RC 21354 (97336) Revised 10/27/99 Computer Science / Mathematics

IBM Research Report

On the Expected Incremental Cost of a Minimum Assignment

Don Coppersmith, Gregory B. Sorkin

IBM T.J. Watson Research Center
P.O. Box 218
Yorktown Heights, New York 10598

Research Division
Almaden · Austin · Beijing · Haifa · T.J. Watson · Tokyo · Zurich

This page intentionally left (almost) blank.

ON THE EXPECTED INCREMENTAL COST OF A MINIMUM ASSIGNMENT

DON COPPERSMITH AND GREGORY B. SORKIN

Abstract

The random assignment problem is to choose a minimum-cost matching in a complete bipartite graph whose edge weights are chosen randomly from some distribution, such as the exponential distribution with parameter 1. When choosing a perfect matching in the complete $n \times n$ bipartite graph, it has been conjectured that the expected cost is $\sum_{i=1}^{n} 1/i^2$, tending to $\pi^2/6$ in the limit. A subsequent, stronger conjecture is that the expectation of a minimum-cost matching of cardinality k in a complete $m \times n$ bipartite graph is $F(k,m,n) \equiv \sum_{i,j \geq 0,\ i+j<k} \frac{1}{(m-i)(n-j)}$.

In this note we show that, under certain hypotheses, the cost of augmenting a minimum $(m-1)$-assignment in an $(m-1) \times n$ bipartite graph, to a minimum m-assignment in an $m \times n$ bipartite graph, is equal to $F(m,m,n) - F(m-1,m-1,n)$. However, the hypotheses required are such that this result, while intriguing, does not provide a proof of the above conjectures.

1. Motivation

Let A be a complete $m \times n$ bipartite graph whose edge weights (or costs; we use the terms interchangeably) are chosen independently from the exponential distribution with mean 1. Equivalently, we may regard A as an $m \times n$ matrix with i.i.d. exponential entries. A k-assignment is a choice of k elements of the matrix A, all lying in different rows and different columns; the minimum k-assignment is the one in which the sum of the chosen elements is minimized.

Where $\mathbb{E}A^\star(k,m,n)$ denotes the expectation of a minimum-cost k-assignment, it was conjectured by Parisi [Par98] that

$$\mathbb{E}A^\star(n,n,n) = \sum_{i=1}^{n} 1/i^2.$$

In [CS] we made the following conjecture, and showed that it agrees with Parisi's when $k = m = n$.

Conjecture 1.

$$\mathbb{E}A^\star(k,m,n) = \sum_{i,j \geq 0,\ i+j<k} \frac{1}{(m-i)(n-j)}.$$

Date: 23 July 1999.

We proved the conjecture in the cases $k \leq 3$ and $k = m = n = 4$, and it has since been proved for $k \leq 4$, $k = m = 5$, and $k = m = n = 6$ [AS].

The purpose of this note, a follow-up to [CS], is to demonstrate a construction which might be used to give an inductive proof of Conjecture 1 when $k = m$. However, since we have not been able to make this connection (because we cannot see a way to enforce the hypotheses Theorem 3), the present work is a hammer in search of a nail.

Lemma 2. *The minimum assignment problem corresponds to the linear program, and its dual, shown in Figure 1.*

Proof. The LP is a relaxation in which the selection variables x_{ij} are permitted to come from the interval $[0, 1]$ rather than the set $\{0, 1\}$, so we need only show that it has an integral optimal solution. For the usual assignment problem, with $k = m = n$, one argues that this is so because the LP is a special case of the transportation problem, but this is not true of our LP.

However, it is still true here that the extreme points are integral, because the constraint matrix is unimodular. The constraint matrix has the form

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & & & & & & & & \\ & & & & 1 & 1 & 1 & 1 & & & & \\ & & & & & & & & 1 & 1 & 1 & 1 \\ & & I & & & & I & & & & I & \end{pmatrix}$$

consisting of a first row with mn 1's, a second group of m rows each with n 1's, and a third group of rows formed from m identity matrices, each $n \times n$.

