On a property of the projector matrices and its application to the canonical representation of matrix functions.

By E. EGERVÁRY in Budapest.

Notations.

$a, b, c, \ldots$ scalars	$\mathbf{A}^* =$ transposed of $\mathbf{A}$
$\mathbf{a}, \mathbf{b}, \mathbf{c}, \ldots$ column vectors	$\lvert\mathbf{A}\rvert =$ determinant of $\mathbf{A}$
$\mathbf{a}^*, \mathbf{b}^*, \mathbf{c}^*, \ldots$ row vectors	$\mathbf{E} = [\delta_{ij}] =$ unit matrix
$\mathbf{A}, \mathbf{B}, \mathbf{C}, \ldots$ matrices	$\langle a_1, a_2, \ldots, a_n \rangle =$ diagonal matrix

1. Let $\mathbf{A}$ be a non singular square matrix of order n, and suppose that the adjoint of $\lambda\mathbf{E}-\mathbf{A}$ is divisible by the discriminant of $|\lambda\mathbf{E}-\mathbf{A}| = D(\lambda) = \prod_1^s (\lambda-\lambda_k)^{\alpha_k}$ $(s \leqq n, \Sigma\alpha_k = n)$, i. e. that the roots $\lambda_1, \lambda_2, \ldots, \lambda_s$ of the minimum-equation $\Delta(\lambda) = \prod_1^s (\lambda-\lambda_k) = 0$ of $\mathbf{A}$ are all distinct[1]).

It is known that in this case there is a system $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ of linearly independent right eigenvectors satisfying $\mathbf{A}\mathbf{v}_k = \lambda_k\mathbf{v}_k$, and a system $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n$ of linearly independent left eigenvectors satisfying $\mathbf{w}_k^*\mathbf{A} = \lambda_k\mathbf{w}_k^*$, such that the systems $\mathbf{v}_k, \mathbf{w}_k$ are reciprocal, i. e. they satisfy $\mathbf{w}_k^*\mathbf{v}_l = \delta_{kl}$. Furthermore it is known that the transformation of the matrix $\mathbf{A}$ to the basis $\mathbf{v}_1, \ldots, \mathbf{v}_n$ or to the basis $\mathbf{w}_1, \ldots, \mathbf{w}_n$ reduces it to the diagonal form $\langle\lambda_1, \ldots, \lambda_2, \ldots, \lambda_s, \ldots\rangle$.

In view of the importance of these theorems in the applications (transformation of a quadric to its principal axes, introduction of normal coordinates in a dynamical system, matrix solution of a system of linear differential equations etc.) it seems to be desirable to have a concise method for the computation of the reciprocal systems of eigenvectors.

If the eigenvalues λ_k are all distinct, then the solutions of the equations $\mathbf{A}\mathbf{v} = \lambda_k\mathbf{v}$ resp. $\mathbf{w}^*\mathbf{A} = \lambda_k\mathbf{w}^*$ are, apart from a scalar factor, uniquely deter-

[1]) If $\mathbf{A}^* = \overline{\mathbf{A}}$ i. e. if $\mathbf{A}$ is hermitian then the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_s$ are all real and the divisibility of $\operatorname{adj}(\mathbf{E}\lambda-\mathbf{A})$ by $(D(\lambda), D'(\lambda))$ is an immediate consequence of the identity $\sum_p \sum_q D_{pq}(\lambda)\overline{D}_{pq}(\lambda) = D'(\lambda)^2 - D(\lambda)D''(\lambda)$.

mined, hence in this case the computation of the eigenvectors by the solution of the corresponding system of linear equations is comparatively simple.

In case of multiple eigenvalues the usual way of computation of the eigenvectors is an awkward and lengthy business[2]). First of all one has to find the rank of $\lambda_k \mathbf{E}-\mathbf{A}$ and a complete system of solutions of $(\lambda_k \mathbf{E}-\mathbf{A})\mathbf{v}=0$ as well as of $\mathbf{w}^*(\lambda_k \mathbf{E}-\mathbf{A})=0$, and afterwards these two systems must be biorthogonalised and normalised.

