TR-CS-6482

Hassan AÏT-KACI

On a necessary and sufficient condition for Doubly Stochastic Matrices: An Algorithmic Proof.

Standford University
March 1977.

This paper is a short mathematical note dealing with the following problem. Defining the sign pattern of a matrix of order n to be the set of its entries that are equal to zero, we would like to find necessary and sufficient conditions for any matrix to have the same sign pattern as a doubly stochastic matrix of same order. That is, if A is a matrix of order n, we would like to find out about the existence of a matrix D such that:

(i) $\sum_{i=1}^{n} d_{ij} = 1$ for all $j = 1, \ldots, n$;

$\sum_{j=1}^{n} d_{ij} = 1$ for all $i = 1, \ldots, n$;

(ii) $a_{ij} = 0$ if and only if $d_{ij} = 0$.

As a matter of fact, this problem is solved as a corollary of the famous maximum flow-minimum cut theorem due to Ford and Fulkerson (see ref. [2]) giving for condition the existence of a maximum flow in a clever network representation. However, in the case where existence of a doubly stochastic matrix is proven, the Ford Fulkerson corollary cannot exhibit one. We would like to propose a graph theoretic proof of these necessary and sufficient conditions that can, in the positive case, come up with a doubly stochastic matrix such as desired. The method uses a labeling algorithm.

We will, after expressing the graph theoretic representation, give a few interesting and useful properties of such notions as bipartite graphs, matchings, etc., then we will proceed to presenting the central theorem proof as a labeling algorithm.

We can associate to a matrix M of order n a unique bipartite graph: $G_M = (I, J; E_M)$ defined as follows

$$(1) \qquad I = J = \{1, 2, \dots, n\} \; ;$$

$$(2) \qquad E_M \subset I \times J \; ;$$

$$(3) \qquad (i, j) \in E_M \iff M_{ij} \neq 0 \qquad i \in I, \; j \in J .$$

That is, the bipartite graph constituted on one side by the row indices, on the other side by the column indices, and whose edges correspond to non zero entries in the matrix M.

P1. proposition: Given a matrix A of order n, each non zero monomial in the development of its determinant characterizes a perfect matching (i.e. a matching of size n) in its associated bipartite graph.

check: It is easy to see this property when we notice that each monomial of A is of the form

$$(4) \qquad \pm \, a_{1i_1} a_{2i_2} \cdots a_{ni_n}$$

where $\{i_1, i_2, \dots, i_n\}$ is a permutation of the set $\{1, 2, \dots, n\}$. Thus, a monomial of A is non zero if and only if $a_{ji_j} \neq 0$ for all $j \in J$. In other words, if and only if there is a matching $\{(1, i_1), (2, i_2), \dots, (n, i_n)\}$ in the associated bipartite graph of A.

definition: Given two bipartite graph G_A and G_B, associated respectively to two $n \times n$ matrices, we say that a matching of G_A is a matching of G_B if the edges of the matching in G_A figure also as edges in G_B.

P2. Proposition: Let A and B be two $n \times n$ matrices. If A and B have the same sign pattern, then all perfect matchings of G_A are in G_B, and conversely.

Proof: Assume A and B have same sign pattern.

case 1: A has no matching of size n. By P1, this means that all its monomials are zero. In other words, for all permutations $\{i_1, i_2, \dots, i_n\}$ of $\{1, 2, \dots, n\}$, we have $a_{1i_1} a_{2i_2} \dots a_{ni_n} = 0$; i.e., for all permutation, there exists $j \in J$ such that $a_{ji_j} = 0$. Since A and B have same sign pattern, this is equivalent to $b_{ji_j} = 0$ for some j in all permutations. Hence, B has no perfect matching.

Case 2: A has at least one perfect matching. By P1, there exists a permutation $\{i_1, i_2, \dots, i_n\}$ such that $a_{1i_1} a_{2i_2} \dots a_{ni_n} \neq 0$. For the same permutation, $b_{1i_1} b_{2i_2} \dots b_{ni_n} \neq 0$. Hence, B has the same perfect matching.

