Atti Accademia Nazionale dei Lincei

Classe Scienze Fisiche Matematiche Naturali

Rendiconti

A. Duane Porter

Orthogonal similarity for skew matrices in GF(q)

Atti della Accademia Nazionale dei Lincei. Classe di Scienze Fisiche, Matematiche e Naturali. Rendiconti, Serie 8, Vol. **42** (1967), n.6, p. 757–762.

Accademia Nazionale dei Lincei

<http://www.bdim.eu/item?id=RLINA_1967_8_42_6_757_0>

Atti della Accademia Nazionale dei Lincei. Classe di Scienze Fisiche, Matematiche e Naturali. Rendiconti, Accademia Nazionale dei Lincei, 1967.

Algebra. — *Orthogonal similarity for skew matrices in* GF(q). Nota di A. DUANE PORTER, presentata (*) dal Socio B. SEGRE.

RIASSUNTO. — Si assegna una forma canonica per le matrici quadrate su di un campo di Galois, simile a quella ben nota relativa alle matrici sul campo reale rispetto al gruppo ortogonale.

It is well-known that over the real field every $n \times n$ skew matrix of rank $2m$ is orthogonally similar to a matrix of the form $D = \operatorname{diag}[D_1, \cdots, D_m, R]$, where $R = \operatorname{diag}(0, \cdots, 0)$ is $(n - 2m) \times (n - 2m)$, and

$$D_j = \begin{bmatrix} 0 & b_j \\ -b_j & 0 \end{bmatrix} \quad , \quad 1 \leq j \leq m,$$

the nonzero eigenvalues of A being $\pm b_j i$, $i^2 = -1$. This canonical form is not valid over a finite field since (1) the inner product of a nonzero vector with itself may be zero and (2) even if a given vector has a nonzero inner product with itself, one may not be able to normalize the vector since not all elements of a finite field have square roots in the field. In view of the use of canonical forms to solve certain problems involving matric equations over GF(q), [1], [2], [3], [4], [5], [6], it would seem desirable to have a simple canonical form under orthogonal similarity available. The purpose of this paper is to obtain such a form for certain skew matrices.

In order to obtain sufficient conditions for a canonical form which is simple enough to be useful, and which resembles the form in the real case, some natural conditions are placed on the matrix. The rather surprising result, as noted in Th. 3.1, is that these conditions are also necessary for the canonical form.

2. NOTATION AND PRELIMINARIES.—Let $F = GF(q)$ be the finite field of $q = p^r$ elements, p odd, and θ an element of GF(q^2) such that $\theta \notin F$, but $\theta^2 = u \in F$. Then $d = a + b\theta \in GF(q^2)$ if $a, b \in F$, and we denote GF(q^2) by $F(\theta)$. It is clear that the identical field GF(q^2) is obtained regardless of our choice of θ satisfying the above conditions. By the *conjugate* of d, we mean $\bar{d} = a - b\theta$.

A matrix $B = (b_{ij})$, $b_{ij} \in F$, will be called *skew* if $B = B' = (b_{ij})'$ where the prime denotes transpose. If $A = (a_{ij})$, $a_{ij} \in F(\theta)$, by the *conjugate transpose* of A, we mean the matrix $\bar{A}' = (\bar{a}_{ij})'$. A square matrix P with elements from F is said to be *orthogonal* if $P'P = I$ = identity matrix. Two matrices A and B are said to be *orthogonally similar* over F if there is an orthogonal matrix P over F such that $P'AP = B$.

(*) Nella seduta del 21 giugno 1967.

If $\alpha = \text{col}(a_1, \cdots, a_n)$, $\beta = \text{col}(b_1, \cdots, b_n)$, $a_i, b_i \in F(\theta)$, $1 \leq i \leq n$, we define the *inner product* of α and β by $\alpha^* \beta = \bar{a}_1 b_1 + \cdots + \bar{a}_n b_n$. Clearly, if any $a_i \in F$, then $\bar{a}_i = a_i$ for these elements. The two vectors α, β will be called *orthogonal* if $\alpha^* \beta = \beta^* \alpha = 0$, and α is *normal* if $\alpha^* \alpha = 1$ = unity of F. A set of vectors $\alpha_1, \cdots, \alpha_n$ are *linearly independent* if $a_1 \alpha_1 + \cdots + a_n \alpha_n$ = zero vector holds only for scalars $a_1 = a_2 = \cdots = a_n = 0$.

