COMPOSITIO MATHEMATICA

RUFUS OLDENBURGER

Factorability of general symmetric matrices

Compositio Mathematica, tome 7 (1940), p. 223-228
<http://www.numdam.org/item?id=CM_1940__7__223_0>

© Foundation Compositio Mathematica, 1940, tous droits réservés.

L'accès aux archives de la revue « Compositio Mathematica » (http://http://www.compositio.nl/) implique l'accord avec les conditions générales d'utilisation (http://www.numdam.org/conditions). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM
Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
http://www.numdam.org/

Factorability of general symmetric matrices

by

Rufus Oldenburger

Chicago

---

1. *Introduction.* The well-known theorem that a quadratic form $Q = a_{ij}x_ix_j$ $[a_{ij}=a_{ji}]$ of rank r is equivalent to a form $\lambda_1 y_1^2 + \lambda_2 y_2^2 + \ldots + \lambda_r y_r^2$ with diagonal matrix is the same as the statement that the matrix $A = (a_{ij})$ of Q can be „factored" into $B'DB$, where D is the diagonal matrix

$$\left\| \begin{matrix} \lambda_1 & & & O \\ & \lambda_2 & & \\ & & \ddots & \\ O & & & \lambda_r \end{matrix} \right\|,$$

B' denotes the transpose of B, and B is a matrix of rank r with r rows. If we write $B = (b_{\alpha i}) = (b_{\alpha j})$, we have

$$A = \left(\sum_{\alpha=1}^{r} \lambda_\alpha b_{\alpha i} b_{\alpha j} \right).$$

In the present paper *we are concerned with the problem of „factorability" of a general symmetric matrix* $(a_{ij\ldots m})$ *into a form*

$$\left(\sum_{\alpha=1}^{\sigma} \lambda_\alpha b_{\alpha i} b_{\alpha j} \cdots b_{\alpha m} \right), \tag{1.1}$$

where σ is finite. If A factors as in (1.1) the associated form $a_{ij\ldots m}\, x_i x_j \cdots x_m$ can be written as a linear combination of powers of linear forms. Such linear combinations are useful in treating some of the classical problems of algebra [1]).

2. *Definitions.* We shall say that a matrix $A = (a_{ij\ldots m})$ is *p-way* if it has p indices $i, j, \ldots, m$. If each index ranges over $1, 2, \ldots, n$, we say that A is of *order n*. In the introduction and in what follows the term symmetric matrix refers to a matrix

---

[1]) R. Oldenburger, Representation and equivalence of forms [Proceedings Nat. Acad. Sci. **24** (1938), 193—198].

for which the values of the elements are unchanged under permutation of the subscripts. If a matrix A can be written as (1.1) with elements in a field K, we shall say that A is *factorable with respect to K.*

3. *Factorability.* In the following theorem, the term „order" of K refers to the number of elements in the field K.

THEOREM 3.1. *The class of symmetric p-way matrices factorable with respect to a field K is identical with the class of all symmetric p-way matrices if and only if K is of order p or more.*

We shall sketch the proof of Theorem 3.1 leaving out some of the more complicated details.

A p-way matrix $A = (a_{ij\ldots m})$ of order n is factorable if and only if there exist elements λ_α, $b_{\alpha i}$ $[\alpha = 1, 2, \ldots, \sigma;$ $i = 1, 2, \ldots, n]$ such that the following equations are satisfied:

$$\sum_{\alpha=1}^{\sigma} \lambda_\alpha b_{\alpha i} b_{\alpha j} \cdots b_{\alpha m} = a_{ij\ldots m}. \tag{3.1}$$

