Combinatorics, Probability and Computing (2004) 13, 103–113. © 2004 Cambridge University Press
DOI: 10.1017/S0963548303005819 Printed in the United Kingdom

Random Assignment with Integer Costs

ROBERT PARVIAINEN

Dept. of Mathematics, Uppsala University, PO Box 480, S-75106 Uppsala, Sweden
(e-mail: robert@math.uu.se)

Received 21 June 2001; revised 16 December 2002

The random assignment problem is to minimize the cost of an assignment in an $n \times n$ matrix of random costs. In this paper we study the problem for some integer-valued cost distributions. We consider both uniform distributions on $1, 2, \ldots, m$, for $m = n$ or n^2, and random permutations of $1, 2, \ldots, n$ for each row, or of $1, 2, \ldots, n^2$ for the whole matrix. We find the limit of the expected cost for the 'n^2' cases, and prove bounds for the 'n' cases. This is done by simple coupling arguments together with recent results of Aldous for the continuous case. We also present a simulation study of these cases.

1. Introduction

In the assignment problem we choose n elements from an $n \times n$ matrix $\boldsymbol{C}$ of costs, one element from each row and each column, in such a way that the total cost is minimized. In other words, we are looking for a permutation π that minimizes

$$Z = \sum_{i=1}^{n} C_{i\pi(i)}.$$

If we let the elements of $\boldsymbol{C}$ be random variables, we have the random assignment problem. Traditionally, the random costs have been independent, identically distributed, with the exponential or the uniform distribution.

When the costs are i.i.d. exponential (mean 1) there are strong conjectures for the more general case of k-assignment from an $m \times n$ cost matrix. Let $Z^*(k, m, n)$ denote the minimal cost. Mézard and Parisi [13, 14] conjectured that

$$\lim_{n\to\infty} E(Z^*(n, n, n)) = \pi^2/6.$$

This was proved by Aldous [2]. Parisi [16] has also conjectured that

$$E(Z^*(n, n, n)) = \sum_{i=1}^{n} \frac{1}{i^2},$$

https://doi.org/10.1017/S0963548303005819 Published online by Cambridge University Press

which was improved by Coppersmith and Sorkin [4] to

$$E(Z^*(k,m,n)) = \sum_{i+j<k} \frac{1}{(m-i)(n-j)}.$$

The last conjecture was proved by Alm and Sorkin [3] for $k \leqslant 4, k = m = 5,$ and $k = m = n = 6$. Linusson and Wästlund [12] extended this to $k \leqslant 6,$ and $k = m = n = 7$.

1.1. Discrete variants

We will study four discrete variants of the random assignment problem.

Case I. Each row in $\boldsymbol{C}$ is an independent random permutation of $\{1, 2, \ldots, n\}$, chosen uniformly from the set of all permutations.

Case II. Each element in $\boldsymbol{C}$ is an independent random number, chosen uniformly from $\{1, 2, \ldots, n\}$.

Case III. $\boldsymbol{C}$ is a random permutation of $\{1, 2, \ldots, n^2\}$ chosen uniformly.

Case IV. Each element in $\boldsymbol{C}$ is an independent random number, chosen uniformly from $\{1, 2, \ldots, n^2\}$.

In the 'n' cases I and II we normalize by n, and in the 'n^2' cases III and IV by n^2, thus considering the problem of minimizing

$$Z = \frac{1}{n} \sum_i C_{i\pi(i)} \text{ or } Z = \frac{1}{n^2} \sum_i C_{i\pi(i)}.$$

The (random) minimal costs will be denoted by $Z_i^*(n)$ for the four discrete cases, and by $Z_c^*(n)$ in the case of continuous costs.

In [2], Aldous proves the following theorems, valid for any nonnegative continuous distribution, such that the density of the independent costs have value 1 at 0. Let π denote the permutation giving an optimal assignment.

Theorem 1.1.

$$\lim_{n\to\infty} EZ_c^*(n) = \frac{\pi^2}{6}.$$

Theorem 1.2. *$nC_{i\pi(i)}$ converges in distribution. The limit distribution has density*

$$h(x) = \frac{e^{-x}(e^{-x} - 1 + x)}{(1 - e^{-x})^2}, \quad 0 \leqslant x < \infty.$$

Theorem 1.3.

$$\lim_{n\to\infty} P(C_{i\pi(i)} \textit{ is the kth-smallest element of the ith row in } \mathbf{C}) = 2^{-k}.$$

Remark. In a simulation study in [15], Olin noted that, even for dimensions as small as $n = 50$, the row rank distribution is surprisingly close to the above.

https://doi.org/10.1017/S0963548303005819 Published online by Cambridge University Press

2. Coupling arguments

In this section we will prove the following theorem.