P3. Theorem (ref. [2], pp. 105-106): If D is a doubly stochastic matrix, at least one of its monomials is non-zero.

This theorem is a corollary of the König's theorem. By P1, it is equivalent to stating that there is at least one perfect matching in the bipartite graph associated to a doubly stochastic matrix.

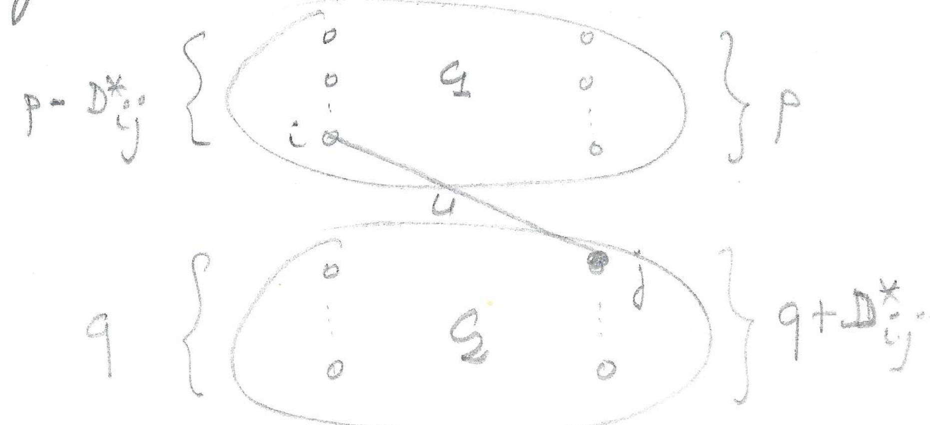
Definition: Given a matrix $M \in \mathbb{R}^{n\times n}$, we define its reduced form $M^* \in \mathbb{R}^{m\times m}$, $m \le n$ to be the matrix obtained from M as follows:

for all entries alone in their row and column (i.e., $\forall M_{ij} \neq 0$ such that $m_{kj} = m_{il} = 0$ $\forall k \neq i, \forall l \neq j$) the corresponding row and column are suppressed.

note: if M is such that M^* is vacuous, then M has exactly one non zero element per row and per column. In this case, it is trivial to find a doubly stochastic matrix with same sign pattern.

P4. Lemma: Let D be a doubly stochastic matrix, then the bipartite graph associated to the reduced matrix D^* is such that every edge belongs to a cycle.

proof: Suppose that in the bipartite graph associated to D^*, there is one edge that $\notin$ cycle $u = (i,j)$. Therefore u must be an isthmus. We have $D^*_{ij} \neq 0$ from the definition of the associated graph. The edge u separates the graph into two components G_1 and G_2 as shown in the figure.



D^* is of order $m \le n$ and doubly stochastic as only isolated entries per row and per column (thus equal to one) were suppressed in D. The sum of its entries along its row and columns must then be equal to the order of D^*. Hence,

$$\begin{cases} (p - D^*_{ij}) + q = m & \text{along the rows;} \\ p + (q + D^*_{ij}) = m & \text{along the columns.} \end{cases}$$

where

$$p = \sum_{l \in J \cap C_1} \sum_{k \in P(l)} D^*_{kl}$$

$$q = \sum_{k \in I \cap C_2} \sum_{l \in S(k)} D^*_{kl}$$

where $P(l)$ is the set of elements of I linked to l; and $S(k)$ is the set of elements of J linked to k.

It hence comes $D^*_{ij} = 0$. A contradiction. Therefore each edge belongs to a cycle. Q.E.D.

Note: In what follows we will consider only a matrix A such that $A^* = A$. Indeed, showing that A has same sign pattern than a doubly stochastic matrix is equivalent to showing that this is the case for A^*, and then completing A^* back to its $n \times n$ form by adding 1's at those entries that were suppressed.

