AD-764 564

A SADDLEPOINT THEOREM FOR SELF-DUAL LINEAR SYSTEMS

Marjorie L. Stein

Wisconsin University

Prepared for:

Army Research Office-Durham

June 1973

THE UNIVERSITY OF WISCONSIN

A SADDLEPOINT THEOREM FOR SELF-DUAL LINEAR SYSTEMS

Marjorie L. Stein

MRC Technical Summary Report # 1311
June 1973

MATHEMATICS RESEARCH CENTER

Address:

Mathematics Research Center
University of Wisconsin-Madison
Madison, Wisconsin 53706
U.S.A.

THE UNIVERSITY OF WISCONSIN-MADISON
MATHEMATICS RESEARCH CENTER

Contract No. DA-31-124-ARO-D-462

A SADDLEPOINT THEOREM FOR SELF-DUAL LINEAR SYSTEMS

Marjorie L. Stein

MRC Technical Summary Report #1311
June 1973

Received November 9, 1972

Madison, Wisconsin 53706

AR 70-31 Unclassified

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Mathematics Research Center University of Wisconsin, Madison, Wis. 53706	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP None

3. REPORT TITLE

A SADDLEPOINT THEOREM FOR SELF-DUAL LINEAR SYSTEMS

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Summary Report: no specific reporting period.

5. AUTHOR(S) (First name, middle initial, last name)

Marjorie L. Stein

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
June 1973	26	15

8a. CONTRACT OR GRANT NO. Contract No. DA-31-124-ARO-D-462 b. PROJECT NO. None c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) #1311 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report) None

10. DISTRIBUTION STATEMENT

Distribution of this document is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
None	Army Research Office-Durham, N.C.

13. ABSTRACT

Let A be a skew matrix of order n over an ordered field. There is a finite class of skew matrices $\bar{A}$ such that $XA = Y$ and $\bar{X}\bar{A} = \bar{Y}$ have have the same solution sets, where $\bar{x}_i = y_i$ and $\bar{y}_i = x_i$ for some indices i (perhaps none) and $\bar{x}_i = x_i$, $\bar{y}_i = y_i$ for the remaining i. We show that for each index h, $1 \leqq h \leqq n$, there exists an $\bar{A}$ such that $\bar{a}_{hj} \geqq 0$ for all j. A proof of von Neumann's Minimax Theorem for symmetric games is one application of the theory.

DD FORM 1 NOV 65 1473

Unclassified

Security Classification

AGO 5696A

ABSTRACT

Let A be a skew matrix of order n over an ordered field. There is a finite class of skew matrices $\bar{A}$ such that $XA = Y$ and $\bar{X}\bar{A} = \bar{Y}$ have the same solution sets, where $\bar{x}_i = y_i$ and $\bar{y}_i = x_i$ for some indices i (perhaps none) and $\bar{x}_i = x_i$, $\bar{y}_i = y_i$ for the remaining i. We show that for each index h, $1 \leq h \leq n$, there exists an $\bar{A}$ such that $\bar{a}_{hj} \geqq 0$ for all j. A proof of von Neumann's Minimax Theorem for symmetric games is one application of the theory.

A SADDLEPOINT THEOREM FOR SELF-DUAL LINEAR SYSTEMS*

Marjorie L. Stein

A finite class of skew coefficient matrices over an ordered field can be associated with a self-dual linear system. Pivot techniques allow us to transform a given matrix of the class into another while maintaining the same sets of solutions. The exchange of two independent variables in the system for two corresponding dependent ones maintains skewsymmetry.

Our main theorem says that in a finite number of pivot steps we can transform any skew matrix into a skew matrix with a saddlepoint — that is, a nonnegative row and corresponding nonpositive column. A tableau form introduced by A. W. Tucker permits an immediate translation of this result about matrices into a statement about types of solutions that must occur in the associated self-dual linear system. This theorem and its corollaries yield an elementary proof of the Minimax Theorem for symmetric games.

* This paper contains portions of the author's doctoral dissertation [9], written at Princeton University under the supervision of Professor A. W. Tucker.

Sponsored by the United States Army under Contract No. DA-31-124-ARO-D-462.

§1. Pivoting in Skew Matrices

Dual systems of linear equations can be represented by the following tableau form of A. W. Tucker [12]:

$$
\begin{array}{c|ccc|l}
 & v_1 & \cdots & v_n & \\
\hline
x_1 & a_{11} & \cdots & a_{1n} & = -u_1 \\
\vdots & & M & & \vdots \\
x_m & a_{m1} & \cdots & a_{mn} & = -u_m \\
\hline
 & =y_1 & \cdots & =y_n &
\end{array}
\qquad \text{or} \qquad
\begin{array}{c|c|l}
 & V & \\
\hline
X & M & = -U \; . \\
\hline
 & = Y &
\end{array}
$$

Here M is an $m \times n$ matrix, X is $1 \times m$, Y is $1 \times n$, U is $m \times 1$ and V is $n \times 1$. This format expresses both a column system:

$$
\left.
\begin{array}{c}
x_1 a_{11} + \cdots + x_m a_{m1} = y_1 \\
\vdots \\
x_1 a_{1n} + \cdots + x_m a_{mn} = y_n
\end{array}
\right\} \; XM = Y
$$

and a row system:

$$
\left.
\begin{array}{c}
a_{11} v_1 + \cdots + a_{1n} v_n = -u_1 \\
\vdots \\
a_{m1} v_1 + \cdots + a_{mn} v_n = -u_m
\end{array}
\right\} \; MV = -U \; .
$$

The systems are always dual in the sense of orthogonality, for

$$XU + YV = X(-MV) + (XM)V = 0 \quad .$$

Definition 1: A column solution for the tableau is a real $1 \times (m+n)$ matrix $Z = (X, Y)$ such that $XM = Y$. A row solution is a real $1 \times (m+n)$ matrix $W = (U^T, V^T)$ such that $MV = -U$.

