2.

Notes on the Jordan canonical form theorem.

2.1. Throughout, $F \in \{\mathbf{R}, \mathbf{C}\}$. We fix a vector space V over F and a linear operator $T : V \to V$.

2.2. If V is finite-dimensional, then a choice of an ordered basis for V allows us to express T in terms of a corresponding square matrix $M = (M_{j,i})_{j,i}$. If the basis is $e_1, \ldots, e_n$, then the matrix entries $M_{j,i}$ are scalars such that

$$Te_i = \sum_j M_{j,i} e_j \qquad \text{for all } i, j.$$

If we change to a new basis, then we get a new matrix. The change-of-basis formula shows that if M, M' are matrices for T with respect to two different bases, then $M' = PMP^{-1}$ for some invertible matrix P. In this case we say that M and M' are *conjugate*. This defines an equivalence relation on square matrices known as conjugacy.

2.3. We say that a linear subspace $W \subseteq V$ is T*-stable* if and only if T maps W into itself. That is, $Tw \in W$ for all $w \in W$. In this case, T defines a linear operator $T|_W : W \to W$, simply by restricting its domain and target.

Note that V and the zero subspace $\{\vec{0}\}$ are always T-stable. These are the trivial cases. The presence of a nontrivial T-stable subspace corresponds to the existence of a *block-triangular* matrix for T. Similarly, a direct-sum decomposition $V = W + W'$, in which W, W' are both nontrivial and T-stable, corresponds to the existence of a *block-diagonal* matrix for T. These terms are best illustrated through examples.

Example 2.1. Consider the following matrices, where $*$ means an arbitrary scalar:

$$\begin{pmatrix} * & * & * \\ 0 & * & * \\ 0 & * & * \end{pmatrix}, \quad \begin{pmatrix} * & 0 & 0 \\ * & * & * \\ * & * & * \end{pmatrix}, \quad \begin{pmatrix} * & 0 & 0 \\ 0 & * & * \\ 0 & * & * \end{pmatrix}.$$

Convince yourself that if T has a matrix of the first form, with respect to some ordered basis (e_1, e_2, e_3), then the line Fe_1 is T-stable; if it has a matrix of the second form, then the plane $Fe_2 + Fe_3$ is T-stable; and if it has a matrix of the third form, then both subspaces are T-stable.

2.4. A matrix is called *diagonal* if and only if its only nonzero entries are diagonal entries: *i.e.*, those of the form $M_{j,j}$. (There can still be zeros on the diagonal, too.) These are the simplest matrices to understand.

We say that T is *diagonalizable* if and only if we can find a basis in which the matrix of T is diagonal. Similarly, we say that a square matrix is diagonalizable if and only if it is conjugate to a diagonal matrix.

Example 2.2. Consider the 2×2 matrices

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Convince yourself that the first is always diagonalizable; the second is diagonalizable when $F = \mathbf{C}$; and the third is never diagonalizable.

2.5. There is a matrix-free interpretation of what it means for a linear operator to be diagonalizable. First, define a *eigenline* for T to be a 1-dimensional T-stable linear subspace. Then T is diagonalizable if and only if V is a sum of eigenlines for T.

An *eigenvector* is a vector whose span is an eigenline. Equivalently, this is a vector v satisfying two conditions:

$$v \neq \vec{0} \quad \text{and} \quad Tv = \lambda v \text{ for some scalar } \lambda \in F.$$

(The first condition means the span is 1-dimensional, not 0-dimensional; the second means the span is T-stable.) Above, λ is called the *eigenvalue* corresponding to the eigenline. It is possible for λ to be 0.

For any λ, we define the *λ-eigenspace of T* to be the linear subspace of vectors $v \in V$ such that $Tv = \lambda v$. Equivalently, it is the sum of all eigenlines for T with eigenvalue λ. It is also $\ker(T - \lambda \mathrm{Id}_V)$.

