Petr Mandl

Decomposable non-negative matrices in a dynamic programming problem

Czechoslovak Mathematical Journal, Vol. 20 (1970), No. 3, 504–510

Persistent URL: `http://dml.cz/dmlcz/100978`

Czechoslovak Mathematical Journal, 20 (95) 1970, Praha

DECOMPOSABLE NON-NEGATIVE MATRICES IN A DYNAMIC PROGRAMMING PROBLEM

PETR MANDL, Praha

(Received July 24, 1969)

1. In control theory a linear multistage decision process is described by a set $S = \{M(u) : u \in U\}$ of matrices depending on a control parameter u. $M(u)$ is the transformation of the state vector $\mathbf{v}$ for the selected value u of the parameter. Thus, if $\mathbf{v}_n$ is the n-th state of the system and u_n the n-th chosen parameter value, then $\mathbf{v}_{n+1} = M(u_n)\,\mathbf{v}_n$. In [2] E. SENETA and the present author studied the minimal convergence radius of the series

$$\sum_{n=0}^{\infty} \mathbf{v}_n z^n \tag{1}$$

in the case when S is a set of indecomposable non-negative matrices. The reciprocal value of the radius is the maximal growth rate attainable by the system. The present paper deals with the general case in which the convergence radiuses need not be the same for all components.

Bold types denote r-dimensional column vectors, e.g., $\mathbf{x} = (x^1, \ldots, x^r)'$. The inequality $\mathbf{x} \geqq \mathbf{y}$ is to be understood componentwise, $\|x\| = \sup_{i=1,\ldots,r} |x^i|$. Capital letters M, N, O denote matrices. We write $M = \|m_{ij}\|_{i,j=1}^r$, $M^n = \|m_{ij}^{(n)}\|_{i,j=1}^r$; $M^0 = \|\delta_{ij}\|_{i,j=1}^r$. $i \underset{M}{\rightarrow} j$ means $m_{ij}^{(n)} > 0$ for some n.

Throughout the paper S will be a non-void set of non-negative $r \times r$ matrices which fulfils the following assumptions.

1. S is closed and bounded.

2. There exists a closed subset $\tilde{S} \subset S$ such that for arbitrary $\mathbf{e} \geqq \mathbf{0}$, $M, N \in S$ an $O \in \tilde{S}$ can be found for which

$$O\mathbf{e} \geqq M\mathbf{e}\,, \quad O\mathbf{e} \geqq N\mathbf{e}\,.$$

Whenever Assumption 2 holds it holds for $\tilde{S} = S$. $\tilde{S}$ will be called a basis of S. It is appropriate to call Assumption 2 *the dynamic programming assumption*, since assumptions of this type are typical for dynamic programming problems. Thus we assume that S is a compact set of non-negative matrices satisfying the dynamic programming assumption. Theorem 1 states that its basis contains a uniformly optimal matrix. According to Theorem 2, the growth rate of the powers of a uniformly optimal matrix cannot be surpassed by multiplying matrices from the convex closure of S.

2. For $M \in S$ denote by $R_{ij}(M)$ the convergence radius of $\sum_{n=0}^{\infty} m_{ij}^{(n)} z^n$. Further let

$$R_i(M) = \inf_{j=1,\dots,r} R_{ij}(M), \quad R(M) = \inf_{i=1,\dots,r} R_i(M),$$

$$\hat{R}_i = \inf_{M \in S} R_i(M), \quad \hat{R} = \inf_{i=1,\dots,r} \hat{R}_i = \inf_{M \in S} R(M).$$

To establish the Theorems concerning the quantities $\hat{R}_i$ we shall need some auxiliary assertions. Let J denote the $r \times r$ matrix consisting of ones. For $\varepsilon > 0$ introduce the sets

$$S_\varepsilon = \{M + \varepsilon J : M \in S\} \tag{2}$$

and denote $\hat{R}_\varepsilon = \inf_{N \in S_\varepsilon} R(N)$. S_ε is a compact set of indecomposable matrices satisfying the dynamic programming assumption. Hence, (see [2], Theorem 3.1) S_ε posseses a global $\hat{R}_\varepsilon$-subinvariant vector. I.e., there exists a strictly positive vector $\mathbf{y}_\varepsilon$ satisfying

$$\hat{R}_\varepsilon N \mathbf{y}_\varepsilon \leqq \mathbf{y}_\varepsilon, \quad N \in S_\varepsilon. \tag{3}$$

Moreover, for $N \in S_\varepsilon$, $\hat{R}_\varepsilon N \mathbf{y}_\varepsilon = \mathbf{y}_\varepsilon$ if and only if $R(N) = \hat{R}_\varepsilon$.

