11. Canonical Forms

Given a matrix A, the effect of a sequence of row-operations on A is to produce UA where U is invertible. Under this "row-equivalence" operation the best that can be achieved is the reduced row-echelon form for A. If column operations are also allowed, the result is UAV where both U and V are invertible, and the best outcome under this "equivalence" operation is called the Smith canonical form of A (Theorem 2.5.3). There are other kinds of operations on a matrix and, in many cases, there is a "canonical" best possible result.

If A is square, the most important operation of this sort is arguably "similarity" wherein A is carried to $U^{-1}AU$ where U is invertible. In this case we say that matrices A and B are *similar*, and write $A \sim B$, when $B = U^{-1}AU$ for some invertible matrix U. Under similarity the canonical matrices, called *Jordan canonical matrices*, are block triangular with upper triangular "Jordan" blocks on the main diagonal. In this short chapter we are going to define these Jordan blocks and prove that every matrix is similar to a Jordan canonical matrix.

Here is the key to the method. Let $T : V \to V$ be an operator on an n-dimensional vector space V, and suppose that we can find an ordered basis B of B so that the matrix $M_B(T)$ is as simple as possible. Then, if B_0 is *any* ordered basis of V, the matrices $M_B(T)$ and $M_{B_0}(T)$ are similar; that is,

$$M_B(T) = P^{-1}M_{B_0}(T)P \quad \text{for some invertible matrix } P$$

Moreover, $P = P_{B_0 \leftarrow B}$ is easily computed from the bases B and D (Theorem 9.2.3). This, combined with the invariant subspaces and direct sums studied in Section 9.3, enables us to calculate the Jordan canonical form of any square matrix A. Along the way we derive an explicit construction of an invertible matrix P such that $P^{-1}AP$ is block triangular.

This technique is important in many ways. For example, if we want to diagonalize an $n \times n$ matrix A, let $T_A : \mathbb{R}^n \to \mathbb{R}^n$ be the operator given by $T_A(\mathbf{x}) = A\mathbf{x}$ or all $\mathbf{x}$ in $\mathbb{R}^n$, and look for a basis B of $\mathbb{R}^n$ such that $M_B(T_A)$ is diagonal. If $B_0 = E$ is the standard basis of $\mathbb{R}^n$, then $M_E(T_A) = A$, so

$$P^{-1}AP = P^{-1}M_E(T_A)P = M_B(T_A)$$

and we have diagonalized A. Thus the "algebraic" problem of finding an invertible matrix P such that $P^{-1}AP$ is diagonal is converted into the "geometric" problem of finding a basis B such that $M_B(T_A)$ is diagonal. This change of perspective is one of the most important techniques in linear algebra.

11.1 Block Triangular Form

We have shown (Theorem 8.2.5) that any $n \times n$ matrix A with every eigenvalue real is orthogonally similar to an upper triangular matrix U. The following theorem shows that U can be chosen in a special way.

Theorem 11.1.1: Block Triangulation Theorem

Let A be an $n \times n$ matrix with every eigenvalue real and let

$$c_A(x) = (x-\lambda_1)^{m_1}(x-\lambda_2)^{m_2}\cdots(x-\lambda_k)^{m_k}$$

where $\lambda_1, \lambda_2, \dots, \lambda_k$ are the distinct eigenvalues of A. Then an invertible matrix P exists such that

$$P^{-1}AP = \left[\begin{array}{ccccc} U_1 & 0 & 0 & \cdots & 0 \\ 0 & U_2 & 0 & \cdots & 0 \\ 0 & 0 & U_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & U_k \end{array}\right]$$

where, for each i, U_i is an $m_i \times m_i$ upper triangular matrix with every entry on the main diagonal equal to λ_i.

The proof is given at the end of this section. For now, we focus on a method for *finding* the matrix P. The key concept is as follows.

Definition 11.1 Generalized Eigenspaces

If A is as in Theorem 11.1.1, the ***generalized eigenspace*** *$G_{\lambda_i}(A)$ is defined by*

$$G_{\lambda_i}(A) = \text{null}\,[(\lambda_i I - A)^{m_i}]$$

where m_i is the multiplicity of λ_i.

Observe that the eigenspace $E_{\lambda_i}(A) = \text{null}\,(\lambda_i I - A)$ is a subspace of $G_{\lambda_i}(A)$. We need three technical results.

Lemma 11.1.1

Using the notation of Theorem 11.1.1, we have $\dim\,[G_{\lambda_i}(A)] = m_i$.

Proof. Write $A_i = (\lambda_i I - A)^{m_i}$ for convenience and let P be as in Theorem 11.1.1. The spaces $G_{\lambda_i}(A) = \text{null}\,(A_i)$ and $\text{null}\,(P^{-1}A_iP)$ are isomorphic via $\mathbf{x} \leftrightarrow P^{-1}\mathbf{x}$, so we show $\dim\,[\text{null}\,(P^{-1}A_iP)] = m_i$. Now $P^{-1}A_iP = (\lambda_i I - P^{-1}AP)^{m_i}$. If we use the block form in Theorem 11.1.1, this becomes

$$\begin{aligned} P^{-1}A_iP &= \left[\begin{array}{cccc} \lambda_i I - U_1 & 0 & \cdots & 0 \\ 0 & \lambda_i I - U_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_i I - U_k \end{array}\right]^{m_i} \\ &= \left[\begin{array}{cccc} (\lambda_i I - U_1)^{m_i} & 0 & \cdots & 0 \\ 0 & (\lambda_i I - U_2)^{m_i} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & (\lambda_i I - U_k)^{m_i} \end{array}\right] \end{aligned}$$

The matrix $(\lambda_i I - U_j)^{m_i}$ is invertible if $j \neq i$ and zero if $j = i$ (because then U_i is an $m_i \times m_i$ upper triangular matrix with each entry on the main diagonal equal to λ_i). It follows that $m_i = \dim [\text{null}(P^{-1}A_i P)]$, as required. □

Lemma 11.1.2

If P is as in Theorem 11.1.1, denote the columns of P as follows:

$$\mathbf{p}_{11}, \mathbf{p}_{12}, \dots, \mathbf{p}_{1m_1}; \quad \mathbf{p}_{21}, \mathbf{p}_{22}, \dots, \mathbf{p}_{2m_2}; \quad \dots; \quad \mathbf{p}_{k1}, \mathbf{p}_{k2}, \dots, \mathbf{p}_{km_k}$$

Then $\{\mathbf{p}_{i1}, \mathbf{p}_{i2}, \dots, \mathbf{p}_{im_i}\}$ is a basis of $G_{\lambda_i}(A)$.

