Optim Lett (2011) 5:153–164
DOI 10.1007/s11590-010-0198-6

ORIGINAL PAPER

On optimality of a polynomial algorithm for random linear multidimensional assignment problem

Pavlo A. Krokhmal

Received: 7 April 2010 / Accepted: 22 April 2010 / Published online: 6 May 2010

Abstract We demonstrate that the linear multidimensional assignment problem with iid random costs is polynomially ε-approximable almost surely (a. s.) via a simple greedy heuristic, for a broad range of probability distributions of the assignment costs. Specifically, conditions on discrete and continuous distributions of the cost coefficients, including distributions with unbounded support, have been established that guarantee convergence to unity in the a. s. sense of the cost ratio between the greedy solution and optimal solution. The corresponding convergence rates have been determined.

Keywords Multidimensional assignment problem · Greedy heuristic · Approximability · Convergence almost surely

1 Introduction

The linear multidimensional assignment problem (LMAP) is a higher-dimensional generalization of the well-known two-dimensional, or linear assignment problem (LAP) (see, e.g., [1,2]). A graph-theoretic formulation of the LAP of cardinality n presents it as finding a minimum-cost perfect matching on a balanced bipartite graph with $2n$ vertices, provided that the cost of matching is defined as the sum of edge costs. Similarly, a d-dimensional LMAP of cardinality n can be formulated as finding a perfect matching on a balanced d-partite d-uniform hypergraph with dn vertices, such that the sum of the costs of hyperedges in the matching is minimized (see, among others, [3,4] for general references on hypergraphs). If the cost of hyperedge $(i_1, \ldots, i_d)$,

This study was supported in part by AFOSR grant FA9550-09-1-0088.

P. A. Krokhmal (✉)
Department of Mechanical and Industrial Engineering, University of Iowa,
3131 Seamans Center, Iowa City, IA 52242, USA
e-mail: krokhmal@engineering.uiowa.edu

where $i_1, \ldots, i_d \in \{1, \ldots, n\}$, is given by $\phi_{i_1 \cdots i_d}$, the mathematical programming formulation of the LMAP of dimensionality d and cardinality n reads as

$$
\begin{aligned}
\min\ & \sum_{i_1=1}^{n} \cdots \sum_{i_d=1}^{n} \phi_{i_1 \cdots i_d} x_{i_1 \cdots i_d} \\
\text{s. t.}\ & \sum_{i_2=1}^{n} \cdots \sum_{i_d=1}^{n} x_{i_1 \cdots i_d} = 1, && i_1 = 1, \ldots, n, \\
& \sum_{i_1=1}^{n} \cdots \sum_{i_{k-1}=1}^{n} \sum_{i_{k+1}=1}^{n} \cdots \sum_{i_d=1}^{n} x_{i_1 \cdots i_d} = 1, && i_k = 1, \ldots, n, \\
& && k = 2, \ldots, d-1, \\
& \sum_{i_1=1}^{n} \cdots \sum_{i_{d-1}=1}^{n} x_{i_1 \cdots i_d} = 1, && i_d = 1, \ldots, n, \\
& x_{i_1 \cdots i_d} \in \{0, 1\}, && i_k = 1, \ldots, n,\ k = 1, \ldots, d,
\end{aligned}
\tag{1}
$$

where $x_{i_1 \cdots i_d} = 1$ if the hyperedge $(i_1, \ldots, i_d)$ is included in the matching, and $x_{i_1 \cdots i_d} = 0$ otherwise. It is easy to see that a feasible solution of (1) is given by n hyperedges $\left(i_1^{(r)}, \ldots, i_d^{(r)}\right), r = 1, \ldots, n$, such that in each dimension k the set $\left\{i_k^{(1)}, \ldots, i_k^{(n)}\right\}$ is a permutation of $\{1, \ldots, n\}$. Hence, problem (1) admits the following geometric interpretation: given a d-dimensional matrix $\Phi = \{\phi_{i_1 \cdots i_d}\} \in \mathbb{R}^{n^d}$, find such a permutation of its "rows" and "columns" in all dimensions that the sum of the diagonal elements is minimized, and thus is also known in the literature as "axial" multidimensional, or multi-index assignment problem.

The LMAP (1) was first introduced by Pierskalla [5], and has found numerous applications in the areas of data association, image recognition, multi-sensor multi-target tracking, peer-to-peer satellite refueling, and so on (for a detailed discussion of the properties and applications of the LMAP, see, for example, [2,6,7]).