We will say a matrix or vector is *over* F or $F(\theta)$ if its elements are from F or $F(\theta)$, respectively.

3. A SKEW CANONICAL FORM.—We now state the main theorem of this paper.

THEOREM 3.1.—*Let* A *be an* $n \times n$ *skew matrix over* F *of rank* $2m$. *Let* $D = diag[D_1, \cdots, D_m, R]$, *where* $R = diag[0, \cdots, 0]$ *is* $n - 2m \times n - 2m$, *and for* $1 \leq j \leq m$

$$D_j = \begin{bmatrix} 0 & b_j \\ b_j u & 0 \end{bmatrix} \quad , \quad b_j \in F \quad , \quad \theta^2 = u.$$

Then A *is orthogonally similar over* F *to* D *if and only if the following conditions hold:*

(1) all nonzero eigenvalues of A occur in conjugate pairs $\pm b_j \theta$, $b_j \in F$;

(2) if r = arbitrary eigenvalue of A of multiplicity k_r, then the dimension of the null space of $rI - A$ is k_r;

(3) if $n \neq 2m$ then the null space of A has an orthonormal basis;

(4) if $r = b\theta$ = arbitrary nonzero eigenvalue of A, then there exists an orthogonal basis $\alpha_1, \cdots, \alpha_t$ of the null space of $rI - A$ with the property that if α = arbitrary one of $\alpha_1, \cdots, \alpha_t$ and we define $\gamma = \alpha + \bar{\alpha}$ and $\delta = \theta(\alpha - \bar{\alpha})$, then for those nonzero γ, δ, we have $\gamma^* \gamma = c^2$, $0 \neq c \in F$, and $\delta^* \delta = d^2$, $0 \neq d \in F$.

We first prove the sufficiency of the conditions, and to facilitate our discussion, we note the following lemmas.

LEMMA 3.2.—*Let* A *be a skew matrix over* F *which satisfies* (1) *of Th. 3.1 Then the following are valid*:

(1) if α is an eigenvector of A corresponding to eigenvalue r, then α is an eigenvector of A′ corresponding to $-r$, and $\bar{\alpha}$ is an eigenvector of A corresponding to $-r$;

(2) if $\alpha_1, \cdots, \alpha_t$ form a basis for the null space of $rI - A$, then $\bar{\alpha}_1, \cdots, \bar{\alpha}_t$ form a basis for the null space of $\bar{r}I - A$;

(3) eigenvectors of A corresponding to distinct eigenvalues are orthogonal.

With the added condition (1) in the above lemma, the proofs of the various parts are like the corresponding proofs in the real case and will not be repeated. Similarly, we may state

LEMMA 3.3.—*If* $\alpha_1, \cdots, \alpha_t$ *are mutually orthogonal vectors over* F *with nonzero inner product, they are linearly independent.*

LEMMA 3.4.—*Let* A *be a skew matrix over* F *which satisfies* (1) *and* (4) *of Theorem 3.1. Then*

(1) γ and δ are nonzero vectors over F;
(2) $A\gamma = b\delta$; $A\delta = bu\gamma$;
(3) $A'\gamma = -b\delta$; $A'\delta = -bu\gamma$;
(4) $\gamma'\delta = \delta'\gamma = 0$.