Theorem 2.1. *Let* $EZ_i^* = \lim_{n\to\infty} EZ_i^*(n)$. *Then*

$$\frac{\pi^2}{6} \leqslant EZ_1^* \leqslant 2, \quad \frac{\pi^2}{6} + \frac{12}{24} \leqslant EZ_2^* < \frac{\pi^2}{6} + \frac{13}{24},$$
$$EZ_3^* = \frac{\pi^2}{6}, \quad EZ_4^* = \frac{\pi^2}{6}.$$

The idea is to compare a discrete case of the problem with the case of (continuous) uniform costs. We want to generate matrices for both cases simultaneously, such that an optimal assignment for one matrix is close to optimal for the other.

When we say that π is optimal for the matrix $\boldsymbol{C}$, we mean that π is a permutation giving an optimal assignment for the random assignment problem, with cost matrix $\boldsymbol{C}$.

2.1. Case I

Let $\boldsymbol{U}$ be an $n \times n$ matrix of i.i.d. U$(0,n)$ (uniform on $(0,n)$) random variables. It will be convenient to denote the rows of $\boldsymbol{U}$ by $\boldsymbol{U}^{(i)}$. We want to use $\boldsymbol{U}$ to get an independent random permutation for each row. To achieve this coupling, we can use the ranks within each row of the matrix $\boldsymbol{U}$. If we let

$$P_i(j) = \text{rank}\left(U_j^{(i)}\right),$$

each P_i will be an independent random permutation, chosen uniformly from the set of all permutations. By Theorem 1.3 we have, for i in the optimal assignment for $\boldsymbol{U}$,

$$\lim_{n\to\infty} P\left(\text{rank}\left(U_j^{(i)}\right) = k\right) = 2^{-k}.$$

If π is the optimal assignment for $\boldsymbol{U}$, this gives

$$\lim_{n\to\infty} EZ_1^*(n) \leqslant \lim_{n\to\infty} E\left(\frac{1}{n}\sum_{i=1}^{n} P_i(\pi(i))\right) = \lim_{n\to\infty} E\left(\text{rank}\left(U_j^{(\pi(i))}\right)\right) = 2.$$

Remark. That $\lim_{n\to\infty} E(\text{rank}\,(U_j^{(\pi(i))})) = 2$ is not *a priori* clear, but it follows from the uniform integrability of $Z_1^*(n)$, which in turn follows from the uniform integrability of $Z_c^*(n)$ (Aldous [1]).

For a lower bound, assume that, for $1 \leqslant i \leqslant n$, P_i is a random permutation of $\{1,2,\ldots,n\}$, and that $\boldsymbol{V}$ is an $n \times n$ matrix with i.i.d. U$(0,n)$ random variables as elements. We will now use the permutations P_i to rearrange the rows of $\boldsymbol{V}$. This will give us another matrix, $\boldsymbol{U}$, also with i.i.d. U$(0,n)$ elements, such that U_{ij} is close to $P_i(j)$. The rows of the random matrix $\boldsymbol{U}$ will have the same ordering as the corresponding row of P. To be precise, let

$$U_j^{(i)} = V_{(P_i(j))}^{(i)} = \text{the } P_i(j)\text{th-smallest element in row } i \text{ of } \boldsymbol{V}.$$

