T. R. NO. 138

138

Technical Report

O SOLVE VAN DER WAERDEN CONJECTURE
ON DOUBLY STOCHASTIC MATRICES

bу

M. Raghavachari

Mb-738 IIM



INDIAN INSTITUTE OF MANAGEMENT AHMEDABAD

A GEOMETRIC PROGRAMMING APPROACH TO SOLVE VAN DER WAERDEN CONJECTURE ON DOUBLY STOCHASTIC MATRICES

bу

M. Raghavachari



T. R. No. 138 November 1976

INDIAN INSTITUTE OF MANAGEMENT AHMEDABAD

ABSTRACT

Let \mathcal{L}_n be the set of all doubly stochastic square matrices of order n i.e. the set of all n x n matrices with non-negative entries with row and column sums equal to unity. The permanent of an n x n matrix $\Lambda = (a_{ij})$ is defined by

$$P(A) = \sum_{\zeta \in S_n} \prod_{i=1}^n a_i \zeta(i)$$

where S_n is the symmetric group of order n. van der Waerden conjectured that $P(\Lambda) > n$!/nⁿ for all $\Lambda \in \mathcal{D}_n$ with equality occurring if and only if $\Lambda = J_n$, where J_n is the matrix all of whose entries are equal to 1/n.

The validity of this conjecture has been shown for a few values of n and for general n under certain assumptions. In this paper the problem of finding the minimum of the permanent of a doubly stochastic matrix has been formulated as a reversed geometric program with a single constraint and an equivalent dual formulation is given. A related problem of reversed homogeneous posynomial programming problem is also studied.

A Geometric Programming approach to solve van der Waerden Conjecture on Doubly Stochastic matrices

Ву

M. Raghavachari Indian Institute of Management Ahmedabad

Let \mathcal{L}_n be the set of all doubly stochastic square matrices of order n i.e. the set of all n x n matrices with non-negative entries with row and column sums equal to unity. The permanent of an n x n matrix $A = (a_{ij})$ is defined by

$$P(\Lambda) = \sum_{\gamma \in S_n} \frac{1}{i=1} \quad a_{i,\gamma(i)}$$
 (1)

where S_n is the symmetric group of order n. van der Waerden (4) conjectured that P(A) > n !/n for all $Ac \mathcal{F}_n$ with equality occurring if and only if $A = J_n$, where J_n is the matrix all of whose entries are equal to 1/n.

The validity of this conjecture has been shown for a few values of n and for general n under certain assumptions. See, for example D. London [2] and Marcus

and Newman [3]. In this paper the problem of finding the minimum of the permanent of a doubly stochastic matrix has been formulated as a reversed geometric program with a single constraint and an equivalent dual formulation is given. A related problem of reversed homogeneous posynomial programming problem is also studied.

2. A theorem on a class of reversed homogeneous posynomial Programs

Let $P(t_1, \ldots, t_N)$ be a posynomial i.e. of the form

$$P(t_1, ..., t_N) = \sum_{j=1}^{M} c_j L_j (t_1, ..., t_N)$$

with

$$L_j(t_1, \ldots, t_N) = \prod_{i=1}^{N} t_i^{b_{ij}}$$

and $c_j>0$ for $j=1,2,\ldots,M$. $(M\gg N)$. b_{ij} 's are nonnegative integers. Let us further assume that $P(t_1,\ldots,t_N)$ is homogenous of degree n, i.e. $\sum\limits_{p=1}^{N}b_{pj}=n$ for each p=1

 $j = 1, \dots, M$. Further assume that $b_{11} = b_{22} = b_{NN} = n$.

Consider the geometric programming problem : Find

inf
$$P(t_1, t_2, \dots, t_N) = \lambda_1$$

subject to

$$\sum_{j=1}^{N} t_j = 1$$
(I)

Problem (I) is a 'reversed' geometric programming problem. This is because in (I) the constraint $\sum_{j=1}^{N} t_j = 1$ can be changed to $\sum_{j=1}^{N} t_j > 1$ without changing (I). Note that (I) is a minimization problem with $c_j > 0$ and b_{ij} 's are > 0. Consider therefore the problem I(a): Find

inf
$$P(t_1, ... t_N)$$

subject to

$$\sum_{j=1}^{N} t_{j} \geqslant 1$$

$$t_1>0$$
 , $t_2>0$, $t_N>0$

We state the following theorem:

Theorem 1: Suppose that there exists a point with $t_1^*>0$, ... $t_N^*>0$ at which the infimum of Problem I(a) is attained.