Proof. It suffices by Lemma 11.1.1 to show that each $\mathbf{p}_{ij}$ is *in* $G_{\lambda_i}(A)$. Write the matrix in Theorem 11.1.1 as $P^{-1}AP = \text{diag}(U_1, U_2, \dots, U_k)$. Then

$$AP = P \, \text{diag}(U_1, U_2, \dots, U_k)$$

Comparing columns gives, successively:

$$\begin{aligned}
A\mathbf{p}_{11} &= \lambda_1 \mathbf{p}_{11}, & \text{so } (\lambda_1 I - A)\mathbf{p}_{11} &= \mathbf{0} \\
A\mathbf{p}_{12} &= u\mathbf{p}_{11} + \lambda_1 \mathbf{p}_{12}, & \text{so } (\lambda_1 I - A)^2\mathbf{p}_{12} &= \mathbf{0} \\
A\mathbf{p}_{13} &= w\mathbf{p}_{11} + v\mathbf{p}_{12} + \lambda_1 \mathbf{p}_{13} & \text{so } (\lambda_1 I - A)^3\mathbf{p}_{13} &= \mathbf{0} \\
&\vdots & &\vdots
\end{aligned}$$

where u, v, w are in $\mathbb{R}$. In general, $(\lambda_1 I - A)^j \mathbf{p}_{1j} = \mathbf{0}$ for $j = 1, 2, \dots, m_1$, so $\mathbf{p}_{1j}$ is in $G_{\lambda_i}(A)$. Similarly, $\mathbf{p}_{ij}$ is in $G_{\lambda_i}(A)$ for each i and j. □

Lemma 11.1.3

If B_i is any basis of $G_{\lambda_i}(A)$, then $B = B_1 \cup B_2 \cup \cdots \cup B_k$ is a basis of $\mathbb{R}^n$.

Proof. It suffices by Lemma 11.1.1 to show that B is independent. If a linear combination from B vanishes, let $\mathbf{x}_i$ be the sum of the terms from B_i. Then $\mathbf{x}_1 + \cdots + \mathbf{x}_k = \mathbf{0}$. But $\mathbf{x}_i = \sum_j r_{ij}\mathbf{p}_{ij}$ by Lemma 11.1.2, so $\sum_{i,j} r_{ij}\mathbf{p}_{ij} = \mathbf{0}$. Hence each $\mathbf{x}_i = \mathbf{0}$, so each coefficient in $\mathbf{x}_i$ is zero. □

Lemma 11.1.2 suggests an algorithm for finding the matrix P in Theorem 11.1.1. Observe that there is an ascending chain of subspaces leading from $E_{\lambda_i}(A)$ to $G_{\lambda_i}(A)$:

$$E_{\lambda_i}(A) = \text{null}[(\lambda_i I - A)] \subseteq \text{null}[(\lambda_i I - A)^2] \subseteq \cdots \subseteq \text{null}[(\lambda_i I - A)^{m_i}] = G_{\lambda_i}(A)$$

We construct a basis for $G_{\lambda_i}(A)$ by climbing up this chain.

Triangulation Algorithm

Suppose A has characteristic polynomial

$$c_A(x) = (x-\lambda_1)^{m_1}(x-\lambda_2)^{m_2}\cdots(x-\lambda_k)^{m_k}$$

1. *Choose a basis of* $\text{null}\,[(\lambda_1 I - A)]$; *enlarge it by adding vectors (possibly none) to a basis of* $\text{null}\,[(\lambda_1 I - A)^2]$; *enlarge that to a basis of* $\text{null}\,[(\lambda_1 I - A)^3]$, *and so on. Continue to obtain an ordered basis* $\{\mathbf{p}_{11}, \mathbf{p}_{12}, \dots, \mathbf{p}_{1m_1}\}$ *of* $G_{\lambda_1}(A)$.
2. *As in* (1) *choose a basis* $\{\mathbf{p}_{i1}, \mathbf{p}_{i2}, \dots, \mathbf{p}_{im_i}\}$ *of* $G_{\lambda_i}(A)$ *for each* i.
3. *Let* $P = \left[\ \mathbf{p}_{11}\mathbf{p}_{12}\cdots\mathbf{p}_{1m_1};\ \ \mathbf{p}_{21}\mathbf{p}_{22}\cdots\mathbf{p}_{2m_2};\ \ \cdots;\ \ \mathbf{p}_{k1}\mathbf{p}_{k2}\cdots\mathbf{p}_{km_k}\ \right]$ *be the matrix with these basis vectors (in order) as columns.*

Then $P^{-1}AP = \text{diag}\,(U_1, U_2, \dots, U_k)$ *as in Theorem 11.1.1.*

Proof. Lemma 11.1.3 guarantees that $B = \{\mathbf{p}_{11}, \dots, \mathbf{p}_{km_1}\}$ is a basis of $\mathbb{R}^n$, and Theorem 9.2.4 shows that $P^{-1}AP = M_B(T_A)$. Now $G_{\lambda_i}(A)$ is T_A-invariant for each i because

$$(\lambda_i I - A)^{m_i}\mathbf{x} = \mathbf{0} \quad \text{implies} \quad (\lambda_i I - A)^{m_i}(A\mathbf{x}) = A(\lambda_i I - A)^{m_i}\mathbf{x} = \mathbf{0}$$