Note that, if k is chosen independently of $\boldsymbol{V}$, then

$$E\left(k - V_{(k)}^{(i)}\right) = k - nk/(n+1) = k/(n+1),$$

https://doi.org/10.1017/S0963548303005819 Published online by Cambridge University Press

since $V_{(k)}^{(i)}/n$ is the kth-order variable of n independent U(0, 1) variables, and therefore is Beta$(k, n+1-k)$-distributed (see, for example, Chapter IV in [8]). Assume that π, the optimal assignment for P, is given. Then generate V independently of P, and the corresponding U. Then, for $i \in \pi$ we still have

$$E\left(P_i(\pi(i)) - U_{\pi(i)}^{(i)}\right) > 0,$$

and the lower bound follows:

$$\begin{aligned} EZ_1^*(n) &= E\left(\frac{1}{n}\sum_{i=1}^{n} P_i(\pi(i))\right) = E\left(\frac{1}{n}\sum_{i=1}^{n}\left(U_{\pi(i)}^{(i)} + P_i(\pi(i)) - U_{\pi(i)}^{(i)}\right)\right) \\ &\geqslant E\left(\frac{1}{n}\sum_{i=1}^{n} U_{\pi(i)}^{(i)}\right) \geqslant EZ_c^*(n) \to \frac{\pi^2}{6}. \end{aligned}$$

2.2. Case II

Let U_{ij} be i.i.d. U$(0, n)$. To get i.i.d. random variables from the discrete uniform distribution on $\{1, 2, \ldots, n\}$, we can simply take the integer part of U_{ij} and add 1. Let

$$Y_{ij} = [U_{ij}] + 1,$$

where $[x]$ denotes the integer part of x. Then Y_{ij} are i.i.d. with the desired distribution, and the differences $Y_{ij} - U_{ij}$ are uniform on $(0, 1)$. Assume that π is an optimal assignment for $\boldsymbol{Y}$. Then we still have

$$Y_{1\pi(1)} - U_{1\pi(1)} \in \mathrm{U}(0,1) \text{ and } E\left(Y_{1\pi(1)} - U_{1\pi(1)}\right) = 1/2,$$

and for the lower bound of $EZ_2^*(n)$,

$$EZ_2^*(n) = E\left(\frac{1}{n}\sum_{i=1}^{n} Y_{i\pi(i)}\right) = E\left(\frac{1}{n}\sum_{i=1}^{n} U_{i\pi(i)}\right) + \frac{1}{2} \geqslant EZ_c^*(n) + \frac{1}{2}.$$

Now for the other direction. Assume that π is the optimal assignment for $\boldsymbol{U}$. Svante Janson [9] has calculated the expectation of the fractional part of one element in the optimal assignment, $\{U_{i\pi(i)}\} = U_{i\pi(i)} - [U_{i\pi(i)}]$, with respect to the limit distribution, given by Theorem 1.2, finding

$$\lim_{n\to\infty} E\left(U_{i\pi(i)} - \left[U_{i\pi(i)}\right]\right) = \int_0^\infty \{x\}h(x)\,dx = \frac{1}{2} - \frac{1}{24} + \sum_{k=1}^{\infty} \frac{\pi^2}{\sinh^2(2\pi^2 k)} = \frac{11}{24} + c,$$

where $c \approx 2.83 \cdot 10^{-16}$. Let Z_2^π be the cost of $\boldsymbol{Y}$ given by the assignment π.

$$\begin{aligned} \lim_{n\to\infty} E\left(Z_c^* - Z_2^\pi\right) &= \lim_{n\to\infty} E\left(\frac{1}{n}\sum_{i=1}^{n} U_{i\pi(i)} - Y_{i\pi(i)}\right) \\ &= \lim_{n\to\infty} E\left(U_{1\pi(1)} - \left[U_{1\pi(1)}\right]\right) - 1 = -\frac{13}{24} + c. \end{aligned}$$

Therefore

$$EZ_2^* \leqslant EZ_2^\pi = EZ_c^\pi + \frac{13}{24} - c.$$

https://doi.org/10.1017/S0963548303005819 Published online by Cambridge University Press

2.3. Case III

The lower bound is similar to the first case, but for ease of notation we consider a vector of n^2 elements instead of an $n \times n$ matrix.

Given a random permutation P of $\{1, 2, \ldots, n^2\}$, and a vector $\boldsymbol{V}$ of n^2 i.i.d. $\mathrm{U}(0, n^2)$ random variables, let U_i be the $P(i)$th-smallest element of $\boldsymbol{V}$, that is, $U_i = V_{(P(i))}$. This coupling is used to get the lower bound.

