Operations Research Letters 42 (2014) 367–373

Contents lists available at ScienceDirect

Operations Research Letters

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

A continuous knapsack problem with separable convex utilities: Approximation algorithms and applications

Retsef Levi, Georgia Perakis *, Gonzalo Romero

Sloan School of Management, Massachusetts Institute of Technology, E62-565, Cambridge, MA 02142, United States

ARTICLE INFO

Article history:
Received 6 September 2013
Received in revised form
12 June 2014
Accepted 13 June 2014
Available online 20 June 2014

Keywords:
Knapsack
Subsidies
Cournot competition

ABSTRACT

We study a continuous knapsack problem with separable convex utilities. We show that the problem is NP-hard, and provide two simple algorithms that have worst-case performance guarantees. We consider as an application a novel subsidy allocation problem in the presence of market competition, subject to a budget constraint and upper bounds on the amount allocated to each firm, where the objective is to minimize the market price of a good.

1. Introduction

We study a continuous knapsack problem, where the objective is to maximize the sum of separable convex utility functions. We denote this problem by (CKP). Beyond general methods for concave minimization, see for example [1], there is not much literature on this class of problems. An exception is [10], and their algorithm to find local minima. A comprehensive review of the related nonlinear knapsack problem literature is presented in [3]; however, in most cases, the objective function considered in this literature is concave. On the other hand, for any given tolerance $\epsilon > 0$, a fully polynomial time approximation scheme (FPTAS) is an algorithm that generates a solution which is within a factor $(1 - \epsilon)$ of being optimal, while the running time of the algorithm is polynomial in the problem size and $1/\epsilon$. Burke et al. [4] provide a tailored FPTAS for a minimization variant of a continuous knapsack problem, in the context of allocating procurement to suppliers. The knapsack problem we study here is a maximization problem, hence the results from [4] do not apply. Finally, [7] develop a general purpose FPTAS for a class of stochastic dynamic programs, which applies to general nonlinear knapsack problems. In contrast, our main goal in this paper is to study the performance of simple algorithms for problem (CKP), as well as to introduce a novel application of continuous knapsack problems into a subsidy allocation problem in the presence of market competition, subject to a budget constraint and upper bounds on the amount allocated to each firm, where the objective is to minimize the market price of a good.

The main contributions of this paper are two-fold. First, we develop two algorithms that are computationally and conceptually simple, such that they can be used in practical applications. We show that these algorithms have good worst-case performance guarantees for problem (CKP). Moreover, we identify special settings where these simple policies are actually optimal. Second, we show that problem (CKP) characterizes a novel subsidy allocation problem, and that the simple algorithms that we develop admit a practical interpretation.

2. Problem formulation

Consider n items indexed by $i \in \{1, \ldots, n\}$. For each i, let x_i be the non-negative quantity of item i, and let $f_i(x_i)$ be the resulting reward. Moreover, $f_i(x_i)$ is assumed to be convex. The quantity of item i cannot exceed a given upper bound u_i, and the total amount of all items is bounded by the capacity of the knapsack, denoted by B. Moreover, both B and u_i are assumed to be integers. We are interested in the following continuous knapsack problem

$$\begin{aligned} & \max \quad F(x) \equiv \sum_{i=1}^{n} f_i(x_i) \\ (\mathrm{CKP}) \quad & \text{s.t.} \quad \sum_{i=1}^{n} x_i \le B \\ & \qquad\; 0 \le x_i \le u_i \; \forall\, i. \end{aligned}$$

The objective function is convex over the feasible set, which is a bounded polyhedron. Therefore, the existence of an extreme point

* Corresponding author.
E-mail addresses: retsef@mit.edu (R. Levi), georgiap@mit.edu (G. Perakis), gromeroy@mit.edu (G. Romero).

http://dx.doi.org/10.1016/j.orl.2014.06.007
0167-6377/

optimal solution follows from concave minimization theory, see for example [1].

The next one is our first result

Proposition 1. *Problem (CKP) is NP-hard.*

Proof. The proof is a reduction from the subset sum problem, which is well known to be NP-complete, see [8].

Consider an arbitrary instance of the subset sum problem, where given a set of n positive integers $\{u_1, u_2, \ldots, u_n\}$, and a positive integer B, the question is whether there exists a subset $J \subseteq \{u_1, u_2, \ldots, u_n\}$ that sums to B.

Now consider the following instance of problem (CKP): let u_i be the upper bound on x_i for each item i, B be the capacity of the knapsack and $f(x_i) = x_i(x_i - u_i) + x_i$ be the convex reward for each item i. It follows that this instance of problem (CKP) can be written as

$$\begin{aligned} \max \quad & \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} x_i(u_i - x_i) \\ \text{s.t.} \quad & \sum_{i=1}^{n} x_i \le B \\ & 0 \le x_i \le u_i \; \forall\, i \in \{1, \ldots, n\}. \end{aligned}$$

Note that B is an upper bound on the optimal objective value of this problem. Moreover, this upper bound is attained if and only if there exists a subset $J \subseteq \{u_1, u_2, \ldots, u_n\}$ that sums to B.

Hence, if we can solve problem (CKP) in polynomial time, it follows that we can solve the subset sum problem in polynomial time.

The proof of Proposition 1 is in the same spirit of [12], who shows the NP-hardness of non-convex quadratic programming, among other problems.

We now make a couple of remarks that will make the rest of the exposition clearer.

Remark 1. There is no loss of generality in assuming that, for each $i \in \{1, \ldots, n\}$, the functions $f_i(x_i)$ are positive and non decreasing.

Specifically, we can pre-process the problem data replacing $f_i(x_i)$ by the amount $\max\{f_i(x_i), f_i(0)\}$, for each i and x_i, obtaining non decreasing functions without changing the problem. Similarly, by adding a constant $K > \min_i\{f_i(0)\}$ to each of the functions $f_i(x_i)$ we obtain positive functions.