2. In the present paper I wish to indicate a concise, straightforward method for the computation of the reciprocal systems of eigenvectors, based on a remarkable property of the projector matrices which, as far as I am aware, does not seem to have been noticed hitherto.

A square matrix $\mathbf{P}$ is a projector if it satisfies the equation $\mathbf{P}^2=\mathbf{P}$. If $\mathbf{P}$ is a non singular projector then we have obviously $\mathbf{P}=\mathbf{E}$. Therefore we confine ourselves to the consideration of a singular projector whose rank $\varrho(\mathbf{P})$ satisfies $1 \leqq \varrho(\mathbf{P}) \leqq n-1$.

Any matrix $\mathbf{A}$ with n rows and m columns having the rank r can be represented in a form making its rank intuitive[3]):

$$(1)\qquad \mathbf{A}=\mathbf{B}\mathbf{C}=\begin{bmatrix} b_{11} & \dots & b_{1r} \\ b_{21} & \dots & b_{2r} \\ \vdots & & \vdots \\ b_{n1} & \dots & b_{nr} \end{bmatrix}\begin{bmatrix} c_{11} & \dots & c_{1m} \\ \vdots & & \\ c_{r1} & \dots & c_{rm} \end{bmatrix}=[\mathbf{b}_1 \dots \mathbf{b}_r]\begin{bmatrix} \mathbf{c}_1^* \\ \vdots \\ \mathbf{c}_r^* \end{bmatrix}=\sum_{k=1}^{r} \mathbf{b}_k \mathbf{c}_k^*$$

in which $\mathbf{B}$ is a matrix with r linearly distinct column vectors and $\mathbf{C}$ is a matrix with r linearly distinct row vectors. The simplest way of this factorisation of a given matrix, which necessitates only the computation of second order determinants, is the following one. If $\varrho(\mathbf{A}) \geqq 1$, then there is at least one non vanishing element $a_{\beta\gamma}$. Then the difference

$$a_{\beta\gamma}\begin{bmatrix} a_{11} & \dots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nm} \end{bmatrix}-\begin{bmatrix} a_{1\gamma} \\ a_{2\gamma} \\ \vdots \\ a_{n\gamma} \end{bmatrix}[a_{\beta 1}\, a_{\beta 2} \dots a_{\beta n}]=\begin{array}{c} \gamma \\ \left[\begin{matrix} a'_{11} \dots 0 \dots a'_{1m} \\ \vdots \qquad\quad \vdots \\ 0 \;\dots 0 \dots 0 \\ \vdots \qquad\quad \vdots \\ a'_{n1} \dots 0 \dots a'_{nm} \end{matrix}\right] \end{array}\begin{matrix} \\ \\ (\beta \\ \\ \\ \end{matrix}=\mathbf{A}'$$

is a matrix with one 0-column and one 0-row, while its further elements a'_{ij} are second order determinants of $\mathbf{A}$. Applying now the same process to $\mathbf{A}'$ we arrive to a matrix with two 0-columns and two 0-rows whose further

---

[2]) Compare f.i. the examples given in W. E. Jung, *Matrizen und Determinanten* (Leipzig, 1951), pp. 97—99.

[3]) In case of square matrices this theorem is stated without proof in R. A. Frazer, W. J. Duncan, A. R. Collar, *Elementary matrices* (Cambridge, 1938), p. 20. See also А. И. Мальцев, Основы линейной алгебры (Москва, 1948), p. 117. Another method for the factorization of non-singular matrices is given in R. Zurmühl, Zur numerischen Auflösung linearer Gleichungssysteme nach dem Matrizenverfahren von Banachiewicz, *Zeitschrift f. angew. Math. u. Mech.*, 29 (1949), pp. 76—84.

elements are second order determinants of $\mathbf{A}'$, hence proportional to the third-order determinants of $\mathbf{A}$[4]). After r steps we arrive to a matrix $\mathbf{A}^{(r)}$ whose elements are proportional to the $r+1$-th order determinants of $\mathbf{A}$ which vanish by assumption and the factorisation is finished.