Proof: (1). Since $\gamma = \alpha + \bar{\alpha}$, then γ is over F. If $\gamma = 0$, then all elements of α must be of the form $a_i\theta$, $1 \leq i \leq n$, and $a_i \in F$. But then $A\alpha = r\alpha = b\theta\alpha \neq 0$ leads to a contradiction since all elements of $A\alpha$ would be of the form $c\theta$ while elements of $b\theta\alpha$ would all be from F. Hence, $\gamma \neq 0$. Since $\delta = \theta(\alpha - \bar{\alpha})$ also has elements from F, a similar argument shows $\delta \neq 0$, so that (1) is established.

(2) The following two equalities establish this part.
$$A\gamma = A(\alpha + \bar{\alpha}) = r\alpha + \bar{r}\bar{\alpha} = r(\alpha - \bar{\alpha}) = b\theta(\alpha - \bar{\alpha}) = b\delta,$$
$$A\delta = \theta A(\alpha - \bar{\alpha}) = \theta(r\alpha - \bar{r}\bar{\alpha}) = \theta r(\alpha + \bar{\alpha}) = b\,\theta^2(\alpha + \bar{\alpha}) = bu\gamma.$$

(3) In view of Lemma 3.2 (1) $A'\alpha = \bar{r}\alpha$ so that $A'\bar{\alpha} = r\bar{\alpha}$, and
$$A'\gamma = A'(\alpha + \bar{\alpha}) = \bar{r}\alpha + r\bar{\alpha} = -r(\alpha - \bar{\alpha}) = -b\theta(\alpha - \bar{\alpha}) = -b\delta,$$
$$A'\delta = \theta A'(\alpha - \bar{\alpha}) = \theta(\bar{r}\alpha - r\bar{\alpha}) = -\theta r(\alpha + \bar{\alpha}) = -bu\gamma.$$

(4) Since $\gamma'\delta \in F$, $(\gamma'\delta)' = \gamma'\delta$ so that $\gamma'\delta = \delta'\gamma$.
Also $\gamma' A\gamma = \gamma'(A\gamma) = \gamma' b\delta = b\gamma'\delta$
$$= (\gamma' A)\gamma = (A'\gamma)'\gamma = (-b\delta)'\gamma = -b\gamma'\gamma = -b\gamma'\delta.$$

Clearly, since $b \neq 0$, the above yields $\gamma'\delta = 0$.

We now construct an orthogonal matrix P over F such that $P'AP = D =$ matrix of Th. 3.1. Let $\pm b\theta$, $b \neq 0$, be a pair of nonzero eigenvalues of A and denote them by r and $\bar{r}$. Let $\alpha_1, \cdots, \alpha_t$ be an orthogonal basis of the null space of $rI - A$ which exist by (4) of the theorem. Clearly, $\bar{\alpha}_1, \cdots, \bar{\alpha}_t$ are orthogonal so by Lem 3.2 (2), they form an orthogonal basis for the null space of $\bar{r}I - A$. If we let α be any one of $\alpha_1, \cdots, \alpha_t$ and $\bar{\alpha}$ its conjugate, we may use these vectors to form γ and δ as in condition (4), Also, if we note Lemma 3.4. (4), and let $\gamma_j = \gamma/c$, $\delta_j = \delta/c$, then γ_j, δ_j are normal, orthogonal, and over F. Hence, we may obtain a set $\gamma_1, \delta_1, \cdots, \gamma_t, \delta_t$ of $2t$ normal vectors over F such that $\gamma_j' \delta_j = 0$, $1 \leq j \leq t$.

Each pair of conjugate null spaces of A will have sets of basis vectors as described above, and since rank $A = 2m$, if we form the union of these orthogonal bases sets, we obtain a set of $2m$ vectors

$$\alpha_1, \bar{\alpha}_1, \cdots, \alpha_m, \bar{\alpha}_m. \tag{3.5}$$

In view of Lemma 3.2 (3) and the choice of the α_i, the above vectors are mutually orthogonal. We can replace each pair $\alpha_j, \bar{\alpha}_j$ by γ_j, δ_j as constructed above and so obtain a set of $2m$ normal vectors with each pair γ_j, δ_j orthogonal,

$$\gamma_1, \delta_1, \cdots, \gamma_m, \delta_m. \tag{3.6}$$

We let $\beta_1, \cdots, \beta_s$, $s \geq 0$, be the orthonormal basis of A which exists by Th. 3.1 (3), and consider the set of normal vectors.

$$\delta_1, \gamma_1, \cdots, \delta_m, \gamma_m, \beta_1, \cdots, \beta_s. \tag{3.7}$$

The number of vectors in (3.7) equals the sum of the dimensions of the null spaces of $rI - A$ for all eigenvectors r of A. Hence, by condition (2) of the theorem, there are n vectors in the above set.