If variables and coefficients take values in some commutative field F, then assuming nondegeneracy, the column system generates an m-dimensional subspace and the row system an n-dimensional subspace of the (m+n)-space F^{m+n} determined by the vectors $Z = (X, Y)$ and $W = (U^T, V^T)$. In an ordered field, the subspaces spanned by the row and column systems are not only orthogonal but also complementary in that the entire (m+n)-space is determined by their direct sum.

Since this paper is concerned with linear inequalities and notions of positivity, we necessarily restrict ourselves to ordered fields and therefore to complementary orthogonal systems. We specifically use the real numbers because applications such as matrix games require numerical solutions to practical problems. The results and methods are nevertheless easily generalized to arbitrary ordered fields.

Throughout this paper we use the symbols $\oplus$ and $\ominus$ respectively to denote nonnegative and nonpositive elements. Matrix inequalities hold componentwise. Thus $M > 0$ means each component of M is positive, $M \geqq 0$ means each component of M is nonnegative, and $M = 0$ means M is the zero or trivial matrix. We adopt the special convention that $M \underline{\geq} 0$ means $M \geqq 0$ and M is nontrivial.

From the tableau it is easy to see that information about the sign patterns in the coefficient matrix M often provides a clue to the types of solutions we can expect to find. For example, if M has a nonnegative first row, then we can obtain a column solution $Z \underline{\geq} 0$ by setting $x_1 > 0$ and $x_i = 0$ for $i > 1$.

We temporarily restrict our investigations to linear systems with skew coefficient matrices over ordered fields, denoting such matrices by the letter A .

Definition 2: A matrix A is skew when $A = -A^T$. A skew matrix is necessarily square, and we say its order is n if it is $n \times n$. Over an ordered field the diagonal entries are always zero.

We present the tableau of a skew matrix in the following abbreviated form:

$$\begin{array}{cc} & X^T \\ X & \boxed{\quad A \quad} = -Y^T \\ & = Y \end{array} .$$

By negative transposition it is clear that any column solution (X, Y) is also a row solution when written (Y, X) and conversely. Hence for the skew case we are justified in dispensing with $W = (U^T, V^T)$ as the row solution, and we say that the dual linear systems are isomorphic or that either system is self-dual.

__Theorem 1:__ A matrix A of order n (over an ordered field) is skew if and only if $XAX^T = 0$ for all $1 \times n$ matrices X.

__Proof:__ If A is skew, then for all $1 \times n$ matrices X we have $X(-A^T)X^T = XAX^T = (XAX^T)^T = XA^TX^T$. Hence $XAX^T = 0$.

Conversely, suppose $XAX^T = 0$ for all $1 \times n$ matrices X. Given an integer i, choose X' such that $x'_i = 1$ and $x'_k = 0$ for $k \neq i$. Then for each i, $X'A(X')^T = a_{ii} = 0$. Now given distinct integers i and j, choose X'' such that $x''_i = x''_j = 1$ and $x''_k = 0$ for $k \neq i, j$. Then

$$0 = X''A(X'')^T = (a_{i1} + a_{j1}, \ldots, a_{in} + a_{jn})(X'')^T$$

$$= a_{ii} + a_{ji} + a_{ij} + a_{jj}$$

$$= a_{ji} + a_{ij} .$$

Hence $a_{ji} = -a_{ij}$ for each i, j where $1 \leqq i \leqq n$ and $1 \leqq j \leqq n$, which means $A = -A^T$. ■

We should like to transform a given tableau into another one with a different pattern of signs in the coefficient matrix while simultaneously maintaining the original solution sets, up to permutations of solution components. Exchanging some independent variable x_i for some dependent variable y_j will serve the purpose; for such an exchange corresponds precisely to solving a column equation for the chosen independent variable, substituting that expression into the remaining equations and simplifying the results. The analogous statement holds for the row system.

The quadratic form XAX^T associated with the skew matrix A is simply the inner product YX^T. Theorem 1 says $YX^T = y_1x_1 + \ldots + y_nx_n = 0$.

Consequently we can preserve this inner product — and hence skewsymmetry in the matrix — while exchanging pairs of variables provided the exchange of x_i for y_j is concomitant with the exchange of x_j for y_i. Note that if $i = j$ the variable x_i does not appear in the expression for y_i since its coefficient is $a_{ii} = 0$. The exchange operation can be defined only for pairs of variables where $i \neq j$ and $a_{ij} = -a_{ji} \neq 0$.

<u>Definition 3</u>: A <u>principal pair pivot</u> or <u>P-pivot</u> in the tableau of a self-dual linear system is a simultaneous exchange of variables x_i with y_j and x_j with y_i, provided $a_{ij} = -a_{ji} \neq 0$, and a permutation that exchanges row i with row j and column i with column j while fixing the remaining rows and columns.

To see what happens in the skew tableau we consider an example of order $n = 4$.