Remark 2. There is no loss of generality in assuming that, for each $i \in \{1, \ldots, n\}$, $u_i \le B$.

Specifically, if any upper bound u_i is larger than the capacity B, then it follows that any feasible solution will allocate at most B to item i. Hence, we can pre-process the data and replace u_i by $\min\{B, u_i\}$, for each i, without changing the problem.

Having established the NP-hardness of problem (CKP), we now focus on simple algorithms with a guaranteed performance, and their practical interpretation.

2.1. *A simple 1/2-approximation algorithm*

We next describe a 1/2-approximation algorithm for problem (CKP). Specifically, we will show that intuitive ideas perform well in this model. Namely, the best solution between (i) allocating the capacity greedily to the items with the *fastest rate of increase* in their utility function, and (ii) allocating the capacity greedily to the items with the *largest absolute increase* in their utility function, attains an objective value that is at most half the value of the optimal objective value.

This algorithm is a generalization of the well known 1/2-approximation algorithm for the 0/1 knapsack problem. The latter is attained by the best solution between greedily picking the objects by decreasing ratio of profit to size, and picking the most profitable object, see for example [15].

Consider first idea (i). We denote the resulting solution by $\mathbf{x}^{\text{rate}}$. Essentially, $\mathbf{x}^{\text{rate}}$ is the result of a greedy procedure with respect to $\frac{f_i(u_i) - f_i(0)}{u_i}$, which is the rate of increase in the utility function of item i, assuming that x_i is set to its upper bound.

Algorithm 1 Compute $\mathbf{x}^{\text{rate}}$

$\mathbf{x}^{\text{rate}} \Leftarrow \vec{0}$
Let $\lambda_i = \frac{f_i(u_i) - f_i(0)}{u_i}$, for each i
Sort indexes by decreasing λ_i
Find $\hat{i}$ s.t. $\sum_{i=1}^{\hat{i}-1} u_i \le B$ and $\sum_{i=1}^{\hat{i}} u_i > B$
$x_i^{\text{rate}} \Leftarrow u_i$, for each $i \le \hat{i}$
$x_{\hat{i}}^{\text{rate}} \Leftarrow B - \sum_{i=1}^{\hat{i}-1} u_i$

On the other hand, consider idea (ii). We denote the resulting feasible solution by $\mathbf{x}^{\max}$. Essentially, $\mathbf{x}^{\max}$ is the result of a greedy procedure with respect to $f_i(\min(u_i, \tilde{B}))$, which is the absolute increase in the utility function of item i, when allocating the minimum between the remaining capacity $\tilde{B}$, and its upper bound. In case of a tie, $f_i(0)$ is used as a tie-breaker.

Algorithm 2 Compute $\mathbf{x}^{\max}$

$\tilde{B} \Leftarrow B$
$\mathbf{x}^{\max} \Leftarrow \vec{0}$
while $\tilde{B} > 0$ **do**
  Let $S_1 = \left\{ i \mid f_i\left(\min\left(u_i, \tilde{B}\right)\right) \ge f_j\left(\min\left(u_j, \tilde{B}\right)\right), \text{ for each } j : \mathbf{x}_j^{\max} = 0 \right\}$
  Let $S_2 = \left\{ i \in S_1 \mid f_i(0) \le f_j(0), \text{ for each } j \in S_1 \right\}$
  Select $i \in S_2$
  $\tilde{B} \Leftarrow \tilde{B} - \min\left(u_i, \tilde{B}\right)$
  $x_i^{\max} \Leftarrow \min\left(u_i, \tilde{B}\right)$
end while

It is not hard to see that each algorithm, considered separately, can be made to perform arbitrarily bad. Examples drawn from a 0/1 knapsack problem are sufficient.

In order to show a worst-case performance guarantee for problem (CKP), we need an upper bound on its optimal objective value, as provided in the following proposition.

Proposition 2. *Let* $\mathbf{x}^*$ *be an optimal solution to problem (CKP). Algorithm* 1 *provides the following upper bound,*

$$F(\mathbf{x}^*) \le F(\mathbf{x}^{\text{rate}}) + f_{\hat{i}}(0) - f_{\hat{i}}(x_{\hat{i}}^{\text{rate}}) + \frac{f_{\hat{i}}(u_{\hat{i}}) - f_{\hat{i}}(0)}{u_{\hat{i}}} x_{\hat{i}}^{\text{rate}}$$

where $F(x) = \sum_{i=1}^{n} f_i(x_i)$.

Proof. Let us relax the knapsack constraint in problem (CKP) with an associated Lagrange multiplier λ, to obtain the following relaxed optimization problem,

$$\begin{aligned} \max \quad & \lambda B + \sum_{i=1}^{n} (f_i(x_i) - \lambda x_i) \\ \text{s.t.} \quad & 0 \le x_i \le u_i \; \forall\, i. \end{aligned}$$

The resulting problem is separable in the variables x_i. Specifically, for each variable it maximizes a convex function over a closed interval. It follows that the optimal solution is attained at one of the extremes of the interval. For any fixed Lagrangian multiplier λ, let $G(\lambda)$ denote the optimal objective value of the relaxed problem. Namely, $G(\lambda) = \lambda B + \sum_{i=1}^{n} \max(f_i(0), f_i(u_i) - \lambda u_i)$. From duality theory, it follows that, for any $\lambda \geq 0$, $G(\lambda)$ is an upper bound for the optimal objective value of problem (CKP), see for example [2]. Moreover, the best possible upper bound can be computed from the following program,

$$\begin{aligned} \min_{\lambda} \quad & G(\lambda) = \lambda B + \sum_{i=1}^{n} \max(f_i(0), f_i(u_i) - \lambda u_i) \\ \text{s.t.} \quad & \lambda \geq 0. \end{aligned}$$

The objective function of this problem is piecewise linear and convex. Hence, it can be solved by trying out the values of λ where the slope of the objective function changes. In particular, Algorithm 1 solves this problem. The optimal Lagrange multiplier is $\lambda_{\hat{i}} = \frac{f_{\hat{i}}(u_{\hat{i}}) - f_{\hat{i}}(0)}{u_{\hat{i}}}$, where $\hat{i}$ was defined in Algorithm 1 as being such that $\sum_{i=1}^{\hat{i}-1} u_i \leq B$ and $\sum_{i=1}^{\hat{i}} u_i > B$.