By Theorem 9.3.7 (and induction), we have

$$P^{-1}AP = M_B(T_A) = \text{diag}\,(U_1, U_2, \dots, U_k)$$

where U_i is the matrix of the restriction of T_A to $G_{\lambda_i}(A)$, and it remains to show that U_i has the desired upper triangular form. Given s, let $\mathbf{p}_{ij}$ be a basis vector in $\text{null}\,[(\lambda_i I - A)^{s+1}]$. Then $(\lambda_i I - A)\mathbf{p}_{ij}$ is in $\text{null}\,[(\lambda_i I - A)^s]$, and therefore is a linear combination of the basis vectors $\mathbf{p}_{it}$ coming *before* $\mathbf{p}_{ij}$. Hence

$$T_A(\mathbf{p}_{ij}) = A\mathbf{p}_{ij} = \lambda_i\mathbf{p}_{ij} - (\lambda_i I - A)\mathbf{p}_{ij}$$

shows that the column of U_i corresponding to $\mathbf{p}_{ij}$ has λ_i on the main diagonal and zeros below the main diagonal. This is what we wanted. □

Example 11.1.1

If $A = \left[\begin{array}{rrrr} 2 & 0 & 0 & 1 \\ 0 & 2 & 0 & -1 \\ -1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{array}\right]$, find P such that $P^{-1}AP$ is block triangular.

Solution. $c_A(x) = \det\,[xI - A] = (x-2)^4$, so $\lambda_1 = 2$ is the only eigenvalue and we are in the case $k = 1$ of Theorem 11.1.1. Compute:

$$(2I - A) = \left[\begin{array}{rrrr} 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \qquad (2I - A)^2 = \left[\begin{array}{rrrr} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2 \\ 0 & 0 & 0 & 0 \end{array}\right] \qquad (2I - A)^3 = 0$$

By gaussian elimination find a basis $\{\mathbf{p}_{11}, \mathbf{p}_{12}\}$ of $\text{null}\,(2I - A)$; then extend in any way to a basis $\{\mathbf{p}_{11}, \mathbf{p}_{12}, \mathbf{p}_{13}\}$ of $\text{null}\,[(2I - A)^2]$; and finally get a basis $\{\mathbf{p}_{11}, \mathbf{p}_{12}, \mathbf{p}_{13}, \mathbf{p}_{14}\}$ of $\text{null}\,[(2I - A)^3] = \mathbb{R}^4$. One choice is

$$\mathbf{p}_{11} = \left[\begin{array}{c} 1 \\ 1 \\ 0 \\ 0 \end{array}\right] \quad \mathbf{p}_{12} = \left[\begin{array}{c} 0 \\ 0 \\ 1 \\ 0 \end{array}\right] \quad \mathbf{p}_{13} = \left[\begin{array}{c} 0 \\ 1 \\ 0 \\ 0 \end{array}\right] \quad \mathbf{p}_{14} = \left[\begin{array}{c} 0 \\ 0 \\ 0 \\ 1 \end{array}\right]$$

Hence $P = \left[\begin{array}{cccc} \mathbf{p}_{11} & \mathbf{p}_{12} & \mathbf{p}_{13} & \mathbf{p}_{14} \end{array}\right] = \left[\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right]$ gives $P^{-1}AP = \left[\begin{array}{rrrr} 2 & 0 & 0 & 1 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & -2 \\ 0 & 0 & 0 & 2 \end{array}\right]$

Example 11.1.2

If $A = \left[\begin{array}{rrrr} 2 & 0 & 1 & 1 \\ 3 & 5 & 4 & 1 \\ -4 & -3 & -3 & -1 \\ 1 & 0 & 1 & 2 \end{array}\right]$, find P such that $P^{-1}AP$ is block triangular.

Solution. The eigenvalues are $\lambda_1 = 1$ and $\lambda_2 = 2$ because

$$\begin{aligned}
c_A(x) &= \left|\begin{array}{cccc} x-2 & 0 & -1 & -1 \\ -3 & x-5 & -4 & -1 \\ 4 & 3 & x+3 & 1 \\ -1 & 0 & -1 & x-2 \end{array}\right| = \left|\begin{array}{cccc} x-1 & 0 & 0 & -x+1 \\ -3 & x-5 & -4 & -1 \\ 4 & 3 & x+3 & 1 \\ -1 & 0 & -1 & x-2 \end{array}\right| \\
&= \left|\begin{array}{cccc} x-1 & 0 & 0 & 0 \\ -3 & x-5 & -4 & -4 \\ 4 & 3 & x+3 & 5 \\ -1 & 0 & -1 & x-3 \end{array}\right| = (x-1)\left|\begin{array}{ccc} x-5 & -4 & -4 \\ 3 & x+3 & 5 \\ 0 & -1 & x-3 \end{array}\right| \\
&= (x-1)\left|\begin{array}{ccc} x-5 & -4 & 0 \\ 3 & x+3 & -x+2 \\ 0 & -1 & x-2 \end{array}\right| = (x-1)\left|\begin{array}{ccc} x-5 & -4 & 0 \\ 3 & x+2 & 0 \\ 0 & -1 & x-2 \end{array}\right| \\
&= (x-1)(x-2)\left|\begin{array}{cc} x-5 & -4 \\ 3 & x+2 \end{array}\right| = (x-1)^2(x-2)^2
\end{aligned}$$

By solving equations, we find $\text{null}\,(I - A) = \text{span}\,\{\mathbf{p}_{11}\}$ and $\text{null}\,(I - A)^2 = \text{span}\,\{\mathbf{p}_{11}, \mathbf{p}_{12}\}$ where

$$\mathbf{p}_{11} = \left[\begin{array}{r} 1 \\ 1 \\ -2 \\ 1 \end{array}\right] \quad \mathbf{p}_{12} = \left[\begin{array}{r} 0 \\ 3 \\ -4 \\ 1 \end{array}\right]$$