Theorem: Let A be a reduced $n \times n$ matrix. A has same sign pattern than a doubly stochastic matrix if and only if all edges in the associated bipartite graph belong to cycles.

proof: For the necessity, assume A has same sign pattern than a doubly stochastic matrix D. Then, by P4, we simply conclude our fact. Conversely, let G_A be the graph associated to A. There is a perfect matching in G_A, since if there were none, by P2 and P3, we could immediately deduce that there is no doubly stochastic matrix with same sign pattern. Let $G_A^{(0)}$ be the graph G_A whose edges are assigned weights in the following manner: on the edges of the perfect matching the weight is equal to one, and all the other edges have a weight equal to zero. The following procedure is going to try to distribute weights along cycles by successively adding and subtracting the same quantity to the existing weights in a way to preserve the property of being doubly stochastic. Since all edges are on cycles and since the number of edges is finite, we eventually obtain a doubly stochastic distribution of weights; hence, a doubly stochastic matrix with same sign pattern as the initial one.

Algorithm to get a doubly stochastic matrix, given its S.P.

Notes : (1) For each node $i \in I$, $S(i)$ is the set of the successors of i ($S(i) \subseteq J$), and $US(i)$ is the set of the unlabeled successors of i ($US(i) \subseteq S(i)$); Accordingly, for each node $j \in J$, $P(j)$ is the set of the predecessors of j ($P(j) \subseteq I$), and $UP(j)$ is the set of the unlabeled predecessors of j ($UP(j) \subseteq P(j)$).

(2) For each node $x \in I \cup J$, label(x) stands for the label of x.

(3) w_{ij} is the weight on edge (i,j).

Algorithm

Step 0 : For all $i \in I$ and all $j \in J$ do $US(i) := S(i)$, $UP(j) := P(j)$; $k := 1$; goto Step 1;

Step 1 : if $w_{ij} \neq 0$ forall $(i,j) \in E$ then Terminate
else let $i_0 \in I$ such that $\exists j_0 \in J$, $w_{i_0 j_0} = 0$;
label(i_0) := $*$; label(j_0) := i_0; $UP(j_0) := UP(j_0) - \{i_0\}$;
$l := j_0$; goto step 2;

Step 2 :

(2.1) : if $UP(l) = \phi$ then [if $l = j_0$ then goto (2.4)
else l := label(label(l))
goto (2.1);

else let $r \in I$: $w_{rl} = \max_{i \in UP(l)} (w_{il})$;

label$(r) := l$; $UP(l) := UP(l) - \{r\}$; goto (2.2);

(2.2): if $US(r) = \phi$ then $\left[\begin{array}{l} r := label(label(r)) \\ \text{if } r = i_0 \text{ then goto (2.4)} \\ \text{else goto (2.2).} \end{array}\right.$

else let $s \in J : w_{rs} = \min\limits_{j \in US(r)} (w_{rj})$;

$label(s) := r$, $US(r) := US(r) - \{s\}$; goto (2.3);

(2.3): if $i_0 \in P(s)$ then goto Step3
else $l := s$; goto Step 2;

(2.4): Stop: $(i_0 j_0)$ is an isthmus;

Step 3: (3.1): $i := i_0$; $j := s$;

(3.2): $w_{ij} := w_{ij} - \frac{1}{2k}$; $i := label(j)$;

$w_{ij} := w_{ij} + \frac{1}{2k}$;

