Mathematical Programming Study 6 (1976) 188–201.
North-Holland Publishing Company

COMPUTATION OF THE EIGENPROJECTION OF A NONNEGATIVE MATRIX AT ITS SPECTRAL RADIUS*

Uriel G. ROTHBLUM**
Yale University, New Haven, Conn., U.S.A.

Received 16 July 1975
Revised manuscript received 10 December 1975

In this paper we give a general representation for a projection in terms of its range and the range of its adjoint projection. By combining this representation with recent results of the author on the structure of the algebraic eigenspace of a nonnegative matrix corresponding to its spectral radius, we develop a computational method to find the eigenprojection of a nonnegative matrix at its spectral radius. The results are illustrated by giving a closed formula for computing the limiting matrix of a stochastic matrix.

1. Introduction

In this paper we give a representation of projections in terms of their range and the range of their adjoint projection. We then apply this representation to develop a computational method to find the eigenprojections of nonnegative matrices at their spectral radius.

After introducing a few notational conventions in Section 2, we state the representation theorem in Section 3. It is shown that if E is a projection and the columns of a matrix X (resp., Y) form a basis of range E (resp., range E^*), then $E = X(Y^*X)^{-1}Y^*$. We then illustrate this representation for orthogonal projections and obtain some known results which are frequently used in estimation problems in statistics and in projection methods in nonlinear programming.

* Parts of the research reported in this paper are based on the author's Ph.D. dissertation submitted to the Department of Operations Research at Stanford University. Research at Stanford University was supported by NSF Grant GK–18339 and ONR Contract N00014–67–A–0112–0050. Further research at the Courant Institute was supported by NSF Grant GP–37069.

** My deepest thanks are given to my teacher and advisor, Professor Arthur F. Veinott, Jr., for his guidance and advice, for his insight and perspective, during the preparation of my Ph.D. dissertation.

Let P be a square matrix; the *eigenprojection* of P at an eigenvalue λ is the unique projection E whose range (resp., the range of its adjoint) is the algebraic eigenspace of P (resp., P^*) at λ (resp., λ^*). In Section 4 we apply the representation theorem to eigenprojections and then, in Section 5, develop an efficient method to compute the eigenprojection of a nonnegative matrix P at its spectral radius. We show how to construct, in a simple way, a matrix X (resp., Y) whose columns form a basis of the algebraic eigenspace of P (resp., P^*) at λ (resp., λ^*) such that Y^*X is triangular. Finally, in Section 6 we apply our results to stochastic matrices.

2. Notational conventions

Let $\mathbf{C}^S$ (resp., $\mathbf{R}^S$) be the S dimensional complex (resp., real) space. In Section 3 and 4 we consider the general complex case, whereas in Sections 5 and 6 our attention is restricted to real spaces. Coordinates of matrices and vectors will be denoted by subscripts. For a finite set J, let $|J|$ stand for the number of elements in J.

Given a matrix $B \in \mathbf{C}^{S\times M}$, let B^* be its *adjoint matrix* (i.e., complex conjugate transpose of B). In particular, if B is a scalar (resp., column vector), B^* is the complex conjugate (resp., complex conjugate row vector) of B. Throughout the paper lower case letters are used for (column) vectors. We will use the standard notation null B and range B for $\{x \in \mathbf{C}^M \mid Bx = 0\}$ and $\{Bx \mid x \in \mathbf{C}^M\}$, respectively. Let $J \subseteq \{1,\ldots,S\}$ and $K \subseteq \{1,\ldots,M\}$. Then by $B_{JK} \in \mathbf{C}^{|J|\times|K|}$ we denote the corresponding submatrix of rows and columns of B. If $J = K$, let $B_J \equiv B_{JJ}$. For $x \in C^S$ and $J \subseteq \{1,\ldots,S\}$ we denote by $x_J \in C^{|J|}$ the corresponding subvector of x. We say that the matrix B is *nonnegative* (resp., *positive*) written $B \geqslant 0$ (resp., $B \gg 0$) if all of its coordinates are nonnegative (resp., positive), We say that B is semipositive, written $B > 0$, if $B \geqslant 0$ and $B \neq 0$. Similar definitions apply to vectors. The matrix B is called a *projection* if B is square and $B^2 = B$. We say that B is a *projection on M along N* (where $M, N \subseteq \mathbf{C}^S$ are subspaces) if range $B = M$ and null $B = N$. The *spectral radius* of a square matrix B, $r(B)$, is the biggest modulus of an eigenvalue of B.