This is a system of linear equations in the λ's. Due to the symmetry of A many equations are repeated in (3.1). When we expand $(x_1+x_2+\ldots+x_n)^p$ we obtain a sum

$$\sum_{i=1}^{N} a_i f_i(x),$$

where the a_i are integers, and the f_i are distinct power products of degree p in the x_j $[j=1, 2, \ldots, n]$. We let b_i denote the set of elements $(b_{i1}, b_{i2}, \ldots, b_{in})$ for each i in the set $1, 2, \ldots, \sigma$. The system of equations (3.1) for $\sigma = N$ is then equivalent to the set

$$\sum_{\alpha=1}^{N} f_\beta(b_\alpha)\lambda_\alpha = y_\beta \qquad (\beta=1, 2, \ldots, N), \tag{3.2}$$

where $y_1, y_2, \ldots, y_n$ are equal in some order to the elements of A. We assume that $(y_1, \ldots, y_n)$ is not the zero vector, since then A is trivial. If we can prove that we can choose the b_α in K so that the determinant

$$|D| = |f_\beta(b_\alpha)|$$

is not zero, there exist solutions for the λ's in (3.2), and A is factorable.

We write the matrix D as the matrix $(M_{\varrho\alpha})$ $[\varrho=1, 2, \ldots, n;$ $\alpha=1, 2, \ldots, N]$ where $M_{\varrho\alpha}$ is the minor of D composed of power products $f_\beta(b_\alpha)$ which contain $b_{\alpha\varrho}$ as a factor, and no $b_{\alpha\sigma}$ where $\sigma > \varrho$. The $M_{\varrho\alpha}$ are minors with one column only. We let t_ϱ denote the number of elements (rows) in $M_{\varrho\alpha}$. We construct minors $N_{\varrho\sigma}$ of D $[\varrho, \sigma = 1, 2, \ldots, n]$ such that $N_{\varrho\sigma}$ is the matrix $(M_{\varrho\alpha})$ composed of the columns $M_{\varrho,\alpha}$ where α ranges over the values $g_\sigma + 1, g_\sigma + 2, \ldots, g_{\sigma+1}$, and g_σ is given by

$$g_1 = 0;\ g_\sigma = \sum_{i=1}^{\sigma-1} t_i.$$

The matrix D is then given by $(N_{\varrho\sigma})$ $[\varrho, \sigma = 1, 2, \ldots, n]$. We set $b_{\alpha i} = 0$ in D when α is in the range $g_\sigma + 1, g_\sigma + 2, \ldots, g_{\sigma+1}$, and i in the range $\sigma + 1, \sigma + 2, \ldots, n$. That is, we set each $b_{\alpha i}$ equal to zero that occurs in $N_{\sigma+1,\sigma}, N_{\sigma+2,\sigma}, \ldots, N_{n\sigma}$ and not in $N_{1\sigma}, N_{2\sigma}, \ldots, N_{\sigma\sigma}$, so that we obtain

$$D = \left\| \begin{matrix} N_{11} & N_{12} & \cdots & N_{1,n-1} & N_{1n} \\ O & N_{22} & \cdots & N_{2,n-1} & N_{2n} \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ O & O & \cdots & O & N_{nn} \end{matrix} \right\|.$$

The minor $N_{\sigma\sigma}$ is square and contains only elements $b_{\alpha\lambda}$, where $\lambda \leqq \sigma$. We take $b_{\alpha\sigma} = 1$ for α in the range $g_\sigma + 1, g_\sigma + 2, \ldots, g_{\sigma+1}$. The minor $N_{\sigma\sigma}$ is now, with possibly a rearrangement of rows, of the form

$$\| c_h^g\, d_h^r \cdots f_h^s \| \qquad \text{(column index is } h\text{)},$$

where $h = 1, 2, \ldots, t_\sigma$, and $g, r, \ldots, s$ are $\sigma - 1$ non-negative integral exponents satisfying the inequality

$$(3.3) \qquad g + r + \cdots + s \leqq p - 1.$$

It is understood that $c_h^0, d_h^0, \ldots, f_h^0$ denote 1 for each h. The distinct sets of exponents $(g, r, \ldots, s)$ satisfying (3.3) are evidently in 1-1 correspondence with the integers in the range of h. We set h in 1-1 correspondence with sets $(i, j, \ldots, m)$ of $\sigma - 1$ non-negative integers $i, j, \ldots, m$ subject to the restriction.

