Discrete Mathematics 21 (1978) 297–308.

A GEOMETRICAL INTERPRETATION OF THE HUNGARIAN METHOD

Hans Joachim SCHMID

Mathematisches Institut, Universität Erlangen–Nürnberg, Bismarckstr. 1 1/2, 8520 Erlangen, Federal Republic of Germany

Received 21 June 1977
Revised 8 August 1977

In this paper a geometrical interpretation of the Hungarian method will be given. This special algorithm to solve the dual transportation problem is not restricted to the edges of the convex polyhedron of feasible solutions. Each covering-step can be considered as a determination of a direction of steepest descent, each reduction-step as movement along that direction to a boundary point of the polyhedron. The dimension of the face that will be crossed depends on the covering that is chosen.

1. Introduction

The Hungarian method is a well-known solution of the dual transportation problem, based on a combinatorial theorem of the Hungarians König and Egerváry. Kuhn [5] used a constructive proof of this theorem to solve the assignment problem. This procedure led to combinatorial solutions of the transportation problem (for references see [3]). The main difficulty in these algorithms was the determination of a minimal covering. Many variants were studied solving the covering problem by primal-dual means (e.g. [2]). By the pseudo-Boolean approach to this problem (see [4]) the Hungarian method was reduced to a pure dual algorithm.

In spite of the great number of articles on this algorithm no geometrical interpretation—analogous to that of the simplex method in [1, chapter 7]—was made, as far as the author knows. Using the compact form of the Hungarian method in [4] and a characterization of the convex polyhedron of feasible solutions we shall give such an interpretation. This paper is based on results from [7] which will be simplified and slightly extended.

2. The transportation problem

Let $a=(a_1, a_2, \ldots, a_m)$ and $b=(b_1, b_2, \ldots, b_n)$ be given, satisfying $\sum_{i=1}^m a_i = \sum_{j=1}^n b_j = r$, where a_i, b_j are positive integers, and let $C=(c_{ij})$ be an $m \times n$ matrix with nonnegative integer entries.

The (uncapacitated) transportation problem is a linear programming problem of the following type:

Problem P. Find an $m\times n$ matrix $X=(x_{ij})$ such that

$$\sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}x_{ij}=\min, \tag{2.1}$$

and

$$x_{ij}\geqslant 0, \qquad i=1,2,\ldots,m; \quad j=1,2,\ldots,n, \tag{2.2}$$

$$\sum_{i=1}^{m} x_{ij}=b_j, \qquad j=1,2,\ldots,n, \tag{2.3}$$

$$\sum_{j=1}^{n} x_{ij}=a_i, \qquad i=1,2,\ldots,m. \tag{2.4}$$

By definition, the dual problem is

Problem D. Find $(u_1,u_2,\ldots,u_m)$ and $(v_1,v_2,\ldots,v_n)$ such that

$$\sum_{i=1}^{m} a_iu_i+\sum_{j=1}^{n} b_jv_j=\max, \tag{2.5}$$

and

$$u_i+v_j\leqslant c_{ij}, \qquad i=1,2,\ldots,m; \quad j=1,2,\ldots,n. \tag{2.6}$$

From the duality theorem of linear programming, cf. [1], it follows that either both problems have no optimal solution or they have the same objective value. In the latter case both problems have optimal solutions. For our purpose the dual problem fits better in an equivalent form. We put

$$\alpha_{ij}=c_{ij}-u_i-v_j \quad \text{for all } i \text{ and } j \tag{2.7}$$

and rewrite Problem D in matrix form.