The *orthogonal complement* of a given subspace $M \subseteq \mathbf{C}^S$ is defined by $M^\perp \equiv \{x \mid y^*x = 0 \text{ for all } y \in M\}$. The *dimension* of M will be denoted dim M.

The vector $(1,1,\ldots,1)$ will always be denoted by e. The dimension of this vector will always be clear from the content. A vector x is called a *probability* *vector* if $x \geq 0$ and $e^*x = 1$. Finally, given a sequence $\{a_0, a_i, \ldots\}$, we say that $\lim_{N\to\infty} a_N = a \ \ (C, 1)$, if $\lim_{N\to\infty} \sum_{i=0}^{N} a_i/N = a$.

3. A representation theorem for projections

In this section we develop a representation for projections in terms of their range and the range of their corresponding adjoint projection. We then illustrate some important special cases of this representation.

It is well known that a projection is uniquely determined by its range together with its null space. Moreover, if M and N are subspaces of $\mathbf{C}^S$, then there exists a projection on M along N if and only if $\dim M + \dim N = S$ and $M \cap N = \{0\}$, and such a projection is unique. Recalling that for every square matrix A, $\operatorname{range} A^* = (\operatorname{null} A)^\perp$, we see that a projection is uniquely determined by its range and the range of its adjoint projection. In addition, if M and N are subspaces of $\mathbf{C}^S$, then there exists a projection E with $\operatorname{range} E = M$ and $\operatorname{range} E^* = N$ if and only if $\dim M = \dim N$ and $M \cap N^\perp = \{0\}$.

The range of a projection is a subspace. We have seen that the dimension of this subspace is the same as that of the adjoint projection. Let E be a given nonvanishing projection and let $x_1, \ldots, x_m$ (resp., $y_1, \ldots, y_m$) be a basis for $\operatorname{range} E$ (resp., $\operatorname{range} E^*$). Let X (resp., Y) be the $S \times m$ matrix whose j-th column is x_j (resp., y_j). Obviously

$$\operatorname{range} E = \{Xz \mid z \in \mathbf{C}^m\}, \tag{3.1}$$

and

$$\operatorname{range} E^* = \{Yt \mid t \in \mathbf{C}^m\}. \tag{3.2}$$

We next show that Y^*X is a nonsingular $m \times m$ matrix. If $Y^*Xz = 0$, then for every $t \in \mathbf{C}^m$, $t^*Y^*Xz = 0$. By (3.2), this means that $Xz \in (\operatorname{range} E^*)^\perp = \operatorname{null} E$, and by (3.1) $Xz \in \operatorname{range} E$. Hence $Xz = EXz = 0$. The independence of the columns of X implies that $z = 0$.

The next result gives a representation of a projection in terms of its range and the range of its adjoint projection. This result appears in [6, p. 25]. We repeat the proof for completeness.

Theorem 3.1. *Let E be a given nonvanishing projection and let $X = (x_1, \ldots, x_m)$ and $Y = (y_1, \ldots, y_m)$ where the columns of X and Y form bases of* $\operatorname{range} E$ *and* $\operatorname{range} E^*$, *respectively. Then*

$$E = X(Y^*X)^{-1}Y^*. \tag{3.3}$$

Proof. Let $z \in \mathbf{C}^S$. Obviously, $Ez \in \text{range}\, E$; so by (3.1) there exists a vector $w \in \mathbf{C}^m$ such that $Ez = Xw$. To determine w we remark that $E^*Y = Y$. Hence $Y^*z = Y^*Ez = Y^*Xw$. Since Y^*X is nonsingular, $w = (Y^*X)^{-1}Y^*z$, and therefore $Ez = X(Y^*X)^{-1}Y^*z$. Since z was chosen arbitrarily, the proof of Theorem 3.1 is completed.

We remark that a special case of the above theorem and the method of proof were given by Debreu and Herstein [1].

An immediate result of the above representation theorem is that a projection is real if and only if its range as well as the range of its adjoint projection have real bases.

A projection E is called *orthogonal* if $(\text{range}\, E)^\perp = \text{null}\, E$. It is easily seen that a projection E is orthogonal if and only if it is *hermitian*, i.e., $E = E^*$. This follows immediately from the observations that $(\text{range}\, E)^\perp = \text{null}\, E^*$, $\text{null}\, E = (\text{range}\, E^*)^\perp$ and the fact that a projection is uniquely determined by its range and null space.