If k is chosen independently of $\boldsymbol{V}$, $V_{(k)}/n^2$ is $\mathrm{Beta}(k, n^2 + 1 - k)$-distributed, and

$$E\left(k - V_{(k)}\right) = \frac{k}{n^2 + 1}.$$

Assume that π, the optimal assignment for P, is given. Then generate V independently of P, and the corresponding U. Then, for $i \in \pi$ we still have

$$\frac{1}{n^2 + 1} \leqslant E\left(P(i) - V_{(P(i))}\right) \leqslant \frac{n^2}{n^2 + 1},$$

and,

$$EZ_3^*(n) = E\left(\frac{1}{n^2}\sum_{i \in \pi} U_i + P(i) - U_i\right) \geqslant EZ_c^*(n) + \frac{1}{n(n^2 + 1)} \to \frac{\pi^2}{6}.$$

For the upper bound, suppose that $\boldsymbol{U}$ is given. Then the ranks $P_i = \text{rank } U_i$ will be a random permutation of $\{1, 2, \ldots, n^2\}$. If π is optimal for $\boldsymbol{U}$,

$$Z_3^*(n) \leqslant \frac{1}{n^2}\sum_{i \in \pi} P_i = Z_c^*(n) + \frac{1}{n^2}\sum_{i \in \pi}(P_i - U_i).$$

To bound the differences, the following result [7, 11] will be useful. There exist constants C and C' such that, if $A_n = \{\max_{i \in \pi} U_i \leqslant Cn\log(n)\}$ is the event that the value of the largest element in the optimal assignment is less than $Cn\log(n)$, then

$$P\left(A_n^c\right) \leqslant C'\frac{1}{n^3}.$$

Next, we need a result on the Kolmogorov–Smirnov statistic. Given m independent $\mathrm{U}(0, b)$ variables $\boldsymbol{V} = \{V_1, V_2 \ldots V_m\}$, let

$$D_m = \sup_x |\hat{F}_m(x) - F(x)|,$$

where $\hat{F}_m(x) = \frac{1}{m}\sum_{k=1}^{m} 1_{\{V_k \leqslant x\}}$ is the empirical distribution function, and $F(x) = x/b$ is the distribution function. Note that, since $\hat{F}_m(V_i) = \frac{1}{m}\text{rank } V_i$,

$$\max_{1 \leqslant i \leqslant m}\left|\frac{\text{rank } V_i}{m} - \frac{V_i}{b}\right| \leqslant D_m.$$

A key element for the upper bound is the following result: $\tilde{D}_m = \sqrt{m}D_m$ converges in distribution to $\max_{0 \leqslant t \leqslant 1} |B_t|$, where B_t is a Brownian bridge (see, for example, Chapter 7.8 in [6]).

Choose $b = Cn\log(n)$, and let $\boldsymbol{V}$ be the subset of elements from $\boldsymbol{U}$ with $U_i \leqslant b$. The size of $\boldsymbol{V}$ is $\mathrm{Bin}(n^2, b/n^2) \approx \mathrm{N}(b, b)$, so if we choose $m = b + n$, with high probability

https://doi.org/10.1017/S0963548303005819 Published online by Cambridge University Press

(approximately $\Phi(n/\sqrt{Cn\log(n)})$), the size is less than m. Let B_n denote this event. Given B_n, add independent $\mathrm{U}(0,b)$ variables to $\boldsymbol{V}$, so that $\boldsymbol{V}$ is a set of m independent $\mathrm{U}(0,b)$ variables. Given A_n, the set $\boldsymbol{V}$ includes all elements in the optimal assignment π for $\boldsymbol{U}$.

Therefore, conditionally on $A_n \cap B_n$,

$$\begin{aligned}\frac{1}{n}\left(\text{rank } U_{\pi(1)} - U_{\pi(1)}\right) &= \frac{m}{n}\left(\frac{\text{rank } U_{\pi(1)}}{m} - \frac{U_{\pi(1)}}{b}\right) + \frac{m}{n}\frac{m-b}{mb}U_{\pi(1)} \\ &\leqslant \frac{\sqrt{m}}{n}\tilde{D}_m + \frac{U_{\pi(1)}}{b} = \sqrt{\frac{C\log(n)+1}{n}}\tilde{D}_m + \frac{U_{\pi(1)}}{Cn\log(n)}.\end{aligned}$$

In the unlikely event $(A_n \cap B_n)^c$, it suffices to use the trivial bound

$$\frac{1}{n}\left(\text{rank } U_{\pi(1)} - U_{\pi(1)}\right) \leqslant n.$$

Note that $P(A_n^c) \leqslant C'\frac{1}{n^3}$, and $P(B_n^c) \leqslant C''\frac{1}{x}e^{-x^2/2}$, where $x = \sqrt{n/C\log(n)}$.