Then we have

$$\frac{\partial P(t_1, \dots t_N)}{\partial t_p} = n P(t_1, \dots t_N)$$

$$\frac{t}{\partial t_p} = t^* \qquad p=1,2,\dots N$$

where $\underset{\sim}{t} = (t_1, \dots t_N)$

Proof of Theorem 1 : First we state and prove a lemma.

Lemma: Under the assumptions of Theorem 1, consider the Problem: (II). Find

inf
$$P(t_1, \dots t_N) = \lambda_2$$

subject to

(II)
$$\left(\frac{t_1}{\alpha_1}\right)^{-\alpha_1} \quad \left(\frac{t_2}{\alpha_2}\right)^{\alpha_2} \quad \dots \quad \left(\frac{t_N}{\alpha_N}\right)^{-\alpha_N} \leqslant 1$$

$$\alpha_1 > 0 \; ; \; t_1 > 0 \qquad \qquad i=1,2,\ldots N$$

$$\begin{array}{c}
N \\
\Sigma \alpha_{i} = 1 \\
\mathbf{i} = 1
\end{array}$$

Then $\lambda_1 = \lambda_2$ with λ_1 defined in I(a)

<u>Proof of Lemma</u>: From (2) and the geometric inequality we have

$$1 \leqslant \left(\frac{\mathbf{t}_{1}}{\alpha_{1}}\right)^{\alpha_{1}} \quad \left(\frac{\mathbf{t}_{2}}{\alpha_{2}}\right)^{\alpha_{2}} \quad \dots \quad \left(\frac{\mathbf{t}_{N}}{\alpha_{N}}\right)^{\alpha_{N}} \quad \leqslant \quad \sum_{\mathbf{j}=1}^{N} \mathbf{t}_{\mathbf{j}}$$

Thus a feasible solution to (II) is also feasible for I(a). Thus $\frac{\lambda_1}{\lambda_2} > \frac{\lambda_1}{\lambda_1}$. If (t_1, \dots, t_N) is feasible for I(a), by choosing $\alpha_i = t_i / \sum_{j=1}^N t_j$, $i = 1, 2, \dots N$, we have

$$\left(\frac{\mathbf{t}_{1}}{\alpha_{1}}\right)^{-\alpha_{1}} \qquad \left(\frac{\mathbf{t}_{2}}{\alpha_{2}}\right)^{-\alpha_{2}} \qquad \cdots \qquad \left(\frac{\mathbf{t}_{N}}{\alpha_{N}}\right)^{-\alpha_{N-}} = \left[\Sigma \ \mathbf{t}_{j}\right]^{-1} \leqslant 1$$

This implies by continuity $\lambda_2 \leqslant \lambda_1$. Hence $\lambda_1 = \lambda_2$ and this proves the lemma.

Under the assumptions of the theorem, the proof of the lemma shows that the infimum im (II) can be replaced by minimum.

Denote α_1 α_2 α_N by β . For a given vector

of α 's such $\alpha_j > 0$ and $\sum_{j=1}^{N} \alpha_j = 1$, (II) becomes a standard posynomial program. The dual to this geometric program is:

maximize
$$h(u) = \left(\frac{u_1}{c_1}\right)^{-u_1} \left(\frac{u_2}{c_2}\right)^{-u_2} \cdots \left(\frac{u_M}{c_M}\right)^{-u_M} \beta^{\nabla}$$
 (3)

s.t.
$$u_1 + u_2 + \dots + u_{N_i} = 1$$
 (4)

$$\Sigma b_{pj} u_{j} - \alpha_{p} v = 0$$
, $p = 1, 2, ... N$
 $u_{j} \ge 0 \quad v \ge 0$ $j = 1, ... M$ (5)

Since $\sum_{p=1}^{N} b_{pj} = n$, for each j=1,2, ...M, the constraints (5) imply on adding that $n \sum_{j=1}^{M} u_j = v$

Thus v = n by (4). The dual problem therefore reduces to

maximise h (u) =
$$\begin{pmatrix} M \\ j=1 \end{pmatrix} \begin{pmatrix} u_j \\ c_j \end{pmatrix}^u b$$
 β^n

s.t.
$$\sum_{j=1}^{M} u_{j} = 1$$

$$u_{j} > 0 \quad j=1, ... M$$
 (7)

For $\sum_{i=1}^{N} \alpha_i = 1$, the equalities (6) when added yield $\sum_{j=1}^{N} u_j = 1$

and hence the constraint (4) can be dropped.