Plugging in the optimal Lagrange multiplier $\lambda_{\hat{i}}$ in $G(\lambda)$, results in the best possible upper bound from this relaxation. Without loss of generality, set $x_{\hat{i}} = 0$, then

$$\begin{aligned} G(\lambda_{\hat{i}}) &= \sum_{i=1}^{\hat{i}-1} f_i(u_i) + \sum_{i=\hat{i}}^{n} f_i(0) + \lambda_{\hat{i}} \underbrace{\left(B - \sum_{i=1}^{\hat{i}-1} u_i\right)}_{=x_{\hat{i}}^{\text{rate}}} \\ &= F(\mathbf{x}^{\text{rate}}) + f_{\hat{i}}(0) - f_{\hat{i}}(x_{\hat{i}}^{\text{rate}}) + \underbrace{\frac{f_{\hat{i}}(u_{\hat{i}}) - f_{\hat{i}}(0)}{u_{\hat{i}}}}_{\lambda_{\hat{i}}} x_{\hat{i}}^{\text{rate}}. \end{aligned}$$

The second equality follows from adding and subtracting the term $f_{\hat{i}}(x_{\hat{i}}^{\text{rate}})$. □

The next theorem is the main result in this section.

Theorem 1. *Let $\mathbf{x}^*$ be an optimal solution to problem (CKP). Let $\mathbf{x}^{\text{rate}}$ be the solution computed by Algorithm 1, and $\mathbf{x}^{\max}$ be the solution computed by Algorithm 2. Then*

$$\frac{\max\left(F(\mathbf{x}^{\text{rate}}), F(\mathbf{x}^{\max})\right)}{F(\mathbf{x}^*)} \geq \frac{1}{2}.$$

Proof. To make the notation clearer, define $\bar{f} = \max_i \{f_i(u_i)\}$. Note that,

$$\begin{aligned} &\frac{\max\left(F(\mathbf{x}^{\text{rate}}), F(\mathbf{x}^{\max})\right)}{F(\mathbf{x}^*)} \\ &\geq \frac{\max\left(F(\mathbf{x}^{\text{rate}}), \bar{f}\right)}{F(\mathbf{x}^*)} \\ &\geq \frac{\max\left(F(\mathbf{x}^{\text{rate}}), \bar{f}\right)}{F(\mathbf{x}^{\text{rate}}) + \underbrace{f_{\hat{i}}(0) - f_{\hat{i}}(x_{\hat{i}}^{\text{rate}})}_{\leq 0} + \frac{f_{\hat{i}}(u_{\hat{i}}) - f_{\hat{i}}(0)}{u_{\hat{i}}} x_{\hat{i}}^{\text{rate}}} \\ &\geq \frac{\max\left(F(\mathbf{x}^{\text{rate}}), \bar{f}\right)}{F(\mathbf{x}^{\text{rate}}) + \frac{f_{\hat{i}}(u_{\hat{i}})}{u_{\hat{i}}} x_{\hat{i}}^{\text{rate}}} \\ &\geq \frac{1}{1 + \underbrace{\frac{f_{\hat{i}}(u_{\hat{i}})}{\bar{f}}}_{\leq 1} \underbrace{\frac{x_{\hat{i}}^{\text{rate}}}{u_{\hat{i}}}}_{\leq 1}} \\ &\geq \frac{1}{2}. \end{aligned}$$

The first inequality follows from Remark 1, and the definitions of $\bar{f}$ and $\mathbf{x}^{\max}$. Specifically, they imply $\bar{f} \leq F(\mathbf{x}^{\max})$. The second inequality follows from Proposition 2, while the third inequality follows from Remark 1. The fourth inequality follows from the definition of $\bar{f}$. □

To conclude this section, the following lemma identifies three cases that can be solved in polynomial time. Specifically, Algorithm 1 solves problem (CKP) exactly if the utility functions of each item are affine, namely $f_i(x_i) = a_i + b_i x_i$ for each i, for some $a_i > 0$, $b_i > 0$. Algorithm 1 also solves problem (CKP) exactly if any number of items, ordered by fastest rate of increase in their utility function, fill the knapsack exactly. Additionally, if the upper bounds are uniform, then problem (CKP) can be solved by applying Algorithm 1 n times.

Lemma 1. *If the functions $f_i(x_i)$ are affine, or if the first $(\hat{i}-1)$ indexes sorted by decreasing value of $\lambda_i = \frac{f_i(u_i) - f_i(0)}{u_i}$ fill the knapsack exactly, for some value of $\hat{i}$, then $\mathbf{x}^{\text{rate}}$ is the optimal solution to problem (CKP).*

On the other hand, if the upper bounds on the allocation to each index are uniform, namely $u_i = u$ for each i, then problem (CKP) can be solved in polynomial time.

Proof. The first statement in the lemma is a direct consequence of Proposition 2. Specifically, if $x_{\hat{i}}^{\text{rate}} = 0$, then it follows that $F(\mathbf{x}^*) \leq F(\mathbf{x}^{\text{rate}})$, hence $\mathbf{x}^{\text{rate}}$ is optimal. This holds in both cases in the first statement of the lemma.