<u>Example 1</u>: The tableau

$$
\begin{array}{c|cccc|c}
 & x_1 & x_2 & x_3 & x_4 & \\
\hline
x_1 & 0 & r & p^* & t & = -y_1 \\
x_2 & -r & 0 & q & s & = -y_2 \\
x_3 & -p^* & -q & 0 & u & = -y_3 \\
x_4 & -t & -s & -u & 0 & = -y_4 \\
\hline
 & = y_1 & = y_2 & = y_3 & = y_4 &
\end{array}
$$

represents a self-dual system of four linear equations:

 #1311

(1) $$-x_2r - x_3p - x_4t = y_1$$

(2) $$x_1r - x_3q - x_4s = y_2$$

(3) $$x_1p + x_2q - x_4u = y_3$$

(4) $$x_1t + x_2s + x_3u = y_4 \ .$$

A P-pivot on the starred entries $p, -p \neq 0$ in the matrix corresponds to solving equations (3) and (1) for x_1 and x_3 respectively, substituting the resulting expressions into equations (2) and (4), simplifying the results and switching positions of equations (1) and (3) in both systems. Thus

$$x_3 = -y_1p^{-1} - x_2rp^{-1} - x_4tp^{-1} \text{ from equation (1) },$$

and $$x_1 = y_3p^{-1} - x_2qp^{-1} - x_4up^{-1} \text{ from equation (3) }.$$

Substitution of these expressions into equations (2) and (4) and simplification yield:

$$y_2 = (y_3p^{-1} - x_2qp^{-1} + x_4up^{-1})r - (-y_1p^{-1} - x_2rp^{-1} - x_4tp^{-1})q - x_4s$$
$$= y_3p^{-1}r + y_1p^{-1}q - x_4[s - p^{-1}(tq + ru)]$$

and $$y_4 = (y_3p^{-1} - x_2qp^{-1} + x_4up^{-1})t + x_2s + (-y_1p^{-1} - x_2rp^{-1} - x_4tp^{-1})u$$
$$= y_3p^{-1}t - y_1p^{-1}u + x_2[s - p^{-1}(tq + ru)] \ .$$

The resulting new tableau is:

	y_1	x_2	y_3	x_4	
y_1	0	$p^{-1}q$	$-p^{-1}$	$-p^{-1}u$	$= -x_1$
x_2	$-qp^{-1}$	0	$-rp^{-1}$	$[s-p^{-1}(tq+ru)]$	$= -y_2$
y_3	p^{-1}	$p^{-1}r$	0	$p^{-1}t$	$= -x_3$
x_4	up^{-1}	$-[s-p^{-1}(tq+ru)]$	$-tp^{-1}$	0	$= -y_4$.
	$= x_1$	$= y_2$	$= x_3$	$= y_4$	

Note that the marginal variables remain consecutively numbered. It is also evident that the tableau form with minus signs in the right-hand margin ensures that a single calculation suffices for both linear systems.

From the preceding example it is easy to generalize the notion of P-pivot to demonstrate its operation in arbitrarily large finite self-dual systems. Given the skew matrix A, a P-pivot on a_{ij}, $a_{ji} \neq 0$ in the tableau yields a skew matrix $\bar{A}$ with the self-dual solution $(\bar{X}, \bar{Y})$ such that $\bar{x}_i = y_i$, $\bar{x}_j = y_j$, $\bar{y}_i = x_i$, $\bar{y}_j = x_j$ and $\bar{x}_k = x_k$, $\bar{y}_k = y_k$ for all $k \neq i, j$. The new matrix $\bar{A}$ is given by the following entries:

$$\bar{a}_{ij} = -a_{ij}^{-1} \quad \text{for } i \neq j$$

$$\bar{a}_{ih} = a_{ij}^{-1} a_{hj} \quad \text{for } h \neq i, j$$

$$\bar{a}_{kj} = -a_{ij}^{-1} a_{ik} \quad \text{for } k \neq i, j$$

$$\bar{a}_{kh} = a_{kh} - a_{ij}^{-1}(a_{ih}a_{kj} - a_{ki}a_{jh}) \quad \text{for } k \neq i,\ h \neq j$$

$$\bar{a}_{ii} = a_{ij}^{-1} a_{ii} = 0 \ .$$

To check that the described matrix transformation is in fact what occurs requires only elementary algebraic manipulations. The new matrix or tableau is called a P-pivot transform of the initial one. Note that the P-pivot transform has precisely the same solution set as the original tableau in the sense that values assigned to the variables in one tableau will hold for the same variables in the other even though positions of variables may have changed. Observe also that the P-pivot operation can be defined as above for the matrix alone, without reference to the marginal variables. For this reason we often say we P-pivot on a_{ij}, a_{ji} in the matrix A.

We define one more operation on the tableau.

Definition 4: A principal permutation or P-permutation in a square tableau is a permutation of columns occurring simultaneously with the same permutation of the corresponding rows.

It is clear that P-permutations as well as P-pivots preserve both skew-symmetry and solution sets. We use the word "principal" in each case because the term is traditionally applied to a submatrix obtained by choosing certain rows and the correspondingly-labeled columns from a given matrix. We occasionally use the term P-steps to describe a sequence of P-permutations or P-pivots or both.

The following theorem is a consequence of the definitions.

Theorem 2: Up to P-permutations, the maximum number of (skew) matrices that can be obtained by P-pivots in a skew matrix of order n is 2^{n-1}.