So, with $\Delta_n = \text{rank } U_{\pi(1)} - U_{\pi(1)}$,

$$\begin{aligned}E\frac{\Delta_n}{n} &= E\left(\frac{\Delta_n}{n}\middle| A_n \cap B_n\right)P(A_n \cap B_n) + E\left(\frac{\Delta_n}{n}\middle| A_n^c \cup B_n^c\right)P\left(A_n^c \cup B_n^c\right) \\ &\leqslant \sqrt{\frac{C\log(n)+1}{n}}E\tilde{D}_m + \frac{EU_{\pi(1)}}{Cn\log(n)} + nP\left(A_n^c\right) + nP\left(B_n^c\right) \to 0.\end{aligned}$$

2.4. Case IV

As in the second case, the coupling is achieved by taking the integer part of uniform variables. Given the i.i.d. uniform $(0,n^2)$ variables U_{ij}, define X_{ij} and Y_{ij} by

$$X_{ij} = [U_{ij}], \quad Y_{ij} = X_{ij} + 1.$$

If π is optimal for $\boldsymbol{Y}$,

$$Z_4^*(n) = \frac{1}{n^2}\left(Y_{1\pi(1)} + \cdots + Y_{n\pi(n)}\right) \geqslant \frac{1}{n^2}\left(U_{1\pi(1)} + \cdots + U_{n\pi(n)}\right) \geqslant Z_c^*(n).$$

If π is optimal for $\boldsymbol{U}$,

$$Z_c^*(n) = \frac{1}{n^2}\left(U_{1\pi(1)} + \cdots + U_{n\pi(n)}\right) \geqslant \frac{1}{n^2}\left(X_{1\pi(1)} + \cdots + X_{n\pi(n)}\right) \geqslant Z_4^*(n) - \frac{1}{n}.$$

Combining this, by letting n tend to infinity we get

$$EZ_4^* = \lim_{n\to\infty} EZ_4^*(n) = \frac{\pi^2}{6}.$$

3. Simulation

The primary purpose of the simulation study is of course to estimate the expected minimal cost. Besides that, we look at the variance of the minimal cost, as well as the row rank distribution.

To solve the realizations, we used an algorithm by Jonker and Volgenant [10], which performed well in a recent survey [5]. The algorithm has time complexity $O(n^3)$. Beside the dimension, the running time also depends on the size of the matrix elements, which makes

https://doi.org/10.1017/S0963548303005819 Published online by Cambridge University Press

Table 1. Simulation results, case I

dim. n	s	mean	std. dev.	s.e. mean
100	40000	1.80990	0.07976	0.0003988
250	40000	1.82313	0.05129	0.0002564
500	40000	1.82716	0.03643	0.0001822
1000	30000	1.82929	0.02587	0.0001494
2000	20000	1.83024	0.01846	0.0001305
3000	20000	1.83055	0.01506	0.0001065
4000	20000	1.83056	0.01299	0.0000919
5000	20000	1.83062	0.01171	0.0000828
6000	20000	1.83068	0.01053	0.0000745
7000	10000	1.83076	0.00980	0.0000980
8000	8000	1.83075	0.00929	0.0001039
9000	8000	1.83075	0.00861	0.0000962
10000	8000	1.83071	0.00823	0.0000921

Table 2. Simulation results, case II

dim. n	s	mean	std. dev.	s.e. mean
100	40000	2.13618	0.12697	0.0006349
250	40000	2.15438	0.08163	0.0004081
500	40000	2.16055	0.05813	0.0002907
1000	30000	2.16336	0.04121	0.0002379
2000	20000	2.16508	0.02923	0.0002067
3000	20000	2.16554	0.02400	0.0001697
4000	20000	2.16575	0.02077	0.0001469
5000	20000	2.16560	0.01852	0.0001310
6000	10000	2.16595	0.01693	0.0001197
7000	10000	2.16581	0.01555	0.0001555
8000	8000	2.16588	0.01449	0.0001621
9000	8000	2.16540	0.01386	0.0001550
10000	8000	2.16563	0.01313	0.0001469

the simulations of cases III and IV more time-consuming. An instance of dimension 1000 is solved in less than a second for all cases. For cases I and II and dimension 10000, it takes about 90 seconds to generate a matrix and solve the problem on a standard personal computer.