Since $\lambda_1 = 1$ has multiplicity 2 as a root of $c_A(x)$, $\dim G_{\lambda_1}(A) = 2$ by Lemma 11.1.1. Since $\mathbf{p}_{11}$ and $\mathbf{p}_{12}$ both lie in $G_{\lambda_1}(A)$, we have $G_{\lambda_1}(A) = \text{span}\,\{\mathbf{p}_{11}, \mathbf{p}_{12}\}$. Turning to $\lambda_2 = 2$, we find that null $(2I-A)$ = span $\{\mathbf{p}_{21}\}$ and null $[(2I-A)^2]$ = span $\{\mathbf{p}_{21},\ \mathbf{p}_{22}\}$ where

$$\mathbf{p}_{21} = \left[\begin{array}{r} 1 \\ 0 \\ -1 \\ 1 \end{array}\right] \quad \text{and} \quad \mathbf{p}_{22} = \left[\begin{array}{r} 0 \\ -4 \\ 3 \\ 0 \end{array}\right]$$

Again, dim $G_{\lambda_2}(A) = 2$ as λ_2 has multiplicity 2, so $G_{\lambda_2}(A)$ = span $\{\mathbf{p}_{21},\ \mathbf{p}_{22}\}$. Hence

$$P = \left[\begin{array}{rrrr} 1 & 0 & 1 & 0 \\ 1 & 3 & 0 & -4 \\ -2 & -4 & -1 & 3 \\ 1 & 1 & 1 & 0 \end{array}\right] \text{ gives } P^{-1}AP = \left[\begin{array}{rrrr} 1 & -3 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 3 \\ 0 & 0 & 0 & 2 \end{array}\right].$$

If $p(x)$ is a polynomial and A is an $n\times n$ matrix, then $p(A)$ is also an $n\times n$ matrix if we interpret $A^0 = I_n$. For example, if $p(x) = x^2-2x+3$, then $p(A) = A^2-2A+3I$. Theorem 11.1.1 provides another proof of the Cayley-Hamilton theorem (see also Theorem 8.7.10). As before, let $c_A(x)$ denote the characteristic polynomial of A.

Theorem 11.1.2: Cayley-Hamilton Theorem

If A is a square matrix with every eigenvalue real, then $c_A(A) = 0$.

Proof. As in Theorem 11.1.1, write $c_A(x) = (x-\lambda_1)^{m_1}\cdots(x-\lambda_k)^{m_k} = \Pi_{i=1}^{k}(x-\lambda_i)^{m_i}$, and write

$$P^{-1}AP = D = \text{diag}\,(U_1,\ \dots,\ U_k)$$

Hence

$$c_A(U_i) = \Pi_{i=1}^{k}(U_i-\lambda_i I_{m_i})^{m_i} = 0 \text{ for each } i$$

because the factor $(U_i-\lambda_i I_{m_i})^{m_i} = 0$. In fact $U_i-\lambda_i I_{m_i}$ is $m_i\times m_i$ and has zeros on the main diagonal. But then

$$\begin{aligned} P^{-1}c_A(A)P = c_A(D) &= c_A[\text{diag}\,(U_1,\ \dots,\ U_k)] \\ &= \text{diag}\,[c_A(U_1),\ \dots,\ c_A(U_k)] \\ &= 0 \end{aligned}$$

It follows that $c_A(A) = 0$. □

Example 11.1.3

If $A = \left[\begin{array}{rr} 1 & 3 \\ -1 & 2 \end{array}\right]$, then $c_A(x) = \det \left[\begin{array}{cc} x-1 & -3 \\ 1 & x-2 \end{array}\right] = x^2-3x+5$. Then
$c_A(A) = A^2-3A+5I_2 = \left[\begin{array}{rr} -2 & 9 \\ -3 & 1 \end{array}\right] - \left[\begin{array}{rr} 3 & 9 \\ -3 & 6 \end{array}\right] + \left[\begin{array}{rr} 5 & 0 \\ 0 & 5 \end{array}\right] = \left[\begin{array}{rr} 0 & 0 \\ 0 & 0 \end{array}\right]$.

Theorem 11.1.1 will be refined even further in the next section.

Proof of Theorem 11.1.1

The proof of Theorem 11.1.1 requires the following simple fact about bases, the proof of which we leave to the reader.

Lemma 11.1.4

If $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a basis of a vector space V, so also is $\{\mathbf{v}_1 + s\mathbf{v}_2, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ for any scalar s.

Proof of Theorem 11.1.1. Let A be as in Theorem 11.1.1, and let $T = T_A : \mathbb{R}^n \to \mathbb{R}^n$ be the matrix transformation induced by A. For convenience, call a matrix a λ-m-ut matrix if it is an $m \times m$ upper triangular matrix and every diagonal entry equals λ. Then we must find a basis B of $\mathbb{R}^n$ such that $M_B(T) = \text{diag}(U_1, U_2, \ldots, U_k)$ where U_i is a λ_i-m_i-ut matrix for each i. We proceed by induction on n. If $n = 1$, take $B = \{\mathbf{v}\}$ where $\mathbf{v}$ is any eigenvector of T.

If $n > 1$, let $\mathbf{v}_1$ be a λ_1-eigenvector of T, and let $B_0 = \{\mathbf{v}_1, \mathbf{w}_1, \ldots, \mathbf{w}_{n-1}\}$ be any basis of $\mathbb{R}^n$ containing $\mathbf{v}_1$. Then (see Lemma 5.5.2)

$$M_{B_0}(T) = \left[\begin{array}{cc} \lambda_1 & X \\ 0 & A_1 \end{array} \right]$$

in block form where A_1 is $(n-1) \times (n-1)$. Moreover, A and $M_{B0}(T)$ are similar, so

$$c_A(x) = c_{M_{B_0}(T)}(x) = (x - \lambda_1)c_{A_1}(x)$$

Hence $c_{A_1}(x) = (x - \lambda_1)^{m_1 - 1}(x - \lambda_2)^{m_2} \cdots (x - \lambda_k)^{m_k}$ so (by induction) let

$$Q^{-1}A_1Q = \text{diag}(Z_1, U_2, \ldots, U_k)$$

where Z_1 is a λ_1-$(m_1 - 1)$-ut matrix and U_i is a λ_i-m_i-ut matrix for each $i > 1$.