Assume now that $u_i = u$ for each i. Each extreme point solution is characterized by one fractional variable x_i, which gets an allocation $\left(B - \lfloor \frac{B}{u} \rfloor u\right)$, while $\lfloor \frac{B}{u} \rfloor$ other variables get an allocation u, and all the remaining variables get no allocation. From the first statement in the lemma, it follows that we can try each variable as the fractional variable, allocating $\left(B - \lfloor \frac{B}{u} \rfloor\right)$ to it; and then use Algorithm 1 to optimally solve the problem of allocating the remaining capacity $\lfloor \frac{B}{u} \rfloor u$, among the remaining variables. This follows because the first $\lfloor \frac{B}{u} \rfloor$ indexes sorted by decreasing value of $\lambda_i = \frac{f_i(u_i) - f_i(0)}{u_i}$ fill this modified knapsack exactly. In conclusion, in this case problem (CKP) can be solved by running Algorithm 1 n times. □

2.2. An $(1 - e^{-1})$-approximation algorithm

In this section we present an $\left(1 - e^{-1}\right)$-approximation algorithm for problem (CKP), where $\left(1 - e^{-1}\right) \approx 0.632$. We denote the resulting solution by $\mathbf{x}^{\text{seq}}$. In this algorithm we enumerate all the solutions that allocate capacity to 3 items or less. Then, for each of these solutions we allocate the remaining capacity, if any, greedily to the remaining indexes with the *fastest rate of increase* in their utility function. In that sense, Algorithm 3 is an extension of Algorithm 1. It captures that among the two simple rules we have considered, the fastest rate of increase rule is the most powerful. Specifically, it is enough to consider all solutions that allocate capacity to 3 items or less, to rule out all the cases where the largest absolute increase rule was important to define the worst-case guarantee.

To the best of our knowledge, this is the first time that these ideas have been used in a continuous optimization setting, like our continuous knapsack problem (CKP). Similar ideas have been used

Algorithm 3 Compute $\mathbf{x}^{\mathrm{seq}}$

Consider all sequences of 3 different indexes and allocate the capacity in this order
Let $\tilde{B}$ be the remaining capacity
for Each sequence **do**
  if $\tilde{B} > 0$ **then**
    Apply Algorithm 1 to the rest of the indexes with capacity $\tilde{B}$
  end if
end for

before in inherently discrete settings, such as solving a budgeted maximum coverage problem in [9], and maximizing a submodular set function subject to a knapsack constraint in [14].

Theorem 2. *Let $\mathbf{x}^*$ be an optimal solution to problem (CKP). Let $\mathbf{x}^{\mathrm{seq}}$ be the solution computed by Algorithm 3. Then,*

$$\frac{F(\mathbf{x}^{\mathrm{seq}})}{F(\mathbf{x}^*)} \geq \left(1 - e^{-1}\right).$$

Proof. If $\mathbf{x}^*$ allocates all the capacity to 3 or less indexes, then we must have $\mathbf{x}^{\mathrm{seq}} = \mathbf{x}^*$ by enumeration. Therefore, assume that $\mathbf{x}^*$ allocates capacity to 4 or more firms. Let

$$\bar{S} \equiv \{i : x_i^* = u_i\} = \{i_1, i_2, \ldots, i_{|S|}\} \tag{1}$$

be the set of indexes for which their allocation attains their upper bound u_i. Assume, without loss of generality, that $\bar{S}$ is ordered such that $f_{i_1}(u_{i_1}) \geq f_{i_2}(u_{i_2}) \geq \cdots \geq f_{i_{|S|}}(u_{i_{|S|}})$. Let $Y \subset \bar{S}$ be the set including the first 3 indexes in $\bar{S}$. Namely, $Y = \{i_1, i_2, i_3\}$. Let B_Y be the remaining capacity, after allocating the capacity to the indexes in Y in this order. Define $\hat{\mathbf{x}}$ to be such that,

$$\hat{x}_i = \begin{cases} u_i & \text{if } i \in Y \\ 0 & \text{otherwise.} \end{cases}$$

Let $\mathbf{x}^{\mathrm{seq}}$ be the solution generated by completing $\hat{\mathbf{x}}$ using Algorithm 1, considering all indexes except those in Y, and an initial capacity B_Y. In fact, Algorithm 3 considers $\mathbf{x}^{\mathrm{seq}}$ as one of its candidate solutions, therefore it outputs a solution at least as good. We will show that $\mathbf{x}^{\mathrm{seq}}$ has a worst-case performance guarantee of $\left(1 - e^{-1}\right)$ for problem (CKP).

Define $H(\mathbf{x}) = F(\mathbf{x}) - F(\hat{\mathbf{x}})$. Assume, without loss of generality, that indexes are numbered such that $1 = i_1, 2 = i_2, 3 = i_3$, and then in decreasing order according to λ_i, for each $i \in I \setminus Y$. Let $\mathbf{x}^i$ be such that

$$\mathbf{x}_j^i = \begin{cases} u_j & \text{if } j \leq i \\ 0 & \text{otherwise.} \end{cases}$$

Note that $\mathbf{x}^3 = \hat{\mathbf{x}}$. Additionally, let $B_i = \sum_{j=1}^{i} u_j$. Let $\rho_k = \lambda_{i+1}$, for each $k = B_i + 1, \ldots, B_{i+1}$, $i \geq 3$, and $\rho_k = 0$, for each $k \leq B_Y$. From Proposition 2 it follows that, for every $i \geq 3$

$$\begin{aligned} H(\mathbf{x}^*) &\equiv F(\mathbf{x}^*) - F(\hat{\mathbf{x}}) \\ &\leq \sum_{j=1}^{i} f_j(u_j) + \sum_{j=i+1}^{n} f_j(0) + \lambda_{i+1}(B - B_i) - F(\hat{\mathbf{x}}) \\ &= F(\mathbf{x}^i) + \lambda_{i+1}(B - B_i) - F(\hat{\mathbf{x}}) \\ &= H(\mathbf{x}^i) + \lambda_{i+1}(B - B_i) \\ &\leq H(\mathbf{x}^i) + \lambda_{i+1}(B - B_Y). \end{aligned} \tag{2}$$