Thus for given $\alpha_i > 0$, $\sum_{i=1}^{N} \alpha_i = 1$, (III) is equivalent to

maximize
$$h(\underline{u}) = \begin{pmatrix} M & \left(\frac{u_j}{c_j}\right)^{-u_j} \end{pmatrix} \beta^n$$
s.t
$$\sum_{j=1}^{M} b_{pj} u_j - n\alpha_p = 0, \quad p=1,2,\dots, N.$$
(IV)
$$u_j > 0 \qquad j=1,2,\dots, M.$$

The primal program (II) is super consistent as defined in Duffin, Peterson and Zener (P 80, 1). This is to say that there exists at least one vector (t_1, t_2, \dots, t_N) with $t_j > 0$, $j=1,2,\dots,N$ such that β $t_1^{-\alpha}1$ $t_2^{-\alpha}2$ \dots $t_N^{-\alpha}N$ (1.

We can take, for example all $t_j = 2$, j = 1,2, ... N.

Then
$$\beta t_1 \qquad t_2 \qquad \cdots t_N = \beta (2^{-1}) = \alpha_1^{\alpha_1} \alpha_2^{\alpha_2}$$

$$\cdots \qquad a_N^{\alpha_N} / 2 \leqslant 1/2$$
Since
$$\alpha_j > 0 \qquad \alpha_1^{\alpha_1} \alpha_2^{\alpha_2} \qquad \alpha_N^{\alpha_N} = 1$$

$$E\alpha_s = 1$$

Further we can find a point $u = (u_1, u_2, ..., u_M)$ with positive components that satisfies the constraints (6).

Take for example uj's such that

$$u_{j} = \rho_{for j} = (N+1), \dots M$$

$$nu_{p} = n\alpha_{p} - \rho_{j=N+1}^{M} b_{pj}, p=1,2, \dots, N$$

$$\rho$$
 >0 is chosen such that
$$\rho < \min_{\mathbf{p}} \left(\begin{array}{c} \mathbf{m} \\ \mathbf{n} \\ \mathbf{p} \end{array} \right)$$

$$\mathbf{p} \neq \sum_{\mathbf{j}=\mathbf{N+1}}^{\mathbf{M}} \mathbf{b}_{\mathbf{p}\mathbf{j}} > 0$$

By Theorem 2, p 120 of Duffin, Peterson and Zener [1], the primal function $P(t_1, \ldots, t_N)$ attains its constrained minimum value at a point, say (t_1^*, \ldots, t_N^*) satisfying the constraint of (II). The conditions of the first duality theorem of geometric programming (Theorem 1. P.117 of Duffin, Peterson and Zener [1] are satisfied. Hence we can assert that

(i) In the dual program (IV), the dual function h (u) attains its constrained maximum value at a point (u₁*, ..., u_M*) satisfying (6) and (7).

(ii) The constrained maximum value of the dual function is equal to the constrained minimum value of the primal function. $h(\underline{u}^*) = P(t_1^*, t_2^*, \dots, t_N^*)$ (8)

(iii)
$$h(u^*) u_j^* = c_j t_1^* t_2^* \cdots t_N^*; j = 1,2, \dots M$$
 (9)

We have $h(u^*) > 0$ so that from (9) we have $u_j^*/c_j = \begin{pmatrix} N & *^b ij \\ I & 1 \end{pmatrix} / h(u^*); j = 1,2,...M$ (10)

$$= \left(\frac{\prod_{i=1}^{N} t_{1}^{b_{ij}}}{\prod_{i=1}^{N} t_{1}^{b_{ij}}}\right) P(t_{1}^{*}, \dots, t_{N}^{*}), j=1,2, \dots M$$
 (11)