If $P = \left[\begin{array}{cc} 1 & 0 \\ 0 & Q \end{array} \right]$, then $P^{-1}MB_0(T) = \left[\begin{array}{cc} \lambda_1 & XQ \\ 0 & Q^{-1}A_1Q \end{array} \right] = A'$, say. Hence $A' \sim M_{B_0}(T) \sim A$ so by Theorem 9.2.4(2) there is a basis B of $\mathbb{R}^n$ such that $M_{B_1}(T_A) = A'$, that is $M_{B_1}(T) = A'$. Hence $M_{B_1}(T)$ takes the block form

$$M_{B_1}(T) = \left[\begin{array}{cc} \lambda_1 & XQ \\ 0 & \text{diag}(Z_1, U_2, \ldots, U_k) \end{array} \right] = \left[\begin{array}{cc|ccc} \lambda_1 & X_1 & & Y & \\ 0 & Z_1 & 0 & 0 & 0 \\ \hline & & U_2 & \cdots & 0 \\ & 0 & \vdots & & \vdots \\ & & 0 & \cdots & U_k \end{array} \right] \tag{11.1}$$

If we write $U_1 = \left[\begin{array}{cc} \lambda_1 & X_1 \\ 0 & Z_1 \end{array} \right]$, the basis B_1 fulfills our needs except that the row matrix Y may not be zero.

We remedy this defect as follows. Observe that the first vector in the basis B_1 is a λ_1 eigenvector of T, which we continue to denote as $\mathbf{v}_1$. The idea is to add suitable scalar multiples of $\mathbf{v}_1$ to the other vectors in B_1. This results in a new basis by Lemma 11.1.4, and the multiples can be chosen so that the new matrix of T is the same as (11.1) except that $Y = 0$. Let $\{\mathbf{w}_1, \ldots, \mathbf{w}_{m_2}\}$ be the vectors in B_1 corresponding to λ_2

(giving rise to U_2 in (11.1)). Write

$$U_2 = \begin{bmatrix} \lambda_2 & u_{12} & u_{13} & \cdots & u_{1m_2} \\ 0 & \lambda_2 & u_{23} & \cdots & u_{2m_2} \\ 0 & 0 & \lambda_2 & \cdots & u_{3m_2} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_2 \end{bmatrix} \quad \text{and} \quad Y = \begin{bmatrix} y_1 & y_2 & \cdots & y_{m_2} \end{bmatrix}$$

We first replace $\mathbf{w}_1$ by $\mathbf{w}_1' = \mathbf{w}_1 + s\mathbf{v}_1$ where s is to be determined. Then (11.1) gives

$$\begin{aligned} T(\mathbf{w}_1') &= T(\mathbf{w}_1) + sT(\mathbf{v}_1) \\ &= (y_1\mathbf{v}_1 + \lambda_2\mathbf{w}_1) + s\lambda_1\mathbf{v}_1 \\ &= y_1\mathbf{v}_1 + \lambda_2(\mathbf{w}_1' - s\mathbf{v}_1) + s\lambda_1\mathbf{v}_1 \\ &= \lambda_2\mathbf{w}_1' + [(y_1 - s(\lambda_2 - \lambda_1)]\mathbf{v}_1 \end{aligned}$$

Because $\lambda_2 \neq \lambda_1$ we can choose s such that $T(\mathbf{w}_1') = \lambda_2\mathbf{w}_1'$. Similarly, let $\mathbf{w}_2' = \mathbf{w}_2 + t\mathbf{v}_1$ where t is to be chosen. Then, as before,

$$\begin{aligned} T(\mathbf{w}_2') &= T(\mathbf{w}_2) + tT(\mathbf{v}_1) \\ &= (y_2\mathbf{v}_1 + u_{12}\mathbf{w}_1 + \lambda_2\mathbf{w}_2) + t\lambda_1\mathbf{v}_1 \\ &= u_{12}\mathbf{w}_1' + \lambda_2\mathbf{w}_2' + [(y_2 - u_{12}s) - t(\lambda_2 - \lambda_1)]\mathbf{v}_1 \end{aligned}$$

Again, t can be chosen so that $T(\mathbf{w}_2') = u_{12}\mathbf{w}_1' + \lambda_2\mathbf{w}_2'$. Continue in this way to eliminate $y_1, \dots, y_{m_2}$. This procedure also works for $\lambda_3, \lambda_4, \dots$ and so produces a new basis B such that $M_B(T)$ is as in (11.1) but with $Y = 0$. □

Exercises for 11.1

Exercise 11.1.1 In each case, find a matrix P such that $P^{-1}AP$ is in block triangular form as in Theorem 11.1.1.

a. $A = \begin{bmatrix} 2 & 3 & 2 \\ -1 & -1 & -1 \\ 1 & 2 & 2 \end{bmatrix}$ b. $A = \begin{bmatrix} -5 & 3 & 1 \\ -4 & 2 & 1 \\ -4 & 3 & 0 \end{bmatrix}$

c. $A = \begin{bmatrix} 0 & 1 & 1 \\ 2 & 3 & 6 \\ -1 & -1 & -2 \end{bmatrix}$ d. $A = \begin{bmatrix} -3 & -1 & 0 \\ 4 & -1 & 3 \\ 4 & -2 & 4 \end{bmatrix}$

e. $A = \begin{bmatrix} -1 & -1 & -1 & 0 \\ 3 & 2 & 3 & -1 \\ 2 & 1 & 3 & -1 \\ 2 & 1 & 4 & -2 \end{bmatrix}$

f. $A = \begin{bmatrix} -3 & 6 & 3 & 2 \\ -2 & 3 & 2 & 2 \\ -1 & 3 & 0 & 1 \\ -1 & 1 & 2 & 0 \end{bmatrix}$

Exercise 11.1.2 Show that the following conditions are equivalent for a linear operator T on a finite dimensional space V.

1. $M_B(T)$ is upper triangular for some ordered basis B of E.

2. A basis $\{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ of V exists such that, for each i, $T(\mathbf{b}_i)$ is a linear combination of $\mathbf{b}_1, \dots, \mathbf{b}_i$.

3. There exist T-invariant subspaces
$$V_1 \subseteq V_2 \subseteq \cdots \subseteq V_n = V$$
such that $\dim V_i = i$ for each i.