Lemma 1.

$$\lim_{\varepsilon \to 0} \hat{R}_\varepsilon = \hat{R}.$$

Proof. Obviously $\hat{R}_\varepsilon$ increases as $\varepsilon \to 0$ and $\hat{R}_\varepsilon \leqq \hat{R}$. Thus, the limit

$$\lim_{\varepsilon \to 0} \hat{R}_\varepsilon = \bar{R} \leqq \hat{R}$$

exists. Assume $\bar{R} < \infty$. Take $\mathbf{y}_\varepsilon$ satisfying (3), $\mathbf{y}_\varepsilon \geqq 0$, $\|\mathbf{y}_\varepsilon\| = 1$. Further, let N_ε be such that $\hat{R}_\varepsilon N_\varepsilon \mathbf{y}_\varepsilon = \mathbf{y}_\varepsilon$. For a suitable sequence ε_n, $n = 1, 2, \dots$, $\varepsilon_n \to 0$,

$$\lim_{n \to \infty} \mathbf{y}_{\varepsilon_n} = \mathbf{x}_0 \neq 0, \quad \lim_{n \to \infty} N_{\varepsilon_n} = {}_0M.$$

We have

$$\bar{R}\,{}_0M \mathbf{x}_0 = \mathbf{x}_0, \quad (\bar{R}\,{}_0M)^n \mathbf{x}_0 = \mathbf{x}_0,$$

and hence,

$$\sum_{n=0}^{\infty} \bar{R}^n \left(\sum_j {}_0 m_{ij}^{(n)} \mathbf{x}_0^j\right) = \infty \tag{4}$$

for all i, for which $x_0^i > 0$. This implies $\hat{R} \leqq \bar{R}$.

Let $\mathbf{x}_0$ be the same as in the proof of Lemma 1. Introduce the set

$$(0) = \{i : x_0^i > 0\} .$$

Note also that

$$\sup_S \hat{R} M \mathbf{x}_0 = \mathbf{x}_0 .$$

Lemma 2. *Let* $\hat{R} < \infty$. *Then*

a)

$$N\mathbf{x}_0 = \sup_S M\mathbf{x}_0 \tag{5}$$

only if

$$R_i(N) = \hat{R} = \hat{R}_i \quad \textit{for} \quad i \in (0) .$$

b) *If* $j \underset{M}{\rightarrow} i$ *for some* $M \in S$, $i \in (0)$, *then* $j \in (0)$.

Proof. To establish a) it is sufficient to repeat the reasoning from the end of the preceding proof since (5) implies

$$\hat{R} N \mathbf{x}_0 = \mathbf{x}_0 .$$

Further, let $M \in S$, $i \in (0)$, $j \underset{M}{\rightarrow} i$. There exists an n such that $m_{ji}^{(n)} > 0$. Then

$$0 < \hat{R}^n \sum_k m_{jk}^{(n)} x_0^k \leqq x_0^j .$$

Thus $j \in (0)$.

Theorem 1. *There exists an* $\hat{M} \in \tilde{S}$ *for which*

$$R_i(\hat{M}) = \hat{R}_i , \quad i = 1, \ldots, r .$$

Proof. The case $\hat{R} = \infty$ is trivial. Assume $\hat{R} < \infty$. For the purpose of the proof a sequence of sets beginning with (0) and a sequence of vectors beginning with $\mathbf{x}_0$ will be constructed. If (0) includes all the indices i with $\hat{R}_i < \infty$, the sequences contain only (0) and $\mathbf{x}_0$. Otherwise set

$$[1] = \{1, \ldots, r\} - (0) .$$

For $M \in S$ denote ${}_1M = \|m_{ij}\|_{i,j\in[1]}$. According to Lemma 2b, $m_{ij} = 0$ for $i \in [1]$, $j \in (0)$, and hence,

$$m_{ij}^{(n)} = \sum_{k_1\in[1]} \dots \sum_{k_{n-1}\in[1]} m_{ik_1}m_{k_1k_2} \dots m_{k_{n-1}j}, \quad n = 2, 3, \dots, \quad i, j \in [1],$$