Since $b_{11} = b_{22} = \cdots = b_{NN} = n$ and $b_{ij} = 0$ for

 $i \neq j$ and $i = 1, 2, \ldots, N, j=1, 2, \ldots, N$ we have

$$u_p^*/c_p = t_p^{*n} /h(u^*); p=1,2,...N$$
 (12)

Hence
$$t_p^* = \left\{ h (u^*) u_p^* / c_p \right\}^{1/n}$$
, $p=1,2,..., N$

From (10) and (12) we have therefore

$$u_{j}^{*}/c_{j} = \begin{bmatrix} N & (u_{p}^{*}/c_{p})^{b_{p}j/n} \\ \frac{1}{p-1} & (u_{p}^{*}/c_{p})^{b_{p}j/n} \end{bmatrix} h(u_{p}^{*}) p=1 \quad p_{j}^{N} \quad h(u_{p}^{*})$$

$$= \underbrace{\pi}_{p=1}^{N} \left(\frac{u_p}{c_p} \right)^{b_p j/n} \qquad j = (N+1), \dots, M$$

since
$$\sum_{p=1}^{N} b_{pj} = n$$
. Since $u_p^*/c_p = \sum_{p=1}^{N} \left(\frac{u_p^*}{c_p^*}\right)^{\frac{b_{pj}}{n}}$ for $p=1,...,$

we can state generally that

$$u_{j}^{*}/c_{j} = \prod_{p=1}^{N} \left(\frac{u_{p}^{*}}{c_{p}^{*}}\right)^{p} j^{/n} \qquad j=1, \dots M \qquad (13)$$

From (III) we have

$$h (u^*) = \prod_{j=1}^{M} \left(\frac{u_j}{c_j} \right)^{u_j} \beta^n$$

$$= \beta^n \prod_{p=1}^{N} \left(\frac{u_p}{c_p} \right)^{-\frac{N}{j=1}} u_j^* \frac{b_{pj}}{n}$$

$$= \beta^n \prod_{p=1}^{N} \left(\frac{u_p}{c_p} \right)^{-\alpha} p \quad \text{by (6)}$$
Thus $h(u^*) = \prod_{p=1}^{N} \left(\frac{u_p}{c_p} \right)^{\alpha} \alpha_p$

$$: \frac{\mathbb{N}}{p=1} \left[\left(\frac{\mathbf{t}_{p}^{*n}}{h(\mathbf{u}^{*})} \right)^{-\alpha} p \quad n\alpha \\ \alpha p \quad 1 \right]$$

=
$$h(u^*)$$
 $\prod_{p=1}^{N} \left(\frac{\alpha_p}{t_p^*}\right)^{n\alpha_p}$ since $\sum_{p=1}^{N} \alpha_p = 1$.

Since
$$h(\underline{u}^*) > 0$$
 we have $\prod_{p=1}^{N} \left(\frac{\alpha_p}{t_p^*}\right)^{n\alpha_p} = 1$

Hence
$$1 = \prod_{p=1}^{N} \left(\frac{\alpha_p}{t_p}\right)^{\alpha_p} = \beta t_1^* \quad \bullet \bullet \bullet \quad t_p^*$$

Further, substituting u_j^* as given by (11) in (6), we have

$$\sum_{j=1}^{M} b_{pj} \left(c_{j} \prod t_{i}^{*} \right) / P(t_{1}^{*}, \dots, t_{N}^{*}) - n\alpha_{p} = 0, p = 1, \dots, N.$$

Using the definition of $P(t_1, \cdot \cdot \cdot, t_N)$ this can be rewritten as

$$\frac{\partial P(t_{1}, \dots, t_{N})}{\partial t_{p}} = \frac{n\alpha_{p}}{t_{p}^{*}} \qquad P(t_{1}^{*}, \dots, t_{N}^{*}) = 0,$$

$$t_{t=t_{N}^{*}}$$

$$p=1, \dots, N. \quad (15)$$

Since the infimum is attained in (II) and also because of (14) there exists an optimal solution to (II) with $\alpha_p = t_p$, p=1, ... N. For the choice of α_p : $\alpha_p = t_p$, p=1, ... N_i is optimal for (II) and hence we have from (15) that

$$\frac{\partial P(t_1, \dots, t_N)}{\partial t_p} = n P(t_1^*, \dots, t_N^*)$$

$$t = t^*$$

$$t = t^*$$

The theorem is proved.