Exercise 11.1.3 If A is an $n \times n$ invertible matrix, show that $A^{-1} = r_0 I + r_1 A + \cdots + r_{n-1}A^{n-1}$ for some scalars $r_0, r_1, \ldots, r_{n-1}$. [*Hint*: Cayley-Hamilton theorem.]

Exercise 11.1.4 If $T : V \to V$ is a linear operator where V is finite dimensional, show that $c_T(T) = 0$. [*Hint*: Exercise 9.1.26.]

Exercise 11.1.5 Define $T : \mathbf{P} \to \mathbf{P}$ by $T[p(x)] = xp(x)$. Show that:

a. T is linear and $f(T)[p(x)] = f(x)p(x)$ for all polynomials $f(x)$.

b. Conclude that $f(T) \neq 0$ for all nonzero polynomials $f(x)$. [See Exercise 11.1.4.]

11.2 The Jordan Canonical Form

Two $m \times n$ matrices A and B are called row-equivalent if A can be carried to B using row operations and, equivalently, if $B = UA$ for some invertible matrix U. We know (Theorem 2.6.4) that each $m \times n$ matrix is row-equivalent to a unique matrix in reduced row-echelon form, and we say that these reduced row-echelon matrices are *canonical forms* for $m \times n$ matrices using row operations. If we allow column operations as well, then $A \to UAV = \left[\begin{array}{cc} I_r & 0 \\ 0 & 0 \end{array}\right]$ for invertible U and V, and the canonical forms are the matrices $\left[\begin{array}{cc} I_r & 0 \\ 0 & 0 \end{array}\right]$ where r is the rank (this is the Smith normal form and is discussed in Theorem 2.6.3). In this section, we discover the canonical forms for square matrices under similarity: $A \to P^{-1}AP$.

If A is an $n \times n$ matrix with distinct real eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$, we saw in Theorem 11.1.1 that A is similar to a block triangular matrix; more precisely, an invertible matrix P exists such that

$$P^{-1}AP = \left[\begin{array}{cccc} U_1 & 0 & \cdots & 0 \\ 0 & U_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & U_k \end{array}\right] = \operatorname{diag}(U_1, U_2, \ldots, U_k) \tag{11.2}$$

where, for each i, U_i is upper triangular with λ_i repeated on the main diagonal. The Jordan canonical form is a refinement of this theorem. The proof we gave of (11.2) is matrix theoretic because we wanted to give an algorithm for actually finding the matrix P. However, we are going to employ abstract methods here. Consequently, we reformulate Theorem 11.1.1 as follows:

Theorem 11.2.1

Let $T : V \to V$ be a linear operator where $\dim V = n$. Assume that $\lambda_1, \lambda_2, \ldots, \lambda_k$ are the distinct eigenvalues of T, and that the λ_i are all real. Then there exists a basis F of V such that $M_F(T) = \operatorname{diag}(U_1, U_2, \ldots, U_k)$ where, for each i, U_i is square, upper triangular, with λ_i repeated on the main diagonal.

Proof. Choose any basis $B = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$ of V and write $A = M_B(T)$. Since A has the same eigenvalues as T, Theorem 11.1.1 shows that an invertible matrix P exists such that $P^{-1}AP = \text{diag}(U_1, U_2, \dots, U_k)$ where the U_i are as in the statement of the Theorem. If $\mathbf{p}_j$ denotes column j of P and $C_B : V \to \mathbb{R}^n$ is the coordinate isomorphism, let $\mathbf{f}_j = C_B^{-1}(\mathbf{p}_j)$ for each j. Then $F = \{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n\}$ is a basis of V and $C_B(\mathbf{f}_j) = \mathbf{p}_j$ for each j. This means that $P_{B \leftarrow F} = \left[C_B(\mathbf{f}_j)\right] = \left[\mathbf{p}_j\right] = P$, and hence (by Theorem 9.2.2) that $P_{F \leftarrow B} = P^{-1}$. With this, column j of $M_F(T)$ is

$$C_F(T(\mathbf{f}_j)) = P_{F \leftarrow B} C_B(T(\mathbf{f}_j)) = P^{-1} M_B(T) C_B(\mathbf{f}_j) = P^{-1} A \mathbf{p}_j$$

for all j. Hence

$$M_F(T) = \left[C_F(T(\mathbf{f}_j))\right] = \left[P^{-1} A \mathbf{p}_j\right] = P^{-1} A \left[\mathbf{p}_j\right] = P^{-1}AP = \text{diag}(U_1, U_2, \dots, U_k)$$

as required. □

Definition 11.2 Jordan Blocks

*If $n \geq 1$, define the **Jordan block** $J_n(\lambda)$ to be the $n \times n$ matrix with λs on the main diagonal, 1s on the diagonal above, and 0s elsewhere. We take $J_1(\lambda) = [\lambda]$.*

Hence

$$J_1(\lambda) = [\lambda], \quad J_2(\lambda) = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}, \quad J_3(\lambda) = \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}, \quad J_4(\lambda) = \begin{bmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{bmatrix}, \quad \dots$$

We are going to show that Theorem 11.2.1 holds with each block U_i replaced by Jordan blocks corresponding to eigenvalues. It turns out that the whole thing hinges on the case $\lambda = 0$. An operator T is called **nilpotent** if $T^m = 0$ for some $m \geq 1$, and in this case $\lambda = 0$ for every eigenvalue λ of T. Moreover, the converse holds by Theorem 11.1.1. Hence the following lemma is crucial.

Lemma 11.2.1

Let $T : V \to V$ be a linear operator where $\dim V = n$, and assume that T is nilpotent; that is, $T^m = 0$ for some $m \geq 1$. Then V has a basis B such that

$$M_B(T) = \text{diag}(J_1, J_2, \dots, J_k)$$

where each J_i is a Jordan block corresponding to $\lambda = 0$.[1]

A proof is given at the end of this section.

[1] The converse is true too: If $M_B(T)$ has this form for some basis B of V, then T is nilpotent.