Proof: The net result of a finite sequence of P-pivots in the tableau corresponds precisely to choosing pairs of independent variables in the column system. Since we must pick an even number of variables, our choice of x_1 depends on whether we have picked an odd number of x's from the remaining $n - 1$ possibilities. That is, the total set of choices for n variables is completely determined by the subsets of $n - 1$ of the variables. Thus the number of choices is the number of subsets of an $(n-1)$-set, which is 2^{n-1}.

It remains to show that a particular choice of independent variables corresponds within P-permutations to precisely one coefficient matrix. For each set X of independent variables and each integer i, $1 \leqq i \leqq n$, we have a solution (X, Y) for which $x_i = 1$, $x_k = 0$ if $k \neq i$, and hence $Y = (a_{i1}, \ldots, a_{in})$. Thus up to P-permutations any independent set of variables uniquely determines the rows of the matrix. ∎

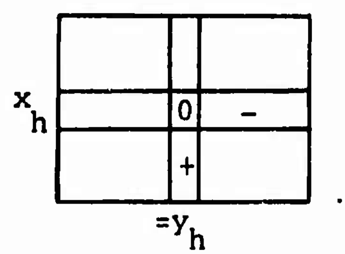
§2. The Saddlepoint Theorem

Given a skew matrix A of order n over an ordered field, we are considering the finite class of skew matrices $\bar{A}$ such that $XA = Y$ and $\bar{X}\bar{A} = \bar{Y}$ have the same solutions, where $\bar{x}_i = y_i$ and $\bar{y}_i = x_i$ for some indices i (perhaps none) and $\bar{x}_i = x_i$, $\bar{y}_i = y_i$ for the remaining indices i, $1 \leqq i \leqq n$.

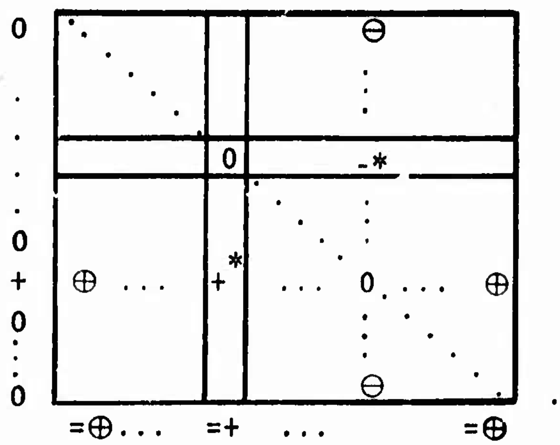
Definition 5: [7] A nontrivial solution (X, Y) of $XA = Y$ is elementary if the only other solutions having the same zero components as (X, Y) are of the form (kX, kY) for scalars $k \neq 0$.

Lemma 1: Let A be a skew matrix of order n and let h be an index, $1 \leqq h \leqq n$. There exists a nonnegative elementary solution of $XA = Y$ such that $x_h + y_h > 0$ if and only if there exists a skew matrix $\bar{A}$ having the same solutions as A and such that $\bar{a}_{hj} \geqq 0$ for all j, $1 \leqq j \leqq n$.

Proof: Given h, suppose there exists a nonnegative elementary solution of $XA = Y$ such that $x_h + y_h > 0$. Then either $x_h = 0$ and $y_h > 0$ or $x_h > 0$ and $y_h = 0$. If $a_{hj} \geqq 0$ for all j, there is nothing more to do. We therefore assume there is at least one index $j \neq h$ such that $a_{hj} < 0$, giving the schema:



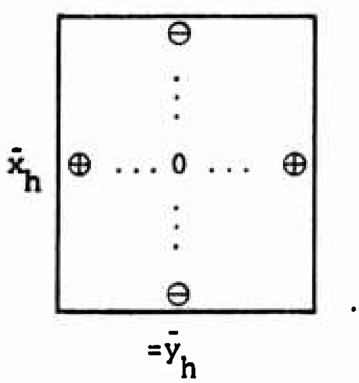
Suppose $x_h = 0$ and $y_h > 0$. Then there exists an index $t \neq h$ such that $x_t > 0$ and $a_{th} > 0$. If x_t can become arbitrarily large without decreasing any y_j, then we must have $a_{tj} \geqq 0$ for all j, $1 \leqq j \leqq n$. In that case a P-pivot on a_{th}, a_{ht} yields an $\bar{A}$ with $\bar{a}_{hj} \geqq 0$ for all j . Otherwise we can increase x_t until we obtain $y_k = 0$ for some $a_{tk} < 0$ and $k \neq h$, with $y_j \geqq 0$ for $j \neq k$. A P-pivot on a_{tk}, a_{kt} will exchange x_t with y_k and x_k with y_t. Even if $x_k = 0$ the P-pivot will increase the number of zeros in the left margin by at least one because $x_t > 0$ implies $y_t = 0$, and already $y_k = 0$. In fewer than n steps we reach the form



A P-pivot on the starred entries yields an $\bar{A}$ with $\bar{a}_{hj} \geqq 0$ for all j.

Suppose alternatively that $x_h > 0$ and $y_h = 0$. A P-pivot on some $a_{hi} < 0$ and $a_{ih} > 0$ yields an $\bar{A}$ with $\bar{x}_h = 0$ and $\bar{y}_h > 0$, which is the previous case.