In contrast to the LAP that is polynomially solvable, the LMAP with $d \geq 3$ is generally NP-complete, a fact first established by Karp [8] for the three-dimensional ($d = 3$) assignment problem (1). For a discussion of approximation properties of the LMAP (1), see Spieksma [9]; in particular, Crama and Spieksma [10] have shown that even in the case when costs $\phi_{i_1 i_2 i_3}$ of the 3-dimensional LMAP are decomposable, i.e., $\phi_{i_1 i_2 i_3} = d_{i_1 i_2} + d_{i_2 i_3} + d_{i_1 i_3}$, there is no polynomial algorithm that yields an ε-approximate solution for any $\varepsilon > 0$.

Exact and heuristic algorithms for three- and higher-dimensional LMAPs were proposed in [11–17], and others. In particular, a number of heuristics for the three-dimensional LMAP were introduced in Balas and Saltzman [11]; on randomly generated problems, some of these heuristics yielded solutions very close to optimality [17]. These observations find a theoretical substantiation in Kravtsov [18], who has demonstrated that if the assignment costs in (1) are iid random variables from a discrete distribution satisfying certain properties, a simple greedy algorithm produces asymptotically optimal solutions with high probability, when the cardinality n of the LMAP (1) increases infinitely.

In this work, we strengthen and generalize the results of [18], by showing that a greedy algorithm produces ε-approximate solutions of random LMAP *almost surely* (a. s.), or, in other words, that the cost of the greedy solution converges strongly to the optimal cost. Further, we extend the analysis to random LMAPs whose costs are continuously distributed, including distributions with unbounded support sets.

Results concerning asymptotic optimality of heuristic algorithms on randomly generated problems are well known in the context of other hard combinatorial optimization problems, such as the quadratic assignment problem (QAP), which is also known to be NP-complete and non-approximable (see, among others, [2,19], and references therein). In the case of random QAP, asymptotic optimality of heuristic solution methods is a manifestation of the fact that for instances of random QAP large enough, *all* its feasible solutions are asymptotically optimal [20]. Moreover, an entire class of combinatorial optimization problems exists that shares this property with the QAP [21,22]. The LMAP (1), however, does *not* belong to this class; recent investigations of asymptotic behavior of random LMAPs [23,24] entail that only a vanishingly small fraction the feasible set of a random LMAP is ε-optimal.

2 A greedy algorithm for LMAP with discrete iid random costs

Algorithm 1 describes the greedy heuristic for the LMAP that is in the focus of this work. The heuristic starts by finding the smallest hyperedge cost of the LMAP (1), which we denote by $\phi_{i_1^{(1)}\cdots i_d^{(1)}}$, and then removing from the cost matrix Φ the cost elements $\phi_{i_1\cdots i_\ell^{(1)}\cdots i_d}$ for each $\ell \in \{1,\ldots,d\}$, i.e., the costs of the hyperedges that are not feasible with respect to the smallest-cost hyperedge $(i_1^{(1)},\ldots,i_d^{(1)})$. Then, the procedure is repeated, and upon finding the smallest hyperedge cost $\phi_{i_1^{(2)}\cdots i_d^{(2)}}$ in the

Algorithm 1 A greedy heuristic for LMAP (1)

input: Cost matrix $\Phi = \left\{\phi_{i_1\cdots i_d} \mid (i_1,\ldots,i_d) \in \{1,\ldots,n\}^d\right\} \in \mathbb{R}^{n^d}$

initialize: $\widetilde{Z}_n := 0$; for each $\ell \in \{1,\ldots,d\}$ define set $\mathcal{N}_\ell := \{1,\ldots,n\}$

for $k := 1$ **to** n **do**

  define a submatrix $\Phi^{(k)} \in \mathbb{R}^{(n-k+1)^d}$ of the cost matrix Φ as $\Phi^{(k)} := \left\{\phi_{i_1\cdots i_d} \mid (i_1,\ldots,i_d) \in \mathcal{N}_1 \times \cdots \times \mathcal{N}_d\right\}$

  find the smallest element $\phi_{i_1^{(k)}\cdots i_d^{(k)}}$ of the submatrix $\Phi^{(k)}$: $\left(i_1^{(k)},\ldots,i_d^{(k)}\right) \in \arg\min\left\{\phi_{i_1\cdots i_d} \in \Phi^{(k)}\right\}$

  let $\widetilde{Z}_n := \widetilde{Z}_n + \phi_{i_1^{(k)}\cdots i_d^{(k)}}$

  for each $\ell \in \{1,\ldots,d\}$ update the set $\mathcal{N}_\ell := \mathcal{N}_\ell \backslash \{i_\ell^{(k)}\}$

end for

for each $k \in \{1,\ldots,n\}$ define $\tilde{x}_{i_1^{(k)}\cdots i_d^{(k)}} := 1$ and $\tilde{x}_{i_1\cdots i_d} := 0$ for all other $(i_1,\ldots,i_d)$

output: A feasible solution $\tilde{x}_{i_1\cdots i_d}$ of LMAP (1) and its cost $\widetilde{Z}_n$

reduced cost array Φ, the costs of hyperedges that are infeasible with respect to the hyperedges $(i_1^{(k)}, \ldots, i_d^{(k)})$, $k = 1, 2$, are discarded, and so on. After n steps, n costs $\phi_{i_1^{(k)} \cdots i_d^{(k)}}$ are obtained, which have the property that for each $\ell = 1, \ldots, d$ the indices $\{i_\ell^{(1)}, \ldots, i_\ell^{(n)}\}$ are all different, i.e., a feasible solution of (1) is found.

