MATH. 513. JORDAN FORM

Let $A_1, \ldots, A_k$ be square matrices of size $n_1, \ldots, n_k$, respectively with entries in a field F. We define the matrix $A_1 \oplus \ldots \oplus A_k$ of size $n = n_1 + \ldots + n_k$ as the block matrix

$$\begin{pmatrix} \boxed{A_1} & 0 & 0 & \ldots & 0 \\ 0 & \boxed{A_2} & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \ldots & \ldots & 0 & \boxed{A_k} \end{pmatrix}$$

It is called the *direct sum* of the matrices $A_1, \ldots, A_k$. A matrix of the form

$$\begin{pmatrix} \lambda & 1 & 0 & \ldots & 0 \\ 0 & \lambda & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \ldots & \ldots & \lambda & 1 \\ 0 & \ldots & \ldots & 0 & \lambda \end{pmatrix}$$

is called a *Jordan block*. If k is its size, it is denoted by $J_k(\lambda)$. A direct sum

$$J = J_{k_1} \oplus \ldots \oplus J_{k_r}(\lambda_r)$$

of Jordan blocks is called a *Jordan matrix*.

Theorem. *Let $T : V \to V$ be a linear operator in a finite-dimensional vector space over a field F. Assume that the characteristic polynomial of T is a product of linear polynimials. Then there exists a basis $\mathcal{E}$ in V such that $[T]_{\mathcal{E}}$ is a Jordan matrix.*

Corollary. *Let $A \in M_n(F)$. Assume that its characteristic polynomial is a product of linear polynomials. Then there exists a Jordan matrix J and an invertible matrix C such that*

$$A = CJC^{-1}.$$

Notice that the Jordan matrix J (which is called a *Jordan form* of A) is not defined uniquely. For example, we can permute its Jordan blocks. Otherwise the matrix J is defined uniquely (see Problem 7). On the other hand, there are many choices for C. We have seen this already in the diagonalization process.

What is good about it? We have, as in the case when A is diagonalizable,

$$A^N = CJ^NC^{-1}.$$

So, if we can compute J^N, we can compute A^N. It follows from the matrix multiplication that

$$(A_1 \oplus \ldots \oplus A_k)^N = A_1^N \oplus \ldots \oplus A_k^N.$$

Thus it is enough to learn how to raise a Jordan block in Nth power. First consider the case when $\lambda = 0$. We have

$$J_k(0)^2 = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \vdots & \vdots & 0 & 1 \\ 0 & \ldots & \ldots & 0 & 0 \end{pmatrix}^2 = \begin{pmatrix} 0 & 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \ldots & \ldots & 0 & 0 & 1 \\ 0 & \ldots & \ldots & 0 & 0 & 0 \\ 0 & \ldots & \ldots & 0 & 0 & 0 \end{pmatrix}$$

We see that the ones go to the right until they disappear. Continuing in this way, we see that

$$J_k(0)^k = 0.$$

Now we have

$$J_k(\lambda)^N = (\lambda I_n + J_k(0))^N =$$

$$\lambda^N I_n + \binom{N}{1}\lambda^{N-1}J_k(0) + \ldots + \binom{N}{i}\lambda^{N-i}J_k(0)^i + \ldots + \lambda\binom{N}{N-1}J_k(0)^{N-1} + J_k(0)^N. \qquad (1)$$

This is proved in the same way as one proves the Newton formula:

$$(a+b)^N = \sum_{i=0}^{N}\binom{N}{i}a^{n-i}b^i.$$

We look at the product of N factors $(a+b)\ldots(a+b)$. To get a monomial $a^{n-i}b^i$ we choose i brackets from which we will take b. The number of choices is $\binom{N}{i}$.

Notice that in formula (1), the powers $J_k(0)^i$ are equal to zero as soon as $i \geq k$.

So we get

$$J_k(\lambda)^N = \begin{pmatrix} \lambda^N & \binom{N}{1}\lambda^{N-1} & \binom{N}{2}\lambda^{N-2} & \ldots & \ldots & \binom{N}{k-1}\lambda^{N-k+1} \\ 0 & \lambda^N & \binom{N}{1}\lambda^{N-1} & \binom{N}{2}\lambda^{N-2} & \ldots & \binom{N}{k-2}\lambda^{N-k+2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \lambda^N & \binom{N}{1}\lambda^{N-1} \\ 0 & 0 & \ldots & \ldots & 0 & \lambda^N \end{pmatrix},$$

where, by definition $\binom{N}{m} = 0$ if $N < m$.

