Centrum voor Wiskunde en Informatica
Centre for Mathematics and Computer Science

S.A. van de Geer, L. Stougie

# On a rate of convergence of the multiknapsack value function

Department of Mathematical Statistics Report MS-R8812 July

# On a Rate of Convergence of the Multiknapsack Value Function

Sara van de Geer
*Centre for Mathematics and Computer Science*
*P.O. Box 4079, 1009 AB Amsterdam, The Netherlands*

Leen Stougie
*University of Amsterdam*
*Jodenbreestraat 23, 1011 NH Amsterdam*
*The Netherlands*

In [MEANTI et al. 1988] an almost sure asymptotic characterization has been derived for the optimal solution of the knapsack capacities, when the profit and requirement coefficients of items to be selected from are random variables. In this paper we establish a rate of convergence for this process using results from the theory of empirical processes.

*Keywords and phrases:* Multiknapsack value function, Lagrangean relaxation, asymptotic characterization, uniform strong law of large numbers, uniform law of the iterated logarithm, Vapnik-Chervonenkis class.
*1980 Mathematical Subject Classification:* 60G50, 90C10.

## 1. INTRODUCTION

In [MEANTI et al. 1988] the optimal value of the *multiknapsack problem* is studied as a function of the knapsack capacities. They consider the following model. Let $a_{ij}$ be the *amount of space* in the $i$-th knapsack ($i=1,..,m$) *required* by the $j$-th item ($j=1,...,n$). Item $j$ yields a *profit* $c_j$ ($j=1,...,n$) upon inclusion. The $i$-th knapsack has *capacity* $b_i$ ($i=1,...,m$). The multiknapsack problem is formulated as:

$$\begin{aligned} &\max \sum_{j=1}^{n} c_j\, x_j \\ &\text{s.t.} \sum_{j=1}^{n} a_{ij}\, x_j \leq b_i \quad (i=1,...,m) \qquad \text{(MK)} \\ &\qquad x_j \in \{0,1\} \quad (j=1,...,n). \end{aligned}$$

MEANTI et al. have shown that if the coefficients $c_j$ and $a_{ij}$ ($j=1,...,n$, $i=1,...,m$) are generated by an appropriate random mechanism, then the sequence of optimal values of (MK), properly normalized, converges *with probability one* (wp1) to a function of the $b_i$'s, as $n$ goes to infinity and $m$ remains fixed. A crucial step in their proof of this result is the derivation of a *uniform strong law of large numbers*, using theory of convergence of convex functions.

We will show in this paper that results from *empirical process theory* can be applied to reprove this result. More interestingly, the application of empirical process theory allows for a rather straightforward derivation of a *rate of convergence* by establishing a *law of the iterated logarithm.* These results are presented in Section 3.

First, in Section 2 we give an outline of [MEANTI at al. 1988] and indicate where our results fit in.

In Section 4 we discuss the interest of such results and the role that application of empirical process theory may play in this field of research.

Report MS-R8812
Centre for Mathematics and Computer Science
P.O. Box 4079, 1009 AB Amsterdam, The Netherlands

## 2. Convergence of the multiknapsack value function.

Meanti et al. assume that the profit coefficients $c_j$, $j=1,...,n$, are i.i.d. random variables with finite expectation, and that the vectors of requirement coefficients $a_j=(a_{1j},\ldots,a_{mj})^T$, $j=1,...,n$, are i.i.d. random vectors with finite expectations. The profit coefficients and requirement coefficients are independent from each other. Let $b_i=n\beta_i$, $i=1,...,m$, for $\beta=(\beta_1,\ldots,\beta_m)^T\in V:=\{\beta: 0\leq\beta_i\leq Ea_{i1}, i=1,...,m\}$. The asymptotic behaviour of the optimal value $z_n^I$ of (MK) is established as a function of $\beta$.