Obviously, the described greedy heuristic for LMAP runs in $O(n^{d+1})$ time. The next lemma provides a foundation for the subsequent probabilistic analysis of the greedy heuristic and is a strengthening of the corresponding result in [18].

Lemma 1 *Consider a set $\mathcal{S}_n$ of cardinality $|\mathcal{S}_n| = \kappa_n$ whose elements are iid random variables distributed uniformly over ρ_n values $a_n < \cdots < b_n$. Assume that ρ_n and κ_n increase with n such that the following series converges:*

$$\sum_n \rho_n \mathrm{e}^{-\frac{\kappa_n}{\rho_n}} < \infty.$$

Then, for n sufficiently large, the set $\mathcal{S}_n$ contains the minimum element a_n almost surely (a.s.)

Proof To verify the statement of the lemma, it is convenient to think about the set $\mathcal{S}_n$ in terms of randomly distributing $|\mathcal{S}_n| = \kappa_n$ different objects into ρ_n boxes. Then, define $\mathcal{A}_n$ as the event that $\mathcal{S}_n$ contains the smallest element a_n, whence

$$\mathsf{P}\{\bar{\mathcal{A}}_n\} \le \mathsf{P}\{\text{at least one box is empty}\} = 1 - \mathsf{P}\{\mathcal{B}_n\}, \tag{2}$$

where $\mathcal{B}_n$ is the event that there are no empty boxes, for which it holds (see, e.g., [25]):

$$\mathsf{P}\{\mathcal{B}_n\} = \sum_{i=0}^{\rho_n} (-1)^i \binom{\rho_n}{i} \left(1 - \frac{i}{\rho_n}\right)^{\kappa_n}.$$

If ρ_n and κ_n increase with n such that the quantity $\lambda_n = \rho_n \mathrm{e}^{-\frac{\kappa_n}{\rho_n}}$ is bounded, that it can be shown that each summand in the above sum is asymptotically equal to $(-\lambda_n)^i / i!$, whereby

$$\mathsf{P}\{\mathcal{B}_n\} \to \mathrm{e}^{-\lambda_n}, \quad n \to \infty.$$

Thus, for n sufficiently large, the probability that the set $\mathcal{S}_n$ does not contain the smallest element a_n of the distribution can be bounded as

$$\mathsf{P}\left\{\bar{\mathcal{A}}_n\right\} \le 1 - \mathsf{P}\{\mathcal{B}_n\} \approx 1 - \mathrm{e}^{-\lambda_n} = \lambda_n + O\left(\lambda_n^2\right).$$

Since by the conditions of the Lemma, $\sum_n \lambda_n < \infty$, from the Borel–Cantelli lemma, we immediately have that $\mathsf{P}\{\bar{\mathcal{A}}_n \text{ i.o.}\} = 0 \quad \Leftrightarrow \quad \mathsf{P}\{\mathcal{A}_n \text{ ev.}\} = 1.$ □

Assuming that the assignment costs of the LMAP (1) are positive, a feasible solution with cost $\widetilde{Z}_n$ is an ε-approximation of the optimal cost Z_n^* of LMAP of cardinality n if it satisfies

$$\widetilde{Z}_n \leq Z_n^*(1+\varepsilon). \tag{3}$$

The next theorem establishes the conditions on the discrete distribution of assignment costs in (1) under which the greedy algorithm delivers an ε-approximation of the optimal cost of the LMAP, or, more precisely, the ratio of the greedy solution cost to the optimal cost approaches unity almost surely.