$$m_{ij}^{(n)} = 0, \quad i \in [1], \quad j \in (0).$$

Consequently, $R_i({}_1M) = R_i(M)$, $i \in [1]$, $M \in S$. A fortiori

$$\hat{R}_i = \inf_{M\in S} R_i({}_1M), \quad i \in [1].$$

Let

$${}_1\hat{R} = \inf_{i\in[1]} \hat{R}_i < \infty.$$

We can apply the reasoning made in this Section to the sets

$${}_1S_\varepsilon = \{{}_1M + \varepsilon_1 J : M \in S\}$$

and construct a non-zero vector $\bar{\mathbf{x}}_1 = (\bar{x}_1^i, i \in [1])'$ satisfying

$$\sup_S {}_1\hat{R}\, {}_1M\bar{\mathbf{x}}_1 = \bar{\mathbf{x}}_1.$$

For i belonging to the set

$$(1) = \{i : \bar{x}_1^i > 0\},$$

we have then $\hat{R}_i = {}_1\hat{R}$. Define finally $\mathbf{x}_1 = (x_1^1, \dots, x_1^r)'$ by setting

$$x_1^i = \bar{x}_1^i \quad \text{for} \quad i \in [1], \quad x_1^i = 0 \quad \text{for} \quad i \notin [1].$$

Proceeding in the same manner until we exhaust all the indices i with $\hat{R}_i < \infty$ we arrive at a sequence $(0), (1), \dots, (l)$ of subsets of $\{1, \dots, r\}$ and a sequence of r-dimensional vectors $\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_l$ with the properties:

a) $i \in (h)$, $j \in (k)$, $i \underset{M}{\to} j$ for some $M \in S$ only if $h \leqq k$.

b) For $N \in S$,

$$\sum_j n_{ij}x_k^j = \sup_S \sum_j m_{ij}x_k^j, \quad i \in (k), \tag{6}$$

only if $R_i(N) = \hat{R}_i$, $i \in (k)$.

Introduce

$$\mathbf{z}_\varepsilon = \mathbf{x}_0 + \varepsilon\mathbf{x}_1 + \varepsilon^2\mathbf{x}_2 + \dots + \varepsilon^l\mathbf{x}_l.$$

In accordance with Assumption 2 there exists an $N(\varepsilon) \in \tilde{S}$ for which

$$N(\varepsilon)\, \mathbf{z}_\varepsilon = \sup_S M\mathbf{z}_\varepsilon\,.$$

For $i \in (k)$ we have

$$\sum_j n_{ij}(\varepsilon)\, z_\varepsilon^j = \varepsilon^k \sum_j n_{ij}(\varepsilon)\, x_k^j + o(\varepsilon^{k+1})\,.$$

Hence,

$$\sum_j n_{ij}(\varepsilon)\, x_k^j = \sup_S \sum_j m_{ij} x_k^j + o(\varepsilon)\,. \tag{7}$$

Let $\hat{M} \in \tilde{S}$ be such that

$$\lim_{n\to\infty} N(\varepsilon_n) = \hat{M}$$

for a suitable sequence ε_n, $n = 1, 2, \ldots$, $\varepsilon_n \to 0$. From (7) it follows that (6) holds with $n_{ij} = \hat{m}_{ij}$ for $k = 1, \ldots, l$. Consequently, $R_i(\hat{M}) = \hat{R}_i$ for $i \in (0) \cup \ldots \cup (l)$. Moreover, $R_j(\hat{M}) = \infty$ whenever $\hat{R}_j = \infty$. The Theorem is established.

Denote by $C(S)$ the convex closure of S.

Theorem 2. *Let M_n, $n = 0, 1, \ldots$ be an arbitrary infinite sequence of matrices from $C(S)$ and let $\mathbf{e} \geqq 0$. Define $\mathbf{v}_n$, $n = 0, 1, \ldots$, recursively by*

$$\mathbf{v}_0 = \mathbf{e}\,, \quad \mathbf{v}_{n+1} = M_n \mathbf{v}_n\,, \quad n = 0, 1, \ldots\,.$$

Then, for $i = 1, \ldots, r$, the convergence radius of

$$\sum_{n=0}^{\infty} v_n^i z^n \tag{8}$$

is greater or equal $\hat{R}_i$.