2.6. Example 2.2 shows that some linear operators are not diagonalizable. However, it turns out that as long as $F = \mathbf{C}$ and V is finite-dimensional, we can always decompose V as a direct sum of T-stable subspaces that are as close as possible to being eigenspaces. This is the matrix-free meaning of the Jordan canonical form theorem. The first step toward the full theorem is the following result.

Theorem 2.3. *If $F = \mathbf{C}$ and V is finite-dimensional and nonzero, then T has* <u>*some*</u> *eigenline in V.*

The key idea in the proof is to form new operators by plugging T into polynomials. Namely, for any polynomial $p(z) = a_0 + a_1 z + \cdots + a_d z^d \in \mathbf{C}[z]$, we set

$$p(T) = a_0 \mathrm{Id}_V + a_1 T + \cdots + a_d T^d,$$

where addition and scalar multiplication of linear operators is defined pointwise, and T^i means the ith iterate of T. We will omit Id_V from the notation, where convenient.

Proof sketch. Set $n = \dim V > 0$. Pick $v \neq \vec{0}$.

Since $v, Tv, \ldots, T^n v$ is a set of $n + 1$ vectors, they must be linearly dependent. So there exist some $a_0, a_1, \ldots, a_n \in F$, not all zero, such that

$$(a_0 + a_1 T + \cdots + a_n T^n)v = \vec{0}.$$

That is, there exists *some* nonzero polynomial $p(z)$ such that $p(T)v = \vec{0}$. (Note that this particular step did not require $F = \mathbf{C}$.)

Now pick $p(z)$ of *minimal* degree among such polynomials. Since $v \neq \vec{0}$, we know that $p(z)$ cannot be a constant polynomial. So by the fundamental theorem of algebra, it has some root λ. That is,

$$p(z) = (z - \lambda)q(z) \quad \text{for some polynomial } q(z).$$

So $(T - \lambda)q(T)v = p(T)v = \vec{0}$. So as long as $q(T)v$ is nonzero, it is an eigenvector for T with eigenvalue λ. But q is a nonzero polynomial, because p is, and $\deg(q) < \deg(p)$, so $q(T)v = \vec{0}$ would contradict the minimality of p. □

Example 2.4. The conclusion of the theorem need not hold...

- ...if $F = \mathbf{R}$. Check that if T is a rotation in $\mathbf{R}^2$ by an angle that is not a multiple of π radians, then T has no eigenlines.
- ...if V is infinite-dimensional. Check that if T is the shift operator on the vector space of infinite sequences given by $T(x_1, x_2, x_3, \ldots) = (0, x_1, x_2, \ldots)$, then T has no eigenlines.

2.7. In terms of matrices, the previous theorem shows that any square matrix over $\mathbf{C}$ is conjugate to one of the form

$$\begin{pmatrix} * & * & \cdots & * \\ 0 & * & \cdots & * \\ 0 & \vdots & \ddots & \vdots \\ 0 & * & \cdots & * \end{pmatrix}.$$

To see this, pick an eigenvector e_1 for the underlying linear operator, extend it to an ordered basis $e_1, e_2, \ldots, e_n$ for the whole vector space, then observe that the linear operator takes the form above with respect to the new basis. This matrix interpretation of Theorem 2.3 motivates the following improvement:

Theorem 2.5. *If $F = \mathbf{C}$ and V is finite-dimensional, then there is some basis for V in which the matrix of T is upper-triangular:*

$$\begin{pmatrix} * & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & * \end{pmatrix}.$$

The key idea in the proof is that for any scalar λ, the linear subspace $\operatorname{im}(T - \lambda) \subseteq V$ is T-stable. More generally, we have a "stability lemma":

$(*)$ for any $p(z), q(z) \in F[z]$, the subspace $\operatorname{im}(p(T))$ is $q(T)$-stable.