In case of a hermitian projector we have $\mathbf{P}=\mathbf{P}^2=\mathbf{P}\bar{\mathbf{P}}^*$, therefore $\mathbf{P}$ is positive semidefinite and one can select a sequence of principal minors $a_{\beta_1\beta_1}$, $\begin{vmatrix} a_{\beta_1\beta_1} & a_{\beta_1\beta_2} \\ a_{\beta_1\beta_2} & a_{\beta_2\beta_2} \end{vmatrix}, \ldots, \begin{vmatrix} a_{\beta_1\beta_1} & \cdots & \\ & & a_{\beta_r\beta_r} \end{vmatrix}$ which are positive. Choosing these as starting elements in the former process of reduction the two factor matrices will be hermitically conjugated.

3. If the matrix $\mathbf{A}$ is a projector then the former representation can be characterised in a more precise way:

Theorem. *If a projector matrix* $\mathbf{P}$ *of order* n *and rank* r *is represented in the form*

$$\mathbf{P}=[\mathbf{v}_1\mathbf{v}_2\ldots\mathbf{v}_r]\begin{bmatrix}\mathbf{w}_1^*\\ \mathbf{w}_2^*\\ \vdots\\ \mathbf{w}_r^*\end{bmatrix}=\sum_{k=1}^{r}\mathbf{v}_k\mathbf{w}_k^*, \tag{2}$$

then the two systems of vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ *and* $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_r$ *occurring in* (2) *are automatically biorthogonalised and normalised i. e. they satisfy the relations*

$$[\mathbf{w}_i^*\mathbf{v}_j]=\mathbf{E} \quad or \quad \mathbf{w}_i^*\mathbf{v}_j=\delta_{ij}.$$

Proof. The projector $\mathbf{P}$ satisfies the equation $\mathbf{P}^2-\mathbf{P}=0$; using the representation (2) this can be written as follows:

$$[\mathbf{v}_1\mathbf{v}_2\ldots\mathbf{v}_r]\left\{\begin{bmatrix}\mathbf{w}_1^*\\ \mathbf{w}_2^*\\ \vdots\\ \mathbf{w}_r^*\end{bmatrix}[\mathbf{v}_1\mathbf{v}_2\ldots\mathbf{v}_r]-\mathbf{E}\right\}\begin{bmatrix}\mathbf{w}_1^*\\ \mathbf{w}_2^*\\ \vdots\\ \mathbf{w}_r^*\end{bmatrix}=0.$$

But the first and the last factors on the left side of this matrix equation have the maximal rank r, hence the middle-factor in the bracket must vanish[5]), i. e.

$$\begin{bmatrix}\mathbf{w}_1^*\\ \mathbf{w}_2^*\\ \vdots\\ \mathbf{w}_r^*\end{bmatrix}[\mathbf{v}_1\mathbf{v}_2\ldots\mathbf{v}_r]=\mathbf{E}.$$

Q. e. d.

---

[4]) This is an immediate consequence of the well-known identity $A_{ik}A_{jl}-A_{il}A_{jk}=A.A_{\substack{ij\\kl}}$.

[5]) This follows immediately by considering the minor matrix of the product, whose right and left factors are not singular r-th order square matrices.

Corollary. A hermitian projector matrix $\mathbf{P}$ of order n and rank r can be represented in the form

$$\mathbf{P} = [\mathbf{v}_1 \mathbf{v}_2 \ldots \mathbf{v}_r] \begin{Vmatrix} \bar{\mathbf{v}}_1^* \\ \bar{\mathbf{v}}_2^* \\ \vdots \\ \bar{\mathbf{v}}_r^* \end{Vmatrix}$$

where the vectors $\mathbf{v}_1 \mathbf{v}_2 \ldots \mathbf{v}_r$ form a unitary set, i. e. they satisfy the relations

$$\bar{\mathbf{v}}_i^* \mathbf{v}_j = \delta_{ij}.$$

4. Suppose now that the matrix $\mathbf{A}$ satisfies the assumptions made in **1** and consider the matrix $f(\mathbf{A})$ where $f(\lambda)$ is an arbitrary scalar polynomial[6]) of λ. Denoting by $L_k(\lambda)$ the Lagrangian interpolation polynomials defined by the conditions