Problem $\mathbf{D}_m$. Find an $m\times n$ matrix $A=(\alpha_{ij})$ such that

$$\varphi(A)=\frac{1}{r}\sum_{i=1}^{m}\sum_{j=1}^{n} \alpha_{ij}a_ib_j=\min, \tag{2.8}$$

and

$$\alpha_{ij}\geqslant 0, \qquad i=1,2,\ldots,m; \quad j=1,2,\ldots,n, \tag{2.9}$$

$$\alpha_{11}+\alpha_{ij}-\alpha_{i1}-\alpha_{1j}=c_{11}+c_{ij}-c_{i1}-c_{1j}, \quad i=2,3,\ldots,m; \quad j=2,3,\ldots,n. \tag{2.10}$$

To show that (2.7) transforms Problem D into Problem D_m and vice versa, we remark that (2.6) is equivalent to (2.9). Matrices $A=(\alpha_{ij})$ satisfying (2.7) are generated by subtracting real values from all elements of each row and each column of C. Hence such matrices are characterized by

$$a_{ij} + a_{kl} - a_{il} - a_{kj} = c_{ij} + c_{kl} - c_{il} - c_{kj} \quad \text{for all } i, j, k, l. \tag{2.11}$$

These equations have an $(m+n-1)$ parameter solution set which enables us to solve back for u_i, v_j and to reduce (2.11) to (2.10). Finally a short computation leads to

$$\varphi(A) = \frac{1}{r} \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} a_i b_j - \left(\sum_{i=1}^{m} u_i a_i + \sum_{j=1}^{n} b_j v_j \right). \tag{2.12}$$

Hence a maximum of (2.5) yields a minimal value of (2.8) and a minimum of (2.8) yields a maximal value of (2.5). We use Problem D_m to state the dual problem, since geometrical properties of the feasible solutions of the dual problem can be easily recognized in matrix form. On the other hand, an optimal solution $X = (x_{ij})$ of Problem P can be derived at once from an optimal solution $A = (a_{ij})$ of Problem D_m. We have to set $x_{ij} = 0$ whenever $a_{ij} > 0$ and to determine the rest of the variables x_{ij} via (2.2), (2.3) and (2.4). The forthcoming results can be translated back to the usual dual variables of Problem D.

3. Characterization of feasible solutions

Let $(\mathscr{L}, \|\cdot\|)$ be the normed linear space of real $m \times n$ matrices over the field of real numbers with

$$\|A\| = \max\{|a_{ij}|: i = 1, 2, \ldots, m; j = 1, 2, \ldots, n\} \tag{3.1}$$

for $A \in \mathscr{L}$. Let Problem D_m be given with a fixed matrix C and fixed vectors a, b. The set of feasible solutions of Problem D_m consists of matrices $A \in \mathscr{L}$ satisfying (2.9) and (2.10). We denote $\mathcal{M} = \mathcal{M}(C, a, b)$ the set of those feasible solutions A satisfying

$$\varphi(A) \leq \varphi(C). \tag{3.2}$$

It is easily checked that $\mathcal{M}$ is a convex polyhedron in $(\mathscr{L}, \|\cdot\|)$. We assume that C itself is not an optimal solution of Problem D_m, since otherwise a geometrical interpretation is trivial. Hence any algorithm solving Problem D_m determines matrices from

$$\mathcal{M}_+ = \{A \in \mathcal{M}: \varphi(A) < \varphi(C)\}.$$

$\mathcal{M}_+$ is a convex subset of $\mathcal{M}$ and each vertex of $\mathcal{M}_+$ is vertex of $\mathcal{M}$. We recall some fundamental definitions from the theory of convex polyhedra. The intersection of a polyhedron with some boundary hyperplanes is called face. Hence $A \in \mathcal{M}$ is an element of a face, if $a_{ij} = 0$ for at least one pair (i, j) or if $\varphi(A) = \varphi(C)$. The dimension of a face $\mathscr{S}$ is defined to be the dimension of the smallest linear manifold containing $\mathcal{S}$. Faces of dimension 0 (1) are vertices (edges) of the polyhedron. For $A \in \mathcal{M}_+$ the face of smallest dimension containing A is defined by

$$\mathcal{S}_A = \{B \in \mathcal{M}: b_{ij} = 0, \quad \text{whenever } a_{ij} = 0\}.$$

We shall now characterize the dimension of $\mathcal{S}_A$ by properties of A. This enables us to give a geometrical characterization for algorithms solving Problem D_m.