3.1. Results

3.1.1. Mean. The results are summarized in Tables 1–4 and Figure 1. The number of realizations is denoted by s in the tables.

For cases I and II we simulated problems with dimension n up to 10000. The number of realizations varies between 40000 and 4000. We see that the estimated means stabilize quite fast. The difference between dimension 2000 and 10000 is of order 10^{-4}, the same order as the standard error (the sample standard deviation divided by the square root of the number of observations). The differences between high dimensions may therefore be ascribed to random fluctuations.

https://doi.org/10.1017/S0963548303005819 Published online by Cambridge University Press

Table 3. Simulation results, case III

dim. n	s	mean	std. dev.	s.e. mean
100	40000	1.60254	0.09058	0.0004529
250	40000	1.62781	0.05871	0.0002936
500	20000	1.63629	0.04167	0.0002946
1000	20000	1.64064	0.02975	0.0002103
2000	10000	1.64274	0.02087	0.0002087
3000	10000	1.64366	0.01704	0.0001704
4000	4000	1.64375	0.01476	0.0002334
5000	4000	1.64424	0.01324	0.0002093
6000	4000	1.64450	0.01222	0.0001932

Table 4. Simulation results, case IV

dim. n	s	mean	std. dev.	s.e. mean
100	40000	1.61806	0.12961	0.0006481
250	40000	1.63331	0.08310	0.0004155
500	20000	1.63948	0.05913	0.0004181
1000	20000	1.64231	0.04228	0.0002990
2000	10000	1.64349	0.02980	0.0002980
3000	10000	1.64406	0.02434	0.0002434
4000	4000	1.64381	0.02080	0.0003288
5000	4000	1.64419	0.01861	0.0002943
6000	4000	1.64487	0.01695	0.0002679

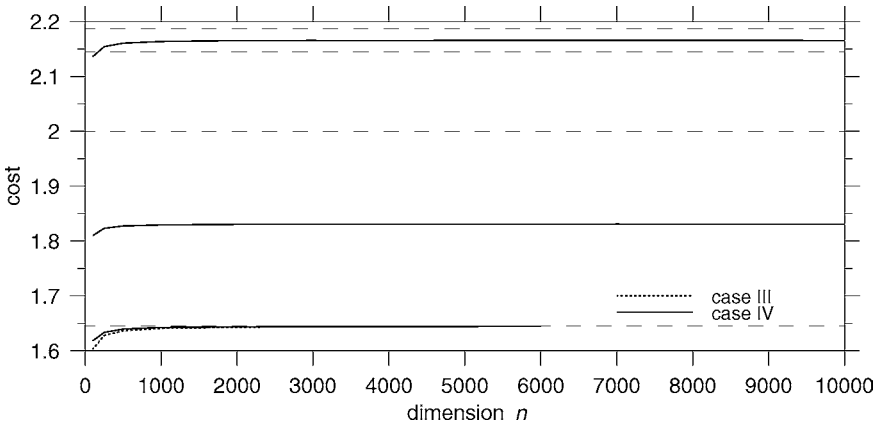


Figure 1. Estimated mean

The 'n^2' cases III and IV behave as expected. The means increase nicely towards $\pi^2/6$, with case IV slightly ahead. Since these cases are more time-consuming, and the limit is known to be $\pi^2/6$, we were content with simulations up to dimension 6000.

3.1.2. Variance. Alm and Sorkin [3] conjecture that the variance in the exponential case is $2/n + O(\log n/n^2)$. It is natural to suspect the same behaviour in all our four cases. Figure 2 shows n times the estimated variance plotted against n. It is interesting to note that the variance in the permutation cases is about half of that in the independent cases.

https://doi.org/10.1017/S0963548303005819 Published online by Cambridge University Press