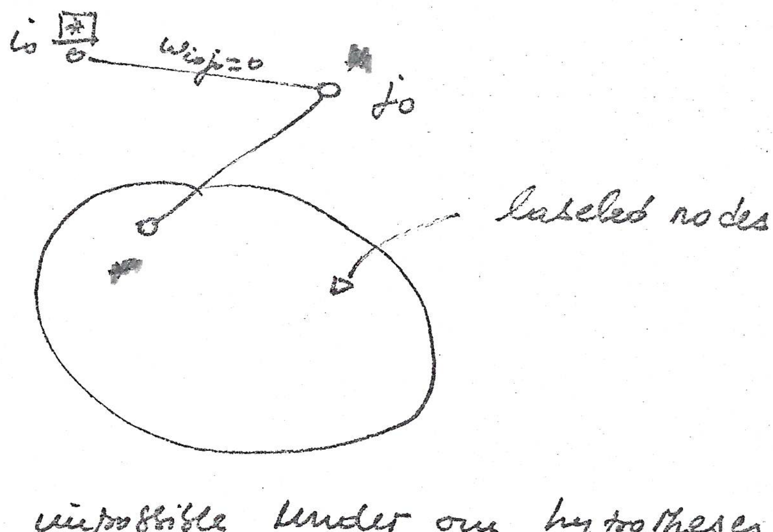
if $i = i_0$ then goto Step4
else $j := label(i)$; goto (3.2);

Step4: For all $i \in I$, and all $j \in J$ do $US(i) := S(i)$, $UP(j) := P(j)$;
$k := k + 1$;
goto Step 1;

---

1. <u>proof of termination</u>: The general procedure in this algorithm is to find an even cycle given a zero edge. First, we clearly see that, provided the loops (2.1) and (2.2) are finite, the bigger loop <u>step 2</u> is always finite since we add two edges at least at each iteration. So we only have to prove that loops (2.1) and (2.2) cannot be infinite.

<u>loop (2.1)</u>: At the first iteration in the step 2 we do not reach into the loop (2.1); so the first time we get into it, we have already checked at least a path from i_0, and if we go back to j_0 that means that $i_0 j_0$ is an isthmus:



And this is impossible under our hypotheses;

<u>loop (2.2)</u>: The reasoning is similar.

Therefore, this algorithm either terminates, or produces an isthmus.

2. <u>proof of legality</u>: We have to show that the way we

choose the nodes r and s at each iteration is correct

Since we know that the graph is such that each node of J has at least one non zero-weighted edge coming in, since we choose S by taking the zero edges by priority when labeling from I, to J, and since we start with a zero edge $(i_0 j_0)$, we conclude that we never label from J to I through a zero edge. Therefore the choices are justified.

Now, when tracing back the cycle we substract $\frac{1}{2^k}$ to positive weights. Since, at the beginning, the non zero weights were "1's" and since $1 > \sum_{k=1}^{n} \frac{1}{2^k}$, $\forall n \in \mathbb{N}$ the weights are always positive or zero.

~~Note: we could substract and add $\frac{1}{p^k}$, for any fixed $p \geq 2$; but $p = 2$ seems to be the one which gives a more equilibrated distribution of weights.~~

