A proof of the Jordan normal form theorem

Jordan normal form theorem states that any matrix is similar to a block-diagonal matrix with Jordan blocks on the diagonal. To prove it, we first reformulate it in the following way:

Jordan normal form theorem. For any finite-dimensional vector space V and any linear operator $A\colon V \to V$, there exist

- a decomposition of V
$$V = V_1 \oplus V_2 \oplus \ldots \oplus V_k$$
into a direct sum of invariant subspaces of A;
- a basis $e_1^{(i)}, \ldots, e_{n_i}^{(i)}$ of V_i for each $i = 1, \ldots, k$ such that
$$(A - \lambda_i \operatorname{Id})e_1^{(i)} = 0; (A - \lambda_i \operatorname{Id})e_2^{(i)} \quad = e_1^{(i)}; \ldots (A - \lambda_i \operatorname{Id})e_{n_i}^{(i)} = e_{n_i-1}^{(i)}$$
for some λ_i (which may coincide or be different for different i). Dimensions of these subspaces and coefficients λ_i are determined uniquely up to permutations.

Invariant subspaces, direct sums, and block matrices

Recall the following important definitions.

Definition 1. A subspace U of a vector space V is called an invariant subspace of a linear operator $A\colon V \to V$ if for any $u \in U$ we have $A(u) \in U$.

Definition 2. For any two subspaces U_1 and U_2 of a vector space V, their sum $U_1 + U_2$ is defined as the set of all vectors $u_1 + u_2$, where $u_1 \in U_1$, $u_2 \in U_2$. If $U_1 \cap U_2 = \{0\}$, then the sum of U_1 and U_2 is called direct, and is denoted by $U_1 \oplus U_2$.

Choose any decomposition of V into a direct sum of two subspaces U_1 and U_2. Note that if $u_1 \in U_1$ and $u_2 \in U_2$, then $u_1 + u_2 = 0$ implies $u_1 = u_2 = 0$. Indeed, one can rewrite it as $u_1 = -u_2$ and use the fact that $U_1 \cap U_2 = \{0\}$. This means that any vector of the direct sum can be represented in the form $u_1 + u_2$, where $u_1 \in U_1$ and $u_2 \in U_2$, in a unique way. An immediate consequence of that fact is the following important formula:
$$\dim(U_1 \oplus U_2) = \dim U_1 + \dim U_2.$$
Indeed, it is clear that we can get a basis for the direct sum by joining together bases for summands.

Fix a basis for U_1 and a basis for U_2. Joining them together, we get a basis for $V = U_1 \oplus U_2$. When we write down the matrix of any linear operator with respect to this basis, we get a block-diagonal matrix $\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$; its splitting into blocks correspond to the way our basis is split into two parts.

Let use formulate important facts which are immediate from the definition of a matrix of a linear operator.

1. $A_{12} = 0$ if and only if U_2 is invariant;
2. $A_{21} = 0$ if and only if U_1 is invariant.

Thus, this matrix is block-triangular if and only if one of the subspaces is invariant, and is block-diagonal if and only if both subspaces are invariant. This admits an obvious generalisation for the case of larger number of summands in the direct sum.

Generalised eigenspaces

The first step of the proof is to decompose our vector space into the direct sum of invariant subspaces where our operator has only one eigenvalue.

Definition 3. Let $N_1(\lambda) = \mathrm{Ker}(A - \lambda \,\mathrm{Id})$, $N_2(\lambda) = \mathrm{Ker}(A - \lambda \,\mathrm{Id})^2$, ..., $N_m(\lambda) = \mathrm{Ker}(A - \lambda \,\mathrm{Id})^m$, ... Clearly,

$$N_1(\lambda) \subset N_2(\lambda) \subset \ldots \subset N_m(\lambda) \subset \ldots$$

Since we only work with finite-dimensional vector spaces, this sequence of subspaces cannot be strictly increasing; if $N_i(\lambda) \neq N_{i+1}(\lambda)$, then, obviously, $\dim N_{i+1}(\lambda) \geqslant 1 + \dim N_i(\lambda)$. It follows that for some k we have $N_k(\lambda) = N_{k+1}(\lambda)$.

Lemma 1. We have $N_k(\lambda) = N_{k+1}(\lambda) = N_{k+2}(\lambda) = \ldots$

Let us prove that $N_{k+l}(\lambda) = N_{k+l-1}(\lambda)$ by induction on l. Note that the induction basis (case $l = 1$) follows immediately from our notation.Suppose that $N_{k+l}(\lambda) = N_{k+l-1}(\lambda)$; let us prove $N_{k+l+1}(\lambda) = N_{k+l}(\lambda)$. If we assume that it is false, then there is a vector v such that

$$v \in N_{k+l+1}(\lambda), v \notin N_{k+l}(\lambda),$$

that is

$$(A - \lambda \,\mathrm{Id})^{k+l+1}(v) = 0, (A - \lambda \,\mathrm{Id})^{k+l}(v) \neq 0.$$

Put $w = (A - \lambda \,\mathrm{Id})(v) = 0$. Obviously, we have

$$(A - \lambda \,\mathrm{Id})^{k+l}(w) = 0, (A - \lambda \,\mathrm{Id})^{k+l-1}(w) \neq 0,$$

which contradicts the induction hypothesis, and our statement follows.