Since each β_i is an eigenvector corresponding to zero, by Lemma 3.2 (3), β_i is orthogonal to each vector in the set (3.5), so is orthogonal to a linear combination of these vectors. Hence, each β_i is orthogonal to all vectors in the set (3.6) and so also orthogonal to all vectors except itself in the set (3.7).

Let $\sigma_j =$ either γ_j or δ_j. Since (3.5) consists of mutually orthogonal vectors, α_k, for $k \neq j$, is orthogonal to both α_j and $\bar{\alpha}_j$ so is orthogonal to σ_j. Likewise, α_k is orthogonal to σ_j. Thus, σ_j is clearly orthogonal to $\sigma_k =$ either γ_k or δ_k, since these are linear combinations of α_k and $\bar{\alpha}_k$.

Hence, in view of the above comments, the set (3.7) consists of n normal, mutually orthogonal vectors, so that by Lemma 3.3, they are linearly independent.

We define P to be the matrix with the vectors (3.7) in that order as its columns, then P is nonsingular and orthogonal since $P'P = I$. In view of Lemma 3.4 (2), we have, for $1 \leq j \leq m$, $A\delta_j = ub_j\gamma_j$ and $A\gamma_j = b_j\delta_j$, which may be written as

$$A(\delta_j, \gamma_j) = (\delta_j, \gamma_j)\begin{bmatrix} 0 & b_j \\ b_j u & 0 \end{bmatrix}. \tag{3.8}$$

Also, by choice of β_i, $A\beta_i = 0$, $1 \leq i \leq s$. Combining this with the definition of P and (3.8), we obtain $AP = PD$ or $P'AP = D = \operatorname{diag}[D_1, \cdots, D_m, R]$, with D_i, $1 \leq i \leq m$, and R as stated in the theorem. Hence, the four conditions of Th. 3.1 are sufficient for orthogonal similarity to D which completes the first part of the proof.

We now suppose A is orthogonally similar to D and show the four conditions hold.

The eigenvalues of D are the roots of $|xI - D| = 0$, and $|xI = D| = \operatorname{diag}(E_1, \cdots, E_m, S)$, where $S = \operatorname{diag}(x, \cdots, x)$ is $n - 2m \times n - 2m$, and, for $1 \leq i \leq m$,

$$E_t = \begin{bmatrix} x & -b_i \\ -b_i u & x \end{bmatrix}.$$

Hence, $|xI - D| = (x^2 - b_1^2 u) \cdots (x^2 - b_m^2 u)\, x^{n-2m}$, so the nonzero eigenvalues of D are $\pm b_j\theta$, $1 \leq j \leq m$. Since A and D have the same eigenvalues, (1) is valid.

If $n \neq 2m$, then the $n - 2m$ columns $P_{2m+1}, \cdots, P_n$ of the matrix P, such that $P'AP = D$, are normal and mutually orthogonal. Also, $AP_k = 0$, $2m + 1 \leq k \leq n$, so they form an orthonormal basis for the null space of A so that (3) is valid.