To prove the converse, suppose there exists a skew matrix $\bar{A}$ having the same solutions as A and such that $\bar{a}_{hj} \geqq 0$ for all j, $1 \leqq j \leqq n$. Then we have the following tableau:



 #1311

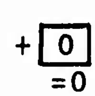
Setting $\bar{x}_i = 0$ for $i \neq h$ and $\bar{x}_h > 0$ gives a solution $(\bar{X}, \bar{Y}) \geq 0$ such that $\bar{x}_h > 0$ and $\bar{y}_h = 0$. Either $\bar{x}_h = x_h > 0$ and $\bar{y}_h = y_h = 0$ or $\bar{x}_h = y_h > 0$ and $\bar{y}_h = x_h = 0$. In either case $x_h + y_h > 0$ and $x_i + y_i \geqq 0$ for all $i \neq h$, where $XA = Y$. The nonnegative solution (X, Y) is elementary because it is determined by the entries $\bar{a}_{hj} \geqq 0$ up to scalar multiples dependent on the value of $\bar{x}_h$. ∎

We can now prove the main existence theorem.

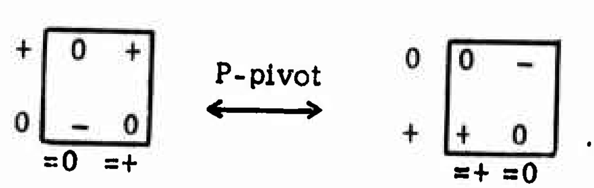
<u>Theorem 3:</u> Given a skew matrix A of order n and an index h, $1 \leqq h \leqq n$, there exists a nonnegative elementary solution of $XA = Y$ such that $x_h + y_h > 0$, and there exists a skew matrix $\bar{A}$ having the same solutions as A and such that $\bar{a}_{hj} \geqq 0$ for all j, $1 \leqq j \leqq n$.

<u>Proof:</u> We use induction on the order n of A. The theorem is clearly true for any trivial A and for $n = 1, 2$.

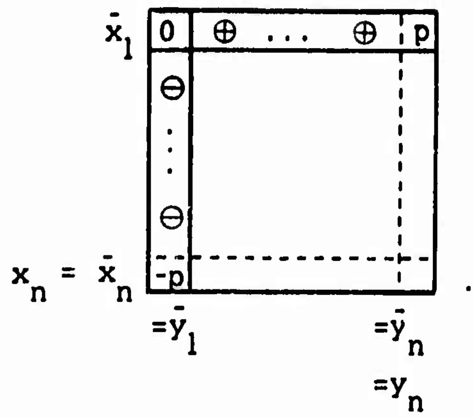
$n = 1$:



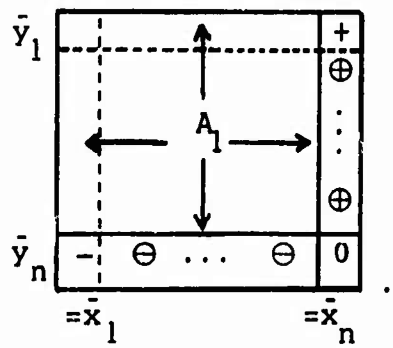
$n = 2$:



As induction hypothesis we assume the theorem is true for skew matrices of order $k < n$. Without loss of generality we choose $h = 1$. That means for a skew matrix A of order n we can reach the following form:



Either $p \geqq 0$ and the theorem is satisfied, or $p < 0$. For $p < 0$ a P-pivot on p, $-p$ yields:



By the induction hypothesis there is a nonnegative elementary solution for the skew submatrix A_1 such that $\bar{x}_1 + \bar{y}_1 > 0$. Setting $\bar{y}_n = 0$ extends the solution to one for A of the desired form. Applying Lemma 1 completes the proof.

■

Definition 6: A zero saddlepoint in a matrix is a sign configuration consisting of a nonnegative row and a nonpositive column. The entry forming their intersection is necessarily zero.

In skew matrices any nonnegative row or nonpositive column determines a zero saddlepoint with the correspondingly-labeled column or row. In view of Definition 6, we call Theorem 3 the Saddlepoint Theorem. It can also be derived as a special case of a theorem on positive semidefinite matrices in T. D. Parsons' thesis [6, pp. 84-85]; but our proof is both elementary and entirely self-contained.

Corollary 1: [Tucker Skew-Symmetric Matrix Theorem]

If A is a skew matrix, then there exists a column solution $Z = (X, Y) \geq 0$ to its tableau such that $X + Y > 0$ and $YX^T = 0$.

Proof: Given a skew matrix A of order n, Theorem 3 says that for each index i, $1 \leqq i \leqq n$, there exists a solution $Z_i = (X, Y) \geq 0$ to the matrix tableau such that either $x_i > 0$ or $y_i > 0$. Because the tableau represents linear systems, the sum of solutions of the type Z_i for a given tableau yields another solution

$$\bar{Z} = (\bar{X}, \bar{Y}) = \sum_{i=1}^{n} Z_i \geq 0$$

such that either $\bar{x}_i > 0$ or $\bar{y}_i > 0$ for each i, $1 \leqq i \leqq n$. By Theorem 1, $YX^T = 0$ for any skew tableau, and the corollary is proved. ■

The Skew-Symmetric Matrix Theorem appeared in 1956 [10]. Goldman and Tucker [3] first indicated its importance in a paper describing some of the theory of linear programming for which it is the cornerstone result.

Corollary 2: [Tucker Complementary Slackness Theorem]

Let M be an $m \times n$ matrix over an ordered field. There exist solutions to the tableau

$$\begin{array}{c|c|l} & V & \\ \hline X & M & = -U \\ \hline & =Y & \end{array}$$

such that $(X, Y) \geqq 0$, $(U^T, V^T) \geqq 0$, $X + U^T > 0$, $V^T + Y > 0$ and $XU + YV = 0$.