For any nonnegative vector of multipliers $\lambda=(\lambda_1,\ldots,\lambda_m)^T$ the optimal value of the *Lagrangean relaxation* of the *linear programming (continuous) relaxation* of (MK) is defined as

$$w_n(\lambda)=\max\{\sum_{i=1}^m\lambda_ib_i+\sum_{j=1}^n(c_j-\sum_{i=1}^m\lambda_ia_{ij})x_j \mid 0\leq x_j\leq 1, j=1,...,n\}$$

$$=\sum_{i=1}^m\lambda_ib_i+\Sigma_{j=1}^n(c_j-\sum_{i=1}^m\lambda_ia_{ij})\,x_j^L(\lambda),$$

where

$$x_j^L(\lambda)=\begin{cases}1 & \text{if } c_j-\sum_{i=1}^m\lambda_ia_{ij}>0\\ 0 & \text{otherwise}\end{cases}\quad, j=1,\ldots,n.$$

Define

$$L_n(\lambda):=\frac{1}{n}w_n(\lambda)=\lambda^T\beta+\sum_{j=1}^n(c_j-\lambda^Ta_j)x_j^L(\lambda),$$

and let $\lambda_n^*$ be a vector such that $L_n(\lambda_n^*)=\min_{\lambda\geq 0}L_n(\lambda)$. Then, Meanti et al. show that

$$L_n(\lambda_n^*)-\frac{m}{n}(\max_{j=1,...,n}c_j)\leq\frac{1}{n}z_n^I\leq L_n(\lambda_n^*)\quad \text{wp1.} \tag{2.1}$$

Let the function $L(\lambda)$ be defined as

$$L(\lambda):=\lambda^T\beta+E(c_1-\lambda^Ta_1)x_1^L(\lambda)$$

and let $\lambda^*$ be a minimizer of $L(\lambda)$. Theorem 3.1. in [Meanti et al. 1988] states that

$$\lim_{n\to\infty}|L_n(\lambda_n^*)-L(\lambda^*)|=0\quad \text{wp1.} \tag{2.2}$$

This, together with (2.1), implies their main result:

$$\lim_{n\to\infty}|\frac{1}{n}z_n^I-L(\lambda^*)|=0\quad \text{wp1.}$$

To prove result (2.2) they show that the strong law of large numbers, which implies that

$$\lim_{n\to\infty}|L_n(\lambda)-L(\lambda)|=0\quad \text{wp1,}$$

holds uniformly over all $\lambda$ in a *compact set* $S\subset\mathbb{R}^m$, using the *convexity* of the functions $L_n(\lambda)$ and $L(\lambda)$. From this (2.2) follows almost immediately, (cf. [Meanti et al. 1988, Proof of Theorem 3.1]). In the following section we show that results from empirical process theory can be applied to reprove the uniform strong law of large numbers and, moreover, to establish a rate of convergence. Specifically, we will present a uniform law of the iterated logarithm that yields

$$(\frac{n}{\log\log n})^{1/2}|L_n(\lambda_n^*)-L(\lambda^*)|=O(1)\quad \text{wp1.}$$

3. RATE OF CONVERGENCE.

Consider the functions $f_\lambda : \mathbf{R}^{m+1} \to \mathbf{R}$, for $\lambda \geqslant 0$, defined as

$$f_\lambda(c,a) = \begin{cases} \lambda^T\beta + (c-\lambda^T a)1_{\{(c,a):c>\lambda^T a\}}(c,a) & , c \geqslant 0,\ a \geqslant 0 \\ 0 & , \text{otherwise} \end{cases},$$

where $1_{\{(c,a):c>\lambda^T a\}}(c,a)$ is the indicator function of the set $\{(c,a):c>\lambda^T a\}$. Then $\mathbf{L}_n(\lambda)$ is the *mean value* of $f_\lambda(c,a)$ over $n$ independent observations $(\mathbf{c}_1,\mathbf{a}_1), \ldots, (\mathbf{c}_n, \mathbf{a}_n)$:

$$\mathbf{L}_n(\lambda) = \frac{1}{n}\sum_{j=1}^{n} f_\lambda(\mathbf{c}_j,\mathbf{a}_j),$$

and $L(\lambda)$ is the *expectation* of $f_\lambda(\mathbf{c},\mathbf{a})$:

$$L(\lambda) = Ef_\lambda(\mathbf{c}_1,\mathbf{a}_1).$$

Let $\mathcal{F}$ be the class of functions $f_\lambda$ made up by all possible vectors $\lambda \geqslant 0$:

$$\mathcal{F} := \{f_\lambda : \lambda \geqslant 0\}.$$

The *graph* of a function $g : \mathbf{R}^d \to \mathbf{R}$ is formulated as

$$\operatorname{graph} g = \{(t,x) \in \mathbf{R}^{d+1} : 0 \leqslant t \leqslant g(x) \vee g(x) \leqslant t \leqslant 0\}.$$

We will present some concepts and results from empirical process theory and show that the class of graphs of the functions in $\mathcal{F}$ has properties that allow direct application of these results.