The reason for the lemma: $p(T)$ and $q(T)$ always commute, because *polynomials* in the same variable always commute, so $q(T)(p(T)v) = p(T)(q(T)v)$ for all vectors v.

Proof sketch. We induct on $\dim V$. In the base case $\dim V = 0$, we have $V = \{\vec{0}\}$, so we're done. Else, by Theorem 2.3, T has some eigenvector v, with some eigenvalue λ. This means $\dim \ker(T - \lambda) > 0$.

Now let $V' = \operatorname{im}(T - \lambda)$. By ($*$), V' is T-stable, so we can form $T' = T|_{V'}$. The hypotheses of our desired theorem still hold with V', T' in place of V, T. But $\dim V' < \dim V$, so by induction, we can find an ordered basis for V' in which the matrix of T' is upper-triangular.

Let $w_1, \ldots, w_m$ be this basis. Extend it in an arbitrary way to an ordered basis $w_1, \ldots, w_m, v_1, \ldots, v_\ell$ for V. We claim that T is upper-triangular with respect to the latter basis. Indeed, sketch the matrix M. Convince yourself that the upper-triangularity of T' implies that the lower-left block of M is zero, while the upper-left block of M is upper-triangular. Meanwhile,

$$Tv_i = (T - \lambda)v_i + \lambda v_i \in V' + \mathbf{C}v_i.$$

Convince yourself that this identity implies the upper-triangularity of the lower-right block of M. □

2.8. Any upper-triangular square matrix is a sum of a diagonal matrix and an upper-triangular matrix with zeros along its diagonal. One can check that the latter kind of matrix is always nilpotent: some power of it is the zero matrix. Similarly, a linear operator is called *nilpotent* if and only if some power of it is the zero operator.

The nilpotent case of Theorem 2.5 says, conversely: If W is a finite-dimensional vector space over $\mathbf{C}$, and $S : W \to W$ is a nilpotent operator, then W has some basis in which the matrix of S is upper-triangular with zeros along its diagonal. To refine Theorem 2.5, we might start by trying to refine this case.

Below, we define the *superdiagonal* of a square matrix M to be its collection of entries of the form $M_{j,j+1}$.

Theorem 2.6. *Let F be arbitrary and V a finite-dimensional vector space over F. If $S : V \to V$ is nilpotent, then there is a basis of V in which the matrix of S only has nonzero entries along its superdiagonal, and these entries are all 1's.*

The following 6×6 matrix has the form specified by the theorem's conclusion. (Empty entries denote 0's.)

$$\begin{pmatrix} 0 & 1 & & & & \\ & 0 & & & & \\ & & 0 & 1 & & \\ & & & 0 & 1 & \\ & & & & 0 & \\ & & & & & 0 \end{pmatrix}$$

This matrix is a *block sum* of three *Jordan blocks*: in order from top to bottom,

$$\begin{pmatrix} 0 & 1 \\ & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & \\ & 0 & 1 \\ & & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \end{pmatrix}.$$

Each Jordan block J corresponds to an S-stable linear subspace $V_J \subseteq V$. To illustrate how, suppose that in the setup above, the ordered basis of W is labeled $e_1, e_2, \ldots, e_6$. Then the three Jordan blocks correspond to $\operatorname{span}(e_1, e_2)$, $\operatorname{span}(e_3, e_4, e_5)$, $\operatorname{span}(e_6)$, in order. The linear operator S sends

$$e_2 \mapsto e_1 \mapsto \vec{0}, \quad e_5 \mapsto e_4 \mapsto e_3 \mapsto \vec{0}, \quad e_6 \mapsto \vec{0}.$$

In general, we define a *Jordan chain* of a nilpotent linear operator $S : V \to V$ to be a sequence of nonzero vectors $v_1, v_2, \ldots, v_\ell \in V$ of maximal length such that S sends

$$v_1 \mapsto v_2 \mapsto \cdots \mapsto v_\ell \mapsto \vec{0}.$$

By construction, S admits a Jordan chain of length ℓ if and only if V admits a basis in which the matrix of S contains a (maximal) $\ell \times \ell$ Jordan block. Thus Theorem 2.6 is equivalent to the following matrix-free statement:

Theorem 2.7. *Let F be arbitrary and V a finite-dimensional vector space over F. If $S : V \to V$ is nilpotent, then there is a basis for V consisting of disjoint Jordan chains of S.*

Proof sketch. Induct on $\dim V$. In the base case $\dim V = 0$, we have $V = \{\vec{0}\}$, so we're done. Else, $\dim \ker(S) > 0$ because S is nilpotent.

Now set $V' = \operatorname{im}(S)$. By $(*)$, V' is S-stable, so we can form $S' = S|_{V'}$. Note that V' remains finite-dimensional and S' remains nilpotent. But $\dim V' < \dim V$, so by induction, we can find a basis for V' consisting of disjoint Jordan chains of S'. We can thus write its elements as

$$w_1, Sw_1, \ldots, S^{d_1} w_1, \quad w_2, Sw_2, \ldots, S^{d_2} w_2, \quad \ldots, \quad w_k, Sw_k, \ldots, S^{d_k} w_k.$$

We need to extend it to a basis for W consisting of disjoint Jordan chains of S.

Since $w_j \in W' = \operatorname{im}(S)$, we can pick some $v_1, \ldots, v_k \in W$ such that $Sv_j = w_j$ for all j. Since the elements $S^{d_1} w_1, \ldots, S^{d_k} w_k$ form a basis for $\ker(S')$, we also know that $k = \dim \ker(S')$. Since $\ker(S') = \ker(S) \cap W'$, a linear subspace of $\ker(S)$, we can pick $u_1, \ldots, u_\ell \in W$ such that

$$S^{d_1} w_1, \ldots, S^{d_k} w_k, u_1, \ldots, u_\ell$$

together form a basis for $\ker(S)$. In this case, $k + \ell = \dim \ker(S)$. So once we append the v_j's and u_i's to our basis of V', we get a set of size

$$\dim \operatorname{im}(S) + k + \ell = \dim \operatorname{im}(S) + \dim \ker(S) = \dim V.$$

Check that the enlarged set is linearly independent. Hence, it forms a basis for V. □

2.9. We can immediately bootstrap Theorem 2.6 to a slightly broader situation. Recall that the λ-eigenspace of a general operator $T : V \to V$ is defined as $\ker(T - \lambda) \subseteq V$. The *generalized λ-eigenspace* of T is defined as

$$\bigcup_{n>0} \ker((T - \lambda)^n).$$

Note that these kernels are nested:

$$\ker(T - \lambda) \subseteq \ker((T - \lambda)^2) \subseteq \ker((T - \lambda)^3) \subseteq \cdots$$

It follows that the generalized λ-eigenspace is indeed a linear subspace of V, and that if V is finite-dimensional, then the sequence of kernels stabilizes after finitely many steps N, so that the generalized eigenspace is just $\ker((T - \lambda)^N)$. A slightly subtler fact:

Lemma 2.8. *If the generalized λ-eigenspace of T is nonzero, then the actual λ-eigenspace of T is also nonzero:* i.e., *λ is actually an eigenvalue for T.*

Proof sketch. Show that there exist $n > 0$ and $v \in V$ such that $(T - \lambda)^{n-1} v \neq \vec{0}$, but $(T - \lambda)^n v = \vec{0}$. Then the desired eigenvector is $(T - \lambda)^{n-1} v$. □

Finally, we can check that the generalized λ-eigenspace of T is the maximal T-stable linear subspace of V on which T restricts to a nilpotent operator. So Theorem 2.6 implies:

Theorem 2.9. *Let F be arbitrary and V a finite-dimensional vector space over F. If V forms a single generalized eigenspace of T, with eigenvalue λ, then there is some basis for V in which the matrix of T only has:*