Before we go to the proof of the Theorem, let us explain how to find J and C. Notice that

$$\operatorname{rank}(J_k(0)) = k-1, \quad \operatorname{rank}(J_k(0)^2) = k-2, \quad \ldots \quad , \operatorname{rank}(J_k(0)^k) = 0.$$

Let us introduce the notion of the *corank* of a matrix $A \in Mat_n(F)$ by setting

$$\boxed{\operatorname{corank}(A) = n - \operatorname{rank}(A)}.$$

Then we see that $\operatorname{corank}(J_k(0))^i = i$ and is equal to the number of the first zero columns. Now, for any Jordan matrix $J = J_{k_1}(\lambda_1) \oplus \ldots J_{k_r}(\lambda_r)$, we have

$$\operatorname{corank}((J - \lambda I_n)^q) = \sum_{i \in I(\lambda)} \operatorname{corank}(J_{k_i}(0)^q),$$

where $I(\lambda)$ is the set of indices i for which $\lambda_i = \lambda$. Let

$$\boxed{d_m(\lambda) = \quad \text{the number of Jordan blocks } J_m(\lambda) \text{ in } J},$$

$$\boxed{c_q(\lambda) = \operatorname{corank}((J - \lambda I_n)^q) = \operatorname{corank}((A - \lambda I_n)^q)}. \tag{2}$$

The last equality follows from the fact that

$$(A - \lambda I_n)^q = (CJC^{-1} - \lambda I_n)^q = (C(J - \lambda I_n)C^{-1})^q = C(J - \lambda I_n)^q C^{-1}.$$

So, $(A - \lambda I_n)^q$ and $(J - \lambda I_n)^q$ are matrices of the same operator and hence have the same rank.

We have

$$c_1(\lambda) = \sum_{m \geq 1} d_m(\lambda),$$

$$c_2(\lambda) - c_1(\lambda) = \sum_{m \geq 2} d_m(\lambda),$$

$$\dots\dots\dots\dots\dots\dots\dots\dots$$

$$c_j(\lambda) - c_{j-1}(\lambda) = \sum_{m \geq j} d_m(\lambda),$$

$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots$$

$$c_n(\lambda) - c_{n-1}(\lambda) = \sum_{m \geq n} d_m(\lambda) = d_n(\lambda).$$

Solving this system for $d_m(\lambda)$ we find

$$\boxed{d_1(\lambda) = 2c_1(\lambda) - c_2(\lambda)},$$

$$\boxed{d_j(\lambda) = -c_{j-1}(\lambda) + 2c_j(\lambda) - c_{j+1}(\lambda)}, \quad j = 2, \dots, n-1, \tag{3}$$

$$\boxed{d_n(\lambda) = c_n(\lambda) - c_{n-1}(\lambda)}.$$

This gives the answer for J.

Remark. For matrices of small size we can list all possible Jordan forms, and then choose the right one by applying formula (2).

Example. Let

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

The characteristic polynomial is $P_A(\lambda) = (-\lambda)^3 + 2(-\lambda)^2$. So the eigenvalues are 0 and 2. Since 0 is a root of multiplicity 2, it must appear twice at the diagonal. Possible Jordan forms are

$$J_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad J_2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Since $\operatorname{corank}(A) = \operatorname{corank}(A - 0I_2) = \operatorname{corank}(J_2 - 0I_2) = 1$ we have to choose $J = J_2$.

Now let us give the answer for C. We have $A = CJC^{-1}$, hence the columns C_i of C form a basis $\mathcal{B}$ in F^n such that the matrix $[T_A]_{\mathcal{B}} = J$. By definition,

$$AC_1 = \lambda_1 C_1, \quad AC_2 = \lambda_1 C_2 + C_1, \quad \dots, \quad AC_{k_1} = \lambda_1 C_{k_1} + C_{k_1 - 1},$$

$$AC_{k_1+1} = \lambda_2 C_{k_1+1}, \quad AC_2 = \lambda_2 C_{k_1+2} + C_{k_1+1}, \quad \ldots, \quad AC_{k_2} = \lambda_2 C_{k_2} + C_{k_2-1}, \ldots$$