Lemma 2. $\mathrm{Ker}(A - \lambda \,\mathrm{Id})^k \cap \mathrm{Im}(A - \lambda \,\mathrm{Id})^k = \{0\}$.

Indeed, assume there is a vector $v \in \operatorname{Ker}(A - \lambda \operatorname{Id})^k \cap \operatorname{Im}(A - \lambda \operatorname{Id})^k$. This means that $(A - \lambda \operatorname{Id})^k(v) = 0$ and that there exists a vector w such that $v = (A - \lambda \operatorname{Id})^k(w)$. It follows that $(A - \lambda \operatorname{Id})^{2k}(w) = 0$, so $w \in \operatorname{Ker}(A - \lambda \operatorname{Id})^{2k} = N_{2k}(\lambda)$. But from the previous lemma we know that $N_{2k}(\lambda) = N_k(\lambda)$, so $w \in \operatorname{Ker}(A - \lambda \operatorname{Id})^k$. Thus, $v = (A - \lambda \operatorname{Id})^k(w) = 0$, which is what we need.

Lemma 3. $V = \operatorname{Ker}(A - \lambda \operatorname{Id})^k \oplus \operatorname{Im}(A - \lambda \operatorname{Id})^k$.

Indeed, consider the direct sum of these two subspaces. It is a subspace of V of dimension $\dim \operatorname{Ker}(A-\lambda \operatorname{Id})^k + \dim \operatorname{Im}(A-\lambda \operatorname{Id})^k$. Let $A' = (A-\lambda \operatorname{Id})^k$. Earlier we proved that for any linear operator its rank and the dimension of its kernel sum up to the dimension of the vector space where it acts. $\operatorname{rk} A' = \dim \operatorname{Im} A'$, so we have

$$\dim \operatorname{Ker}(A - \lambda \operatorname{Id})^k + \dim \operatorname{Im}(A - \lambda \operatorname{Id})^k = \dim \operatorname{Ker} A' + \operatorname{rk} A' = \dim V.$$

A subspace of V whose dimension is equal to $\dim V$ has to coincide with V, so the lemma follows.

Lemma 4. $\operatorname{Ker}(A - \lambda \operatorname{Id})^k$ and $\operatorname{Im}(A - \lambda \operatorname{Id})^k$ are invariant subspaces of A.

Indeed, note that $A(A - \lambda \operatorname{Id}) = (A - \lambda \operatorname{Id})A$, so

– if $(A - \lambda \operatorname{Id})^k(v) = 0$, then $(A - \lambda \operatorname{Id})^k(A(v)) = A(A - \lambda \operatorname{Id})^k(v) = 0$;
– if $v = (A-\lambda \operatorname{Id})^k(w)$, then $A(v) = A(A-\lambda \operatorname{Id})^k(w) = (A-\lambda \operatorname{Id})^k(A(w))$.

To complete this step, we use induction on $\dim V$. Note that on the invariant subspace $\operatorname{Ker}(A - \lambda \operatorname{Id})^k$ the operator A has only one eigenvalue (if $Av = \mu v$ for some $0 \neq v \in \operatorname{Ker}(A - \lambda \operatorname{Id})^k$, then $(A - \lambda \operatorname{Id})v = (\mu - \lambda)v$, and $0 = (A - \lambda \operatorname{Id})^k v = (\mu - \lambda)^k v$, so $\mu = \lambda$), and the dimension of $\operatorname{Im}(A - \lambda \operatorname{Id})^k$ is less than $\dim V$ (if λ is an eigenvalue of A), so we can apply the induction hypothesis for A acting on the vector space $V' = \operatorname{Im}(A - \lambda \operatorname{Id})^k$. This results in the following

Theorem. For any linear operator $A\colon V \to V$ whose (different) eigenvalues are $\lambda_1, \ldots, \lambda_k$, there exist integers $n_1, \ldots, n_k$ such that

$$V = \operatorname{Ker}(A - \lambda_1 \operatorname{Id})^{n_1} \oplus \ldots \oplus \operatorname{Ker}(A - \lambda_k \operatorname{Id})^{n_k}.$$

Normal form for a nilpotent operator

The second step in the proof is to establish the Jordan normal form theorem for the case of an operator $B\colon V \to V$ for which $B^k = 0$ (such operators are called nilpotent). This would basically complete the proof, after we put $B = A - \lambda \operatorname{Id}$ and use the result that we already obtained; we will discuss it more precisely below.

Let us modify slightly the notation we used in the previous section; put $N_1 = \operatorname{Ker} B$, $N_2 = \operatorname{Ker} B^2$, $\ldots$, $N_m = \operatorname{Ker} B^m$, $\ldots$ We have $N_k = N_{k+1} = N_{k+2} = \ldots = V$.