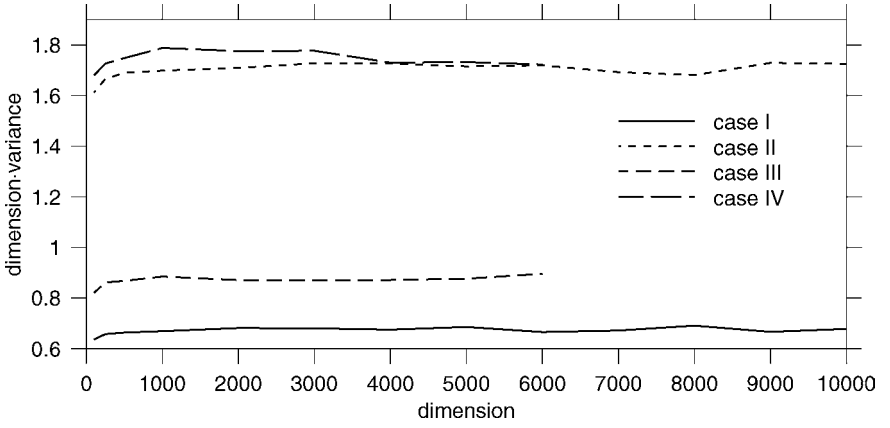


Figure 2. Estimated variance

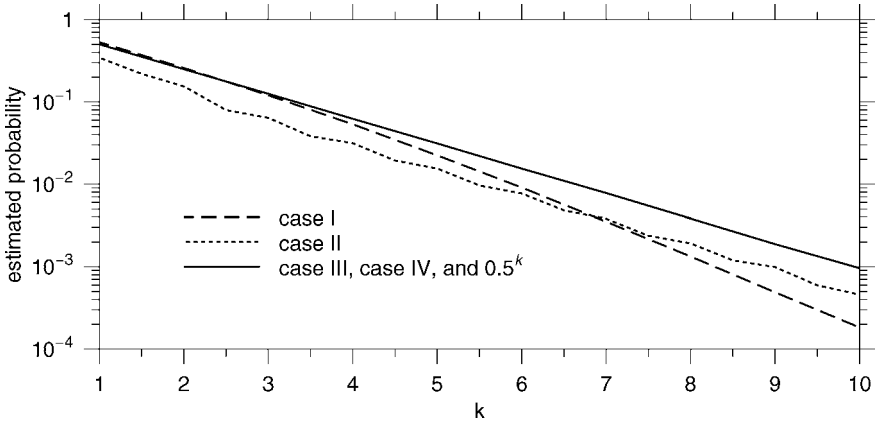


Figure 3. Estimated rank distribution, log-scale

3.1.3. The rank distribution. In the continuous cases, the limiting rank distribution is geometric, with parameter 1/2, as conjectured by Olin [15] and proved by Aldous [2]. For comparison, we generated 1000 matrices of dimension 2000 for cases I–IV, and determined the rank of every element in the optimal assignment given by the program (in the discrete cases, optimal assignments are not unique with probability 1). For case II, if there are more elements in the same row with the same cost as the assignment element, it is given the average rank of these elements. Ties occur also for case IV, although only for a few of every 10000 elements (and even less frequently for larger dimensions). Therefore, in this case, in order to eliminate the fractional ranks and get a smoother plot, we give the element the lowest rank possible.

Figure 3 shows a plot of the estimated rank distributions. Cases III and IV seem to have the same limiting distribution as in the continuous case. This is not surprising, because of the small difference between optimal assignments in these cases and the uniform case.

https://doi.org/10.1017/S0963548303005819 Published online by Cambridge University Press

Table 5. Expected costs for small dimensions

	dimension			
case	1	2	3	4
I	1	$\frac{5}{4}$	$\frac{49}{36}$	$\frac{13259}{9216}$
II	1	$\frac{21}{16}$	$\frac{40}{27}$	$\frac{3426611507}{2147483648}$
III	1	1	$\frac{6047}{5670}$	
IV	1	$\frac{263}{256}$	$\frac{65431}{59054}$	

The rank distribution for case II is hard to interpret, owing to the high frequency of ties. The data suggest that both the integer and fractional ranks (viewed separately) have a geometric distribution, with some extra weight on the smallest values.

For case I the picture looks a bit different. When plotted on a logarithmic scale, we no longer get a straight line, but a slightly concave curve.

4. Exact calculations for small dimensions

In view of Parisi's and Coppersmith and Sorkin's conjectures, it is natural to look for exact formulas for the expected cost for other distributions as well. However, it might be the case that it is the nice properties of the exponential distribution that makes a simple expression possible.