$$\text{degree of } L_k(\lambda) \leqq s-1$$

$$L_k(\lambda_h) = \delta_{k,h} \qquad (k, h = 1, 2, \ldots, s)$$

we have

$$f(\lambda) \equiv \sum_{k=1}^{s} f(\lambda_k) L_k(\lambda) \pmod{\Delta(\lambda)},$$

consequently

(3) $$f(\mathbf{A}) = \sum_{k=1}^{s} f(\lambda_k) L_k(\mathbf{A})\ ^{7}).$$

From

$$(\lambda - \lambda_k) L_k(\lambda) \equiv 0 \ (\text{mod. } \Delta(\lambda))$$

we deduce immediately

(4) $$(\lambda_k \mathbf{E} - \mathbf{A}) L_k(\mathbf{A}) = 0, \qquad L_k(\mathbf{A})(\lambda_k \mathbf{E} - \mathbf{A}) = 0.$$

Hence the column vectors of $L_k(\mathbf{A})$ are right eigenvectors, the row vectors of $L_k(\mathbf{A})$ are left eigenvectors belonging to λ_k. The number of the linearly distinct right and left eigenvectors (the dimension of the right and left eigenmanifolds) belonging to λ_k is evidently equal to the nullity $n - \varrho(L_k(\mathbf{A}))$ of $L_k(\mathbf{A})$.

The polynomials $L_k(\lambda)$ satisfy the identity

$$\sum_{1}^{s} L_k(\lambda) \equiv 1,$$

hence

$$\sum_{1}^{s} L_k(\mathbf{A}) = \mathbf{E},$$

---

[6]) In order to avoid limit-processes we confine ourselves here to the consideration of polynomials.

[7]) The Lagrangian interpolation formula has been used first in matrix-theory by J. J. Sylvester, *Philosophical Magazine*, **16** (1883), 267—269.

consequently

$$\sum_{1}^{s} \varrho(L_k(\mathbf{A})) \geqq \varrho\left(\sum_{1}^{s} L_k(\mathbf{A})\right) = \varrho(\mathbf{E}) = n = \sum_{1}^{s} \alpha_k \text{ [8]}. \tag{5.1}$$

On the other side, for the eigenvalue λ_k of multiplicity α_k we have

$$\varrho(\lambda_k \mathbf{E} - \mathbf{A}) \geqq n - \alpha_k \tag{5.2}$$

and in consequence of (4)

$$\varrho(\lambda_k \mathbf{E} - \mathbf{A}) + \varrho(L_k(\mathbf{A})) \leqq n \text{ [8]}. \tag{5.3}$$

From (5.2) and (5.3) we deduce

$$\varrho(L_k(\mathbf{A})) \leqq \alpha_k, \quad \sum_{1}^{s} \varrho(L_k(\mathbf{A})) \leqq \sum_{1}^{s} \alpha_k. \tag{5.4}$$

Comparison of (5.1) and (5.4) shows immediately that

$$\varrho(L_k(\mathbf{A})) = \alpha_k, \tag{5.5}$$

i. e. the number of the linearly distinct right and left eigenvectors belonging to λ_k is equal to the multiplicity α_k.

5. The construction of the reciprocal systems of right and left eigenvectors will be now based upon the following properties of the Lagrangián matrix polynomials $L_h(\mathbf{A})$:

$$L_k(\lambda) L_h(\lambda) \equiv 0 \pmod{J(\lambda)} \qquad (k \neq h), \tag{I}$$

hence

$$L_k(\mathbf{A}) L_h(\mathbf{A}) = \mathbf{0} \qquad (k \neq h); \tag{6}$$

$$L_k(\lambda)^2 - L_k(\lambda) \equiv 0 \pmod{J(\lambda)}, \tag{II}$$

hence

$$L_k(\mathbf{A})^2 = L_k(\mathbf{A}). \tag{7}$$

The interpretation of the property (I) in the geometry of vector space is obvious by the decomposition (1) of $L_k(\mathbf{A})$ and $L_h(\mathbf{A})$:

$$L_k(\mathbf{A}) L_h(\mathbf{A}) = [\mathbf{v}_{k1} \dots \mathbf{v}_{k\alpha_k}] \left\{ \begin{vmatrix} \mathbf{w}^*_{k1} \\ \vdots \\ \mathbf{w}^*_{k\alpha_k} \end{vmatrix} [\mathbf{v}_{h1} \dots \mathbf{v}_{h\alpha_h}] \right\} \begin{vmatrix} \mathbf{w}^*_{h1} \\ \vdots \\ \mathbf{w}^*_{h\alpha_h} \end{vmatrix} = 0.$$