To make our proof more neat, we shall use the following definition.

Definition 4. For a vector space V and a subspace $U \subset V$, we say that a sequence of vectors $e_1, \ldots, e_l$ is a basis of V relative to U if any vector $v \in V$ can be uniquely represented in the form $c_1e_1 + c_2e_2 + \ldots + c_le_l + u$, where $c_1, \ldots, c_l$ are coefficients, and $u \in U$. In particular, the only linear combination of $e_1, \ldots, e_l$ that belongs to U is the trivial combination (all coefficients are equal to zero).

Example 1. The usual notion of a basis is contained in the new notion of a relative basis: a usual basis of V is a basis relative to $U = \{0\}$.

Definition 5. We say that a sequence of vectors $e_1, \ldots, e_l$ is linearly independent relative to U if the only linear combination of $e_1, \ldots, e_l$ that belongs to U is the trivial combination (all coefficients are equal to zero).

Exercise 1. Any sequence of vectors that is linearly independent relative to U can be extended to a basis relative to U.

Now we are going to prove our statement, constructing a required basis in k steps. First, find a basis of $V = N_k$ relative to N_{k-1}. Let $e_1, \ldots, e_s$ be vectors of this basis.

Lemma 5. The vectors $e_1, \ldots, e_s, B(e_1), \ldots, B(e_s)$ are linearly independent relative to N_{k-2}.

Indeed, assume that

$$c_1e_1 + \ldots + c_se_s + d_1B(e_1) + \ldots + d_sB(e_s) \in N_{k-2}.$$

Since $e_i \in N_k$, we have $B(e_i) \in N_{k-1} \supset N_{k-2}$, so

$$c_1e_1 + \ldots + c_se_s \in -d_1B(e_1) - \ldots - d_sB(e_s) + N_{k-2} \subset N_{k-1},$$

which means that $c_1 = \ldots = c_s = 0$ ($e_1, \ldots, e_s$ form a basis relative to N_{k-1}). Thus,

$$B(d_1e_1 + \ldots + d_se_s) = d_1B(e_1) + \ldots + d_sB(e_s) \in N_{k-2},$$

so

$$d_1e_1 + \ldots + d_se_s \in N_{k-1},$$

and we deduce that $d_1 = \ldots = d_s = 0$ ($e_1, \ldots, e_s$ form a basis relative to N_{k-1}), so the lemma follows.

Now we extend this collection of vectors by vectors $f_1, \ldots, f_t$ which together with $B(e_1), \ldots, B(e_s)$ form a basis of N_{k-1} relative to N_{k-2}. Absolutely analogously one can prove

Lemma 6. The vectors $e_1, \ldots, e_s$, $B(e_1), \ldots, B(e_s)$, $B^2(e_1), \ldots, B^2(e_s)$, $f_1, \ldots, f_t$, $B(f_1), \ldots, B(f_t)$ are linearly independent relative to N_{k-3}.

We continue that extension process until we end up with a usual basis of V of the following form:

$$e_1, \ldots, e_s, B(e_1), \ldots, B(e_s), B^2(e_1), \ldots, B^{k-1}(e_1), \ldots, B^{k-1}(e_s),$$
$$f_1, \ldots, f_t, B(f_1), \ldots, B^{k-2}(f_1), \ldots, B^{k-2}(f_t),$$
$$\ldots,$$
$$g_1, \ldots, g_u,$$

where the first line contains a vector from N_k, a vector from N_{k-1}, ..., a vector from N_1, the second one — a vector from N_{k-1}, a vector from N_{k-2}, ..., a vector from N_1,, ..., the last one — just a vector from N_1.

To get from this basis a Jordan basis, we just re-number the basis vectors. Note that the vectors

$$v_1 = B^{k-1}(e_1), v_2 = B^{k-2}(e_1), \ldots, v_{k-1} = B(e_1), v_k = e_1$$

form a "thread" of vectors for which $B(v_1) = 0$, $B(v_i) = v_{i-1}$ for $i > 1$, which are precisely formulas for the action of a Jordan block matrix. Arranging all vectors in chains like that, we obtain a Jordan basis.

Remark 1. Note that if we denote by m_d the number of Jordan blocks of size d, we have

$$\begin{aligned} m_1 + m_2 + \ldots + m_k &= \dim N_1, \\ m_2 + \ldots + m_k &= \dim N_2 - \dim N_1, \\ &\ldots \\ m_k &= \dim N_k - \dim N_{k-1}, \end{aligned}$$

so the sizes of Jordan blocks are uniquely determined by the properties of our operator.

General case of the Jordan normal form theorem

From the second section we know that V can be decomposed into a direct sum of invariant subspaces $\text{Ker}(A - \lambda_i \text{Id})^{n_i}$. From the third section, changing the notation and putting $B = A - \lambda_i \text{Id}$, we deduce that each of these subspaces can be decomposed into a direct sum of subspaces where A acts by a Jordan block matrix; sizes of blocks can be computed from the dimension data listed above. This completes the proof of the Jordan normal form theorem.