$$(3.4) \qquad i + j + \ldots + m \leqq p - 1.$$

For each set $(i, j, \ldots, m)$ and corresponding h we write

$$c_h = \alpha_i,\ d_h = \alpha_j, \ldots, f_h = \alpha_m,$$

where $\alpha_1, \alpha_2, \ldots, \alpha_{p-1}$ are indeterminates over K and $\alpha_0 = 1$. By this choice of the $c_h, \ldots, f_h$ the minor $N_{\sigma\sigma}$ takes on the form

$$(3.5) \qquad (\alpha_i^g\, \alpha_j^r \ldots \alpha_m^s),$$

where the exponents satisfy (3.3) and (3.4). We remark that the exponents in (3.5) form a multipartite row index of $N_{\sigma\sigma}$, and the subscripts form a multipartite column index of $N_{\sigma\sigma}$. We shall need the following lemma.

LEMMA 3.1. *The matrix* (3.5) *is non-singular if* $\alpha_0(=1)$, α_1, $\alpha_2, \ldots, \alpha_{p-1}$ *are distinct elements in* K.

Lemma 3.1 can be proved by showing that the matrix (3.5) is equivalent to a triangular matrix with diagonal minors of the same form as (3.5) with p replaced by smaller integers. Since (3.5) is non-singular if it is of order 1 [that is, $p = 1$ in (3.3) and (3.4)], it follows by induction that Lemma 3.1 holds. Thus A is factorable if K is of order p or more.

To complete the proof of the theorem we assume that K is of order $\psi < p$. It is obviously necessary to consider only p-way matrices where $p \geqq 3$. We shall exhibit a p-way matrix A of order two which is not factorable with respect to K. We define A to be a p-way symmetric matrix $(a_{ij\ldots m})$ of order 2 whose non vanishing elements are those which have exactly ψ subscripts equal to 1; the non-vanishing elements of A are taken equal to one. We let S denote the subset of the equations (3.1) for which $(i, j, \ldots, m)$ range over the sets of values $(2, 2, \ldots, 2)$, $(2, 2, \ldots, 2, 1)$, $(2, 2, \ldots, 2, 1, 1), \ldots, (2, 2, \ldots, 2, 1, \ldots, 1)$, where there are ψ 1's in the last set. If there is no solution for the λ's in the set S there is no solution for the λ's in (3.1). We assume that there is a positive integer σ, and that there are values λ_α, $b_{\alpha i}$, in K so that S is satisfied. The matrix $T = (b_{\alpha i} b_{\alpha j} \ldots b_{\alpha m})$ of coefficients of the λ's in S is the following $(\psi+1)$ by σ rectangular matrix:

$$\left\| \begin{matrix} b_{12}^{p} & b_{22}^{p} & \cdots & b_{\sigma 2}^{p} \\ b_{12}^{p-1} b_{11} & b_{22}^{p-1} b_{21} & \cdots & b_{\sigma 2}^{p-1} b_{\sigma 1} \\ \cdot & \cdot & \cdots & \cdot \\ b_{12}^{p-\psi} b_{11}^{\psi} & b_{22}^{p-\psi} b_{21}^{\psi} & \cdots & b_{\sigma 2}^{p-\psi} b_{\sigma 1}^{\psi} \end{matrix} \right\| .$$

Since K is of order ψ, it follows from the theory of Vandermonian determinants that each possible $(\psi+1)$-st order minor of T vanishes for each choice of the b's. Thus for a choice of the b's the rank of T is r, where $r < \psi + 1$. The matrix

$$T' = \left\| \begin{matrix} & 0 \\ T & \vdots \\ & 0 \\ & 1 \end{matrix} \right\|$$

obtained by adjoining the column of elements $(a_{2\dots 2})$, $(a_{2\dots 21})$, $\dots$, $(a_{2\dots 21\dots 1})$ of A occurring in S, is the augmented matrix of the set S. Since $r \leqq \psi$, the rank of T is $r+1$. The ranks of T and T' are thus unequal. By the well-known theorem that a system of linear equations has a solution if and only if the rank of the matrix of coefficients equals the rank of the augmented matrix, the set S has no solution for the λ's. Thus A is not factorable. The proof of Theorem 3.1 is now complete.