Theorem 1 *Consider LMAP* (1) *with* $d \geq 3, n \geq 2$, *whose cost coefficients are iid random variables distributed uniformly over* n^α *values*[1] $a_n < \cdots < b_n$, *where* $a_n > 0$ *and* $\alpha > 0$. *Then, there exists a constant* $M > 0$ *such that the greedy algorithm produces a solution with the cost* $\widetilde{Z}_n$, *which for sufficiently large* n *satisfies*

$$\frac{\widetilde{Z}_n}{Z_n^*} - 1 \leq M\left(\frac{b_n}{a_n} - 1\right) n^{\alpha/d-1} \ln^{1/d} n \qquad a.s., \tag{4}$$

where Z_n^* *is the optimal cost of the LMAP.*

Proof From the description of the greedy heuristic, it follows that the cost of the feasible solution can be represented as

$$\widetilde{Z}_n = \sum_{k=1}^{n} \phi_{i_1^{(k)} \cdots i_d^{(k)}} = \sum_{k=1}^{n} \tilde{\phi}_k, \tag{5}$$

where each $\tilde{\phi}_k$ is equal to the smallest element of a submatrix $\Phi^{(k)}$:

$$\tilde{\phi}_k = \min\left\{a \mid a \in \Phi^{(k)}\right\}, \quad k = 1, \ldots, n.$$

In general, the summation in (5) contains terms $\tilde{\phi}_k$ that are either equal to the smallest element a_n of the distribution, or exceed it. Let K_n denote the (random) number of the summands in (5) that are greater than a_n:

$$K_n = \left|\left\{k \mid \tilde{\phi}_k > a_n\right\}\right|. \tag{6}$$

Then, noting that $0 \leq K_n \leq n$, the optimal cost Z_n^* of the LMAP (1) and the cost $\widetilde{Z}_n$ returned by the greedy heuristic can be bounded as

$$na_n \leq Z_n^* \leq \widetilde{Z}_n \leq (n-K_n)a_n + K_n b_n = na_n\left(1 + \frac{K_n}{n}\frac{b_n - a_n}{a_n}\right), \tag{7}$$

[1] Here and in what follows we omit rounding to avoid unnecessary ramifications in exposition.

from which it is easy to see that $\widetilde{Z}_n$ is an ε-approximate solution of (1) by means of the approximation inequality (3) as soon as $\widetilde{Z}_n \le na_n(1+\varepsilon)$. Thus, for some $\varepsilon_n > 0$ consider

$$\mathsf{P}\left\{\widetilde{Z}_n > na_n(1+\varepsilon_n)\right\} \le \mathsf{P}\{K_n > \gamma_n n\},$$

where the inequality follows from (7) provided that γ_n is chosen as

$$\gamma_n = \frac{\varepsilon_n}{b_n/a_n - 1}.$$

Observe that if $K_n > \gamma_n n$ holds, then there exists an integer $\nu \in \{1, \ldots, n\}$ such that $\nu > \gamma_n n$ and the corresponding submatrix of Φ with ν^d elements does not contain elements equal to a_n. Then, from Lemma 1 it follows that for sufficiently large values of n,

$$\mathsf{P}\{K_n > \gamma_n n\} \le \mathsf{P}\left\{ \text{set of size } (\gamma_n n)^d \text{ does not contain } a_n \right\} \le n^\alpha \exp\left\{-\frac{(\gamma_n n)^d}{n^\alpha}\right\}.$$

Choosing the parameter ε_n in the form

$$\varepsilon_n = (2+\alpha)^{1/d}\left(\frac{b_n}{a_n} - 1\right) n^{\alpha/d-1} \ln^{1/d} n,$$

we have that for values of n large enough,

$$\mathsf{P}\{K_n > \gamma_n n\} \le n^{-2},$$

whence expression (4) follows by the Borel–Cantelli lemma. □

Corollary 1.1 *If $\alpha < d$ in (4) and the ratio b_n/a_n satisfies*

$$\frac{b_n}{a_n} = o\left(n^{1-\alpha/d} \ln^{-1/d} n\right), \quad n \gg 1,$$

then for sufficiently large n, the greedy cost $\widetilde{Z}_n$ is an ε-approximation of the optimal cost Z_n^* of random LMAP due to (3) for any $\varepsilon > 0$ almost surely. Put differently, the cost ratio $\widetilde{Z}_n/Z_n^*$ between greedy and optimal solutions converges to unity a. s., with the convergence rate given by (4).

Remark 1.1 The intuition behind Lemma 1 and Theorem 1 is that if the elements of the cost matrix Φ are drawn at random from a sufficiently small set of values, then at each step of the greedy heuristic, the submatrix $\Phi^{(k)}$ will contain the smallest element from that set with sufficiently high probability. This observation can be pressed into service to address the case when the elements of the cost matrix Φ of the LMAP (1) are continuous iid variables, as it is shown next.

3 Greedy heuristic for LMAP with continuous iid costs

In [23,24], it has been demonstrated that if the assignment costs $\phi_{i_1\cdots i_d}$ in LMAP (1) are iid random variables with a continuous distribution F, then asymptotic behavior of the optimal value of random LMAP is controlled by the properties of the distribution F in the vicinity of the left-end point $F^{-1}(0)$ of the support set of the distribution, where

$$F^{-1}(0) = \inf\{t \mid F(t) > 0\}.$$

In view of that, we restrict our discussion to continuous distributions F whose support sets are bounded from above,

$$F^{-1}(1) < +\infty, \quad \text{where} \quad F^{-1}(1) = \sup\{t \mid F(t) < 1\}.$$

The next theorem generalizes the results of the previous section to continuous distributions.