For $A = (a_{ij}) \in \mathcal{M}$ consider the bipartite graph G_A with vertices $i \in \{1, 2, \ldots, m\}$ and $j \in \{1, 2, \ldots, n\}$ and an edge (i, j), whenever $a_{ij} = 0$. G_A is called connected, if for every two vertices k and l from G_A there is a sequence of edges of the form

$$(k, k_1), (k_1, k_2), \ldots, (k_{r-1}, k_r), (k_r, k_{r+1}), \quad \text{where } k_{r+1} = l \quad \text{and} \quad r \geqslant 0.$$

We refer to such a sequence as path from k to l. We call A connected, if G_A is connected. If A is not connected, A can be transformed by permutation of rows and columns to

$$\bar{A} = \begin{pmatrix} A_1 & A_3 \\ A_4 & A_2 \end{pmatrix},$$

where A_1, A_2 contain all zeros of A.

If no vertex of G_A is isolated, we call A non-isolated. A non-isolated matrix contains a zero in each row and column. If a non-isolated A is not connected, A can be transformed by permutation of rows and columns to

$$\bar{A} = \begin{pmatrix} A_1 & & & & \\ & A_2 & & & \\ & & \ddots & & \\ & & & A_{k-1} & \\ & & & & A_k \end{pmatrix},$$

where $A_1, A_2, \ldots, A_k$ are representations of the connected components of G_A. Hence $A_1, A_2, \ldots, A_k$ contain all zeros of A. The number of connected components in G_A will be denoted $\sigma(A)$. If A is connected we have $\sigma(A) = 1$.

Lemma 3.1 [7, 8]. *If $A \in \mathcal{M}_+$ is connected, then $\mathcal{S}_A = \{A\}$.*

Proof. Assume that there is a $B \in \mathcal{S}_A$ with $B \neq A$. From (2.7) and the remark following (2.11) we obtain u_i, v_j such that $b_{ij} = a_{ij} - u_i - v_j$, $i = 1, 2, \ldots, m$; $j = 1, 2, \ldots, n$. Since $B \neq A$, there must be at least one $u_l = \alpha \neq 0$. The connectedness of G_A implies that for each $k \in \{1, 2, \ldots, n\}$ we can find a path from k to l which is represented by a sequence of zeroes of the form

$$a_{k,j_1} = a_{i_1,j_1} = a_{i_1,j_2} = \cdots = a_{i_r,j_{r+1}} = 0, \quad \text{where } j_{r+1} = l \quad \text{and} \quad r \geqslant 0,$$

or

$$a_{k,j} = a_{k,j_1} = a_{i_1,j_1} = \cdots = a_{i_r,j_{r+1}} = 0, \quad \text{where } j_{r+1} = l \quad \text{and} \quad r \geqslant 0.$$

We get $v_k = -\alpha$ for all $k \in \{1, 2, \ldots, n\}$, since $b_{ij} = 0$ whenever $a_{ij} = 0$. For each $i \in \{1, 2, \ldots, m\}$ there must be a $j \in \{1, 2, \ldots, n\}$ such that $a_{ij} = 0$, since A is non-isolated. Hence we get $u_i = \alpha$ for all i. This implies $B = A$, in contradiction to our assumption.

Remark 3.2. The proof of Lemma 3.1 can be used to show that every connected matrix A in $\mathcal{M}$ is a vertex of $\mathcal{M}$.

We shall prove the converse of Lemma 3.1.