Minimum of a Permanent of doubly stochastic matrix and reversed geometric programs

In this section we formulate the problem of finding the minimum of the permanent of a nxn doubly stochastic matrix as a reversed geometric programming problem.

It is well known that a $n \times n$ doubly stochastic matrix A_n can be expressed as a convex combination of the n! permutation matrices $\cdot i_n e_n$

$$A_n = t_1 D_1 + t_2 D_2 + \cdots + t_N D_N$$

with N = n, D_1 , D_2 ... D_N are the permutation matrices

of order n ; $\mathbf{t_j}$, $\mathbf{j=1}$, ... n are such that $\mathbf{t_j} \geqslant \mathbf{0}$ and

 Σ t_{j=1}. The permanent of A_n can then be expressed in j=1 j=1. The permanent of A_n can then be expressed in terms of the t's say $P(t_1, \dots t_N) = \sum_{j=1}^{M} L_j (t_1, \dots t_N)$ with $M = \begin{pmatrix} (x+1) & 1 \\ \end{pmatrix} N$ and $L_j(t_1, \dots t_N) = \prod_{j=1}^{N} t_j$ by if (b_{pj}) satisfy the properties that $\sum_{p=1}^{N} b_{pj} = n$ for each

j=1,... M and they satisfy a further property that

M Σ b_{pj} is the same for each p = 1,2, .. N.

Since $\sum_{p=1}^{N} \sum_{j=1}^{N} b_{pj} = Mn$, this stated property implies

that

$$\overset{M}{\Sigma} \overset{b}{b}_{j=1} = Mn/N \text{ for each } p = 1, ... N.$$
(16)

Further $b_{11} = b_{22} = \cdots = b_{NN} = n$.

Since the permanent is a posynomial belonging to the class of posynomials discussed in section 2, the results proved in that section are applicable here. It is also to be noted that the permanent satisfies the further property (16) which

seems to be important in the settling of van der Waerden Conjecture. Stated in the dual form, corresponding to (II), the problem of finding the minimum of a permanent is given by the following min max problem subject to linear constraints:

minimize maximize
$$g(\underline{u}, \alpha) = \alpha_1 \quad \alpha_2 \quad \alpha_N \quad \prod_{j=1}^{M} (u_j)^j$$

subject to

$$\sum_{j=1}^{M} p_{j} u_{j} - n\alpha_{p} = 0, p = 1, 2, ... N$$
 (17)

$$\sum_{p=1}^{N} \alpha_p = 1 ; \alpha_p > 0 p=1, ... N$$

Note that min $\max_{\alpha} g(\underline{u}, \alpha)$ subject to $\sum_{p=1}^{N} \alpha_p = 1$;

N/n with the optimal solution corresponding to $\alpha_i \equiv \frac{1}{N}$; $u_j \equiv \frac{1}{M}$. This solution corresponds to the doubly stochastic matrix corresponding $t_i = \frac{1}{N}$, i=1,...N which is the J_n matrix

with all elements equal to $\frac{1}{n}$. If the van der Waerden conjecture is true, the structure of the matrix $B = (b_{pj})_{NXM}$ will have to be exploited to show that the constraints (17) could be dropped without changing the dual problem (V).

We shall illustrate the ideas developed so far in verifying the van der Waerden conjecture for n = 2 and in discussing the case n = 3.

for $t_1 > 0$, $t_2 > 0$ and $t_1 + t_2 = 1$. The permanent equals $t_1^2 + t_2^2$. The reversed geometric programming problem is therefore

minimize
$$t_1^2 + t_2^2$$

s.t $t_1 + t_2 > 1$
 $t_1 > 0$ $t_2 > 0$

This problem is equivalent to :

inf
$$t_1^2 + t_2^2$$

$$-\alpha_1 - \alpha_2$$

$$\beta t_1 t_2 \leq 1$$

$$t_1>0, t_2>0$$

with
$$\beta = \alpha_1^{\alpha_2} \quad \alpha_2^{\alpha_2}$$
; $\alpha_1 > 0 \quad \alpha_2 > 0$, $\alpha_1 + \alpha_2 = 1$