- *λ's along its diagonal,*
- *0's and 1's along its superdiagonal,*
- *0's everywhere else.*

We refer to blocks along the matrix diagonal that take the form

$$\begin{pmatrix} \lambda & 1 \\ & \lambda \end{pmatrix}, \quad \begin{pmatrix} \lambda & 1 & \\ & \lambda & 1 \\ & & \lambda \end{pmatrix}, \quad \begin{pmatrix} \lambda \end{pmatrix}, \quad \textit{etc.}$$

as *Jordan blocks with eigenvalue λ*, this terminology being justified by Lemma 2.8.

2.10. Although V need not be a sum of eigenspaces of T, we can say more about generalized eigenspaces of T:

Theorem 2.10. *If $F = \mathbf{C}$ and V is finite-dimensional, then V is a direct sum of generalized eigenspaces of T. That is, there exist a finite list of scalars $\lambda_1, \ldots, \lambda_k \in F$ and a direct-sum decomposition $V = W_1 + \cdots + W_k$ such that W_i is the generalized λ_i-eigenspace of T for all i.*

Proof sketch. Induct on $\dim V$. If $\dim V = 0$, then we're done, Else, by Theorem 2.3, T has an eigenvalue λ. Hence the λ-eigenspace of T is nonzero.

Pick N large enough that the generalized λ-eigenspace of T is exactly $\ker((T-\lambda)^N)$. Let $V' = \mathrm{im}((T-\lambda)^N)$. We claim that $V = V' + \ker((T-\lambda)^N)$, and that this is a direct sum. By the usual dimension formula, it suffices to check that $V' \cap \ker((T-\lambda)^N) = \{\vec{0}\}$. Indeed, if w belongs to this intersection, then $w \in V'$ implies that $w = (T-\lambda)^N v$ for some $v \in V$, while $w \in \ker((T-\lambda)^N)$ implies that $(T-\lambda)^{2N} v = (T-\lambda)^N w = \vec{0}$. Since we chose N large enough that

$$\ker((T-\lambda)^N) = \ker((T-\lambda)^{N+1}) = \cdots = \ker(T-\lambda)^{2N}),$$

we deduce that $w = (T-\lambda)^N v = \vec{0}$.

By ($*$), V' is T-stable, so we can form $T' = T|_{V'}$. Since $\dim V' < \dim V$, the inductive hypothesis says that V' is a direct sum of generalized eigenspaces of T'. Note that λ cannot occur as one of the eigenvalues here, since V' has zero intersection with the generalized λ-eigenspace of T. For all other scalars μ, check that the generalized μ-eigenspace of T' in V' is also the generalized μ-eigenspace of T in V. Check that the sum of all these subspaces along with $\ker((T-\lambda)^N)$ remains a direct sum. ☐

Together, Theorem 2.9 and Theorem 2.10 produce a refinement of Theorem 2.5:

Corollary 2.11 (Jordan Canonical Form)**.** *If $F = \mathbf{C}$ and V is finite-dimensional, then there is some basis for V in which the matrix of T is a block sum of Jordan blocks with possibly-varying, possibly-nonzero eigenvalues.*

Equivalently, every square matrix over $\mathbf{C}$ is conjugate to some block sum of Jordan blocks.

The following is a typical Jordan canonical form matrix, with eigenvalues $\lambda_1, \lambda_2, \lambda_3$:

$$\begin{pmatrix} \lambda_1 & 1 & & & & \\ & \lambda_1 & & & & \\ & & \lambda_2 & 1 & & \\ & & & \lambda_2 & 1 & \\ & & & & \lambda_2 & \\ & & & & & \lambda_3 \end{pmatrix}$$

Note that there are no entries "between" the blocks. However, two or more of the λ_i's could be equal. The matrix is nilpotent precisely when they are all zero.