DEFINITION 1. Let $\mathcal{D}$ be a class of subsets of a space $X$. For $x_1, x_2, \ldots, x_n \in X$ define

$$\Delta^{\mathcal{D}}(x_1, \ldots, x_n) := \operatorname{card}\{D \cap \{x_1, \ldots, x_n\} : D \in \mathcal{D}\}$$

and

$$m^{\mathcal{D}}(n) := \sup\{\Delta^{\mathcal{D}}(x_1, \ldots, x_n) : x_1, \ldots, x_n \in X\}.$$

Note that $m^{\mathcal{D}}(n) \leqslant 2^n$. The class $\mathcal{D}$ is called a *Vapnik-Chervonenkis class* if $m^{\mathcal{D}}(n) < 2^n$ for some $n \geqslant 1$ (cf. [VAPNIK & CHERVONENKIS 1971]).

For classes of functions we have a similar definition based on their graphs.

DEFINITION 2. A class $\mathcal{G}$ of real-valued functions is called a *Vapnik-Chervonenkis graph class* if the graphs of the functions in $\mathcal{G}$ form a Vapnik-Chervonenkis class.

The following theorem from [ALEXANDER 1984] establishes a uniform law of the iterated logarithm for a Vapnik-Chervonenkis graph class of functions.

THEOREM 3.1. *Let* $\mathbf{x}_1, \ldots, \mathbf{x}_n$ *be sequence of i.i.d. random variables taking values in a space* $(X, \mathcal{Q})$ *and let* $\mathcal{G}$ *be a class of measurable real valued functions on* $X$*, such that*
*(i)* $\mathcal{G}$ *is a Vapnik-Chervonenkis graph class and*
*(ii) the functions in* $\mathcal{G}$ *are uniformly bounded.*
*Then, modulo measurability,*

$$\sup_{f \in \mathcal{G}} \left(\frac{n}{\log\log n}\right)^{1/2} \left|\frac{1}{n}\sum_{j=1}^{n} g(\mathbf{x}_j) - Eg(\mathbf{x}_1)\right| = O(1) \text{ wp1.} \qquad \square$$

For application of this theorem in our analysis we have to verify the two conditions (i) and (ii) for the class of functions $\{f_\lambda : \lambda \geqslant 0\}$. To show that the first condition is satisfied is rather straightforward, given the theory on Vapnik-Chervonenkis classes. It is known that a class $\mathcal{D}$ of halfspaces in $\mathbf{R}^{d+1}$ say, i.e. $\mathcal{D} = \{\{x : \theta^T x \geqslant 0\}, \theta \in \mathbf{R}^{d+1}\}$ is a Vapnik-Chervonenkis class (see e.g. [DUDLEY 1984]).

Moreover, the Vapnik-Chervonenkis property is preserved under taking finite unions and intersections [POLLARD 1984, page 18]. Now, the graph of a function $f_\lambda$ is

$$\{\{0 \leqslant t \leqslant \lambda^T \beta\} \cap \{c \leqslant \lambda^T a\} \cap \{c \geqslant 0,\ a \geqslant 0\}\} \cup$$
$$\{\{0 \leqslant t \leqslant \lambda^T \beta + c - \lambda^T a\} \cap \{c > \lambda^T a\} \cap \{c \geqslant 0,\ a \geqslant 0\}\} \cup$$
$$\{\{t = 0\} \cap \{c < 0,\ a < 0\}\},$$

which is clearly a finite union of finite intersections of halfspaces in $\mathbf{R}^{m+2}$.

As for the second condition of Theorem 3.1, we notice that, under the assumption of bounded support of $\mathbf{c}$ and $\mathbf{a}$, the functions $f_\lambda$ are uniformly bounded only for $\lambda$ in a bounded set. In [MEANTI et al. 1988, Lemma 3.1], it is shown that the interesting values of $\lambda$, i.e., those values that are candidates for minimizing $\mathbf{L}_n(\lambda)$ and $L(\lambda)$, are in the set $S := \{\lambda : \lambda^T \beta \leqslant E\mathbf{c}_1 + 1,\ \lambda \geqslant 0\}$. Therefore, we arrive at the following lemma.