Let $\hat{i}$ be the last index with a positive allocation in $\mathbf{x}^{\mathrm{seq}}$. Note that $\hat{i} \geq 4$, therefore $B_{\hat{i}} > B_Y$. Additionally, note that

$$H(\mathbf{x}^i) = \sum_{j=4}^{i} (f_j(u_j) - f_j(0)) = \sum_{j=4}^{i} \lambda_j u_j = \sum_{k=1}^{B_i} \rho_k \quad \forall i \geq 3 \tag{3}$$

where the first equality follows from the definition of $\mathbf{x}^i$ and $H(\mathbf{x})$. The second equality follows from the definition of λ_j. The last equality follows from the definition of ρ_k.

Hence,

$$\min_{i=1,\ldots,\hat{i}-1} \left\{ H(\mathbf{x}^i) + \lambda_{i+1} B_Y \right\} = \min_{i=1,\ldots,\hat{i}} \left\{ \sum_{k=1}^{B_i} \rho_k + \rho_{B_{i+1}} B_Y \right\} \tag{4}$$

$$= \min_{s=1,\ldots,B_{\hat{i}}} \left\{ \sum_{k=1}^{s-1} \rho_k + \rho_s B_Y \right\}. \tag{5}$$

The first equality follows from Eq. (3), and the second equality follows because we are only adding non-negative terms, therefore the minimizer does not change.

It follows that,

$$\begin{aligned} \frac{H(\mathbf{x}^{\hat{i}})}{H(\mathbf{x}^*)} &\geq \frac{H(\mathbf{x}^{\hat{i}})}{\min_{1,\ldots,\hat{i}} \left\{ H(\mathbf{x}^i) + \lambda_{i+1}(B - B_Y) \right\}} \\ &= \frac{\sum_{k=1}^{B_{\hat{i}}} \rho_k}{\min_{s=1,\ldots,B_{\hat{i}}} \left\{ \sum_{k=1}^{s-1} \rho_k + \rho_s B_Y \right\}} \\ &\geq 1 - \left(1 - \frac{1}{B_Y}\right)^{B_{\hat{i}}} \\ &> 1 - e^{\frac{-B_{\hat{i}}}{B_Y}} \\ &> 1 - e^{-1}. \end{aligned} \tag{6}$$

The first inequality follows from Eq. (2). The first equality follows from Eqs. (3) and (4). The second and third inequalities are due to Wolsey, where it is required that both B_Y and $B_{\hat{i}}$ are integers, see [16]. The last inequality follows from $B_{\hat{i}} > B_Y$.

Finally, we conclude that,

$$\begin{aligned} F(\mathbf{x}^{\mathrm{seq}}) &= H(\mathbf{x}^{\mathrm{seq}}) + F(\hat{\mathbf{x}}) \\ &= F(\hat{\mathbf{x}}) + H(\mathbf{x}^{\hat{i}}) - \left(H(\mathbf{x}^{\hat{i}}) - H(\mathbf{x}^{\mathrm{seq}})\right) \\ &= F(\hat{\mathbf{x}}) + H(\mathbf{x}^{\hat{i}}) - \left(F(\mathbf{x}^{\hat{i}}) - F(\mathbf{x}^{\mathrm{seq}})\right) \\ &> F(\hat{\mathbf{x}}) + \left(1 - e^{-1}\right) H(\mathbf{x}^*) - \left(F(\mathbf{x}^{\hat{i}}) - F(\mathbf{x}^{\mathrm{seq}})\right) \\ &= \left(1 - e^{-1}\right) F(\mathbf{x}^*) + e^{-1} F(\hat{\mathbf{x}}) - \left(f_{\hat{i}}(u_{\hat{i}}) - f_{\hat{i}}(x_{\hat{i}})\right) \\ &> \left(1 - e^{-1}\right) F(\mathbf{x}^*) + \frac{1}{3} F(\hat{\mathbf{x}}) - f_{\hat{i}}(u_{\hat{i}}) \\ &> \left(1 - e^{-1}\right) F(\mathbf{x}^*). \end{aligned}$$

The first and third equalities follow from the definition of $H(\mathbf{x})$. The first inequality follows from Eq. (6). The last inequality follows from the definition of Y, and the order of set $\bar{S}$. Specifically,

$$F(\hat{\mathbf{x}}) = \sum_{i=1}^{3} f_i(u_i) + \sum_{i=4}^{n} f_i(0) > \sum_{i=1}^{3} f_i(u_i) > 3 f_{\hat{i}}(u_{\hat{i}})$$

where the last inequality follows from $\frac{1}{3} F(\hat{\mathbf{x}}) - f_{\hat{i}}(u_{\hat{i}}) < 0$. Which follows from $\hat{i} \geq 4$, and the order of the set $\bar{S}$. □

3. Application: allocating technology subsidies to minimize a good's market price

We consider the problem faced by a central planner with the goal of increasing the consumption of a given good, due to the positive societal externalities that it generates. Concrete examples of such goods are vaccines and infectious disease treatments. In order to achieve this goal, she can allocate a given budget in the form of lump sum subsidies among heterogeneous competing firms that produce the good. The introduction of subsidies in the market will induce a demand increase. We assume that the firms do not have the installed capacity to serve all the induced demand, therefore capacity is scarce. We model this by assuming that the firm's marginal costs are increasing. Furthermore, in our model, it is in the best interest of each firm to invest the subsidy to improve the efficiency of its production process, reducing its marginal costs. Therefore, we refer to them as *technology subsidies*. We model the central planner's objective as minimizing the good's market price, therefore increasing its consumption.