Proof: Given M, we form its $n \times m$ negative transpose and two trivial square matrices of orders m and n to construct the expanded tableau A:

$$\begin{array}{c|c|c|l} & x_1 \dots x_m & v_1 \dots v_n & \\ \hline \begin{matrix} x_1 \\ \vdots \\ x_m \end{matrix} & 0 & M & \begin{matrix} = -u_1 \\ \vdots \\ = -u_m \end{matrix} \\ \hline \begin{matrix} v_1 \\ \vdots \\ v_n \end{matrix} & -M^T & 0 & \begin{matrix} = -y_1 \\ \vdots \\ = -y_n \end{matrix} \\ \hline & =u_1 \dots =u_m & =y_1 \dots =y_n & \end{array} \quad \text{or} \quad \begin{array}{c|c|c|l} & X^T & V & \\ \hline X & 0 & M & = -U \\ \hline V^T & -M^T & 0 & = -Y \\ \hline & =U^T & =Y & \end{array}$$

where X is $1 \times m$, Y is $1 \times n$, V is $n \times 1$ and U is $m \times 1$. The labeling of variables in the tableau is justified by the observation that in general any column (row) solution to the tableau of M is a row (column) solution to the tableau of $-M^T$ and vice versa.

Clearly $A = -A^T$, so the expanded matrix is skew. By Corollary 1 there exists a column solution $(X, V^T, U^T, Y) \geq 0$ to A such that $(X, V^T) + (U^T, Y) > 0$ and $(U^T, Y)(X, V^T)^T = 0$. In particular, $X + U^T > 0$ and $V^T + Y > 0$; also $U^T X^T = 0$ and $YV = 0$, and the proof is complete. ∎

Corollary 1 is clearly a self-dual form of Corollary 2. In fact the results are equivalent because we can derive Corollary 1 from Corollary 2 simply by observing that in the self-dual tableau, $X = V^T$ and $Y = U^T$.

The notion of complementary slackness in dual linear systems was introduced with the Complementary Slackness Theorem by Tucker in 1956 [10] and developed further by him in 1968 [14]. The idea is that in each complementary pair of variables x_i and u_i or v_j and y_j, precisely one of them is "slack" — that is, strictly positive.

§3. Application to Matrix Games

Using the results of the previous section, we give a direct elementary proof of John von Neumann's Minimax Theorem [15, p. 153] for the special case of symmetric games.

Any real matrix $G = (g_{ij})$ determines a two-person game in which one player chooses a row i while the other chooses a column j, and the column player pays the row player the amount g_{ij}. If $g_{ij} < 0$ then the row player gives the column player the payoff $|g_{ij}|$. Such a game is called zero-sum because the relation

$$\text{row player's winnings} + \text{column player's winnings} = 0$$

holds for each play or pair of choices consisting of a row and a column. It is clear that in any matrix game the row player is trying to maximize his gains at the same time that the column player attempts to minimize his own losses.

Definition 7: A solution to an $m \times n$ matrix game G is an $(m+n+1)$-tuple $[p_1, \ldots, p_m; q_1, \ldots, q_n; v] = [P; Q^T; v]$ such that

(1) $$\sum_{i=1}^{m} p_i = \sum_{j=1}^{n} q_j = 1, \quad p_i \geqq 0 \text{ and } q_j \geqq 0 \ .$$

(2) $$\sum_{i=1}^{m} p_i g_{ij} \geqq v, \quad 1 \leqq j \leqq n \ .$$

(3) $$\sum_{j=1}^{n} g_{ij} q_j \leqq v, \quad 1 \leqq i \leqq m \ .$$

Condition (1) says simply that $P = (p_1, \ldots, p_m)$ and $Q^T = (q_1, \ldots, q_n)$ are probability vectors. Conditions (2) and (3) correspond to the following tableau:

$$\begin{array}{c|ccc|c}
 & q_1 & \cdots & q_n & \\
\hline
p_1 & & & & \leqq v \\
\vdots & & G & & \vdots \\
p_m & & & & \leqq v \\
\hline
 & \geqq v & \cdots & \geqq v &
\end{array} \ .$$

The number v, which von Neumann showed exists uniquely for any game, is called the value of the game. If $v = 0$ the game is called fair.

The vectors P and Q are the respective mixed strategies for the row and column players. If the row player chooses row i with probability p_i, he can expect to gain at least v regardless of what the column player does. Similarly the column player can expect to lose at most v if he chooses column j with probability q_j .

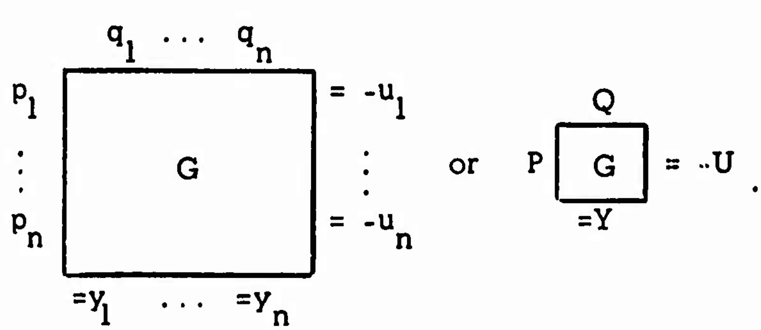
Definition 8: A matrix game is symmetric if its payoff matrix is skew.