If P is an orthogonal projection and $X = (x, \ldots, x_m)$, where the columns of X form a basis of range E, then these columns are also a basis of range E^*. Theorem 3.1 now implies that

$$E = X(X^*X)^{-1}X^*. \tag{3.4}$$

This representation is well known (e.g., [4, Theorem 4.4.1, p. 74]) and is frequently used in estimation problems in statistics and projection methods in nonlinear programming.

We will next extend the concept of orthogonal projections to the case where the scalar product is not necessarily Euclidean. Let V be an $S \times S$ hermitian positive definite matrix. The *V-inner product* of two vectors $x, y \in \mathbf{C}^S$ is defined by $(x, y)_V \equiv x^*Vy$. We require that V be hermitian and positive definite so that $(\cdot, \cdot)_V$ will really be an inner product. A projection E is called *V-orthogonal* if $(x, y)_V = 0$ for every $x \in \text{null}\, E$ and $y \in \text{range}\, E$.

Let V be an $S \times S$ positive definite matrix and E be a V-orthogonal projection. Then

$$\text{range}\, E^* = (\text{null}\, E)^\perp \subseteq \{Vy \mid y \in \text{range}\, E\}. \tag{3.5}$$

Recalling that V is nonsingular (V is positive definite) it follows from dimension calculations that the inclusion in (3.5) can be replaced by equality. The nonsingularity of V now implies that if $x_1, \ldots, x_m$ is a basis of range E, then $Vx_1, \ldots, Vx_m$ form a basis of range E^*. Applying Theorem 3.1 we get that if $X = (x_1, \ldots, x_m)$, then

$$E = X((VX)^*X)^{-1}(VX)^* = X(X^*VX)^{-1}X^*V. \tag{3.6}$$

This representation is well known (e.g., [4, Theorem 4.4.1, p. 74]) and is frequently used in estimation problems in statistics and projection methods in nonlinear programming.

We finally remark that every projection has a representation similar to the one given in (3.6). Let E be a given projection, and let X (resp., Y) be a matrix whose columns form a basis of range E (resp., range E^*). It can easily be seen that there exists a square nonsingular matrix L such that $L^*X = Y$. By Theorem 3.1

$$E = X(Y^*X)^{-1}Y^* = X(X^*LX)^{-1}X^*L. \tag{3.7}$$

If L is hermitian, one can consider the *L-improper inner product* defined by $(x, y)_L \equiv x^*Ly$ (see, [9, Ch. IX, pp. 215–236]). Defining an *L-orthogonal projection* as a projection F for which $(x, y)_L = 0$ for all $x \in \text{range}\, F$ and $y \in \text{null}\, F$, it is easily seen that E is an L-orthogonal projection.

Obviously, one would like to choose a basis of range E such that X^*X (resp., X^*VX) in (3.4) (resp., (3.6)) is the identity. This would save the inversion of a matrix in the corresponding representation of E. We remark that this is possible by a simple Gram–Schmidt orthonormalization of any given basis. This procedure might fail in trying to diagonalize X^*LX in (3.7).

In general, Y^*X in (3.3) can also be diagonalized by replacing X by $X(Y^*X)^{-1}$. It is easily seen that if the columns of X form a basis of range E, then so do the columns of $X(Y^*X)^{-1}$. We remark that this method is not practical for finding a good pair of bases of, respectively, range E and range E^*.

4. Algebraic eigenspaces and eigenprojections

An interesting example of projections which are not necessarily orthogonal are eigenprojections. Let us first define these projections.

Let P be a square complex matrix and let λ be a complex number. The *algebraic eigenspace* of P at λ is defined by

$$N_\lambda(P) \equiv \bigcup_{m=0}^{\infty} \text{null}\,(P - \lambda I)^m.$$

Elements of this space are called *algebraic eigenvectors* of P at λ. The smallest integer n such that $(P-\lambda I)^{n+1}x=0$ implies $(P-\lambda I)^n x=0$ is called the *index* of λ for P, written $\nu_\lambda(P)$. It is known (e.g., [10, p. 27]) that there exists a projection $E_\lambda(P)$, called the *eigenprojection* of P at λ, for which

$$\operatorname{range} E_\lambda(P)=N_\lambda(P) \quad \text{and} \quad \operatorname{range} E_\lambda(P)^*=N_{\lambda^*}(P^*). \tag{4.1}$$

By the discussion at the beginning of Section 3, (4.1) uniquely determines $E_\lambda(P)$. We remark that $E_\lambda(P)\neq 0$ if and only if λ is an eigenvalue of P.