For use in proving (2) and (4), we construct eigenvectors of A as follows: Let P_{2j-1}, P_{2j}, $1 \leq j \leq m$, represent the first $2m$ columns of the matrix P given above. Define

$$(3.9)\qquad \begin{cases} \alpha_j = \frac{1}{2} P_{2j} + \frac{\theta}{2u} P_{2j-1}, \\ \bar{\alpha}_j = \frac{1}{2} P_{2j} - \frac{\theta}{2u} P_{2j-1}. \end{cases}$$

In view of the definitions of P and D, we have $AP_{2j-1} = b_j u P_{2j}$ and $AP_{2j} = bP_{2j-1}$. A direct calculation will show

$$(3.10)\qquad \begin{cases} A\alpha_j = \frac{b_j\theta}{2}\left(P_{2j} + \frac{\theta}{u} P_{2j-1}\right) = b_j \theta \alpha_j, \\ A\bar{\alpha}_j = \frac{b_j\theta}{2}\left(-P_{2j} + \frac{\theta}{u} P_{2j-1}\right) = -b_j \theta \bar{\alpha}_j. \end{cases}$$

Hence, α_j and $\bar{\alpha}_j$ are eigenvectors of A corresponding to $b_j\theta$ and $-b_j\theta$, respectively, for $1 \leq j \leq m$.

Without loss of generality, we may assume the 2×2 blocks appear on the diagonal of D in any order. Of the m pairs of nonzero conjugate eigenvalues of A, suppose t pairs $c_1\theta, -c_1\theta, \cdots, c_t\theta, -c_t\theta$ are distinct, i.e., $c_i \neq c_j$, $1 \leq i \neq j \leq t$. Suppose $c_i\theta$ has multiplicity m_i where it is clear that each k_r of condition (2) equals some m_i. We then assume the diagonal blocks of D have the following order:

$$(3.11)\qquad \begin{cases} b_j = c_1, \quad 1 \leq j \leq m_1, \quad \text{and for} \quad 1 < i \leq t, \\ b_j = c_i, \quad m_1 + \cdots + m_{i-1} < j \leq m_1 + \cdots + m_i. \end{cases}$$

To prove (2), we let $r = b_j\theta = c_i\theta$ be an arbitrary eigenvalue of A multiplicity $k_r = m_i$. Let $s_i = m_1 + \cdots + m_{i-1}$, $s_i + 1 = m_1 + \cdots + m_i$, where $s_i = 1$ if $i = 1$. Then α_j and $\bar{\alpha}_j$, $s_i \leq j \leq s_i + 1$, are eigenvectors of A corresponding to r and $\bar{r}$, respectively. Suppose there are scalars a_{2j-1}, a_{2j}, $s_i \leq j \leq s_i + 1$ such that

$$\sum_{j=s_i}^{s_i+1} (a_{2j-1}\alpha_j + a_{2j}\bar{\alpha}_j) = 0.$$

Upon substitution from (3.9) for α_j and $\bar{\alpha}_j$, and after recombining terms, the above sum may be written as

$$\sum_{j=s_i}^{s_i+1} \left[(a_{2j-1} + a_{2j}) P_{2j} + (a_{2j-1} - a_{2j}) \frac{\theta}{2u} P_{2j-1}\right] = 0.$$

Since P_{2j} and P_{2j-1} are over F, the above yields the following two equalities:

$$\sum_{j=s_i}^{s_i+1} (a_{2j-1} + a_{2j}) P_{2j} = 0 \quad ; \quad \sum_{j=s_i}^{s_i+1} (a_{2j-1} - a_{2j}) P_{2j-1} = 0.$$

But, the colums of P are linearly independent, so that, for $s_i \leq j \leq s_i + 1$, $a_{2j-1} + a_{2j} = 0 = a_{2j-1} - a_{2j}$. Hence, $a_{2j-1} = a_{2j} = 0$ and the set of

$2m_i$ vectors α_j, $\bar{\alpha}_j$, $s_i \leq j \leq s_i + 1$ are linearly independent. Thus, the subset α_j, $s_i \leq j \leq s_i + 1$ is linearly independent so that the dimension of the null space of $rI - A = m_i = k_r$ and (2) has been proven.