The first and last factors have maximal ranks α_k resp. α_h, hence the middle-factor in the bracket must vanish, consequently

$$\mathbf{v}^*_{ki} \mathbf{w}_{hj} = 0 \qquad (k \neq h). \tag{8.1}$$

Thus, we have arrived by this method to the well-known fact that any right eigenvector and any left eigenvector, which belong to different eigenvalues, are orthogonal.

---

[8]) Here we use the following elementary theorems. The rank of a sum cannot exceed the sum of the ranks of the terms and if the product of two square matrices is null the sum of their ranks cannot exceed their order. See f. i. FRAZER, DUNCAN and COLLAR, l. c., p. 23.

The property (II) shows that each Lagrangian matrix polynomial is a projector. But the real counterpiece to the former interpretation of (I) is furnished by the application of the theorem proved in § 3.

If a Lagrangian matrix polynomial of rank α_k is decomposed into two factors

$$L_k(\mathbf{A}) = [\mathbf{v}_{k1} \dots \mathbf{v}_{k\alpha_k}] \begin{bmatrix} \mathbf{w}^*_{k1} \\ \vdots \\ \mathbf{w}_{k\alpha_k} \end{bmatrix} = \sum_{\varkappa=1}^{\alpha_k} \mathbf{v}_{k\varkappa} \mathbf{w}^*_{k\varkappa}, \tag{1. 1}$$

then the vectors $\mathbf{v}_{k1}, \dots, \mathbf{v}_{k\alpha_k}$ *and* $\mathbf{w}_{k1}, \dots, \mathbf{w}_{k\alpha_k}$ *constitute a system of* α_k *linearly distinct right resp. left eigenvectors such that the two systems are automatically biorthogonalised and normalised*:

$$\mathbf{v}^*_{ki} \mathbf{w}_{kj} = \delta_{ij}. \tag{8. 2}$$

The equations (8. 1) and (8. 2) can be united into the single system of equations

$$\mathbf{v}^*_{ki} \mathbf{w}_{hj} = \delta_{kh} \delta_{ij} \qquad \begin{pmatrix} k, h = 1, 2, \dots, s \\ i = 1, 2, \dots, \alpha_k \\ j = 1, 2, \dots, \alpha_h \end{pmatrix}. \tag{8}$$

Hence

$$\begin{aligned} \mathbf{V} &= [\mathbf{v}_{11} \dots \mathbf{v}_{1\alpha_1} \dots \mathbf{v}_{s1} \dots \mathbf{v}_{s\alpha_s}] \\ \mathbf{W} &= [\mathbf{w}_{11} \dots \mathbf{w}_{1\alpha_1} \dots \mathbf{w}_{s1} \dots \mathbf{w}_{s\alpha_s}] \end{aligned} \qquad \left(\sum_1^s \alpha_k = n \right)$$

are reciprocal matrices of order n.

Replacing now in (3) the Lagrangian matrix polynomials by their expressions (1. 1) we arrive at the following representation of $f(\mathbf{A})$:

$$f(\mathbf{A}) = \sum_{k=1}^{s} f(\lambda_k)(\mathbf{v}_{k1}\mathbf{w}^*_{k1} + \cdots + \mathbf{v}_{k\alpha_k}\mathbf{w}^*_{k\alpha_k}) = \mathbf{V} \langle f(\lambda_1) \dots f(\lambda_2) \dots f(\lambda_s) \dots \rangle \mathbf{V}^{-1}$$

In the special case $f(\lambda) \equiv \lambda$ we get from here the diagonal representation of the matrix $\mathbf{A}$ itself.

(Received March 15, 1953.)