Allocating subsidies to producers, rather than to consumers, makes sense when the coordination costs associated with paying to each consumer are larger than the additional benefits generated by impacting consumers directly. This is frequently the case when subsidizing infectious disease treatments in developing countries. One example is the budget of $1.5 billion allocated as lump sum subsidies to producers of the pneumococcal vaccine in 2007, by the Global Alliance for Vaccines and Immunization (GAVI) (see [13]).

To model the market equilibrium, we assume that the market is composed by $n \geq 2$ competing firms. Firms are profit maximizers, and engage in quantity competition, with a linear inverse demand function $P(Q) = \alpha - \frac{1}{\mu}Q$, where $\mu > 0$, and $Q = \sum_{i=1}^{n} q_i$ is the total output produced by all firms at equilibrium. Furthermore, we denote by x_i the technology investment that each firm incurs, in order to become more efficient. Specifically, we assume that firms have a linear marginal cost function of their output q_i, $MC_i = \bar{g}_i(x_i)q_i$, for each i. Note that $\bar{g}_i(x_i) > 0$ is the parameter of the marginal cost function, which captures firm i's efficiency. Specifically, the smaller the value of $\bar{g}_i(x_i)$, the more efficient the firm is. A larger technology investment x_i reduces the value of $\bar{g}_i(x_i)$, at a cost $c_i(x_i)$, with a maximum amount that can be borrowed $\bar{x}_i$. The function $c_i(x_i)$ models the financing cost of firm i. Note that the firms in our model are heterogeneous. The linear demand, and linear marginal costs assumptions are a good approximation, which allow us to obtain closed-form expressions for equilibrium outcomes, and to get insights on the subsidy allocation. Similar assumptions are frequently made by researchers in order to get different insights (see for example [5]).

To simplify the exposition, define $g_i(x_i) \equiv (\bar{g}_i(x_i) + \frac{1}{\mu})$. Adding a constant $\frac{1}{\mu}$ to the marginal cost of each firm will not make a difference in the analysis, therefore we will refer to $g_i(x_i)$ as the marginal cost functions from now on. We make the following assumption on $g_i(x_i)$.

Assumption 1. Assume that $g_i(x_i)$ are continuous, positive and decreasing, for each $i \in \{1, \ldots, n\}$. Moreover, assume $g_i''(x_i)g_i(x_i) \leq 2(g_i'(x_i))^2$ for any technology investment $x_i \geq 0$.

Assumption 1 implies that $\frac{1}{g_i(x_i)}$ is a convex function of the technology investment x_i. To assume that $g_i(x_i)$ is positive decreasing is natural in our setting. On the one hand, it implies that a technology investment x_i cannot increase the cost of production. On the other hand, it implies that no matter how large a technology investment x_i is, the resulting marginal cost cannot be zero or negative. The latter condition in Assumption 1 is not very restrictive. It is satisfied by any concave, positive and decreasing function $g_i(x_i)$. It is also satisfied by convex functions, such as $g_i(x_i) = k_1 e^{-k_2 x_i}$, and $g_i(x_i) = k_1 x_i^{-k_2}$ for any $k_2 \geq 1$.

Similarly, we make the following assumption on the financing cost of firm i.

Assumption 2. Assume that $c_i(x_i)$ are continuous, positive, increasing and convex in $[0, \bar{x}_i]$, for each $i \in \{1, \ldots, n\}$, where $\lim_{x_i \to \bar{x}_i} c_i(x_i) = \infty$.

Moreover, assume $\frac{c_i''(x_i)}{c_i'(x_i)} \geq -\frac{2(g_i'(x_i))^2 - g_i''(x_i)g_i(x_i)}{g_i(x_i)g_i'(x_i)} \geq 0$ for any technology investment $x_i \in [0, \bar{x}_i]$.

Assumption 2 states that the cost of borrowing money increases at an increasing rate for each firm, where $\bar{x}_i$ is the maximum amount that can be borrowed. The latter condition in Assumption 2 is technical, it ensures the existence of the market equilibrium. Intuitively, it states that the financing cost of each firm is convex enough for the profit of each firm to be quasi-concave in the technology investment x_i. Examples of pair of functions $(g_i(x_i), c_i(x_i))$, that satisfy Assumptions 1 and 2, include $\left(k_1 e^{-k_2 x_i}, k_3 e^{k_4 x_i}\right)$, for any $k_4 \geq k_2 \geq 0$, and $(k_1 x_i^{-k_2}, k_3 e^{k_4 x_i})$ for any $k_4 \geq \frac{k_2 - 1}{k_2 + 1} \geq 0$, $k_2 \geq 1$.

Now we characterize the market equilibrium. Let $P(\mathbf{x})$ denote the market price under technology investments $\mathbf{x} \in R^n$. Assuming quantity competition with linear demand allows us to write the following closed form expressions.

Proposition 3. *The equilibrium market price, induced by a technology investment vector* $\mathbf{x}$, *can be written as*

$$P(\mathbf{x}) = \alpha\mu \left(\sum_{i=1}^{n} \frac{1}{g_i(x_i)} + \mu \right)^{-1}. \tag{7}$$

While the market output satisfies $q_i(\mathbf{x}) = P(\mathbf{x})/g_i(x_i)$, *for each firm i.*

Derivations of similar closed form expressions can be found in the transportation and economics literature, therefore the proof is omitted, see for example [11].

From Proposition 3, it follows that the market equilibrium is only a function of the technology investments x_i. Moreover, let $\Pi_i(\mathbf{x})$ be the profit obtained by firm i at the market equilibrium. Note that the revenue of firm i is the market price times its market output. Similarly, the cost of production of firm i is the integral of its linear marginal cost, from zero to its market output. Finally, we also need to consider the financing cost of firm i. It follows that firm i's profit, at the market equilibrium, can be written as $\Pi_i(\mathbf{x}) = P(\mathbf{x})q_i(\mathbf{x}) - g_i(x_i)q_i(\mathbf{x})^2/2 - c_i(x_i)$, where the first term is its revenue, the second term is its production cost, and the third term is its financing cost. Moreover, from Proposition 3 we conclude that $\Pi_i(\mathbf{x}) = P(\mathbf{x})^2/(2g_i(x_i)) - c_i(x_i)$. Assumption 2 allows us to get the following result.