It is a classical result that a matrix is totally unimodular if every group of rows can be partitioned into two, such that when the rows within each group are summed, the difference between the sums is a vector whose entries are all in $\{-1, 0, 1\}$ [GH62]. Here there are just two cases. If the first row is *not* selected, then partition the selected rows into those of the second type and those of the third type; each sum is a 0–1 vector, and the difference is in $\{-1, 0, 1\}^{mn}$. If the first row *is* selected, partition into the first row and all other rows; the sum over the second partition is in $\{0, 1, 2\}^{mn}$, and subtracting the first gives a vector in $\{-1, 0, 1\}^{mn}$. □

The (flawed) motivation for the main result described in this note is the following. Given a matrix A and a corresponding optimal dual solution $(\mathbf{z}, \mathbf{u}, \mathbf{v}) = (z, u_1, \ldots, u_m, v_1, \ldots, v_n)$, consider the matrix A' with the duals "subtracted out", $a'_{ij} = a_{ij} - z + u_i + v_j$. For convenience of phrasing, assume that A is "in general position": that it has a unique minimum k-assignment and no events of probability 0 occur. (But note that even under these assumptions, an optimal *dual* is not usually unique.) Then A' is elementwise non-negative, and its zero elements a'_{ij} correspond to elements a_{ij} belonging to the (unique) minimum k-assignment. (For the special case $k = m = n$, subtraction of a canonical optimal dual was the basis of Karp's upper bound of 2 for the expected cost of a minimum assignment [Kar87].)

It would be nice if, after subtracting an optimal dual from a random matrix A, the residue A' had, outside of its values of 0 on assignment edges, values which

Primal:

minimize $\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}x_{ij}$

subject to

$x_{ij} \geq 0 \; (i = 1 \ldots m, \; j = 1 \ldots n)$

$\sum_{i=1}^{m}\sum_{j=1}^{n} x_{ij} = k$

$\sum_{j=1}^{n} x_{ij} \leq 1 \; (i = 1 \ldots m)$

$\sum_{i=1}^{m} x_{ij} \leq 1 \; (j = 1 \ldots n)$

Dual:

maximize $kz - \sum_{i=1}^{m} u_i - \sum_{j=1}^{n} v_j$

subject to

$z \geq 0; \; u_i \geq 0 \; (i = 1 \ldots m); \; v_j \geq 0 \; (j = 1 \ldots n)$

$a_{ij} - z + u_i + v_j \geq 0 \; (i = 1 \ldots m, \; j = 1 \ldots n)$

FIGURE 1. The (primal) linear program describing the assignment problem for a matrix A, and its dual LP. The variables x_{ij} in basic optimal solutions to the primal LP will have values of 0 or 1 (not fractional), $x_{ij} = 1$ indicating assignment of row i to column j. We call the variables u_i and v_j row duals and column duals.

were independent, mean-1 exponentials.[1] Unfortunately, it is not generally so. One difficulty is that the "subtracted" dual quantity $z - u_i - v_j$ can in fact be negative. For example, consider the matrix shown below, where the entries not shown numerically are large. A minimum 3-assignment, consisting of the diagonal a_{11}, a_{22}, a_{33}, has corresponding optimal row duals u_i shown along its left side, column duals v_j along the top edge, and z at the corner.

7	7	4	0
0	0	a_{12}	a_{13}
2	a_{21}	1	5
0	a_{31}	a_{32}	7

Note that this gives $a'_{21} = a_{21} - 7 + 7 + 2 = a_{21} + 2$. For any fixed $c \geq 0$, and an exponential random variable X, conditional upon $X - c \geq 0$, the random variable $X - c$ is again exponentially distributed; but of course this is not true if $c < 0$. When $z - u_i - v_j$ *adds* to a value a_{ij}, it seems impossible to preserve the exponential distribution.

Nonetheless, such considerations lead us to the following theorem.

[1] For any collection of dual variables $(\mathbf{z}, \mathbf{u}, \mathbf{v})$, an optimal k-assignment for a matrix A is also optimal for the corresponding residual matrix A', and vice-versa; the problem instances A and A' are equivalent up to the additive constant $kz - \sum_{i=1}^{m} u_i - \sum_{j=1}^{n} v_j$.