Theorem 2 *Consider LMAP* (1) *with* $d \geq 3, n \geq 2$*, whose cost coefficients are iid random variables with a continuous distribution F that has a bounded support* $[a, b]$*, where* $a > 0$*. Then, for any* $\alpha > 0$*, there exists a constant* $M > 0$ *such that the greedy algorithm produces a solution with cost* $\widetilde{Z}_n$*, which for sufficiently large n satisfies*

$$\frac{\widetilde{Z}_n}{Z_n^*} - 1 \leq \frac{F^{-1}(n^{-\alpha})}{a} - 1 + Mn^{\alpha/d-1}\ln^{1/d} n \qquad a.s., \tag{8}$$

where Z_n^* *is the optimal cost of the LMAP.*

Proof For a continuous distribution F on $[a, b] \subset \mathbb{R}$, define the sequence $\{\delta_n(k)\}$, $k = 0, \ldots, \rho_n$, as

$$\delta_n(0) = 0, \quad \delta_n(k) = -a + F^{-1}\left(\frac{1}{\rho_n} + F(a + \delta_n(k-1))\right), \quad k = 1, \ldots, \rho_n.$$

The intervals $\mathcal{I}_k = (a + \delta_n(k-1), a + \delta_n(k)]$, $k = 1, \ldots, \rho_n$, partition the set $(a, b]$ into ρ_n equiprobable "bins", such that for any F-distributed random variable X

$$\mathsf{P}\{a + \delta_n(k-1) < X \leq a + \delta_n(k)\} = \frac{1}{\rho_n}, \quad k = 1, \ldots, \rho_n.$$

Then, the elements of the cost matrix Φ can be labeled with ρ_n different labels, in accordance to the "bin" $\mathcal{I}_k$ that the corresponding cost element falls into. Obviously, the labels are independently and identically uniformly distributed. Therefore, taking into account that the elements of the cost matrix Φ that fall into bin $\mathcal{I}_1$ are less than or equal to $a + \delta_n(1)$, the cost $\widetilde{Z}_n$ of the greedy solution of the MAP can be bounded as

$$na \leq Z_n^* \leq \widetilde{Z}_n \leq (n - K_n)(a + \delta_n(1)) + K_n b, \tag{9}$$

where K_n is equal to the number of summands in the cost $\widetilde{Z}_n$ of the greedy solution that do not fall into the bin $\mathcal{I}_1 = (a, a + \delta_n(1)]$. Then, for any fixed $\varepsilon_n > 0$ it holds that

$$\begin{aligned}\mathsf{P}\{\widetilde{Z}_n(\Phi) > na(1+\varepsilon_n)\} &\le \mathsf{P}\{(n-K_n)(a+\delta_n(1)) + K_n b > na(1+\varepsilon_n)\}\\ &= \mathsf{P}\left\{K_n > n\gamma_n\right\},\end{aligned}$$

where

$$\gamma_n = \frac{\varepsilon_n - \delta_n(1)/a}{b/a - 1 - \delta_n(1)/a}.$$

Similarly to the arguments of Theorem 1, $K_n > n\gamma_n$ holds provided that there exists $\nu \in \{1, \ldots, n\}$ such that $\nu > n\gamma_n$ and the corresponding submatrix of size ν^d does not contain elements from the interval $\mathcal{I}_1$, whence

$$\mathsf{P}\left\{K_n > \gamma_n n\right\} \le \rho_n \exp\left\{-\frac{(\gamma_n n)^d}{\rho_n}\right\},$$

Let $\rho_n = n^\alpha$ for some $\alpha > 0$, then, choosing ε_n as

$$\varepsilon_n = \frac{F^{-1}(n^{-\alpha})}{a} - 1 + \left(\frac{b}{a} - \frac{F^{-1}(n^{-\alpha})}{a}\right)(\alpha+2)^{1/d} n^{\alpha/d-1} \ln^{1/d} n,$$

and taking into account that $\delta_n(1) = F^{-1}(n^{-\alpha}) - a = o(1), n \gg 1$, we obtain that

$$\mathsf{P}\{K_n > \gamma_n n\} \le \rho_n^{-2/\alpha} = n^{-2},$$

which verifies statement (8) of the Theorem by virtue of the Borel–Cantelli lemma. □