Lemma 3.3 [7, 8]. *If $A \in \mathcal{M}_+$ is not connected, then A is no vertex of $\mathcal{M}_+$.*

Proof. Since A is not connected, we may assume A to be of the form

$$A = \begin{pmatrix} A_1 & A_3 \\ A_4 & A_2 \end{pmatrix},$$

where A_1 and A_2 contain all zeros of A. Let E_3, E_4 be matrices corresponding to A_3, A_4 and containing only entries equal to 1. We define $B_k = (b_{ij}^{(k)})$ for $k = 1, 2$ by

$$B_1 = A + \alpha_1 \begin{pmatrix} 0 & E_3 \\ -E_4 & 0 \end{pmatrix} \quad \text{and} \quad B_2 = A + \alpha_2 \begin{pmatrix} 0 & -E_3 \\ E_4 & 0 \end{pmatrix},$$

where $\alpha_1, \alpha_2 > 0$, $\alpha_1 \leq \min\{a_{ij}: a_{ij} \text{ in } A_4\}$, and $\alpha_2 \leq \min\{a_{ij}: a_{ij} \text{ in } A_3\}$. B_1, B_2 are feasible solutions. Since $\varphi(A) < \varphi(C)$ and

$$\varphi(B_k) = \varphi(A) + \alpha_k \frac{1}{r} \sum_{i=1}^{m} \sum_{j=1}^{n} \delta_{ij}^{(k)} a_i b_j, \qquad k = 1, 2.$$

we can find α_1, α_2 such that $\varphi(B_k) < \varphi(C)$. This leads to $A = \alpha_2/(\alpha_1 + \alpha_2) B_1 + \alpha_1/(\alpha_1 + \alpha_2) B_2$, $B_1, B_2 \in \mathcal{M}_+$, hence A is no vertex of $\mathcal{M}_+$.

By Lemma 3.1 and Lemma 3.3 the vertices of $\mathcal{M}_+$ are characterized as connected matrices of $\mathcal{M}_+$. This result can be extended to

Theorem 3.4 [7]. *Let $A \in \mathcal{M}_+$ be non-isolated, then $\sigma(A)$ and $\dim \mathcal{S}_A$ are linked by $\sigma(A) - 1 = \dim \mathcal{S}_A$.*

Proof. We assume A in a form

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1k} \\ A_{21} & A_{22} & \cdots & A_{2k} \\ \vdots & \vdots & & \vdots \\ A_{k1} & A_{k2} & \cdots & A_{kk} \end{pmatrix}, \qquad k = \sigma(A),$$

where the A_{ii}'s are connected and contain all zeros of A. Correspondingly we define submatrices B_{ij} for $B \in \mathcal{S}_A$.

Applying Lemma 3.1 to the connected matrix B_{ii}, $i \in \{1, 2, \ldots, k\}$, yields $B_{ii} = A_{ii}$. Since $b_{11} = a_{11}$ Eq. (2.10), applied to A and B, reduces to $a_{ij} - a_{i1} - a_{1j} = b_{ij} - b_{i1} - b_{1j}$, for all i and j. So $B \in \mathscr{S}_A$ is uniquely determined if one element from each submatrix $B_{12}, B_{13}, \ldots, B_{1k}$ is known. We define for $i, j \in \{1, 2, \ldots, k\}$ a matrix E_{ij} corresponding to B_{ij} which contains only entries equal to 1. Now each $B \in \mathscr{S}_A$ can be represented as

$$B = A + \sum_{i=2}^{k} \beta_i C_i, \qquad \beta_i \text{ real}, \tag{3.3}$$

where

$$C_i = \begin{pmatrix} & & & E_{1i} & & & \\ & 0 & & \vdots & & 0 & \\ & & & E_{i-1,i} & & & \\ -E_{i1} & \cdots & -E_{i,i-1} & 0 & -E_{i,i+1} & \cdots & -E_{ik} \\ & & & E_{i+1,i} & & & \\ & 0 & & \vdots & & 0 & \\ & & & E_{ki} & & & \end{pmatrix},$$

$i = 2, 3, \ldots, k$. We can determine some $B \in \mathscr{S}_A$ which can be represented by (3.3) with each $\beta_i \neq 0$. Each element of the linear manifold spanned by $\mathscr{S}_A$ can also be represented by (3.3). Since the C_i's are linearly independent by construction, we have $\dim \mathscr{S}_A = k - 1$.