Theorem 11.2.2: Real Jordan Canonical Form

Let $T : V \to V$ be a linear operator where $\dim V = n$, and assume that $\lambda_1, \lambda_2, \dots, \lambda_m$ are the distinct eigenvalues of T and that the λ_i are all real. Then there exists a basis E of V such that

$$M_E(T) = \operatorname{diag}(U_1, U_2, \dots, U_k)$$

in block form. Moreover, each U_j is itself block diagonal:

$$U_j = \operatorname{diag}(J_1, J_2, \dots, J_k)$$

where each J_i is a Jordan block corresponding to some λ_i.

Proof. Let $E = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ be a basis of V as in Theorem 11.2.1, and assume that U_i is an $n_i \times n_i$ matrix for each i. Let

$$E_1 = \{\mathbf{e}_1, \dots, \mathbf{e}_{n_1}\}, \quad E_2 = \{\mathbf{e}_{n_1+1}, \dots, \mathbf{e}_{n_2}\}, \quad \dots, \quad E_k = \{\mathbf{e}_{n_{k-1}+1}, \dots, \mathbf{e}_{n_k}\}$$

where $n_k = n$, and define $V_i = \operatorname{span}\{E_i\}$ for each i. Because the matrix $M_E(T) = \operatorname{diag}(U_1, U_2, \dots, U_m)$ is block diagonal, it follows that each V_i is T-invariant and $M_{E_i}(T) = U_i$ for each i. Let U_i have λ_i repeated along the main diagonal, and consider the restriction $T : V_i \to V_i$. Then $M_{E_i}(T - \lambda_i I_{n_i})$ is a nilpotent matrix, and hence $(T - \lambda_i I_{n_i})$ is a nilpotent operator on V_i. But then Lemma 11.2.1 shows that V_i has a basis B_i such that $M_{B_i}(T - \lambda_i I_{n_i}) = \operatorname{diag}(K_1, K_2, \dots, K_{t_i})$ where each K_i is a Jordan block corresponding to $\lambda = 0$. Hence

$$\begin{aligned} M_{B_i}(T) &= M_{B_i}(\lambda_i I_{n_i}) + M_{B_i}(T - \lambda_i I_{n_i}) \\ &= \lambda_i I_{n_i} + \operatorname{diag}(K_1, K_2, \dots, K_{t_i}) = \operatorname{diag}(J_1, J_2, \dots, J_k) \end{aligned}$$

where $J_i = \lambda_i I_{f_i} + K_i$ is a Jordan block corresponding to λ_i (where K_i is $f_i \times f_i$). Finally,

$$B = B_1 \cup B_2 \cup \cdots \cup B_k$$

is a basis of V with respect to which T has the desired matrix. □

Corollary 11.2.1

If A is an $n \times n$ matrix with real eigenvalues, an invertible matrix P exists such that $P^{-1}AP = \operatorname{diag}(J_1, J_2, \dots, J_k)$ where each J_i is a Jordan block corresponding to an eigenvalue λ_i.

Proof. Apply Theorem 11.2.2 to the matrix transformation $T_A : \mathbb{R}^n \to \mathbb{R}^n$ to find a basis B of $\mathbb{R}^n$ such that $M_B(T_A)$ has the desired form. If P is the (invertible) $n \times n$ matrix with the vectors of B as its columns, then $P^{-1}AP = M_B(T_A)$ by Theorem 9.2.4. □

Of course if we work over the field $\mathbb{C}$ of complex numbers rather than $\mathbb{R}$, the characteristic polynomial of a (complex) matrix A splits completely as a product of linear factors. The proof of Theorem 11.2.2 goes through to give

Theorem 11.2.3: Jordan Canonical Form[2]

Let $T : V \to V$ be a linear operator where $\dim V = n$, and assume that $\lambda_1, \lambda_2, \dots, \lambda_m$ are the distinct eigenvalues of T. Then there exists a basis F of V such that

$$M_F(T) = \text{diag}(U_1, U_2, \dots, U_k)$$

in block form. Moreover, each U_j is itself block diagonal:

$$U_j = \text{diag}(J_1, J_2, \dots, J_{t_j})$$

where each J_i is a Jordan block corresponding to some λ_i.

Except for the order of the Jordan blocks J_i, the Jordan canonical form is uniquely determined by the operator T. That is, for each eigenvalue λ the number and size of the Jordan blocks corresponding to λ is uniquely determined. Thus, for example, two matrices (or two operators) are similar if and only if they have the same Jordan canonical form. We omit the proof of uniqueness; it is best presented using modules in a course on abstract algebra.

Proof of Lemma 1

Lemma 11.2.1

Let $T : V \to V$ be a linear operator where $\dim V = n$, and assume that T is nilpotent; that is, $T^m = 0$ for some $m \geq 1$. Then V has a basis B such that

$$M_B(T) = \text{diag}(J_1, J_2, \dots, J_k)$$

where each $J_i = J_{n_i}(0)$ is a Jordan block corresponding to $\lambda = 0$.

Proof. The proof proceeds by induction on n. If $n = 1$, then T is a scalar operator, and so $T = 0$ and the lemma holds. If $n \geq 1$, we may assume that $T \neq 0$, so $m \geq 1$ and we may assume that m is chosen such that $T^m = 0$, but $T^{m-1} \neq 0$. Suppose $T^{m-1}\mathbf{u} \neq \mathbf{0}$ for some $\mathbf{u}$ in V.[3]

Claim. $\{\mathbf{u}, T\mathbf{u}, T^2\mathbf{u}, \dots, T^{m-1}\mathbf{u}\}$ is independent.

Proof. Suppose $a_0\mathbf{u} + a_1T\mathbf{u} + a_2T^2\mathbf{u} + \cdots + a_{m-1}T^{m-1}\mathbf{u} = \mathbf{0}$ where each a_i is in $\mathbb{R}$. Since $T^m = 0$, applying T^{m-1} gives $\mathbf{0} = T^{m-1}\mathbf{0} = a_0T^{m-1}\mathbf{u}$, whence $a_0 = 0$. Hence $a_1T\mathbf{u} + a_2T^2\mathbf{u} + \cdots + a_{m-1}T^{m-1}\mathbf{u} = \mathbf{0}$ and applying T^{m-2} gives $a_1 = 0$ in the same way. Continue in this fashion to obtain $a_i = 0$ for each i. This proves the Claim.