Corollary 2.1 *If $\alpha < d$, it follows from (8) that $\widetilde{Z}_n$ represents an ε-optimal solution of random LMAP in the sense (3) for any $\varepsilon > 0$, and the cost ratio $\widetilde{Z}_n/Z_n^*$ converges to unity a. s. It is natural that the value of the parameter $\alpha \in (0, d)$ in (8) is selected based on the properties of F^{-1} at the origin so as to increase the rate of convergence. In particular, if in some neighborhood of 0 the inverse F^{-1} of the distribution F satisfies for some $\nu > 0$*

$$F^{-1}(u) \le a + Lu^\nu, \quad L > 0,\ u \to 0+,$$

then there exists a constant $M_1 > 0$ such that

$$\frac{\widetilde{Z}_n}{Z_n^*} - 1 \le M_1 n^{(1+\nu d)^{-1}-1} \ln^{1/d} n \qquad a.s.$$

Next, we consider the case of a continuous distribution F with support of the form $(-\infty, b]$, where the following bounds on the optimal cost of random LMAP play a key role. Namely, as shown in [24], the optimal value Z_n^* of random LMAP with iid cost coefficients whose distribution has a support unbounded from below, satisfies for sufficiently large n

$$nF^{-1}\left(\frac{1}{n^{d-1}}\right) \le Z_n^* \le nF^{-1}\left(\frac{3\ln n}{n^{d-1}}\right) \qquad \text{a.s.} \tag{10}$$

Expression (10) entails that when support of F is unbounded from below, $F^{-1}(0) = -\infty$, one has that $Z_n^* < 0$ a.s. for large enough n. Note that in this case the approximation condition (3) takes the form

$$\widetilde{Z}_n \le Z_n^*(1-\varepsilon), \quad \varepsilon > 0. \tag{11}$$

Taking into account (11), the following statement holds regarding the quality of the greedy solution to a random LMAP (1).

Theorem 3 *Consider LMAP (1) with $d \ge 3, n \ge 2$ whose cost coefficients are iid random variables with continuous distribution F such that $F^{-1}(0) = -\infty$, $F^{-1}(1) < \infty$. Then, the greedy algorithm produces a solution with cost $\widetilde{Z}_n$ that for sufficiently large n satisfies*

$$1 - \frac{\widetilde{Z}_n}{Z_n^*} \le \left(\frac{d-1}{n}\right)^{1/d} + \frac{F^{-1}\left(\frac{1}{n^{d-1}}\right)}{F^{-1}\left(\frac{3\ln n}{n^{d-1}}\right)} - 1 \qquad a.s., \tag{12}$$

where Z_n^* is the optimal cost of the LMAP.

Proof Similarly to the proof of Theorem 2, let us partition the semi-infinite support $(-\infty, b]$ of the distribution F into ρ_n "bins" $(\alpha_n(k-1), \alpha_n(k)]$ such that

$$\mathsf{P}\{\alpha_n(k-1) < X \le \alpha_n(k)\} = \frac{1}{\rho_n}, \quad k = 1, \ldots, \rho_n,$$

where X is F-distributed random variable, and $\alpha_n(k)$ is defined as

$$\alpha_n(0) = -\infty, \quad \alpha_n(k) = F^{-1}\left(F(\alpha_n(k-1)) + \frac{1}{\rho_n}\right), \quad k = 1, \ldots, \rho_n. \tag{13}$$

Then, similar arguments allow us to construct an upper bound $\widetilde{\widetilde{Z}}_n$ on the greedy cost $\widetilde{Z}_n$ in the form

$$\widetilde{Z}_n \le \widetilde{\widetilde{Z}}_n \equiv (n - K_n)F^{-1}(\rho_n^{-1}) + K_n F^{-1}(1),$$

where N_k is the number of summands in the greedy cost $\widetilde{Z}_n$ that do not fall into the first "bin" $(-\infty, F^{-1}(\rho_n^{-1})]$.

Next, observe that if the optimal cost Z_n^* of the LMAP can be bounded from below and above, e.g.,

$$\underline{Z}_n \le Z_n^* \le \overline{Z}_n,$$

then, given a fixed $\varepsilon > 0$, the greedy solution cost $\widetilde{Z}_n$ satisfies the approximation inequality (11) as soon as the upper bound $\widetilde{\widetilde{Z}}_n$ satisfies

$$\widetilde{\widetilde{Z}}_n - \underline{Z}_n \le -\varepsilon \overline{Z}_n.$$