and so on. We solve the first equation for C_1. This is a homogeneous system of linear equations $(A - \lambda_1 I_n)x = 0$. We take any non-zero solution C_1 of this system for which $(A - \lambda I_n)x = C_1$ is solvable. Then we solve this equation to find C_2. Continuing in this way we find $C_1, \ldots, C_{k_1}$. Then we go to the next block with eigenvalue λ_2. If $\lambda_1 = \lambda_2$, when we solve $(A - \lambda_1 I_n)x = 0$ to find C_{k_1+1}, we choose a solution which is not proportional to C_1. If there are more Jordan blocks with the same eigenvalue λ_1, each time we solve $(A - \lambda_1 I_n)x = 0$ we find a solution which is not a linear combination of the previous solutions. In fact, we find a basis of $\mathrm{N}(A - \lambda_1 I_n)$ and take its elements for the columns $C_1, C_{k_1+1}, \ldots, C_{k_s+1}$ if $\lambda_1 = \ldots = \lambda_s \neq \lambda_i, \quad i > s$. It follows from the proof of the theorem about the existence of a Jordan form that one can always find C defined by the above equations.

Example (cont.) Let us find the matrix C such that $A = CJC^{-1}$. First we solve $Ax = 0$. The space of solutions is one dimensional. The vector $(-1, 1, 0)$ forms a basis in the space of solutions. This will be our column C_1. The second column C_2 is found by solving $AC_2 = C_1$. A solution is $(3, 0, -2)$. It is not a unique solution. We can add to it any solution of $Ax = 0$ to obtain another solution. Finally we find the third column C_3 by solving the equation $(A - 2I_3)C_3 = 0$. A solution is $C_3 = (1, 1, 0)$. So

$$C = \begin{pmatrix} -1 & -3 & 1 \\ 1 & 0 & 1 \\ 0 & 2 & 0 \end{pmatrix}.$$

We leave to the reader to verify that $A = CJC^{-1}$.

Now let us go to the proof of the main theorem.

For any linear operator $T : V \to V$ and a non-negative integer i we denote by T^i the composition of T with itself i-times. By definition, $T^0 = \mathrm{id}_V$.

A linear subspace $L \subset V$ is called *invariant* with respect to T if $T(v) \in L$ for any $v \in L$.

Let

$$L_1 \subset L_2 \subset \ldots \subset L_k \subset \ldots \subset V$$

be a sequence of linear subspaces of V such that each is a subset of the next one. Let

$$L = \bigcup_k L_k$$

be the union of these subspaces (it could be infinitely many of them). I claim that L is a linear subspace of V. Indeed, take $v, w \in L$. Let $v \in L_k, w \in L_m$. Without loss of generality we may assume that $k \leq m$. Since $L_k \subset L_m$, we get $v, w \in L_m$. Since L_m is a linear subspace, we get $v + cw \in L_m \subset L$ for any $c \in F$. Thus L is a linear subspace.

We apply this to the following situation. Let U be a linear operator, and $L_k = N(U^k)$. We have $N(U^k) \subset N(U^{k+1})$ since $N^k(v) = 0$ implies $N(U^{k+1}) = N(N^k(v)) = N(0) = 0$. Thus we have a sequence of linear subspaces

$$N(U) \subset N(U^2) \subset \ldots \subset N(U^k) \subset \ldots \subset V.$$

It follows from above that

$$\bigcup_k N(U^k) = \{v \in V : U^k(v) = 0 \quad \text{for some } k > 0\}.$$

is a linear subspace of V. It is also invariant with respect to U. In fact, if $v \in N(U^k)$ for some k, then $U^{k-1}(U(v)) = U^k(v) = 0$, hence $U(v) \in N(U^{k-1})$ (if $k = 0$, we have $v = 0$ so $U(v) = 0$ belongs to any subspace).

Recall that the eigensubspace of T corresponding to an eigenvalue λ is the kernel of the operator $T - \lambda \mathrm{id}_V$. Define the *generalized eigensubspace* of T corresponding to an eigenvalue λ by

$$V(T, \lambda) = \{v \in V : (T - \lambda \mathrm{id}_V)^i(v) = 0 \quad \text{for some } i > 0\}.$$

Take $U = T - \lambda_V$ in above, we obtain the proof of the following.