From the definition it is clear that the roles of row and column players are interchangeable in symmetric games. For a broader definition that includes other types of matrices see Shapley [8].

Theorem 4: [von Neumann Minimax Theorem for Symmetric Games]

Every two-person zero-sum symmetric game has a solution $[P; Q^T; v]$ for which $v = 0$ and $P = Q^T$.

Proof: We form the tableau for skew G :



By Corollary 1 there exists a column solution $Z = (P, Y) \geq 0$ such that $P + Y > 0$ and $YP^T = 0$. Note that $P \geq 0$, otherwise $P = Y = 0$. Letting $s = \sum_{i=1}^{n} p_i > 0$ permits the construction of $p_i' = s^{-1} p_i$ and $y_i' = s^{-1} y_i$ for each i, $1 \leqq i \leqq n$. This normalization of P yields a probability vector $P' = (p_1', \ldots, p_n')$ such that $\sum_{i=1}^{n} p_i' g_{ij} = y_j' \geqq 0$. Since the system is self-dual, setting $Q^T = P'$ and $U^T = Y'$ gives the desired solution, where $\sum_{i=1}^{n} p_i' g_{ij} \geqq 0$ for $1 \leqq j \leqq n$ and $\sum_{j=1}^{n} g_{ij} q_j \geqq 0$ for $1 \leqq i \leqq n$. Hence $v = 0$, and the symmetric game is fair. ∎

Using Corollary 1, a skewsymmetrization technique of Tucker [13] and the following lemma, we give Tucker's proof of von Neumann's theorem for arbitrary finite matrix games.

Lemma 2: Let M be an $m \times n$ matrix in tableau form and let c, d be scalars. If there exist both a column solution $Z = (X, Y)$ such that $X \geq 0$ and $Y \geqq c$, and a row solution $W = (U^T, V^T)$ such that $V \geq 0$ and $-U \leqq d$, and if $\sum_{i=1}^{m} x_i = \sum_{j=1}^{n} v_j = 1$, then $c \leqq d$.

Proof: For any tableau

$$\begin{array}{c|c|c} & V & \\ \hline X & M & = -U \\ \hline & =Y & \end{array} \qquad \begin{cases} Z = (X, Y) \\ W = (U^T, V^T) \end{cases}$$

$ZW^T = 0$. Thus $XU + YV = 0$, which implies $YV = -XU$. Given the hypotheses of the lemma,

$$c \leqq \min_j y_j = (\min_j y_j)\sum_{j=1}^{n} v_j \leqq YV = -XU \leqq -\sum_{i=1}^{m} x_i (\min_i u_i) = -(\min_i u_i) \leqq d \ . \quad \blacksquare$$

Theorem 5: [The von Neumann Minimax Theorem]

Every two-person zero-sum matrix game has a solution and a unique value.

Proof: [13] If G is an $m \times n$ matrix game we form an $(m+n+2) \times (m+n+2)$ skew tableau A as follows:

$$\begin{array}{c|c|c|c|c|l}
 & p_1 \dots p_m & q_1 \dots q_n & r & t & \\
\hline
p_1 & & & -1 & 1 & = -y_1 \\
\vdots & 0 & G & \vdots & \vdots & \\
p_m & & & -1 & 1 & \\
\hline
q_1 & & & 1 & -1 & \\
\vdots & -G^T & 0 & \vdots & \vdots & \\
q_n & & & 1 & -1 & \\
\hline
r & 1 \ \dots \ 1 & -1 \ \dots \ -1 & 0 & 0 & \\
\hline
t & -1 \ \dots \ -1 & 1 \ \dots \ 1 & 0 & 0 & = -y_{m+n+2} \\
\hline
 & =y_1 \ \dots & \dots & \dots & =y_{m+n+2} &
\end{array}$$

where $P = (p_1, \dots, p_m)$

$Q^T = (q_1, \dots, q_n)$

$Y = (y_1, \dots, y_{m+n+2})$.

By Corollary 1 there exists a solution $(P, Q^T, r, t, Y) \geq 0$ to the self-dual system of A such that $(P, Q^T, r, t) + Y > 0$ and $Y(P, Q^T, r, t)^T = 0$. The last two column equations of the tableau yield:

$$0 \leqq y_{m+n+2} = \sum_{i=1}^{m} p_i - \sum_{j=1}^{n} q_j = -y_{m+n+1} \leqq 0 .$$

Hence $y_{m+n+1} = y_{m+n+2} = 0$, $r > 0$, $t > 0$ and $\sum_{i=1}^{m} p_i = \sum_{j=1}^{n} q_j$. Then either $P = 0$ and $Q^T = 0$ or $P \geq 0$ and $Q^T \geq 0$.

If $P = Q^T = 0$, then $y_j > 0$ for $1 \leqq j \leqq m+n$. But $y_1 = r - t = -y_{m+1}$, which would be impossible. Hence $P \geq 0$ and $Q^T \geq 0$. Let $s = \sum_{i=1}^{m} p_i = \sum_{j=1}^{n} q_j > 0$ and $v = s^{-1}(r-t)$. Then $[s^{-1}P; s^{-1}Q^T; v]$ is a solution to G; for $\sum_{i=1}^{m} s^{-1}p_i = \sum_{j=1}^{n} s^{-1} q_j = 1$ and $s^{-1}PG \geqq s^{-1}(r-t) \geqq s^{-1}Q^TG^T = Gs^{-1}Q$.

