is small, then our problem is just an ordinary linear programming problem, and we think it is best solved by the simplex method for linear programming. In the problems we are concerned with, however, K is vast-finite, but easily of the order of $n!$ with n around 100. We therefore make no effort to write out the data c_k, v_k explicitly for all k, but rather suppose them so defined that, *given* π, the task of then finding c_k, v_k for which the minimum of (1.2) is assumed is reasonably convenient. That task, called "column generation", is central to most schemes for the solution of very large linear programming problems (a good account of which is given by Lasdon [15, Chapter 4]). As we have discussed elsewhere [24], all problems amenable to treatment by the Dantzig–Wolfe decomposition procedure can be cast in the form (1.1), (1.2).

A first encounter with the kind of method reported here was described by Held and Karp [11] in their work on the traveling-salesman problem, for which they devised a function w of the type (1.2) constituting a kind of "pseudo-dual" for that problem; its maximum value provided excellent bounds for their branch-and-bound algorithm. Having tried both a steepest-ascent procedure and a simplex method using column generation for maximizing w, and finding them dis-hearteningly slow, they invented a "subgradient" method which turned out to be highly effective. Subsequently, A.J. Hoffman pointed out that their procedure could be viewed as an application of a method discussed by Agmon [2] and Motzkin and Schoenberg [17] for the solution of linear inequality systems, and Held and Karp presented the procedure from that point of view. We have recently found a sizable body of Russian literature dealing with the theory of subgradient methods, which apparently originated with Shor [23], and has perhaps been best expounded by Poljak [20, 21]; we recommend his work for a discussion of the algorithm in more general spaces and for theoretical results on the rate of convergence.

Despite the mathematical interest of the Agmon–Motzkin–Schoenberg work and that of the Russians, there appears to be little computational experience with the method beyond that reported by Held and Karp. References in the Soviet literature to computational experience are at best cryptic. For example, quoting in full one of the two such references we have found [21]: "The effectiveness of the methods