Proof. We shall show first that the convergence radius of

$$\sum_{n=0}^{\infty} \mathbf{v}_n z^n = \mathbf{e} + \sum_{n=0}^{\infty} \Big(\prod_{j=0}^{n} M_j\Big)\, \mathbf{e} z^n \tag{9}$$

is at least equal $\hat{R}$. Introduce again the classes S_ε, $\varepsilon > 0$, defined by (2) and let $\mathbf{x}_\varepsilon$ fulfil

$$\mathbf{x}_\varepsilon \geqq \mathbf{e}\,, \quad \hat{R}_\varepsilon N \mathbf{x}_\varepsilon \leqq \mathbf{x}_\varepsilon\,, \quad N \in S_\varepsilon\,.$$

Then also

$$\hat{R}_\varepsilon N \mathbf{x}_\varepsilon \leqq \mathbf{x}_\varepsilon \quad \text{for} \quad N \in C(S_\varepsilon)\,. \tag{10}$$

The sequence $\mathbf{v}_n$, $n = 0, 1, \ldots$, is majorized by $\mathbf{w}_n$, $n = 0, 1, \ldots$, satisfying the relations

$$\mathbf{w}_0 = \mathbf{x}_\varepsilon\,, \quad \mathbf{w}_{n+1} = (M_n + \varepsilon J)\,\mathbf{w}_n\,, \quad n = 0, 1, \ldots\,.$$

From (10) one obtains

$$\mathbf{v}_n \leqq \mathbf{w}_n \leqq \hat{R}_\varepsilon^{-n}\mathbf{x}_\varepsilon\,, \quad n = 0, 1, \ldots\,.$$

Consequently, (9) is convergent whenever $|z| < \hat{R}_\varepsilon \to \hat{R}$ for $\varepsilon \to 0$.

Consider now the set

$$[\bar{1}] = \{i : R_i > \hat{R}\}\,.$$

Let $[\bar{1}] \neq 0$. For some j,

$$[\bar{1}] = \{1, \ldots, r\} - [(0) \cup \ldots \cup (j)]$$

where (k), $k = 0, \ldots, j$, were defined in the proof of Theorem 1. For $M \in C(S)$ let ${}_{[\bar{1}]}M$ denote the matrix $\|m_{ij}\|_{ij\in[\bar{1}]}$. Since $m_{ij} = 0$ whenever $i \in [\bar{1}]$, $j \notin [\bar{1}]$, the values of $\mathbf{v}_n^i$, $i \in [\bar{1}]$, $n = 0, 1, \ldots$. remain unchanged if we start from the sequence ${}_{[\bar{1}]}M_n$, $n = 0, 1, \ldots$, instead of M_n, $n = 0, 1, \ldots$ Thus we establish by the above method that the convergence radius of (8) for $i \in [\bar{1}]$ is not less than

$${}_{[\bar{1}]}\hat{R} = \inf\{\hat{R}_i : i \in [\bar{1}]\}\,.$$

Proceeding to the set $[\bar{2}] = \{i : \hat{R}_i > {}_{[\bar{1}]}\hat{R}\}$ and so on until we arrive at a void $[\bar{k}]$ we conclude the proof of the Theorem.

Theorem 3. *Let* $\mathbf{e}$ *have all components positive,* $z > 0$. *The equation*

$$\mathbf{d} = \mathbf{e} + \sup_S zM\mathbf{d} \tag{11}$$

has a positive solution if and only if $z < \hat{R}$. *The solution for* $0 < z < \hat{R}$ *is unique and*

$$\sum_{n=0}^{\infty} z^n M^n \mathbf{e} \leqq \mathbf{d} \quad \textit{for} \quad M \in S\,. \tag{12}$$

We shall not give the proof of Theorem 3, since (11) is simply the Bellman equation for maximizing the left hand side of (12). More detailed proof of a similar theorem was given in [1].

References

[1] *P. Mandl:* An iterative method for maximizing the characteristic root of positive matrices. Rev. roum. math. pures et appl. *12* (1967), 1317—1322.

[2] *P. Mandl, E. Seneta:* The theory of non-negative matrices in a dynamic programming problem. Austral. J. of Stats. *11* (1969), 85—96.

Author's address: Vyšehradská 49, Praha 2, ČSSR (Ústav teorie informace a automatizace ČSAV).