Proposition 4. *The profit of firm i,* $\Pi_i(\mathbf{x})$, *is quasi-concave in* $x_i \in [0, \bar{x}_i]$.

Moreover, the function $P(\mathbf{x})^2/(2g_i(x_i))$ *is quasi-concave in* $x_i \geq 0$, *and attains its maximum when* $q_i(\mathbf{x}) = \alpha\mu/2$.

Proof. We need to check that the derivative of each function changes sign at most once. From Proposition 3 it follows that the profit obtained by firm i at the market equilibrium can be written as

$$\Pi_i(\mathbf{x}) = \frac{\alpha^2\mu^2}{2g_i(x_i)} \left(\sum_{i=1}^{n} \frac{1}{g_i(x_i)} + \mu \right)^{-2} - c_i(x_i).$$

Define $\gamma \equiv \sum_{j=1, j\neq i}^{n} \frac{1}{g_j(x_j)} + \mu$, and note it is constant with respect to x_i. Then, the partial derivative of $\Pi_i(\mathbf{x})$ with respect to x_i is proportional to

$$\gamma - \frac{1}{g_i(x_i)} + \frac{2c_i'(x_i)}{\alpha^2\mu^2 g_i(x_i) g_i'(x_i)} \left(1 + \gamma g_i(x_i)\right). \tag{8}$$

From Assumption 1 the second term in Eq. (8) is increasing in x_i. Therefore, it is enough for the third term in Eq. (8) to be decreasing in x_i, for the partial derivative of $\Pi_i(\mathbf{x})$ with respect to x_i to be increasing. The derivative of the third term in Eq. (8) with respect to x_i is proportional to

$$\begin{aligned}
&\frac{g_i''(x_i)}{g_i'(x_i)} + \frac{g_i'(x_i)}{g_i(x_i)} - \frac{3\alpha\mu g_i'(x_i)}{\sqrt{2}+\alpha\mu g_i(x_i)} - \frac{c_i''(x_i)}{c_i'(x_i)}\\
&\quad < \frac{g_i''(x_i)}{g_i'(x_i)} - \frac{2g_i'(x_i)}{g_i(x_i)} - \frac{c_i''(x_i)}{c_i'(x_i)}\\
&\quad = -\frac{2(g_i'(x_i))^2 - g_i''(x_i)g_i(x_i)}{g_i(x_i)g_i'(x_i)} - \frac{c_i''(x_i)}{c_i'(x_i)}\\
&\quad \le 0.
\end{aligned}$$

The first inequality follows from $g_i'(x_i) < 0$. The last inequality follows from Assumption 2.

Hence, the partial derivative of $\Pi_i(\mathbf{x})$ with respect to x_i is increasing and $\Pi_i(\mathbf{x})$ is quasi-concave in $x_i \in [0, \bar{x}_i]$.

Similarly, the partial derivative of $P(\mathbf{x})^2/(2g_i(x_i))$ is proportional to

$$\left(\sum_{j=1, j\neq i}^{n} \frac{1}{g_j(x_j)} + \mu\right) g_i(x_i) - 1. \tag{9}$$

Which is decreasing in $x_i \ge 0$, therefore the function $P(\mathbf{x})^2/(2g_i(x_i))$ is quasi-concave in $x_i \ge 0$. Moreover, it attains its minimum when we set the expression in Eq. (9) to zero. Namely

$$\left(\sum_{j=1}^{n} \frac{1}{g_j(x_j)} + \mu\right) = \frac{2}{g_i(x_i)}.$$

Or equivalently $q_i(\mathbf{x}) = \frac{\alpha\mu}{2}$. □

Proposition 4 leads to the following theorem.

Theorem 3. *There exists a market equilibrium as a function of the technology investments x_i.*

Moreover, if the financing cost of each firm is zero there is no market equilibrium, as each firm keeps increasing its investment level x_i without bound.

Proof. The strategy set space of each player is $[0, \bar{x}_i]$, a compact and convex set. The profit function $\Pi_i(\mathbf{x})$ is continuous and quasi-concave in $x_i \in [0, \bar{x}_i]$. Hence, the existence of a pure strategy equilibrium follows from the Debreu–Glicksberg–Fan Theorem, see for example [6].

From Proposition 4, it follows that if the financing cost of each firm is zero, then each firm has an incentive to increase its technology investment up to the point where its market output is $q_i(\mathbf{x}) = \alpha\mu/2$. Note that this is the optimal output of a monopolist with no production costs, facing a linear inverse demand function $P(Q) = \alpha - \frac{1}{\mu}Q$. Moreover, this output is unattainable for two or more firms simultaneously. Hence, each firm keeps increasing its investment level x_i without bound. □

Let $\hat{\mathbf{x}}$ be the equilibrium technology investment vector. For simplicity, let us rescale, without loss of generality, the investment levels such that the equilibrium technology investments are denoted by $\hat{x}_i = 0$. We consider the case where the market consumption induced by $\hat{\mathbf{x}}$ is less than what is socially optimal. In this context, the central planner intervenes the market with the objective of minimizing the market price. The central planner invests her budget B, which we assume to be integer, into technology subsidies $x_i \ge 0$ (additional technology investments beyond the equilibrium levels), for each firm i. Note that from Theorem 3 it follows that it is in the best interest of each firm to invest the technology subsidy in becoming more efficient, as this extra technology investment has no cost. We consider the case where the central planner has an integer upper bound, denoted u_i, on the amount of money that she can allocate to each firm i. These upper bounds are motivated by fairness constraints. From the closed form expression given in Eq. (7), it follows that the problem faced by the central planner can be written as,

$$\begin{aligned}
&\min_{\mathbf{x}} \quad P(\mathbf{x}) = \alpha\mu\left(\sum_{i=1}^{n} \frac{1}{g_i(x_i)} + \mu\right)^{-1}\\
\text{(TSAP)}\quad &\text{s.t.} \quad \sum_{i=1}^{n} x_i \le B\\
&\qquad\quad 0 \le x_i \le u_i \;\forall\, i.
\end{aligned}$$