4. *Example.* Let $A = (a_{ij})$ be a symmetric matrix of order 2. Equations (3.2) now become

$$\sum_{\alpha=1}^{3} \lambda_\alpha b_{\alpha 1}^2 = a_{11},$$

$$\sum_{\alpha=1}^{3} \lambda_\alpha b_{\alpha 1} b_{\alpha 2} = a_{12},$$

$$\sum_{\alpha=1}^{3} \lambda_\alpha b_{\alpha 2}^2 = a_{22}.$$

The matrix D is

$$\left\| \begin{matrix} b_{11}^2 & b_{21}^2 & b_{31}^2 \\ b_{11}b_{12} & b_{21}b_{22} & b_{31}b_{32} \\ b_{12}^2 & b_{22}^2 & b_{32}^2 \end{matrix} \right\|.$$

Now

$$D = \left\| \begin{matrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \end{matrix} \right\|,$$

where $M_{1i} = b_{i1}^2$ for $i = 1, 2, 3,$ and

$$M_{2i} = \left\| \begin{matrix} b_{i1} & b_{i2} \\ & b_{i2}^2 \end{matrix} \right\|.$$

We write $N_{11} = M_{11}$; $N_{21} = M_{21}$, $N_{12} = (M_{12}M_{13})$, $N_{22} = (M_{22}M_{23})$, whence

$$D = \left\| \begin{matrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{matrix} \right\|,$$

where N_{11}, N_{22} are square minors of orders 1 and 2, respectively. Setting $b_{12} = 0$, we get

$$D = \left\| \begin{matrix} N_{11} & N_{12} \\ 0 & N_{22} \end{matrix} \right\|.$$

Taking $b_{11} = b_{22} = b_{32} = 1$, we obtain

$$N_{11} = 1,\ N_{22} = \left\| \begin{matrix} 0 & 1 \\ 1 & 0 \end{matrix} \right\| \cdot \left\| \begin{matrix} c_1^0 & c_2^0 \\ c_1^1 & c_2^1 \end{matrix} \right\| .$$

We write $c_1 = \alpha_0$, $c_2 = \alpha_1$, whence the last matrix above becomes

$$\left\| \begin{matrix} \alpha_0^0 & \alpha_1^0 \\ \alpha_0^1 & \alpha_1^1 \end{matrix} \right\| .$$

Taking $\alpha_0 = 1$, and $\alpha_1 \neq 1$, we arrive at a non-singular specialization of D.

5. *Note on the matrix* (3.5). The non-singularity of the matrix (3.5) for distinct α's may be used to give a new proof of the following theorem. The proof is not shorter than existing proofs, but is merely given to illustrate a use of (3.5).

THEOREM 5.1. *Let P be a polynomial of degree p with coefficients in a field K of order $p + 1$ or more. If P is zero for all values of the variables in K, then P is identically zero (that is, all coefficients of P vanish).*

The polynomial $P = P(x, y, \ldots, z)$ can be written as

$$\sum_{r, s, \ldots, t} a_{rs\ldots t} x^r y^s \cdots z^t, \tag{5.1}$$

where $x, y, \ldots, z$ are the variables in P, say n in all, and the summation is over all admissible values of $r, s, \ldots, t$. Let $\alpha_0 = 1$, and $\alpha_0, \alpha_1, \ldots, a_p$ be $p + 1$ distinct elements in K. Let the set $S = (\alpha_i, \alpha_j, \ldots, \alpha_m)$ correspond to the term $a_{ij\ldots m} x^i y^j \cdots z^m$ in (5.1). This correspondence is unique. Substitute the sets of values S for $(x, y, \ldots, z)$ in the equation $P = 0$. We thus obtain the set of linear equations

$$\sum_{r, s, \ldots, t} a_{rs\ldots t} \alpha_i^r \alpha_j^s \cdots \alpha_m^t = 0$$

homogeneous in the a's. Since by Lemma 3.1 the matrix $(\alpha_i^r a_j^s \cdots a_m^t)$ of coefficients is non-singular, the a's vanish.

ARMOUR INSTITUTE OF TECHNOLOGY AND
THE INSTITUTE FOR ADVANCED STUDY.

(Received July 8th, 1939.)