4. The Hungarian method

The Hungarian method, see e.g. [4] and [6], solves the dual transportation problem by constructing in each reduction-step a non-isolated matrix of $\mathscr{M}_+$ such that the objective function φ will be decreased. This is done by use of minimal coverings. The step in which such a covering is determined is called covering-step.

Usually a covering (of the zeros) of a non-isolated matrix $A = (a_{ij}) \in \mathscr{M}$ is a set of row-indices and column-indices containing at least one of the indices i and j, if $a_{ij} = 0$. The sum of the corresponding a_i and b_j is called capacity of the covering. A covering of A with a minimum capacity in respect to all possible coverings of A is called minimal covering of A.

In [4] the indices are represented by Boolean vectors $x = (x_1, x_2, \ldots, x_m)$, $y = (y_1, y_2, \ldots, y_n)$. Such vectors have coefficients $x_i, y_j \in \{0, 1\}$. We put $x_i = 1$, whenever i is in the covering set, and $y_j = 0$, whenever j is in that set. Then x and y determine a covering of A, if

$$y_j \leqslant x_i \quad \text{for all } i \text{ and } j \text{ with } a_{ij} = 0. \tag{4.1}$$

The quantity

$$\sum_{i=1}^{m} a_i x_i + r - \sum_{j=1}^{n} b_j y_j \tag{4.2}$$

is the capacity of the covering. We define a covering as follows. An $m \times n$ matrix $Z = (z_{ij})$ is called a covering of $A = (a_{ij}) \in \mathcal{M}$, if there are Boolean vectors x and y satisfying (4.1) such that

$$z_{ij} = x_i - y_j \quad \text{for all } i \text{ and } j. \tag{4.3}$$

An argument analogous to that following problem Problem D_m shows that a covering Z of A is characterized by

$$0 \leq z_{ij} \quad \text{whenever } a_{ij} = 0, \tag{4.4}$$

$$z_{ij} \in \{-1, 0, 1\} \quad \text{for all } i \text{ and } j, \tag{4.5}$$

$$z_{11} + z_{ij} - z_{i1} - z_{1j} = 0 \quad \text{for all } i \text{ and } j. \tag{4.6}$$

As in (2.12) we find

$$\sum_{i=1}^{m} a_i x_i + r - \sum_{j=1}^{n} b_j y_j = r + \varphi(Z). \tag{4.7}$$

Let A be non-isolated, then Z is a minimal covering of A, if $\varphi(Z) \leq \varphi(Z^*)$ for all coverings Z^* of A.

In our notation the Hungarian method can be derived as follows. We start by constructing a non-isolated matrix from C. This will be done by subtracting the minimum of each row from each element of that row, and thereafter correspondingly for each column. This operation decreases the objective function and leads to a non-isolated matrix $A \in \mathcal{M}$. The algorithm continues with the

Covering-step: A minimal covering Z of A has to be determined. This can be done as proposed in [2]. We shall refer later on to another method which was introduced in [4].

The reduction-step is the constructive proof of

Theorem 4.1 [5]. *Let $A \in \mathcal{M}$ be non-isolated and Z be a minimal covering of A with $\varphi(Z) < 0$. Then there is an $\alpha > 0$ such that $B = A + \alpha Z$ is a non-isolated element of $\mathcal{M}_+$ satisfying $\varphi(B) < \varphi(A)$.*

Proof.

Reduction-step: By permutation of rows and columns a covering Z of A can be transformed into

$$\tilde{Z} = \begin{pmatrix} 0 & -E_3 \\ E_4 & 0 \end{pmatrix},$$

where E_3, E_4 contain only entries equal to one. Hence we may assume A and Z to be of the form

$$A=\begin{pmatrix} A_1 & A_3 \\ A_4 & A_2 \end{pmatrix} \quad \text{and} \quad Z=\begin{pmatrix} 0 & -E_3 \\ E_4 & 0 \end{pmatrix},$$

where A_1, A_2, A_4 contain all zeros of A, and where E_3, E_4 contain only entries equal to one.