LEMMA 3.2.

$$\sup_{\lambda \in S} \left(\frac{n}{\log\log n}\right)^{1/2} |\mathbf{L}_n(\lambda) - L(\lambda)| = O(1) \text{ wp1.} \qquad \square$$

From this lemma the following theorem follows easily.

THEOREM 3.3.

$$\left(\frac{n}{\log\log n}\right)^{1/2} |\mathbf{L}_n(\lambda_n^*) - L(\lambda^*)| = O(1) \text{ wp1.}$$

PROOF. Let $\lambda^*$ be a minimum of $L(\lambda)$. If $L(\lambda^*) \leqslant \mathbf{L}_n(\lambda_n^*)$, then $|\mathbf{L}_n(\lambda_n^*) - L(\lambda^*)| \leqslant \mathbf{L}_n(\lambda^*) - L(\lambda^*)$ since $\lambda_n^*$ minimizes $\mathbf{L}_n(\lambda)$. Otherwise, if $L(\lambda^*) > \mathbf{L}_n(\lambda_n^*)$, then $|\mathbf{L}_n(\lambda_n^*) - L(\lambda^*)| \leqslant L(\lambda_n^*) - \mathbf{L}_n(\lambda_n^*)$. Hence,

$$|\mathbf{L}_n(\lambda_n^*) - L(\lambda^*)| \leqslant \sup_{\lambda \in S} |\mathbf{L}_n(\lambda) - L(\lambda)|.$$

This inequality together with Lemma 3.2 completes the proof. $\square$

Regarding (2.1), this theorem leads easily to the rate of convergence of the normalized optimal value $\mathbf{z}_n^I$ of the multiknapsack problem.

THEOREM 3.4.

$$\left(\frac{n}{\log\log n}\right)^{1/2} \left|\frac{1}{n} \mathbf{z}_n^I - L(\lambda^*)\right| = O(1) \text{ wp1.} \qquad \square$$

## 4. POSTLUDE

We observe that there exist other examples of applications of empirical process theory in the research area of probabilistic value analysis of combinatorial problems. In a way analogous to the one in this paper an almost sure characterization of a covering problem has been established [PIERSMA 1987]. In [RHEE 1986] and [RHEE & TALAGRAND 1988] empirical process theory has been used in probabilistic analyses of the optimal value of respectively a Euclidean matching problem and a median location problem. We also mention the probabilistic value analysis of a minimum flowtime scheduling problem [VAN DE GEER et al.].

In view of this, it seems worthwile to try for insights in the specific structures of combinatorial optimization problems that allow for such applications. That wider applicability is possible is intuitively supported by the fact that theorems from empirical process theory, like Theorem 3.1, heavily depend on combinatorial properties of classes of functions.

Finally, we mention the approach of [POLLARD 1985] towards proving central limit theorems, which appears to be well-suited for value functions expressable as empirical process on Vapnik-Chervonenkis graph classes. This might make it feasible, for instance, to study asymptotic distributional behaviour of the optimal value of the Lagrangean relaxation that we considered in this paper.

## REFERENCES


K.S. Alexander, "Probability inequalities for empirical processes and a law of the iterated logarithm", *Annals of Probability* 12 (1984) 1041-1067.

R.M. Dudley, "A course on empirical processes", *Springer Lecture Notes in Math. (Lectures given at Ecole d'Eté de Probabilités de St. Flour 1982)* (1984) 1-142.

S.A. van de Geer, A. Marchetti Spaccamela, W.T. Rhee, L Stougie, "A probabilistic value analysis of the minimum weighted flowtime scheduling problem", Forthcoming.

M. Meanti, A.H.G. Rinnooy Kan, L. Stougie, C. Vercellis, "A probabilistic analysis of the multiknapsack value function", to appear in *Mathematical Programming.*

N. Piersma, "Empirical process theory and combinatorial optimization: a survey", Thesis, University of Amsterdam (Amsterdam 1987).

D. Pollard, *Convergence of stochastic processes* (Springer, New York, 1984).

D. Pollard, "New ways to prove central limit theorems", *Economic Theory* 1 (1985) 295-314.

W.T. Rhee, "On the optimal value of upward matching", Report Colombia State University (Ohio, 1986).

W.T. Rhee, M. Talagrand, "A concentration inequality for the k-median problem", to be published (1988).

V.N. Vapnik, Y.A. Chervonenkis, "On the uniform convergence of relative frequencies of events to their probabilities", *Theory of Probability and its Applications/SIAM* 16 (1971) 264-280.