For a given α_1 , α_2 , the dual is

maximise
$$h(u_1, u_2) = u_1 - u_2 \beta$$

 $u_1 + u_2 = 1$
s.t. $2u_1 - \alpha_1 v = 0$
 $2u_2 - \alpha_2 v = 0$

Thus
$$\mathbf{v}=2$$
; $\mathbf{u}_1=\alpha_1$, $\mathbf{u}_2=\alpha_2$ for which
$$\mathbf{h}(\mathbf{u}_1,\mathbf{u}_2)=\beta=\alpha_1^{\alpha_1}\alpha_2^{\alpha_2}$$

and this takes minimum value for $\alpha_1 = \alpha_2 = 1/2$ and the corresponding value of $\beta = \frac{1}{2} = \frac{2!}{2^2}$

For the case n = 3, we give below the expression of the permanent in the posynomial form and give the dual problem

which finds the minimum value of the permanent.

Case n = 3

A doubly stochastic matrix has the form

with
$$t_j > 0$$
, $\sum_{j=1}^{6} t_j = 1$

$$= \begin{pmatrix} t_1 + t_2 & t_3 + t_5 & t_4 + t_6 \\ t_3 + t_6 & t_1 + t_4 & t_2 + t_5 \\ t_4 + t_5 & t_2 + t_6 & t_1 + t_3 \end{pmatrix}$$

The permanent is a posynomial:

$$P(t_1, t_6) = t_1^3 + t_2^3 + t_3^3 + t_4^3 + t_5^3 + t_6^3 + t_1^2 t_4$$

$$+ t_1^2 t_3 + t_1^2 t_2 + t_1 t_2^2 + t_2^2 t_6 + t_2^2 t_5$$

Here M = 48; N = 6. The problem of finding the minimum of $P(t_1, \dots, t_6)$ is equivalent to

minimize maximize
$$g(u, \alpha) = \begin{pmatrix} 48 & -u_j \\ \frac{\pi}{j-1} & u_j \end{pmatrix} \beta^3$$
s.t.
$$\sum_{j=1}^{48} u_j = 1$$

$$\sum_{p=1}^{N} \alpha_{p} = 1$$

$$\alpha_{p} > 0 \qquad p = 1, 2, ... N$$

A lower bound for the permanent of a doubly stochastic matrix:

Using the dual formulation (V) we can immediately give a lower bound for the permanent of a n x n doubly stochastic square matrix. For a given α_p

 $\sum_{p=1}^{N} p^{2} = 1, \quad \alpha_{p} > 0 \quad \text{it can be verified that } u_{p} = \alpha_{p} \quad p = 1, \dots N$ and $u_{j} = 0$ for $j = (N+1), \dots$ M is feasible for the constraints (17).

for every $\{\alpha_p\}$ such that $\alpha_p>0$, p=1,... N and $\Sigma\alpha_p=1$.

Thus
$$\min_{\alpha} \max_{u} g(u, \alpha) > \frac{1}{N^{n-1}}$$
. Of course this bound

$$-(n-1)$$

N is smaller than the conjectured bound N/nn.

4. A question was raised whether van der Waerden Conjecture is true if we allow negative entries in the matrix whose row sums and column sums are equal to unity. The following example shows that the permanent of such a matrix can be negative.

Example 1 : Consider the 3 x 3 matrix

The rwo sums and column sums are equal to unity. The permanent is however equal to - 1154 which is less than the conjectured bound.

REFERENCES

- 1. R.J. Duffin, E.L. Peterson and C. Zener Geometric Programming - Theory and Application. John Wiley and Sons, Inc, New York 1967.
- D. London: Some notes on the van der Waerden
 Conjecture, Linear Algebra and its applications 4
 (1971) 155-160.
- 3. Marvin Marcus and Morris Newman: On the minimum of the permanent of a doubly stochastic matrix. Duke Mathematics Journal, 26 (1959) 61-72.
- 4. B.L. van der Waerden, Aufgabe 45 Jber. Deutschen Math. Verein 35 (1926) 117.