To show uniqueness of v, assume $[P_1 : Q_1^T; v_1]$ and $[P_2; Q_2^T: v_2]$ are distinct solutions of the game. Then $v_1 \leqq P_1G$ for one column solution of the tableau and $GQ_2 \leqq v_2$ for one row solution, thus satisfying Lemma 2. Hence $v_1 \leqq v_2$. Similarly $v_2 \leqq v_1$, and therefore $v_1 = v_2$. ■

A survey of alternative methods of proving the Minimax Theorem appears in the article by Kuhn and Tucker on von Neumann's work in game theory [5]. A paper by Gale, Kuhn and Tucker [2] contains further details on other methods of skewsymmetrization that lead to algebraic proofs. For an efficient algorithm that uses linear programming to solve matrix games, see Tucker [11].

Adopting the terminology of Gale [1] we can make fuller use of the complementary slackness principle in solutions of arbitrary $m \times n$ matrix games. Each choice of a single column or row constitutes a pure strategy for the appropriate player. The probability vectors in a solution of the game constitute <u>optimal mixed strategies</u> in that they specify the frequency with which a particular pure strategy should be chosen in order for the player to minimize his expected losses or maximize his expected gains. Gale's idea is to partition the probability vectors for a given solution of the game into their positive and zero components and to call the corresponding pure strategies <u>essential</u> and <u>superfluous</u>.

Thus if we have a solution $[P; Q^T; v]$ to the $m \times n$ game G we can obtain the following tableau by permuting rows and columns:

$$
\begin{array}{r|c|c|l}
 & q_1 \; \dots \; q_h & q_{h+1} \; \dots \; q_n & \\
 & + \; \dots \; + & 0 \; \dots \; 0 & \\
\hline
p_1 = + & & & = -v \\
\vdots & G_1 & & \vdots \\
p_k = + & (k \times h) & & = -v \\
\hline
p_{k+1} = 0 & & & < -v \\
\vdots & & & \vdots \\
p_m = 0 & & & < -v \\
\hline
 & =v \; \dots \; =v & >v \; \dots \; >v &
\end{array}
$$

The rows $k+1, \dots, m$ are "superfluous" strategies in the sense that if the row player chooses one of them while the column player sticks to his "essential" choices $1, \dots, h$ then in time the payoff to the row player will be less than the value v which he could have obtained had he stuck to his own "essential" choices. An analogous statement holds for the "superfluous" strategies $h + 1, \ldots, n$ of the column player. Of course, the set of superfluous strategies for either player may be empty.

In a symmetric game we can always find a symmetric solution — that is, one where $p_i = q_i$ for $1 \leqq i \leqq n$. For such solutions the permutation to display the partition of mixed strategies will be a P-permutation and will yield a principal square submatrix G_1 of G corresponding to the essential strategies.

REFERENCES

1. D. Gale, "Solutions of finite two-person games," Doctoral Dissertation, Department of Mathematics, Princeton University, 1949.

2. D. Gale, H. W. Kuhn and A. W. Tucker, "On symmetric games," in Contributions to the Theory of Games, Vol. I, Annals of Mathematics Studies 24, edited by H. W. Kuhn and A. W. Tucker (Princeton: Princeton University Press) 1950.

3. A. J. Goldman and A. W. Tucker, "Theory of linear programming," in [4], pp. 53-97.

4. H. W. Kuhn and A. W. Tucker (Editors), Linear Inequalities and Related Systems, Annals of Mathematics Studies 38 (Princeton: Princeton University Press) 1956.

5. ________________, "John von Neumann's work in the theory of games and mathematical economics," Bull. Amer. Math. Soc. 64 (1958), 100-122.

6. T. D. Parsons, "A combinatorial approach to convex quadratic programming," Doctoral Dissertation, Department of Mathematics, Princeton University, 1966.

7. R. T. Rockafellar, "The elementary vectors of a subspace of $\mathbb{R}^N$," in Combinatorial Math. and Its Applications, edited by R. C. Bose and T. A. Dowling (Chapel Hill: U. of N. Carolina Press) 1969, pp. 104-127.

8. L. S. Shapley, "Symmetric Games," RAND Memo RM-2476 (Santa Monica, California: RAND Corporation) June 9, 1960.

9. Marjorie L. Stein, "A pivot theory of skew matrices with applications," Doctoral Dissertation, Department of Mathematics, Princeton University, 1972.

10. A. W. Tucker, "Dual systems of homogeneous linear relations," in [4], pp. 3-18.

11. ________________, "Solving a matrix game by linear programming," IBM Journal 4 1960, 507-517.

12. ________________, "Combinatorial theory underlying linear programs," in Recent Advances in Mathematical Programming, edited by R. L. Graves and P. Wolfe (New York: McGraw Hill) 1963, pp. 1-16.

13. ________________, Pivotal Algebra, Lecture notes by T. D. Parsons, Department of Mathematics, Princeton University, 1965.

14. ________________, "Complementary slackness in dual linear subspaces," in Mathematics of the Decision Sciences Part I, Vol. 11 of Lectures in Applied Mathematics, edited by G. B. Dantzig and A. F. Veinott, Jr. (Providence: AMS) 1968, pp. 137-143.

15. John von Neumann and Oskar Morgenstern, Theory of Games and Economic Behavior, 3rd ed. (Princeton: Princeton University Press) 1953.