2. The cost of an alternating path

Theorem 3. *Let A be a an $m \times n$ bipartite graph whose edges weights $a_{11}, \dots, a_{kk}$ are all 0, and whose other edge weights are i.i.d. exponential rv's. Then the expected cost of extending the 0-cost k-assignment to a minimum $(k+1)$-assignment in A is*

$$\Delta(k, m, n) = \sum_{r=0}^{k} \frac{\binom{k}{r}}{\binom{n}{r}} \cdot \frac{1}{(m-k+r)(n-r)}.$$

Proof. It is a standard result in matching theory that an optimal k-matching may be augmented to an optimal $(k+1)$-matching via the shortest alternating path from any of the $m-k$ previously unmatched "left vertices" (rows) to any of the $n-k$ previously unmatched right vertices (columns).

An augmenting, alternating path is uniquely described by its non-matching edges, from $j_0 \in k+1, \dots, m$ to some j_1, from j_1 to j_2, and so forth, until an (unassigned) $j \in \{k+1, \dots, n\}$ is reached. If the path has length ℓ, its cost is $\sum_{r=0}^{\ell-1} a(j_r, j_{r+1})$.

Let D be a directed non-bipartite analogue of A, having the self-loop $d_{mm} = a_{mm}$, having edge weights $d_{ij} = a_{ij}$ for $i \in [m]$ and $j \in [n]$, and having no other edges (in particular no out-edges from $i \in \{m+1 \dots n\}$). In D, distinguish vertices $k+1, \dots, m$ as "sources" and vertices $\{k+1, \dots, n\}$ as "terminals". (Some sources are also terminals.) Then each alternating path in A corresponds to a simple path in D, and the minimum-cost augmenting path in A corresponds to a shortest non-trivial path from a source to a terminal in D.

Find such a path using Dijkstra's algorithm, as follows. Initialize $S = S(0)$ to be the sources $k+1, \dots, m$ and $T = T(0)$ to be $1, \dots, n$. For $i \in S$, let $\mathrm{dist}(i)$ be the weight of the shortest path from any source to i, and let $\mathrm{dist}(S) = \max_{i \in S} \mathrm{dist}(i)$. Initially, for any $i \in k+1, \dots, m$, $\mathrm{dist}(i) = \mathrm{dist}(S(0)) = 0$. We assert that for every r, the weights a_{ij} of edges ij, $i \in S(r)$, $j \in T(r)$, are i.i.d. exponential rv's conditioned upon

$$\mathrm{dist}(i) + a_{ij} \geq \mathrm{dist}(S(r)). \tag{1}$$

At step 0 this means they are not conditioned, which is true by hypothesis. We will show inductively that the assertion (1) continues to hold.

Consider taking step r, which adds a vertex to $S(r)$ and removes the same vertex from $T(r)$, giving $S(r+1)$ and $T(r+1)$. Before this step, by the inductive hypothesis, the weights a_{ij} of edges ij, $i \in S(r)$, $j \in T(r)$, are i.i.d. exponential rv's conditioned upon $\mathrm{dist}(i) + a_{ij} \geq \mathrm{dist}(S)$. (Also, if $\mathrm{dist}(i) + a_{ij}$ were less than $\mathrm{dist}(S)$, vertex j would already be in S, not in T.) Equivalently, the random variables

$$X_{ij} = \mathrm{dist}(i) + a_{ij} - \mathrm{dist}(S)$$

form an i.i.d. exponential collection.

In step r, Dijkstra's algorithm adds to $S(r)$ the vertex $j(r+1) \in T(r)$ which satisfies

$$\text{(2)} \qquad \begin{aligned} \operatorname{dist}(j(r+1)) &= \min_{i \in S,\, j \in T} \{\operatorname{dist}(i) + a_{ij}\}. \\ &= \operatorname{dist}(S) + \min_{i \in S,\, j \in T} \{X_{ij}\} \end{aligned}$$

By definition, $\operatorname{dist}(S(r+1)) = \operatorname{dist}(j(r+1))$.