One can easily see that if P is hermitian and λ is real, then $E_\lambda(P)$ is hermitian and therefore orthogonal. The reverse is not necessarily true. For example, let

$$P=\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}.$$

Since $(P-I)^2=0$, $N_1(P)=\mathbf{C}^2$ and therefore $E_1(P)$ is the identity. Also observe that no $\lambda\neq 1$ is an eigenvalue of P, and therefore $E_\lambda(P)=0$ for all $\lambda\neq 1$. We see that all the eigenprojections of P are hermitian though P is not.

Eigenprojections are important in the computation of the Laurent expansion of the resolvent of a matrix as well as obtaining asymptotic expansions of its powers (see [10, Ch. III]). We will next develop a method to compute these eigenprojections by using Theorem 3.1.

By Dunford and Schwartz [2, VII, 1.2, p. 556] $\nu_\lambda(P)\leq S$. Thus, in finding a basis for $N_\lambda(P)$ (resp., $N_{\lambda^*}(P^*)$) is equivalent to finding a basis of the set of solutions to the homogeneous system

$$(P-\lambda I)^S x=0 \tag{4.2}$$

(resp.,

$$y^*(P-\lambda I)^S=0). \tag{4.3}$$

We point out that if P and λ are real, it suffices to find real bases of the set of real solutions to (4.2) and (4.3).

We next describe a method to find a basis to the solution set of (4.2) which does not require to raise $(P-\lambda I)$ in the S-th power. A similar procedure can be applied toward the solution set of (4.3). First find a basis $x(1),\ldots,x(k_1)$ of $\{x \mid (P-\lambda I)x=0\}$. We next find a basis of $\{x \mid (P-\lambda I)^2x=0\}$. This is done by adding to $x(1),\ldots,x(k_1)$ vectors from

$$\bigcup_{i=1}^{k_1}\{x \mid (P-\lambda I)x=x(i)\}.$$

By continuing this method one can find a basis to the solution set of (4.2). The details are omitted. A more detailed development of the above ideas, for the case when P is nonnegative and λ is its spectral radius, appears in Sections 5 and 6.

5. Eigenprojections of nonnegative matrices corresponding to their spectral radius

In this section we develop a method for computing the eigenprojection of a nonnegative matrix at its spectral radius. We remark that by the Perron–Frobenius Theorem (e.g., [14, p. 46]) the spectral radius of a nonnegative matrix is always one of its eigenvalues; hence the corresponding eigenprojection is never degenerate.

The eigenprojection of a nonnegative matrix P at $r \equiv r(P)$ is of special interest. In particular if $r > 0$ and $\nu_r(P) = 1$, then $\lim_{N\to\infty} r^{-N}P^N = E_r(P)$ $(C, 1)$ (e.g., [12, Section 3]). The most familiar examples of this situation are stochastic matrices (for which $r = 1$) and irreducible matrices. Asymptotic expansions of $r^{-N}P^N$ as $N \to \infty$, in terms of $E_r(P)$, were obtained for matrices P with $\nu_r(P) > 1$ (and still $r > 0$) in [12]. These asymptotic expansions have many applications in economics, the theory of Markov chains and dynamic programming. To quote just a few references see [7, 3, 1, 5, 8, 13 and 15].

If P is a nonnegative matrix with spectral radius r and $\nu_r(P) = 1$, then $E_r(P)$, as the (Cesaro average) limit of a sequence of nonnegative matrices, is also nonnegative. This result is not necessarily true when $\nu_r(P) > 1$. A result which generalizes the nonnegativity of $E_r(P)$ for this case is discussed in [10, Section V, Lemma 7.1].

Two difficulties arise in the attempt to use (3.3) to compute a given projection. First, one has to find bases of range E and range E^*, respectively, and second, one has to invert the matrix Y^*X. We will next show a simple method to find such bases for which Y^*X is triangular. Moreover, for the important case in which $\nu_r(P) = 1$, Y^*X will be diagonal. Obviously this will simplify the inversion required in (3.3). We will illustrate our results for stochastic matrices in Section 6.