Since $\varphi(Z)<0$, there must be at least one element $z_{ij}=-1$. Also, since Z is a minimal covering of A, the submatrices A_1, A_2 are non-isolated. For otherwise, if we could find a column in A_1 or a row in A_2 containing only positive elements, there would exist a covering $\bar{Z}$ of A such that $\varphi(\bar{Z})<\varphi(Z)$. We can determine

$$\alpha=\min\{a_{ij}:\ i,j \text{ such that } z_{ij}=-1\}.$$

Now

$$B=\begin{pmatrix} B_1 & B_3 \\ B_4 & B_2 \end{pmatrix}=A+\alpha Z \tag{4.8}$$

is a feasible solution satisfying $\varphi(B)=\varphi(A)+\alpha\varphi(Z)<\varphi(A)$. Since the zeros in A_1 and A_2 are preserved by the reduction-step (4.8), B_1 and B_2 are non-isolated, hence B is non-isolated.

Actually we only need a covering Z with $\varphi(Z)<0$ to decrease the objective function, but the difference $\varphi(A)-\varphi(B)$ will be maximal, if we take a minimal covering.

The algorithm proceeds with covering-steps and reduction-steps. If we have constructed an $A\in\mathcal{M}$ such that each minimal covering Z of A satisfies $\varphi(Z)=0$, then A is optimal. This criterium—which goes back to a theorem of König and Egerváry (see [5])—can be regarded as a consequence of our geometrical interpretation of minimal coverings, cf. Corollary 5.6.

Since there is only a finite number of possibilities to form a covering Z for $m\times n$ matrices, we have $\varphi(Z)\geq\gamma$ for all coverings Z with $\varphi(Z)<0$, where γ depends on a_i and b_j only. The reduction-step preserves the integer values of the matrices, hence each α chosen in a reduction-step satisfies $\alpha\geq 1$. So the algorithm must obtain an optimum of Problem D_m in a finite number of steps, since $\mathcal{M}$ is bounded.

5. Geometrical interpretation

As a consequence of Theorem 3.4 the matrices generated by the reduction-step need not to be vertices of $\mathcal{M}_+$. We shall now study the precise geometrical movement of the Hungarian method in $\mathcal{M}_+$.

An $m\times n$ matrix $W=(w_{ij})$ is called a normed direction at $A\in\mathcal{M}$, if the following conditions hold,

$$\text{there is an } \alpha>0 \text{ such that } A+\alpha W\in\mathcal{M}, \tag{5.1}$$

$$\|W\|\leq 1. \tag{5.2}$$

From (4.5) and Theorem 4.1 it follows that a covering Z of A with $\varphi(Z) < 0$ is a normed direction at A. Let W be a normed direction at A, we call W a direction of steepest descent, if $\varphi(W) \leq \varphi(W^*)$ for all normed directions W^* at A. Therefore the determination of a direction of steepest descent corresponds to minimizing $\varphi(W^*)$.

The reduction-step can be interpreted geometrically as movement along Z from A to $A+\alpha Z$. To give a concise characterization we introduce the covering number $\mu_Z(A)$ of a covering Z of A.

Let A be a non-isolated matrix in $\mathcal{M}$ with a covering $Z=(z_{ij})$. By permutation of rows and columns we get

$$\bar{A} = \begin{pmatrix} A_1 & A_3 \\ A_4 & A_2 \end{pmatrix}$$

such that all z_{ij} corresponding to A_1 and A_2 are 0, those corresponding to A_3, resp. A_4, are -1, resp. 1. Since A_1 and A_2 are non-isolated, $\sigma(A_1)$ and $\sigma(A_2)$ are defined. Now $\mu_Z(A) = \sigma(A_1) + \sigma(A_2)$ is called the covering number of Z.

