Discrete Applied Mathematics 200 (2016) 161–169

Contents lists available at ScienceDirect

Discrete Applied Mathematics

journal homepage: www.elsevier.com/locate/dam

Simultaneous selection

Wojciech Olszewski [a], Rakesh Vohra [b,*]

[a] Department of Economics, Northwestern University, 2001 Sheridan Rd., Evanston IL 60208-2600, Philippines

[b] Department of Economics & Department of Electrical and Systems Engineering, University of Pennsylvania, Philadelphia, PA 19104, United States

ARTICLE INFO

Article history:
Received 10 June 2014
Received in revised form 8 July 2015
Accepted 10 July 2015
Available online 30 July 2015

Keywords:
Submodularity
Greedy
Selection

ABSTRACT

We generalize Chade and Smith's (2006) simultaneous search problem to a class of discrete optimization problems. More precisely, we study the problem of maximizing a weighted sum of utilities of objects minus the sum of costs of acquiring these objects, given the constraint that the sum of weights cannot exceed the value of some submodular function.

We show that the problem has a simple solutions in the particular case in which the submodular function depends only on the number of objects. Namely, the optimal set of objects can be found by the greedy algorithm. We provide some economic applications of this result. The particular case studied in the present paper, and the particular case studied by Chade and Smith complement one another, but they do not exhaust all instances of our general discrete optimization problem. We also show that in the general case the problem does not have a simple solution.

1. Introduction

Chade and Smith [4] introduce an interesting discrete optimization problem involving a high-school student applying to colleges. The student must choose a subset of colleges. Each application costs a $c > 0$. Studying at different colleges gives different utilities, and the probability of being admitted also varies across colleges. From the set of colleges to which the student is admitted, she or he chooses the one with the highest utility. Chade and Smith point out that the threshold strategy of applying to colleges which give more than a certain utility may not be optimal, and they show that the optimum is attained by a simple *greedy* algorithm.[1]

According to this algorithm, we first put on a list the college we would apply to, if we were allowed to apply to one college only. Then, we add to this list the college that we would apply to, if we were allowed to apply to one more college. We keep adding colleges to the list until the marginal benefit of adding any other college falls below the cost c.

In this paper we are concerned with a more general problem of maximizing a weighted sum of utilities of objects minus the sum of costs of acquiring these objects, given the constraint that the sum of weights assigned to any subset of objects S cannot exceed $f(S)$, where f is a non-decreasing, non-negative submodular function. Call this more general problem the *simultaneous selection problem*, or (*SSP*) in abbreviation.

This generalizes Chade and Smith's problem in two ways. It allows for different objects having different costs. In addition, in Chade and Smith's problem $f(S)$ has a very special form, which is the probability of being admitted by at least one college

* Corresponding author.
E-mail addresses: wo@northwestern.edu (W. Olszewski), rvohra@seas.upenn.edu (R. Vohra).

[1] It should be noted that the optimality of the greedy algorithm for the college application problem is not the only result in Chade and Smith (see Section 2.1).

http://dx.doi.org/10.1016/j.dam.2015.07.010
0166-218X/

from S. Both ways in which Chade and Smith's problem is generalized, allow for new applications. In searching for a job instead of applying to colleges, the application process typically varies across employers. Other functions f also enlarge the set of applications. For example, imagine an inventory planning problem, in which you must order today some goods, which you will later be selling gradually, one per period. In this application, assuming geometric discounting, $f(S)$ is the present value of the stream of 1's over $|S|$ periods.

We show that the (SSP) can be solved by the greedy algorithm in the case in which $f(S)$ depends on $|S|$ only. In particular, the greedy algorithm selects an optimal set in some versions of the applications described in the previous paragraph. Our result and that obtained by Chade and Smith complement one another. Neither subsumes the other, and they also apply to somewhat different settings.

In the absence of additional assumptions on the function f, the (SSP) is NP-hard. In fact it is APX-hard (Feige et al. (2009)). An interesting open question is the design of good approximation algorithms for the (SSP). We show by an example that for any $\varepsilon > 0$, the greedy algorithm may select a set S that returns less than an ε share of the optimum.

2. Simultaneous selection problem

Let f be a nonnegative function defined on all subsets of a finite set N. We assume that f is non-decreasing, i.e.,

$$f(S) \le f(T), \quad \forall S \subset T \subset N,$$

and submodular, i.e.,

$$f(S \cap T) + f(S \cup T) \le f(S) + f(T), \quad \forall S, T \subset N.$$

An equivalent definition of submodularity is

$$f(S \cup i) - f(S) \ge f(T \cup i) - f(T)$$

for all $S \subseteq T$ and $i \notin T$.

Given numbers (utilities) $w_i \ge 0, \forall i \in N$, define the function g on each $T \subseteq N$ by the following optimization problem:

$$g(T) = \max \sum_{i \in T} w_i x_i$$
$$\text{s.t.} \sum_{i \in S} x_i \le f(S), \quad \forall S \subseteq T$$
$$x_i \ge 0, \quad \forall i \in T.$$

This function can be interpreted as the maximum utility that one can obtain by assigning weights x_i to the elements of T, subject to the "resource" constraint that the aggregate weight assigned to any set S cannot exceed $f(S)$.

The submodularity of f implies that g is defined by maximizing a linear function over a polymatroid. Hence, using the well known greedy algorithm for polymatroid optimization, guarantees that g has the following expression. Label the elements of T as $\{1, 2, \ldots, |T|\}$, and order them so that $w_1 \ge w_2 \ge \cdots w_{|T|}$. Then, it is optimal to assign the highest possible weight, $x_1 = f(1)$, to the element with the highest utility, that is, to element 1 ; and next, to assign the highest possible weight subject to the resource constraint, $x_2 = f(1, 2) - f(1)$, to the element with the second-highest utility, that is, to element 2; etc., until assigning weight $x_{|T|} = f(1, 2, \ldots, |T|) - f(1, 2, \ldots, |T| - 1)$ to element $|T|$.[2]

In what follows, we denote by $x(T)$ the vector of optimal weights assigned to set $T \subseteq N$. Hence,

$$g(T) = \sum_{i \in T} w_i x_i(T).$$

It is immediate that function g, defined on all subsets of N, is submodular.

Given numbers (costs) $c_i \ge 0, \forall i \in N$, we define a problem called the simultaneous selection problem (in short, (SSP)) as

$$\max_{T \subseteq N} \left\{ g(T) - \sum_{i \in T} c_i \right\}. \tag{SSP}$$

The number c_i can be interpreted as the cost of including element i to set T. For convenience, set

$$H(T) = g(T) - \sum_{i \in T} c_i.$$

In the following sections, we demonstrate that the (SSP) covers various, basic and seemingly unrelated applications.[3]

---

2 See [6].

3 We note that the (SSP) belongs to a larger class of problems that are obtained by replacing a hard capacity constraint with a penalty term that scales linearly with a violation in the capacity constraint. For an example of this see [1].