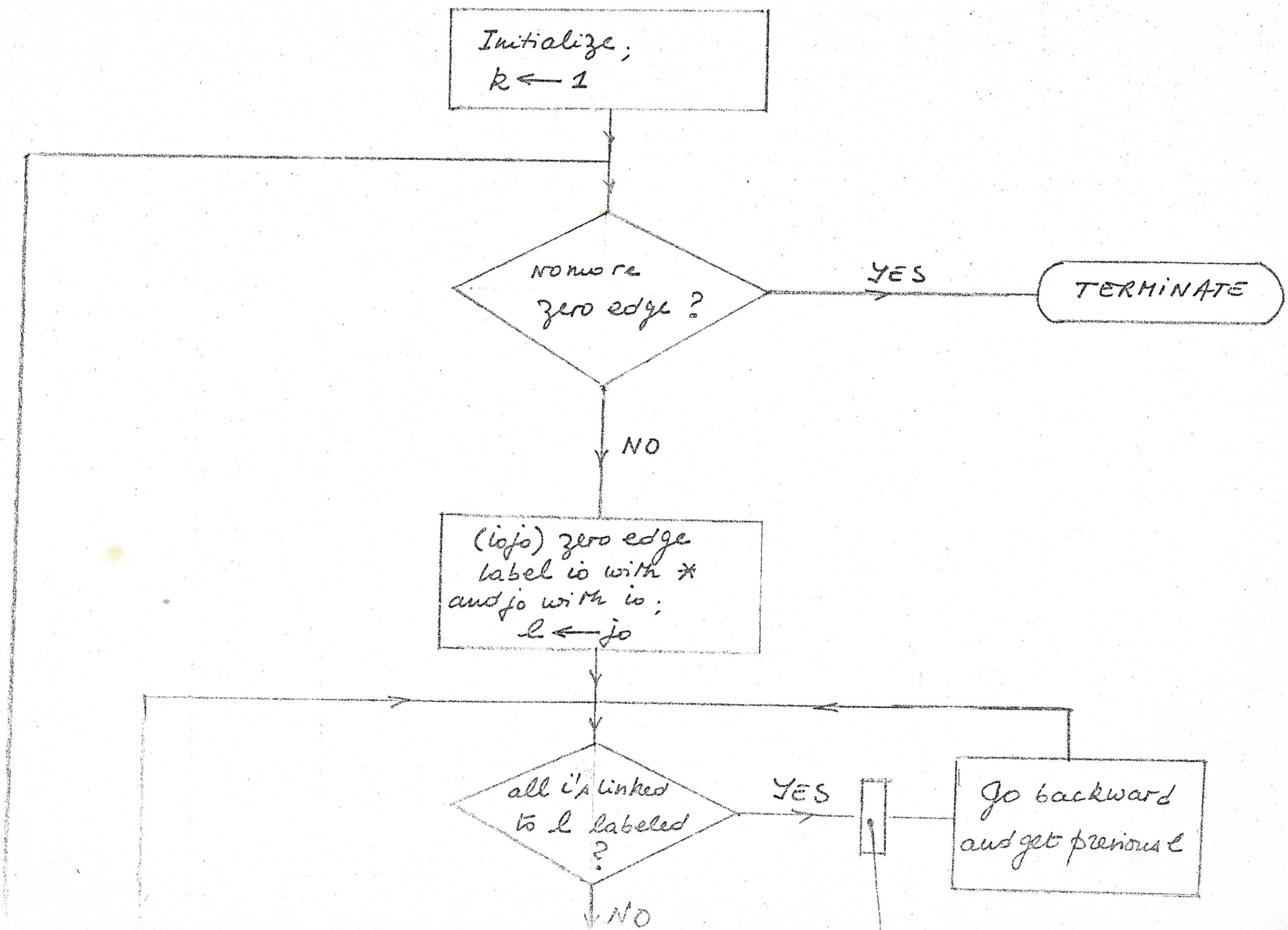
References

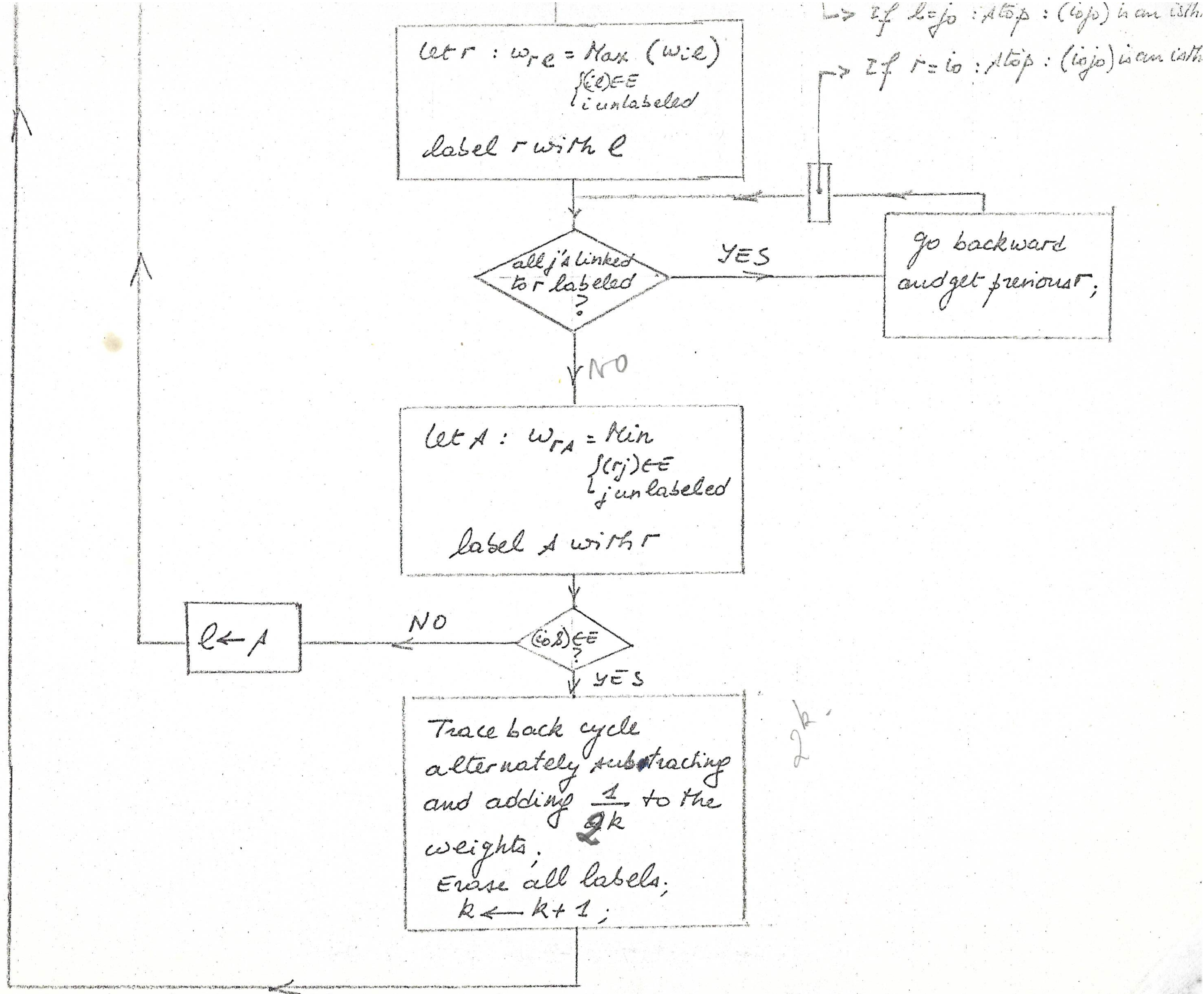
[1] L. R. Ford, Jr, and D. R. Fulkerson, Flows in Networks, Princeton University Press, 1962.

[2] Claude Berge, The Theory of Graphs and its Applications, (London: Methuen & Co), 1962.

FLOW-CHART ~~of the Algorithm to get a doubly stochastic matrix~~

Representation of ~~the~~ labeling procedure.




If $l = j_0$: stop : $(i_0 j_0)$ is an isth
If $r = i_0$: stop : $(i_0 j_0)$ is an isth
let r : $w_{rl} = \mathrm{Max}\,(w_{il})$ $\{(il) \in E,\ i$ unlabeled
label r with l
all j's linked to r labeled ?
YES
go backward and get previous r;
NO
let s : $w_{rs} = \mathrm{Min}$ $\{(rj) \in E,\ j$ unlabeled
label s with r
$(i_0 s) \in E$?
NO
$l \leftarrow s$
YES
Trace back cycle alternately subtracting and adding $\frac{1}{2k}$ to the weights;
Erase all labels;
$k \leftarrow k+1$;
$2k$