The asserted conditioning (1) now holds for edges $uv \in S(r+1) \times T(r+1)$. If $u \in S(r)$, then $uv \in S(r) \times T(r)$ was one of the edges considered in (2), therefore $\operatorname{dist}(u) + a_{uv} \geq \operatorname{dist}(S(r+1))$. Otherwise $u = j(r+1)$, and the edges out of u, never having been examined before, are unconditioned, or equivalently, "conditioned" by $a_{uv} \geq 0$; since $\operatorname{dist}(u) = \operatorname{dist}(j(r+1)) = \operatorname{dist}(S(r+1))$, this is equivalent to (1).

If j is any terminal $k+1, \ldots, n$, then Dijkstra's algorithm terminates. What is the expected distance from the sources to this terminal?

In step r, when the $m-k+r+1$'st vertex is about to be added to $S(r)$, $S(r)$ contains $m-k+r$ vertices (including the original sources $k+1, \ldots, m$), and $T(r)$ contains $n-r$ vertices (the original $1, \ldots, n$ less the vertices shifted into S). Thus, the minimum in (2) is the minimum of $(m-k+r)(n-r)$ i.i.d. exponential rv's, and has expectation $1/[(m-k+r)(n-r)]$. It follows that

$$\mathbb{E}(\operatorname{dist}(S(r+1))) - \mathbb{E}(\operatorname{dist}(S(r))) = 1/[(m-k+r)(n-r)].$$

Dijkstra's algorithm continues until a terminal is hit. Since the vertices added are by construction distinct, but otherwise random (in (2), the residual edge costs X_{ij} out of S are i.i.d.), the probability that all the vertices hit in r steps are among the non-terminals $1, \ldots, k$ (and Dijkstra's algorithm needs to be continued) is $\binom{k}{r}/\binom{n}{r}$. Note that after $k+1$ steps, failure is impossible, as there are only k non-terminal vertices, and this is reflected in $\binom{k}{k+1} = 0$.

The expected distance to the terminal hit can be computed by summing, over r, the probability that a terminal has not yet been hit, times the cost to add the next vertex in this event. The sum is

$$\sum_{r=0}^{k} \frac{\binom{k}{r}}{\binom{n}{r}} \cdot \frac{1}{(m-k+r)(n-r)}.$$

□

The next corollary may be interpreted as saying that the expected cost of augmenting an $(m-1)$-assignment to an m-assignment, in an $m \times n$ bipartite graph, and under the conditions of Theorem 3, is in agreement with Conjecture 1.

Corollary 4. *With $F(k,m,n) \equiv \sum_{i,j \geq 0,\ i+j<k} \frac{1}{(m-i)(n-j)}$ per Conjecture 1, and $\Delta(m-1,m,n)$ as given by Theorem 3, $\Delta(m-1,m,n) = F(m,m,n) - F(m-1,m-1,n)$.*

Proof. To show that

$$\sum_{r=0}^{m-1} \frac{\binom{m-1}{r}}{\binom{n}{r}} \frac{1}{(r+1)(n-r)} = \frac{1}{m} \sum_{i=0}^{m-1} \frac{1}{n-i},$$

it suffices to show equality after multiplication by $n(n-1)\cdots(n-m+1)$, that is, to show that

$$\sum_{r=0}^{m-1} \frac{(m-1)!(n-r-1)!n!}{(m-1-r)!n!(r+1)(n-m)!} = \frac{1}{m}\sum_{i=0}^{m-1} n(n-1)\cdots\widehat{(n-i)}\cdots(n-m+1),$$

where $\widehat{(n-i)}$ indicates absence of this term from the product. Equivalently, and writing N in lieu of n to emphasize the dependence on this parameter, we wish to show that

$$\sum_{r=0}^{m-1} \frac{(m-1)!(N-r-1)\cdots(N-m+1)}{(m-1-r)!(r+1)} = \frac{1}{m}\sum_{i=0}^{m-1} N(N-1)\cdots\widehat{(N-i)}\cdots(N-m+1).$$