For the four discrete cases discussed here, we have calculated the expected cost for dimensions $n = 1, 2, 3$, and for the 'n' cases I and II also for dimension $n = 4$: see Table 5. The calculations were done by brute force: every possible case of costs was generated and the resulting problem solved. A more sophisticated approach would have to be used for higher dimensions, and we have not pursued that in this work.

We did not find any candidates for formulas that give these numbers for given n, and with limits that agree with the simulation results. For case I and dimension n up to 3, both $EZ_1^*(n) = \sum_{k=1}^{n} \frac{1}{k^2}$ and $EZ_1^*(n) = \left(\frac{2n+1}{2n}\right)^{n-1}$ work, but fail for $n = 4$ and in the limit.

5. Conclusions

We have studied four discrete cases of cost distributions for the random assignment problem. We have shown that the limiting expected cost for the 'n^2' cases III and IV is the same as for the continuous case, $\pi^2/6$. The simulation results indicate that also the rank distribution for these cases should be the same as for the continuous case, geometric with mean 2. For the 'n' cases I and II we give bounds for the limiting expected costs; for case II the bounds are quite tight. It would be nice to show that all inequalities in Theorem 2.1 are strict, but we have not succeeded.

Based on the simulation results, the variance is $\Theta(\frac{1}{n})$ for all cases, and larger for the independent cases than for the permutation cases. The data also suggest that the expected

https://doi.org/10.1017/S0963548303005819 Published online by Cambridge University Press

cost is probably monotonic for these cases, as it is believed to be for the exponential and uniform cases.

Acknowledgement

I would like to thank Svante Janson for suggesting both the problem and the coupling approach. I am also grateful to an anonymous referee, for pointing out an erroneous proof and suggesting the Brownian bridge approach, and for other valuable suggestions.

References

[1] Aldous, D. J. (1992) Asymptotics in the random assignment problem. *Probab. Theory Rel. Fields* **93** 507–534.
[2] Aldous, D. J. (2001) The $\zeta(2)$ limit in the random assignment problem. *Random Struct. Alg.* **18** 381–418.
[3] Alm, S. E. and Sorkin, G. B. (2002) Exact expectations and distributions in the random assignment problem. *Combin. Probab. Comput.* **11** 217–248.
[4] Coppersmith, D. and Sorkin, G. B. (1999) Constructive bounds and exact expectations for the random assignment problem. *Random Struct. Alg.* **15** 113–144.
[5] Dell'Amico, M. and Toth, P. (2000) Algorithms and codes for dense assignment problems: The state of the art. *Discrete Appl. Math.* **100** 17–48.
[6] Durrett, R. (1995) *Probability: Theory and Examples*, Duxbury Press.
[7] Frieze, A. M. and Sorkin, G. B. (2001) The probabilistic relationship between the assignment and asymmetric traveling salesman problems. In *Proc. SODA*, pp. 652–660.
[8] Gut, A. (1995) *An Intermediate Course in Probability*, Springer.
[9] Janson, S. (2001) Personal communication.
[10] Jonker, R. and Volgenant, A. A. (1987) A shortest augmenting path algorithm for dense and sparse linear assignment problems. *Computing* **38** 325–340.
[11] Lee, S. and Su, Z. (2002) On the fluctuation in the random assignment problem. *Comm. Korean Math. Soc.* **39** 321–330.
[12] Linusson, S. and Wästlund, J. (2000) A generalization of the random assignment problem. Mathematics arxiv, `http://arxiv.org/ps/math/0006146`.
[13] Mézard, M. and Parisi, G. (1985) Replicas and optimization. *Journal de Physiques Lettres* **46** 771–778.
[14] Mézard, M. and Parisi, G. (1987) On the solution of the random link matching problem. *Journal de Physiques Lettres* **48** 1451–1459.
[15] Olin, B. (1992) Asymptotic properties of random assignment problems. PhD thesis, Kungliga Tekniska Högskolan, Stockholm, Sweden.
[16] Parisi, G. (1998) A conjecture on random bipartite matching. Preprint cond-mat/9801176 at `xxx.lanl.gov`.

https://doi.org/10.1017/S0963548303005819 Published online by Cambridge University Press