Theorem 5.1 [7]. *Let $A \in \mathcal{M}$ be non-isolated with a covering Z satisfying $\varphi(Z) < 0$. Then the reduction-step is a movement from A to $A+\alpha Z$ crossing a face $\mathcal{S}$ of dimension $\mu_Z(A)-1$.*

Proof. We descend from A to $B = A+\alpha Z$ via the matrices $B_\beta = A + \beta Z$, $\beta \in (0, \alpha)$. Since all the zeros of B_β are in the positions corresponding to A_1 and A_2, each B_β belongs to the same face $\mathcal{S} - \mathcal{S}_{B_\beta}$. Applying Theorem 3.4 we obtain $\dim \mathcal{S} = \mu_Z(A) - 1$.

Now we characterize the special normed directions which are chosen by the Hungarian method (working with minimal coverings). We have

Lemma 5.2. *Let A be a matrix of $\mathcal{M}$. If $W=(w_{ij})$ is a normed direction at A, there are vectors $(p_1, p_2, \ldots, p_m)$ and $(q_1, q_2, \ldots, q_n)$ such that*

$$w_{ij} = p_i - q_j \quad \text{for all } i \text{ and } j, \tag{5.3}$$

and,

$$\max\{|p_i - q_j| : \text{for all } i \text{ and } j\} \leq 1. \tag{5.4}$$

Proof. If W is a normed direction, there is an $\alpha > 0$ such that $A+\alpha W$ is an element of $\mathcal{M}$. This yields (5.3). Eq. (5.4) is evident.

Lemma 5.3. *Let W be any normed direction at $A \in \mathcal{M}$. Then there are $(p_1, p_2, \ldots, p_m)$ and $(q_1, q_2, \ldots, q_n)$ such that (5.3), (5.4) are satisfied and $p_i \in [0, 2]$, $q_j \in [-1, 1]$.*

Proof. Let $(p_1^+, p_2^+, \ldots, p_m^+)$ and $(q_1^+, q_2^+, \ldots, q_n^+)$ be arbitrary vectors satisfying (5.3) and (5.4). We determine $\gamma = \min\{p_i^+ : i = 1, 2, \ldots, m\}$ and define $p_i = p_i^+ - \gamma$, $q_j = q_j^+ - \gamma$ for all i and j. Now $(p_1, p_2, \ldots, p_m)$ and $(q_1, q_2, \ldots, q_n)$ satisfy (5.3), (5.4), and $p_i \geq 0$. Let $\delta = \max\{p_i : i = 1, 2, \ldots, m\}$, then all p_i belong to $[0, \delta]$. Eq. (5.4) yields $p_i - 1 \leq q_j \leq p_i + 1$ for all i and j, hence $q_j \in [\delta - 1, 1]$. On the other hand we get $0 \leq \delta \leq 2$. Thus all p_i belong to $[0, 2]$, all q_j to $[-1, 1]$.

Let A be a non-isolated matrix in $\mathcal{M}$. We want to determine a normed direction W of steepest descent at A. Hence we have to minimize

$$\varphi(W) = \sum_{i=1}^{m} a_i p_i - \sum_{j=1}^{n} b_j q_j, \tag{5.5}$$

where p_i, q_j are chosen as in Lemma 5.3. To guarantee $a_{ij} + \alpha(p_i - q_j) \geq 0$ for some $\alpha > 0$, we must have

$$q_j \leq \min\{p_i : i \text{ such that } a_{ij} = 0\}, \quad j = 1, 2, \ldots, n. \tag{5.6}$$

There is no need to decrease q_j below the value on the right hand side of (5.6), since we want to minimize (5.5). So we put

$$q_j = \min\{p_i : i \text{ such that } a_{ij} = 0\} \quad \text{for all } j. \tag{5.7}$$

Theorem 5.4 [7]. *Let A be a non-isolated matrix in $\mathcal{M}$. Each normed direction W of steepest descent at A can be represented by Boolean vectors.*

Proof. We have to minimize (5.5) subject to the restrictions (5.4) and (5.7) with $p_i \in [0, 2]$. If $p_i > 1$, then changing p_i to 1 gives a better value in (5.5), hence $p_i \in [0, 1]$. Now (5.4) is satisfied and it remains the problem to minimize (5.5) with respect to (5.7) subject to $0 \leq p_i \leq 1$, $i = 1, 2, \ldots, m$. The convex polyhedron of feasible solutions of this linear programming problem has a simple structure: All vertices are Boolean, hence there must be an optimal solution which is also Boolean. Eq. (5.7) asserts that $(q_1, q_2, \ldots, q_n)$ is Boolean too.