Before presenting the main results of this section we introduce a few definitions concerning nonnegative matrices. Let P be a square nonnegative matrix. We say that *i has access to j* or that *j has access from i* if for some integer $n \geqslant 0$, $(P^n)_{ij} > 0$. Two states each having access to the other are said to *communicate*. The communication relation is an equivalence relation, hence we may partition the totality of states into *equivalence classes* of communicating states. A class J of P is called *basic* if the spectral radius of P_J equals that of P. We will use the notation $\mathcal{A}(J)$ (resp., $\mathcal{A}^*(J)$) for the set of states having access to (resp., from) some, or equivalently every, state in the class J. We say that a state i has *direct access* to a basic class J if $i \in \mathcal{A}(J)$ and there is no basic class K such that $i \in \mathcal{A}(K)$ and $K \subset \mathcal{A}(J)$. The *height* of a basic class J, $\nu(J)$, is defined as the maximal integer $n \geq 1$, such that there exist distinct basic classes $J_1 = J, J_2, \ldots, J_n$ for which $J_{i+1} \subset \mathcal{A}(J_i)$, $i = 1, \ldots, n-1$.

We next need some properties of the algebraic eigenspace of a nonnegative matrix at its spectral radius. By slightly extending the results of [11, Theorem 3.1, parts 1 and 2] it is easily seen that:

Theorem 5.1. *Let P be a square nonnegative matrix with spectral radius r. Then:*

(1) The dimension of $N_r(P)$ (resp., $N_r(P^*)$) equals the number of basic classes of P. Moreover, for every basic class J, there exists a vector $x(J) \in N_r(P)$ (resp., $y(J) \in N_r(P^*)$) such that

$$\{i \mid x(J)_i \neq 0\} \subseteq \mathcal{A}(J) \quad \textit{and} \quad x(J)_J \neq 0 \tag{5.1}$$

and

$$\{i \mid y(J)_i \neq 0\} \subseteq \mathcal{A}^*(J) \quad \textit{and} \quad y(J)_J \neq 0. \tag{5.2}$$

Every collection $\{x(J)$ (resp., $y(J)) \mid J$ is a basic class$\}$ of vectors in $N_r(P)$ (resp., $N_r(P^*)$) satisfying (5.1) (resp., (5.2)) is a basis of $N_r(P)$ (resp., $N_r(P^*)$).

(2) The index of r for P equals the maximal height of a basic class. In particular, for every basic class J, $\nu(J) = \nu_r(P_{\mathcal{A}(J)})$.

Notice that statement (1) extends the results of [11] since we do not require that our vectors be nonnegative and that $x(J)_{\mathcal{A}(J)} \gg 0$ (resp., $y(J)_{\mathcal{A}^*(J)} \gg 0$).

We will next show an interesting property of bases of the type described in Theorem 5.1.

Theorem 5.2. *Let P be a square nonnegative matrix with spectral radius r and basic class $J_1, \ldots, J_m$, where*

$$J_i \textit{ does not have access to } J_j \quad \textit{for } i, j = 1, \ldots, m \quad \textit{with } i > j. \tag{5.3}$$

For every basic class J, let $x(J)$ (resp., $y(J)$) be a vector in $N_r(P)$ (resp., $N_r(P^*)$) which satisfies (5.1) (*resp.*, (5.2)). *If $X = (x(J_1), \ldots, x(J_m))$ and $Y = (y(J_1), \ldots, y(J_m))$, then Y^*X is upper triangular.*

Proof. If $(Y^*X)_{ij} = y(J_i)^*x(J_j) \neq 0$, then by (5.1) and (5.2), $\mathscr{A}(J_j) \cap \mathscr{A}^*(J_i) \neq \emptyset$. This implies that $J_i \subseteq \mathscr{A}(J_j)$, i.e., J_i has access to J_j. By (5.3) it follows that $i \leq j$, thus proving that $(Y^*X)_{ij} = 0$ whenever $i > j$. This completes the proof of Theorem 5.2.