From Eq. (7) it follows that, in order to minimize $P(x)$, we can equivalently maximize the convex function $\sum_{i=1}^{n} \frac{1}{g(x_i)}$ over a polyhedron. It follows that, in the absence of upper bounds u_i on the amount of money allocated to each firm i, in an optimal solution the whole budget B would be allocated to only one firm. However, this type of solution would increase the market share of the selected firm, and decrease everyone else's, resulting in a highly concentrated market. Recognizing that allocating the whole budget to only one firm can be impractical, it is natural to consider upper bounds on the technology subsidy that can be allocated to each firm.

By defining the convex function $f_i(x_i) \equiv \frac{1}{g_i(x_i)}$, for each i, it follows that the central planner's problem (TSAP) is equivalent to our continuous knapsack problem with separable convex utility functions (CKP). Moreover, any α-approximation algorithm for problem (CKP) leads to a $\frac{1}{\alpha}$-approximation algorithm for problem (TSAP) (note that problem (TSAP) is a minimization problem, while problem (CKP) is a maximization problem). Specifically,

$$\frac{P(\mathbf{x}^{\text{alg}})}{P(\mathbf{x}^*)} = \frac{F(\mathbf{x}^*) + \mu}{F(\mathbf{x}^{\text{alg}}) + \mu} \le \frac{F(\mathbf{x}^*)}{F(\mathbf{x}^{\text{alg}})} \le \frac{1}{\alpha},$$

where the equality follows from Eq. (7) and $F(x) = \sum_{i=1}^{n} f_i(x_i)$. The first inequality follows from $\mu > 0$.

Therefore, the results from previous sections suggest that simple subsidy allocation policies have a good performance guarantee for problem (TSAP). In particular, Theorems 1 and 2 show that simple ideas, like allocating the subsidies greedily to the firms that can increase their efficiency *faster* (Algorithms 1 and 3), or allocating the subsidies greedily to the firms that can increase their efficiency *the most* (Algorithm 2), have a guaranteed performance for this model.

Acknowledgments

We would like to thank the Area Editor, the Associate Editor and the anonymous referee for their insightful comments that have helped us improve the paper, both in terms of content as well as exposition. The research of the first author is partially supported by NSF grant CMMI-0846554 (CAREER Award) and by AFOSR award FA9550-11-1-0150. The research of the second author was partially supported by NSF grants CMMI-1162034, CMMI-0758061, CMMI-0824674 and the MITEI Energy Initiative Seed Fund.

References

[1] H.P. Benson, Concave minimization: theory, applications and algorithms, in: R. Horst, P.M. Pardalos (Eds.), Handbook of Global Optimization, Kluwer, 1995, pp. 43–148.
[2] S. Boyd, L. Vandenberghe, Convex Optimization, Cambridge, 2007, pp. 215–288.
[3] K.M. Bretthauer, B. Shetty, The nonlinear knapsack problem—algorithms and applications, European J. Oper. Res. 92 (2002) 459–472.
[4] G.J. Burke, J. Geunes, H.E. Romeijn, A. Vakharia, Allocating procurement to capacitated suppliers with concave quantity discounts, Oper. Res. Lett. 36 (2008) 103–109.
[5] S. Deo, C.J. Corbett, Cournot competition under yield uncertainty: the case of the US influenza vaccine market, Manuf. Serv. Oper. Manage. 11 (2009) 563–576.

[6] D. Fudenberg, J. Tirole, Game Theory, MIT Press, Cambridge, MA, 1991.
[7] N. Halman, D. Klabjan, C. Li, J. Orlin, D. Simchi-Levi, Fully polynomial time approximation schemes for stochastic dynamic programs, in: Proceedings of the Nineteenth Annual ACM–SIAM Symposium on Discrete Algorithms, Society for Industrial and Applied Mathematics, 2008, pp. 700–709.
[8] R.M. Karp, Reducibility among combinatorial problems, in: R.E. Miller, J.W. Thatcher (Eds.), Complexity of Computer Computations, Plenum Press, New York, 1972, pp. 85–103.
[9] S. Khuller, A. Moss, J. Naor, The budgeted maximum coverage problem, Inform. Process. Lett. 70 (1999) 39–45.
[10] J.J. More, S.A. Vavasis, On the solution of concave knapsack problems, Math. Program. 49 (1991) 397–411.
[11] A. Nagurney, Network Economics: A Variational Inequality Approach, Kluwer Academic Publishers, Cambridge, MA, 1999.
[12] S. Sahni, Computationally related problems, SIAM J. Comput. 3 (1974) 262–279.
[13] C.M. Snyder, W. Begor, E.R. Berndt, Economic perspectives on the advance market commitment for pneumococcal vaccines, Health Aff. 30 (2011) 1508–1517.
[14] M. Sviridenko, A note on maximizing a submodular set function subject to a knapsack constraint, Oper. Res. Lett. 32 (2004) 41–43.
[15] D.P. Williamson, D.B. Shmoys, The Design of Approximation Algorithms, Cambridge University Press, New York, NY, 2011.
[16] L.A. Wolsey, Maximising real-valued submodular functions: primal and dual heuristics for location problems, Math. Oper. Res. 7 (1982) 410–425.