Now define $P = \text{span}\{\mathbf{u}, T\mathbf{u}, T^2\mathbf{u}, \dots, T^{m-1}\mathbf{u}\}$. Then P is a T-invariant subspace (because $T^m = 0$), and $T : P \to P$ is nilpotent with matrix $M_B(T) = J_m(0)$ where $B = \{\mathbf{u}, T\mathbf{u}, T^2\mathbf{u}, \dots, T^{m-1}\mathbf{u}\}$. Hence we are done, by induction, if $V = P \oplus Q$ where Q is T-invariant (then $\dim Q = n - \dim P < n$ because $P \neq 0$,

[2]This was first proved in 1870 by the French mathematician Camille Jordan (1838–1922) in his monumental *Traité des substitutions et des équations algébriques*.

[3]If $S : V \to V$ is an operator, we abbreviate $S(\mathbf{u})$ by $S\mathbf{u}$ for simplicity.

and $T : Q \to Q$ is nilpotent). With this in mind, choose a T-invariant subspace Q of maximal dimension such that $P \cap Q = \{\mathbf{0}\}$.[4] We assume that $V \neq P \oplus Q$ and look for a contradiction.

Choose $\mathbf{x} \in V$ such that $\mathbf{x} \notin P \oplus Q$. Then $T^m\mathbf{x} = \mathbf{0} \in P \oplus Q$ while $T^0\mathbf{x} = \mathbf{x} \notin P \oplus Q$. Hence there exists k, $1 \leq k \leq m$, such that $T^k\mathbf{x} \in P \oplus Q$ but $T^{k-1}\mathbf{x} \notin P \oplus Q$. Write $\mathbf{v} = T^{k-1}\mathbf{x}$, so that

$$\mathbf{v} \notin P \oplus Q \quad \text{and} \quad T\mathbf{v} \in P \oplus Q$$

Let $T\mathbf{v} = \mathbf{p} + \mathbf{q}$ with $\mathbf{p}$ in P and $\mathbf{q}$ in Q. Then $\mathbf{0} = T^{m-1}(T\mathbf{v}) = T^{m-1}\mathbf{p} + T^{m-1}\mathbf{q}$ so, since P and Q are T-invariant, $T^{m-1}\mathbf{p} = -T^{m-1}\mathbf{q} \in P \cap Q = \{\mathbf{0}\}$. Hence

$$T^{m-1}\mathbf{p} = \mathbf{0}$$

Since $\mathbf{p} \in P$ we have $\mathbf{p} = a_0\mathbf{u} + a_1T\mathbf{u} + a_2T^2\mathbf{u} + \cdots + a_{m-1}T^{m-1}\mathbf{u}$ for $a_i \in \mathbb{R}$. Since $T^m = 0$, applying T^{m-1} gives $\mathbf{0} = T^{m-1}\mathbf{p} = a_0T^{m-1}\mathbf{u}$, whence $a_0 = 0$. Thus $\mathbf{p} = T(\mathbf{p}_1)$ where

$$\mathbf{p}_1 = a_1\mathbf{u} + a_2T\mathbf{u} + \cdots + a_{m-1}T^{m-2}\mathbf{u} \in P$$

If we write $\mathbf{v}_1 = \mathbf{v} - \mathbf{p}_1$ we have

$$T(\mathbf{v}_1) = T(\mathbf{v} - \mathbf{p}_1) = T\mathbf{v} - \mathbf{p} = \mathbf{q} \in Q$$

Since $T(Q) \subseteq Q$, it follows that $T(Q + \mathbb{R}\mathbf{v}_1) \subseteq Q \subseteq Q + \mathbb{R}\mathbf{v}_1$. Moreover $\mathbf{v}_1 \notin Q$ (otherwise $\mathbf{v} = \mathbf{v}_1 + \mathbf{p}_1 \in P \oplus Q$, a contradiction). Hence $Q \subset Q + \mathbb{R}\mathbf{v}_1$ so, by the maximality of Q, we have $(Q + \mathbb{R}\mathbf{v}_1) \cap P \neq \{\mathbf{0}\}$, say

$$\mathbf{0} \neq \mathbf{p}_2 = \mathbf{q}_1 + a\mathbf{v}_1 \quad \text{where} \quad \mathbf{p}_2 \in P, \quad \mathbf{q}_1 \in Q, \quad \text{and} \quad a \in \mathbb{R}$$

Thus $a\mathbf{v}_1 = \mathbf{p}_2 - \mathbf{q}_1 \in P \oplus Q$. But since $\mathbf{v}_1 = \mathbf{v} - \mathbf{p}_1$ we have

$$a\mathbf{v} = a\mathbf{v}_1 + a\mathbf{p}_1 \in (P \oplus Q) + P = P \oplus Q$$

Since $\mathbf{v} \notin P \oplus Q$, this implies that $a = 0$. But then $\mathbf{p}_2 = \mathbf{q}_1 \in P \cap Q = \{\mathbf{0}\}$, a contradiction. This completes the proof. □

Exercises for 11.2

Exercise 11.2.1 By direct computation, show that there is no invertible complex matrix C such that

$$C^{-1}\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}C = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Exercise 11.2.2 Show that $\begin{bmatrix} a & 1 & 0 \\ 0 & a & 0 \\ 0 & 0 & b \end{bmatrix}$ is similar to $\begin{bmatrix} b & 0 & 0 \\ 0 & a & 1 \\ 0 & 0 & a \end{bmatrix}$.

Exercise 11.2.3

a. Show that every complex matrix is similar to its transpose.

b. Show every real matrix is similar to its transpose. [*Hint*: Show that $J_k(0)Q = Q[J_k(0)]^T$ where Q is the $k \times k$ matrix with 1s down the "counter diagonal", that is from the $(1, k)$-position to the $(k, 1)$-position.]

[4] Observe that there *is* at least one such subspace: $Q = \{\mathbf{0}\}$.