Remark. The requirement of (5.3) can be obtained, for example, by reordering the basic classes so that

$$\nu(J_1) \leq \nu(J_2) \leq \ldots \leq \nu(J_m). \tag{5.4}$$

Corollary 5.3. *Let P be a square nonnegative matrix with spectral radius r $\nu_r(P) = 1$ and basic classes $J_1, \ldots, J_m$. For every basic class J, let $x(J)$ (resp., $y(J)$) be a vector in $N_r(P)$ (resp., $N_r(P^*)$) which satisfies* (5.1) (*resp.*, (5.2)). *If $X = (x(J_1), \ldots, x(J_m))$ and $Y = (y(J_1), \ldots, y(J_m))$, then Y^*X is a diagonal matrix.*

Proof. Since $\nu_r(P) = 1$ it follows from Theorem 5.1 (compare also [11, Corollary 3.3]) that no basic class has access to any other basic class. This implies that (5.3) is satisfied under every permutation of $J_1, \ldots, J_m$. Thus by Theorem 5.2, Y^*X is upper triangular under every permutation of its rows and corresponding columns. This obviously implies that Y^*X is diagonal, completing the proof of Corollary 5.3.

Let P be a given nonnegative matrix with spectral radius r and basic classes $J_1, \ldots, J_m$ ordered in a way such that (5.4) is satisfied. We will next give a method for constructing vectors $x(J_1), \ldots, x(J_m) \in N_r(P)$ satisfying (5.1), respectively, with $J_1, \ldots, J_m$. This will be done by induction on the height of the basic classes. A similar procedure may be used for finding vectors in $N_r(P^*)$ satisfying (5.2). Our method has a similar underlying idea to the proof of [11, Theorem 3.1, part 1].

Let J be a basic class with height $\nu(J) = \nu$, and assume that for every basic class K with $\nu(K) < \nu$ we found a vector $x(K) \in N_r(P)$ satisfying (5.1) with respect to K. We will show how to construct a vector $x \in N_r(P)$ satisfying (5.1) with respect to J.

By possibly permuting rows and corresponding columns of P we may assume that

$$P = \begin{pmatrix} P_M & P_{ML} & P_{MK} \\ 0 & P_L & P_{LR} \\ 0 & 0 & P_R \end{pmatrix},$$

where L is the set of states having direct access to J, $M \equiv \mathscr{A}(J)\setminus L$ and $R \equiv \{1, \ldots, S\}\setminus \mathscr{A}(J)$. We next introduce two additional notations. Let $\bar{P} \equiv P - rI$ and similarly for $B \subset \{1, \ldots, S\}$, let $\bar{P}_B \equiv P_B - rI$, where I is the identity matrix of the appropriate dimension.

By part (2) of Theorem 5.1 $\nu = \nu_r(P_{\mathscr{A}(J)})$. It now follows that we are looking for two vectors, $z \in R^{|M|}$ and $t \in R^{|L|}$ such that $t_J \neq 0$ and

$$\bar{P}^{\nu}\begin{pmatrix} z \\ t \\ 0 \end{pmatrix} = 0$$

or equivalently that

$$\bar{P}_M^{\nu} z + \sum_{i=0}^{\nu-1} \bar{P}_M^{\nu-1-i} P_{ML} \bar{P}_L^{i} t = 0 \tag{5.5}$$

and

$$\bar{P}_L^{\nu} t = 0. \tag{5.6}$$

By Theorem 5.1 we know that a nontrivial solution to the above system of equations exists. Moreover, since J is the unique basic class of P_L and every state in L has access to J, the above mentioned theorem implies that $\nu_r(P_L) = 1$, $N_r(P_L)$ is one dimensional and is spanned by a positive vector. Thus t satisfies (5.6) if and only if

$$\bar{P}_L t = 0 \tag{5.7}$$

and t is proportional to a positive vector; moreover, t solving (5.7) is unique (up to proportion)[1]. Since (5.5) and (5.6) form a homogeneous system in z and t it follows from the fact that t is proportional to a positive vector, that the requirement $t_J \neq 0$ can be replaced by

$$e^* t \neq 0. \tag{5.8}$$

Substituting (5.7) into (5.5) we see that z solves (5.5) with a vector t which satisfies (5.6), or equivalently (5.7), if and only if

$$\bar{P}_M^{\nu} z + \bar{P}_M^{\nu-1} P_{ML} t = \bar{P}_M^{\nu-1}(\bar{P}_M z + P_{ML} t) = 0. \tag{5.9}$$

[1] The form of the unique solution (up to proportion) of (5.7) is illustrated in (6.1).