Now observe that as functions of N, both sides of the would-be equality above are simply polynomials of degree $m-1$. Thus it suffices to prove equality at any m distinct values of N, and we shall use $N = 0, \dots, m-1$. (So now $N < m$, whereas $n \geq m$ in the original context.) In this range, terms of the left side vanish when $r \leq N-1$, and terms of the right side vanish except when $i = N$, so the desired equality may be rewritten as

$$\sum_{r=N}^{m-1} \frac{(m-1)!(-1)^{m-r-1}(m-1-N)\cdots(r+1-N)}{(m-1-r)!(r+1)} = (-1)^{m-1-N}\frac{1}{m}N!(m-1-N)! \tag{3}$$

Attacking the left side,

$$\begin{aligned}
&\sum_{r=N}^{m-1} \frac{(m-1)!(-1)^{m-r-1}(m-1-N)\cdots(r+1-N)}{(m-1-r)!(r+1)} \\
&\qquad = \sum_{r=N}^{m-1} \frac{(m-1)!(-1)^{m-r-1}(m-1-N)!}{(m-1-r)!(r+1)(r-N)!},
\end{aligned}$$

which, substituting $s = r - N$,

$$\begin{aligned}
&= \sum_{s=0}^{m-1-N} \frac{(m-1)!(-1)^{m-s-N-1}(m-1-N)!}{(m-1-s-N)!(s+N+1)s!} \\
&= (m-1)!\sum_{s=0}^{m-1-N} \binom{m-1-N}{s}\frac{(-1)^{m-s-N-1}}{s+N+1} \\
&= (m-1)!\int_0^1 (1-x)^{m-1-N}x^N dx(-1)^{m-1-N} \\
&= (m-1)!\frac{(m-1-N)!N!}{m!}(-1)^{m-1-N} \\
&= \frac{1}{m}N!(m-1-N)!(-1)^{m-1-N},
\end{aligned}$$

matching the right side of (3) as desired. $\square$

3. Conclusions

On the negative side, Corollary 4, which refers to an increment $\Delta(m-1, m, n)$, fails to generalize to arbitrary increments $\Delta(k, m, n)$. For example, as per the Corollary, we know that $\Delta(3, 4, 6)$ is equal to $F(4, 4, 6) - F(3, 3, 6)$. On the other hand, $\Delta(3, 5, 6)$ is not equal to $F(k, m, n) - F(k', m', n')$ for any values $k, m, n, k', m', n' \in 0, \ldots, 7$.

It is an open question whether Corollary 4 is merely an algebraic coincidence, or whether it might be used to prove the correctness of Conjecture 1 in the case $k = m$. Such a proof would require producing the unbiased edges of Theorem 3 from the biased edges that result when optimal dual variables are subtracted from the original edge weights (see the example on page 3), or from the biases introduced by growing a minimum k-assignment through successive augmentations (each augmentation performed by the Dijkstra's algorithm, as in the proof of Theorem 3).

Alternatively, it is possible that biases could be tolerated, using Theorem 3 to produce a bound on the expected assignment cost, rather than its exact value.

Acknowledgments

We thank Jon Lee and David Williamson for helpful discussions.

References

[AS] Sven Erick Alm and Gregory B. Sorkin. On the expectation and variance of a minimum assignment. in preparation.

[CS] Don Coppersmith and Gregory B. Sorkin. Constructive bounds and exact expectations for the random assignment problem. *Random Structures and Algorithms*. to appear.

[GH62] Alain Ghouila-Houri. Caractérisation des matrices totalement unimodulaires. *C. R. Acad. Sci. Paris*, 254:1192–1194, 1962.

[Kar87] Richard M. Karp. An upper bound on the expected cost of an optimal assignment. In David S. Johnson et al., editors, *Discrete Algorithms and Complexity: Proceedings of the Japan-US Joint Seminar*, volume 15 of *Perspectives in Computing*, pages 1–4. Academic Press, 1987.

[Par98] Giorgio Parisi. A conjecture on random bipartite matching. Physics e-Print archive, http://xxx.lanl.gov/ps/cond-mat/9801176, January 1998.

(Don Coppersmith) IBM T.J. Watson Research Center, Department of Mathematical Sciences, Yorktown Heights NY 10598

E-mail address: copper@watson.ibm.com

(Gregory B. Sorkin) IBM T.J. Watson Research Center, Department of Mathematical Sciences, Yorktown Heights NY 10598

E-mail address: sorkin@watson.ibm.com