Available online at www.sciencedirect.com
ScienceDirect
European Journal of Operational Research 176 (2007) 27–45
EUROPEAN JOURNAL OF OPERATIONAL RESEARCH
www.elsevier.com/locate/ejor

Continuous Optimization

A heuristic algorithm for a chance constrained stochastic program

Concetta A. DePaolo [a,*], David J. Rader Jr. [b,1]

[a] College of Business, Indiana State University, Terre Haute, IN 47809, USA
[b] Department of Mathematics, Rose-Hulman Institute of Technology, Terre Haute, IN 47803, USA

Received 5 August 2004; accepted 18 August 2005
Available online 18 November 2005

Abstract

A chance constrained stochastic program is considered that arises from an application to college enrollments and in which the objective function is the expectation of a linear function of the random variables. When these random variables are independent and normally distributed with mean and variance that are linear in the decision variables, the deterministic equivalent of the problem is a nonconvex nonlinear knapsack problem. The optimal solution to this problem is characterized and a greedy-type heuristic algorithm that exploits this structure is employed. Computational results show that the algorithm performs well, especially when the normal random variables are approximations of binomial random variables.

Keywords: Heuristics; Nonlinear programming; Nonlinear knapsack problem; Chance constraints

1. Introduction and background

We consider a stochastic program with a chance constraint in which, with a given probability, a linear combination of the random variables Y_i is bounded either above or below. Each random variable Y_i is normally distributed with mean and variance that are linear in x_i, a decision variable assumed to be bounded and continuous. The objective is to maximize the expectation of a linear function of the Y_i.

Such a decision problem may arise in higher education applications when an institution wishes to maximize the desirability of a class of students while satisfying a constraint on the number of students who

* Corresponding author. Tel.: +1 812 237 2283; fax: +1 812 237 8133.
E-mail addresses: C-DePaolo@indstate.edu (C.A. DePaolo), david.rader@rose-hulman.edu (D.J. Rader Jr.).
[1] Tel.: +1 812 877 8361; fax: +1 812 877 8883.

0377-2217/$ - see front matter
doi:10.1016/j.ejor.2005.08.008

enroll. In this setting, the decision variables x_i represent the number of students of a given type who are admitted, while the random variable Y_i represents the number of such students who enroll. Similar problems may arise in other applications such as: airline or hotel overbooking problems, where x_i is the number of reservations accepted and Y_i is the number of passengers or guests; marketing, where x_i is the number of solicitations issued and Y_i is the number of responses; survey research, where x_i is the size of the sample and Y_i is the number of responses; and manufacturing, where x_i is the number of items produced and Y_i is the number of nondefective items.

In each of these applications, the random variables can be characterized as binomial random variables dependent on the x_i, and since the number of items is typically large, the binomial random variables are easily approximated by normal random variables. Using this approximation, the deterministic equivalent optimization problem is a nonlinear program with a linear objective and a nonconvex constraint. Optimal solutions to such problems have a special structure that can be exploited by a greedy-type heuristic algorithm that performs very well in practice. This algorithm also performs well in some cases in which the normal random variables do not arise from approximations to binomials.

The deterministic problem can be described as a nonconvex nonlinear knapsack problem. In a review of the literature, Bretthauer and Shetty [5] cite many examples of nonlinear knapsack problems, algorithms and applications. Much of the research involves convex objective functions and linear constraints (for example, [13,16,23,25,28]) or the convex separable case (e.g. [3,4,14,17,22]). Nonconvex problems are not well represented in the literature, although Bretthauer and Shetty mention a few exceptions (e.g. [2,15,24]).

The application of a greedy procedure to find the optimal solution of a continuous linear knapsack problem is well known, and other applications of greedy-type heuristics have also been studied extensively in the literature. However, most have been applied to decision problems with linear constraints. Some references have investigated greedy algorithms with regard to integer knapsack problems (for example, [18,21,27]). In [12], the author investigates a "pseudogreedy" algorithm applied to a continuous nonlinear knapsack problem with linear constraints. It is believed that the problem investigated in the following discussion is not well represented in the literature, as it explores a nonlinear program in which the constraints are neither linear nor convex and applies a greedy heuristic to find an approximate solution.

The organization of the paper is as follows. In Section 2, a description of the application and the resulting optimization problem is provided. In Section 3, the greedy structure of the optimal solution is established, which leads to an "ordering" of the decision variables presented in Section 4. In Section 5, this "ordering" is used to develop a heuristic algorithm for the decision problem, and in Section 6, computational results are presented. Conclusions and future research are presented in Section 7.

2. The college enrollment application

Consider a decision problem arising from the following higher education application. Each year an institution must decide which applicants to admit in order to best achieve the institution's goals, which may include maximizing the quality or desirability of the students who enroll. At the same time, the institution has a limited capacity for students that it can accommodate and may wish to impose bounds on enrollment. Uncertainty arises because admitted students may not enroll.

Finding the best mix of enrolled students to achieve institutional objectives has been dealt with in the higher education literature in a largely qualitative manner. The process of making admissions decisions to achieve the desired mix has been addressed with simple calculations based on the historical probability that an admitted student enrolls, with logistic regression models (e.g. [6,20]) or more recently, with artificial neural networks (e.g. [7]). The literature suggests that this problem has largely been studied in a piecemeal fashion, although there are some exceptions. For example, both [10] and [11] use mathematical programming techniques and consider both admissions decisions and resulting enrollment. Other authors have used