The height of each basic class of P_M is less than $\nu = \nu(J)$, hence by Theorem 5.1, part (2), $\nu_r(P_M) \leq \nu - 1$, which implies that (5.9) is equivalent to requiring that

$$\bar{P}_M z + \bar{P}_{ML} t \in N_r(P_M). \tag{5.10}$$

Let $K_1, \ldots, K_n$ be the basic classes of P_M. It is easily seen that if these classes are considered as basic classes of P then $\nu(K_j) < \nu(J)$ for $j = 1, \ldots, n$. Hence by our assumption, we already found vectors $x(j) \in N_r(P)$ satisfying (5.1) with respect to K_j, $j = 1, \ldots, n$. It follows directly from Theorem 5.1 that $x(1)_M, \ldots, x(n)_M$ form a basis of $N_r(P_M)$. Thus, (5.9) is satisfied if and only if there exist constants $\alpha_1, \ldots, \alpha_n$ such that

$$\bar{P}_M z + P_{ML} t = \sum_{i=1}^{n} \alpha_i x(i)_M. \tag{5.11}$$

We described a method to find a vector $x \in N_r(P)$ satisfying (5.1) with respect to J by finding $z \in \mathbf{R}^{|M|}$, $t \in \mathbf{R}^{|L|}$ and constants $\alpha_1, \ldots, \alpha_n$ which satisfy (5.7), (5.8) and (5.11). This reduced our problem to solving a set of linear equations.

We remark that our arguments imply that a solution to our system of equations exists; moreover, t is uniquely determined by (5.7) and (5.8). Also observe that since $x(1), \ldots, x(n)$ are independent it follows that for every fixed z, $\alpha_1, \ldots, \alpha_n$ are uniquely determined by (5.11). We finally remark that the underlying matrix of our system of equations does not contain high powers of a singular matrix (which causes high round off errors).

6. Applications to stochastic matrices

The purpose of this section is to apply results obtained in previous sections to stochastic matrices. We remark that the spectral radius of a stochastic matrix is always one and its basic classes are precisely its recurrent classes. We will construct a set of vectors whose convex hull spans the set of stationary distributions. We will also give a closed form for the limiting matrix, i.e., for $E_r(P) = \lim_{N\to\infty} P^N$ $(C, 1)$, where P is our given stochastic matrix.

Let P be an $S \times S$ stochastic matrix with recurrent classes $J(1), \ldots, J(m)$ and let $J(0) = \{1, \ldots, S\} \backslash \bigcup_{j=1}^{m} J(j)$. By possibly permuting rows and corresponding columns of P, we may assume that

$$P = \begin{bmatrix} P(0) & P(0,1) & \cdots & P(0,m) \\ & P(1) & \cdots & P(1,m) \\ & & \ddots & \\ & 0 & & \\ & & & P(m) \end{bmatrix},$$

where $P(j) \equiv P_{J(j)}$ and $P(i,j) \equiv P_{J(i),J(j)}$ for $i, j = 0, \ldots, m$.

Observing that $P(j)$, $j = 1, \ldots, m$, is an irreducible stochastic matrix (i.e., ergodic) it follows that $P(j)e = e$ and that there exists a probability vector $q(j) \gg 0$ such that $q(j)^* P(j) = q(j)^*$. For $j = 1, \ldots, m$ let $x(j)$ and $y(j)$ be the vectors defined by

$$x(j)_{J(i)} = \begin{cases} (I - P(0))^{-1} P(0, j)e & \text{if } i = 0, \\ e & \text{if } i = j, \\ 0 & \text{if } i = 1, \ldots, m \text{ with } i \neq j; \end{cases} \tag{6.1}$$

and

$$y(j)_{J(i)} = \begin{cases} 0 & \text{if } i = j, \\ q(j) & \text{if } i = 0, \ldots, m \quad \text{with } i \neq j. \end{cases} \tag{6.2}$$

Under these definitions $x(j) \in N_r(P)$ (resp., $y(j) \in N_r(P^*)$) and satisfies (5.1) (resp., (5.2)) with respect to $J(j)$, for $j = 1, \ldots, m$. It now follows from Theorem 5.1 that $x(1), \ldots, x(m)$ (resp., $y(1), \ldots, y(m)$) form a basis of $N_r(P)$ (resp., $N_r(P^*)$). In particular, this implies the well known fact that the convex combinations of $y(1), \ldots, y(m)$ span the cone of stationary probabilities of P. (Cf. [7, Theorem 6.2.1] where this result is stated for matrices with a unique recurrent class.)