In view of the bounds (10) on the optimal cost of random LMAP due to [24] that hold almost surely for large enough n, define

$$\underline{Z}_n = nF^{-1}\left(\frac{1}{3\rho_n \ln n}\right), \quad \overline{Z}_n = nF^{-1}\left(\frac{1}{\rho_n}\right), \quad \text{where} \quad \rho_n = \frac{n^{d-1}}{3\ln n},$$

and for any fixed $\varepsilon_n > 0$ consider the probability

$$\mathsf{P}\left\{\widetilde{\widetilde{Z}}_n(\Phi) - \underline{Z}_n > -\varepsilon_n \overline{Z}_n\right\} = \mathsf{P}\{K_n > n\gamma_n\}, \tag{14}$$

where

$$\gamma_n = \left(F^{-1}\left(\frac{1}{3\rho_n \ln n}\right) - F^{-1}\left(\frac{1}{\rho_n}\right)(1+\varepsilon_n)\right)\left(F^{-1}(1) - F^{-1}\left(\frac{1}{\rho_n}\right)\right)^{-1}.$$

Following the arguments of Theorems 1 and 2, it can be shown that for sufficiently large values of n the probability in (14) satisfies

$$\mathsf{P}\{K_n > n\gamma_n\} \le \rho_n \exp\left\{-\frac{(n\gamma_n)^d}{\rho_n}\right\}. \tag{15}$$

If one selects the parameter ε_n in the form

$$\begin{aligned}\varepsilon_n &= \frac{F^{-1}\left(\frac{1}{3\rho_n \ln n}\right)}{F^{-1}\left(\rho_n^{-1}\right)} + \left(1 - \frac{F^{-1}(1)}{F^{-1}\left(\rho_n^{-1}\right)}\right)\frac{1}{n}\left(3\rho_n \ln \rho_n\right)^{1/d} - 1 \\ &= \left(\frac{d-1}{n}\right)^{1/d}\left(1 - \frac{\ln(3\ln n)}{(d-1)\ln n}\right)^{1/d} + \frac{F^{-1}\left(\frac{1}{3\rho_n \ln n}\right)}{F^{-1}\left(\rho_n^{-1}\right)} - 1,\end{aligned}$$

inequality (15) implies that the probability in (14) is bounded as

$$\mathsf{P}\{K_n > n\gamma_n\} \le \rho_n^{-2} \le n^{-2}$$

for all large enough values of n, thereby verifying the estimate (12) of approximation quality of the greedy solution by means of the Borel–Cantelli lemma. □

Remark 3.1 In the case when the distribution F is such that

$$F^{-1}\left(\frac{1}{n^{d-1}}\right) \Big/ F^{-1}\left(\frac{3\ln n}{n^{d-1}}\right) \to 1, \quad n \to \infty, \tag{16}$$

Theorem 3 asserts that the solution cost produced by the greedy heuristic is an ε-approximation of the optimal cost of random LMAP, for any $\varepsilon > 0$. Condition (16) holds, for instance, for distributions F whose inverse F^{-1} has a logarithmic singularity at the origin, i.e., when the following asymptotic representation holds in the vicinity of 0:

$$F^{-1}(u) \sim -c_0 \ln^\beta \frac{1}{u}, \quad u \to 0+ \quad \text{for some } \; c_0 > 0, \; \beta > 0.$$

Continuous distributions that satisfy this condition and whose support is bounded from above include, for instance, exponential distribution on $(-\infty, 0]$, truncated normal distribution on $(-\infty, b]$:

$$F(t) = \mathrm{e}^t \mathbf{1}_{(-\infty,0]}(t) + \mathbf{1}_{(0,\infty)}(t), \quad F(t) = \frac{\Phi(t)}{\Phi(b)} \mathbf{1}_{(-\infty,b]}(t) + \mathbf{1}_{(b,\infty)}(t),$$

where $\Phi(t)$ is the standard normal distribution function. Observe that in the case when the inverse F^{-1} of the cost distribution F has, for example, a power singularity at the origin, i.e.,

$$F^{-1}(u) \sim -c_0 u^{-\beta}, \quad u \to 0+, \quad c_0, \beta > 0,$$

the ratio in (16) is unbounded in n, hence no statement can be inferred from Theorem 3 regarding the quality of the greedy solution in this case.

Remark 3.2 According to [24], the asymptotic behavior of the optimal cost of random LMAP (1) is determined completely by the properties of the distribution function F in the vicinity of the left-end point of its support. Nevertheless, the requirement of boundedness from above of the distribution's support, $F^{-1}(1) < \infty$, which is imposed in Theorems 1, 2, 3, is essential for estimating the quality of the greedy solution, as it allows one to obtain an upper bound on the cost of the solution produced by the greedy algorithm.

4 Conclusions

We have demonstrated that a random Linear Multidimensional Assignment Problem, whose assignment costs are iid random variables is polynomially ε-approximable almost surely using a simple greedy heuristic, for a broad range of probability distributions of the assignment costs. Specifically, conditions on discrete and continuous distributions, including distributions with support unbounded from below, have been established that guarantee convergence to unity in the a. s. sense of the cost ratio between the greedy solution and optimal solution. The corresponding convergence rates have also been established.