To prove (4), we let $r = b_j$ be arbitrary, and let α_j, $\bar{\alpha}_j$, $s_i \leq j \leq s_i + 1$ be the basis for the null spaces of $rI - A$ and $\bar{r}I - A$, respectively, where α_j, $\bar{\alpha}_j$ are defined by (3.10) and s_i is as defined above. Let $\gamma = \alpha_j + \bar{\alpha}_j$, and $\delta = \theta(\alpha_j - \bar{\alpha}_j)$ for some fixed j. Clearly, $\gamma = P_{2j}$ and $\delta = P_{2j-1}$ so that $\gamma^* \gamma = P'_{2j} P_{2j} = 1 = 1^2$, and $\delta^* \delta = P'_{2j-1} P_{2j-1} = 1 = 1^2$. Also, since the columns P_{2j-1}, P_{2j} of P are mutually orthogonal, it is easily seen that the set α_j, $s_i \leq j \leq s_i + 1$, is mutually orthogonal. Hence, (4) is established.

4. ANOTHER STATEMENT OF THE THEOREM.—We first prove:

LEMMA 4.1.—*Let* A *be an* $n \times n$ *skew matrix over* F *of rank* $2m$ *with eigenvalues in* $F(\theta)$. *Suppose that to each distinct eigenvalue of* A *corresponds at least one eigenvector with nonzero inner product. Then the nonzero eigenvalues of* A *occur in conjugate pairs* $\pm b_j \theta$, $b_j \in F$, $1 \leq j \leq m$.

Proof: Let $r =$ arbitrary non zero eigenvalue of A. Then there is a vector α over $F(\theta)$ such that $\alpha^* \alpha = a \neq 0$ and $A\alpha = r\alpha$. Thus, $\alpha^* A \alpha = r \alpha^* \alpha$ and $(\alpha^* A \alpha)^* = \alpha^* A' \alpha = \bar{r} \alpha^* \alpha$ so that, since A is skew, $\alpha^* A \alpha = -\bar{r} \alpha^* \alpha$ which implies $r = -\bar{r}$. Hence, if $r = a + b\theta$, then $r = 0 + b\theta = b\theta$, $b \in F$. Taking conjugates in the equation $A\alpha = r\alpha$, we obtain $A\bar{\alpha} = \bar{r}\bar{\alpha}$ so $\bar{r} = -b\theta$ is also an eigenvalue of A. Since rank $A = 2m$, then A has m pairs of eigenvalues $\pm b_j \theta$, $b_j \in F$, which proves the lemma.

In order to check for orthogonal similarity by Th. 3.1, it is necessary to first find all nonzero eigenvalues of the given matrix. This might not be necessary if we can simply establish that the matrix satisfies the conditions of the above lemma. Hence, we restate Th. 3.1 as follows:

THEOREM 4.2. — *Let A be a skew matrix which satisfies the conditions of Lemma 4.1. Then* A *is orthogonally similar over* F *to* D *of Theorem 3.1 if and only if* A *satisfies* (2), (3), (4) *of Theorem 3.1.*

Proof: In view of Lemma 4.1, A satisfies (1) of Th. 3.1, so if A also satisfies (2), (3), (4) then A is orthogonally similar to D. The converse follows immediately from Th. 3.1.

REFERENCES.

[1] L. CARLITZ, *Representations by skew forms in a finite field*, « Archiv Der Nathematik », 5 19–31 (1954).

[2] L. CARLITZ and JOHN H. HODGES, *Representations by hermitian forms in a finite field*, « Duke Mathematical Journal », *22* 393–405 (1955).

[3] JOHN H. HODGES, *Representations by bilinear forms in a finite field*, « Duke Mathematical Journal », *22*, 497–509 (1955).

[4] JOHN H. HODGES, *Exponential sums for skew matrices in a finite field*, « Archiv Der Mathematik », *7*, 116–121 (1956).

[5] JOHN H. HODGES, *Weighted partitions for symmetric matrices in a finite field*, « Math. Zeitschr. », Bd. 66, S. 13–24, 13–24 (1956).

[6] JOHN H. HODGES, *Some matrix equations over a finite field*, « Annali di Matematica », *44*, 245–250 (1957).