Setting $X \equiv (x(1), \ldots, x(m))$ and $Y \equiv (y(1), \ldots y(m))$ it is easily seen that Y^*X is the identity matrix; hence by Theorem 3.1

$$E_r(P) = XY^* = \begin{bmatrix} 0 & A(1) & \cdots & A(m) \\ & E(1) & & \\ & & \ddots & 0 \\ & 0 & & \ddots \\ & & & E(m) \end{bmatrix} \tag{6.3}$$

where $E(j) \equiv eq(j)^*$ and $A(j) \equiv (I - P(0))^{-1}P(0,j)eq(j)^*$, $j = 1, \ldots, m$. We remind the reader that $E_r(P)$ is the absorption limiting matrix of P, i.e, $E_r(P) = \lim_{N\to\infty} P^N$ $(C, 1)$. The representation given in (6.3) for absorbing Markov chains appears in [7, p. 52, Theorem 3.3.7).

We will next interpret (6.3), i.e., we will give a standard way to derive it from probabilistic considerations. We assume that we have a Markov process with transition matrix P, and we will compute the limiting absorption matrix directly. If we start in a recurrent class $J(i)$, then we will never leave it. Thus, the absorption probabilities are obtained by restricting ourselves to this recurrent class, i.e., $(E_r(P))_J = E_r(P_J)$. Observing that P_J is an irreducible matrix with a unique stationary probability vector $q(i)$, it follows that $E_r(P_J) = eq(i)^*$. Thus, for $i = 1, \ldots, m$

$$E_r(P)_{J(i),J(j)} = \begin{cases} eq(i)^* & \text{if} \quad i = j, \\ 0 & \text{if} \quad j = 0, \ldots, m \quad \text{with } i \neq j. \end{cases}$$

We will next find the absorbing probabilities when starting from a transient state i. If we identify all the states of a recurrent class $J(j)$ as one state, it follows from [7, p. 52, Theorem 3.3.7] that the probability of absorption in $J(j)$, given we start in i, is given by

$$\sum_{p\in J(j)} \sum_{k\in J(0)} \sum_{n=0}^{\infty} P(0)^n_{ik} P(0,j)_{kp} = ((I - P(0))^{-1}P(0,j)e)_i.$$

The probability of absorption in a given state $k \in J(j)$, given we started in i, is now obtained by conditioning on absorption in $J(j)$ and observing that once we entered $J(j)$, absorption in state k is given by $(q(j))_k$. Thus, the probability of absorption in state $k \in J(j)$, given we start in state $i \in J(0)$, is given by $(E_r(P))_{ij} = ((I - P(0))^{-1}P(0,j)e)_i(q(j))_k$, which is precisely the expression obtained in (6.3).

References

[1] G. Debreu, and I.N. Herstein, "Nonnegative square matrices", *Econometrica* 21 (1953) 597–607.

[2] N. Dunford and J.T. Schwartz, *Linear operators*, Part I (Interscience, New York, 1958).

[3] D. Gale, *The theory of linear economic models* (McGraw-Hill, New York, 1960).

[4] F.A. Graybill, *Introduction to matrices with applications to statistics* (Wadsworth Publishing Company, Inc., Belmont, Calif., 1969).

[5] S. Karlin, *Mathematical methods and theory of games, programming, and economics*, Vol. I (Addison-Wesley, Reading, Mass., 1959).

[6] T. Kato, *Perturbation theory for linear operators* (Springer, New York, 1966).
[7] J. Kemeny and J. Snell, *Finite Markov chains* (Van Nostrand, Princeton, N.J., 1960).
[8] H. Nikaido, *Convex structure and economic theory* (Academic Press, New York, 1968).
[9] M.C. Pease, *Methods of matrix algebra* (Academic Press, New York, 1965).
[10] U.G. Rothblum, "Multiplicative Markov decision chains", Ph.D. Dissertation, Department of Operations Research, Stanford University, Stanford, Calif. (1974).
[11] U.G. Rothblum, "Algebraic eigenspaces of nonnegative matrices", *Linear Algebra and its Applications* 12 (1975) 281–292.
[12] U.G. Rothblum, "Expansions of sums of matrix powers and resolvents", to appear.
[13] U.G. Rothblum and A.F. Veinott, Jr., "Average-overtaking cumulative optimality for polynomially bounded Markov decision chains", to appear.
[14] R.S. Varga, *Matrix iterative analysis* (Prentice-Hall, Englewood Cliffs, N.J., 1962).
[15] A.F. Veinott, Jr., "Discrete dynamic programming with sensitive discount optimality criteria", *Annals of Mathematical Statistics* 40 (1969) 1635–1660.