From Theorem 5.4 we get two important corollaries.

Corollary 5.5. *Let $A \in \mathcal{M}$ be non-isolated, then a minimal covering Z with $\varphi(Z) < 0$ is a normed direction of steepest descent.*

Corollary 5.6. *Let $A \in \mathcal{M}$ be non-isolated, then A is optimal, if and only if each minimal covering Z of A satisfies $\varphi(Z) = 0$.*

If $x = (x_1, x_2, \ldots, x_m)$ and $y = (y_1, y_2, \ldots, y_n)$ are Boolean vectors, (5.7) can be expressed by

$$y_j = \prod_{i \in I_j} x_i, \tag{5.8}$$

where $I_j = \{i: a_{ij} = 0\}$, $j = 1, 2, \ldots, n$. So we get for (5.5)

$$\varphi(W) = \sum_{i=1}^{m} a_i x_i - \sum_{j=1}^{n} b_j \prod_{i \in I_j} x_i. \tag{5.9}$$

Now the covering-step consists in finding a Boolean vector x which minimizes (5.9); an algorithm solving this problem can be found in [3]. This approach to the covering problem was proposed in [4].

With each covering-step the Hungarian method chooses a normed direction of steepest descent. From this point of view the Hungarian method seems superior to the simplex algorithm which can choose directions of steepest descent being generated by edges of $\mathcal{M}$ only.

6. Hungarian method and simplex algorithm

If we start the Hungarian method with a connected matrix of $\mathcal{M}$ and determine with each covering-step a minimal covering with covering number 2, the Hungarian method is restricted (by Theorem 5.1) to the edges of $\mathcal{M}$. Such an algorithm has been developed in [7].

The determination of these special coverings is performed by using a Boolean algorithm of [3]. Now (5.9) has to be minimized subject to two restrictions. It can be shown that this special version of the Hungarian method is equivalent to the simplex algorithm, applied to Problem D_m. In this case a minimal covering with covering number 2 is equivalent to the pivot column of the simplex tableau. In this sense the Hungarian method is a generalization of the simplex method applied to the dual transportation problem.

Acknowledgement

The author is indebted to an anonymous referee for some useful comments and suggestions.

References

[1] G.B. Dantzig, Linear Programming and Extensions (Princeton University Press, Princeton, NJ, 1963).

[2] L.R. Ford and R. Fulkerson, Flows in Networks (Princeton University Press, Princeton, NJ, 1962).

[3] P.L. Hammer (Ivănescu) and S. Rudeanu, Boolean Methods in Operations Research and Related Areas (Springer, Berlin, 1968).

[4] P.L. Ivănescu and S. Rudeanu, A pseudo-Boolean approach to matching problems in graphs with applications to assignment and transportation problems, in: Théorie des Graphes, Journées Internationales d'Etude à Rome (Dunod, Paris and Gordon and Breach, New York, 1967) 161–175.

[5] H.W. Kuhn, The Hungarian method for the assignment problem, Naval Res. Logist. Quart. 2 (1955) 83–97.
[6] J. Munkres, Algorithms for the assignment and transportation problems, J. Soc. Indust. Appl. Math. 5 (1957) 32–38.
[7] H.J. Schmid, Eine geometrische Deutung der Ungarischen Methode, Dissertation, Universität Tübingen (1974).
[8] Zhu Yong-len, Maximum number of iterations in the dual algorithm of the Kantorovicz-Hitchcock problem in linear programming, Chinese Math.–Acta. 3 (1963) 307–313.