References

1. Papadimitrou, C.H., Steiglitz, K.: Combinatorial Optimization: Algorithms and Complexity. Dover, New York (1998)
2. Burkard, R.E., Dell'Amico, M., Martello, S.: Assignment Problems. SIAM, Philadelphia (2009)
3. Berge, C.: Hypergraphs: Combinatorics of Finite Sets. North-Holland, Amsterdam (1989)
4. Bollobás, B.: Modern Graph Theory. Springer, New York (1998)
5. Pierskalla, W.: The multidimensional assignment problem. Oper. Res. **16**(2), 422–431 (1968)
6. Burkard, R.E., Çela, E.: Quadratic and three-dimensional assignments. In: Dell'Amico, M. (ed.) Annotated Bibliographies in Combinatorial Optimization, pp. 373–391. Wiley, Chichester (1999)
7. Burkard, R.E.: Selected topics on assignment problems. Discret. Appl. Math. **123**(1–3), 257–302 (2002)
8. Karp, R.M.: Reducibility among combinatorial problems. In: Miller, R.E., Thatcher, J.W. (eds.) Complexity of Computer Computations, pp. 85–103. Plenum Press, New York (1972)
9. Spieksma, F.C.R.: Multi index assignment problems: complexity, approximation, applications. In: Pitsoulis, L., Pardalos, P.M. (eds.) Nonlinear Assignment Problems: Algorithms and Applications, pp. 1–12. Kluwer, Dordrecht (2000)
10. Crama, Y., Spieksma, F.C.R.: Approximation algorithms for three-dimensional assignment problems with triangle inequalities. Eur. J. Oper. Res. **60**(3), 273–279 (1992)
11. Balas, E., Saltzman, M.J.: An algorithm for the three-index assignment problem. Oper. Res. **39**(1), 150–161 (1991)
12. Poore, A.B.: Multidimensional assignment formulation of data association problems arising from multitarget and multisensor tracking. Comput. Optim. Appl. **3**(1), 27–54 (1994)
13. Poore, A.B.: A numerical study of some data association problems arising in multitarget tracking. In: Hager, W.W., et al. (eds.) Large Scale Optimization: State of the Art, pp. 339–361. Kluwer, Dordrecht (1994)
14. Poore, A.B., Robertson, A.J.: A new Lagrangian relaxation based algorithm for a class of multidimensional assignment problems. Comput. Optim. Appl. **8**(2), 129–150 (1997)
15. Murphey, R., Pardalos, P., Pitsoulis, L.: A greedy randomized adaptive search procedure for the multitarget multisensor tracking problem. In: Pardalos, P.M., Du, D.-Z. (eds.) Network Design: Connectivity and Facilities Location, volume 40 of DIMACS Series in Discrete Mathematics and Theoretical Computer Science, pp. 277–302, American Mathematical Society (1998)
16. Murphey, R., Pardalos, P., Pitsoulis, L.: A parallel GRASP for the data association multidimensional assignment problem. In: Parallel Processing of Discrete Problems. The IMA Volumes in Mathematics and its Applications, vol. 106, pp. 159–180. Springer, Berlin (1998)
17. Andrijich, S.M., Caccetta, L.: Solving the multisensor data association problem. Nonlinear Anal. **47**(8), 5525–5536 (2001)
18. Kravtsov, V.M.: Polynomial algorithms for finding the asymptotically optimum plan of the multiindex axial assignment problem. Cybern. Syst. Anal. **41**(6), 940–944 (2005)
19. Krokhmal, P., Pardalos, P.M.: Random assignment problems. Eur. J. Oper. Res. **194**(1), 1–17 (2009)
20. Burkard, R.E., Fincke, U.: The asymptotic probabilistic behavior of quadratic sum assignment problems. Zeitschrift für Oper. Res. **27**, 73–81 (1982)
21. Burkard, R.E., Fincke, U.: Probabilistic asymptotic properties of some combinatorial optimization problems. Discret. Appl. Math. **12**(1), 21–29 (1985)
22. Szpankowski, W.: Combinatorial optimization problems for which almost every algorithm is asymptotically optimal!. Optimization **33**(4), 359–367 (1995)
23. Krokhmal, P., Grundel, D., Pardalos, P.: Asymptotic behavior of the expected optimal value of the multidimensional assignment problem. Math. Program. **109**(2–3), 525–551 (2007)
24. Krokhmal, P.A., Pardalos, P.M.: Limiting optimal values and convergence rates in some combinatorial optimization problems on hypergraph matchings. In review (2009)
25. Feller, W.: An Introduction to Probability Theory and Its Applications, vol. 1, 3rd edn. Wiley, New York (1968)