Lemma 1. *$V(T, \lambda)$ is a linear susbspace of V. It is invariant with respect to T.*

Let us restrict the operator T to the invariant subspace $V(T, \lambda)$ (that is consider the same rule for T only applied to vectors from $V(T, \lambda)$). We shall exhibit a basis in $V(T, \lambda)$ such that the matrix of T with respect to this basis is the direct sum of Jordan block matrices with λ at the diagonal.

Notice that the operator $U = T - \lambda \mathrm{id}_V$ when restricted to $V(T, \lambda)$ satisfies $U^m = 0$ for some $m > 0$. In fact, every vector $v \in V(T, \lambda)$ satisfies $U^i(v) = 0$ for some $i > 0$. Choose a basis $v_1, \ldots, v_k$ in $V(T, \lambda)$ and let m be chosen such that $U^m(v_i) = 0$ for all $i = 1, \ldots, k$. This can be done since $U^j(v) = 0$ imlplies $U^s(v) = U^{s-j}(U^j(v)) = 0$ for $s \geq j$. Now writing any $v \in V(T, \lambda)$ as a linear combination of the basis, and using that U^m is a linear operator, we obtain that $U^m(v) = 0$ for all $v \in V(T, \lambda)$.

Let us consider any finite-dimensional vector space W and a linear operator $U : W \to W$ satisfying $U^m = 0$ for some $m \geq 0$ (a linear operator with such property is called a *nilpotent operator*).

Observe that

$$\mathrm{R}(U^{i+1}) \subset \mathrm{R}(U^i).$$

Indeed $U^{i+1}(v) = U^i(U(v))$, so if a vector w is equal to the value of U^{i+1} at some vector v, then it is also equal to the value of U^i at $U(v) \in V(T, \lambda)$. So we have a chain of linear subspaces

$$\{0\} = \mathrm{R}(U^m) \subset \mathrm{R}(U^{m-1}) \subset \ldots \subset \mathrm{R}(U) \subset W. \tag{4}$$

Observe that

$$U(\mathrm{R}(U^i)) = \mathrm{R}(U^{i+1}).$$

To see this, use that $U^{i+1}(v) = U(U^i(v))$, so each vector in $\mathrm{R}(U^{i+1})$ is equal to the value of U at some vector in $\mathrm{R}(U^i)$.

Lemma 2. *Let U be a nilpotent linear operator on a vector space $W \neq \{0\}$. Let m be the minimal positive integer such that $U^m = 0$ (the level of the nilpotency of U). Then all inclusions in (4) are strict.*

Proof. We use induction on m. If $m = 1$, the assertion is obvious since $U = 0\mathrm{id}_V$, hence $\mathrm{R}(U) = \{0\} \neq W$. Assume the assertion is true for linear operators with level of nilpotency $< m$. Take $W_1 = \mathrm{R}(U)$ and let $U' : W_1 \to W_1$ be the restriction of U to W_1. Then $\mathrm{R}(U^k) = \mathrm{R}(U'^{k-1})$ and we have the sequence

$$\{0\} = \mathrm{R}(U^m) \subset \mathrm{R}(U^{m-1}) \subset \ldots \subset \mathrm{R}(U) = W_1 \subset W,$$

equal to the sequence

$$\{0\} = \mathrm{R}(U'^{m-1}) \subset \mathrm{R}(U'^{m-2}) \subset \ldots \subset \mathrm{R}(U') \subset W_1 \subset W.$$

Since U is not invertible (because $U^m = 0$), we have $\dim \mathrm{R}(U) = \dim W - \dim N(U) < \dim W$. So, $\mathrm{R}(U) \neq W$. By induction, starting from W_1, all the inclusions are strict. Thus all the inclusions are strict.

Let us go back to our situation when $U = T - \lambda \mathrm{id}_V$ restricted to $V(T, \lambda)$. Let $n_1 = \dim \mathrm{R}(U^{m-1})$. Since $R(U^m) = 0$, U sends all vectors from $\mathrm{R}(U^{m-1})$ to $\{0\}$. Let $v_1^{(1)}, \ldots, v_{n_1}^{(1)}$ be a basis of this space. Since $U : \mathrm{R}(U^{m-2}) \to \mathrm{R}(U^{m-1})$ is surjective, we can find $v_1^{(2)}, \ldots, v_{n_1}^{(2)}$ in $\mathrm{R}(U^{m-2})$ with

$$U(v_i^{(2)}) = v_i^{(1)}, \quad i = 1, \ldots, n_1.$$

I claim that

$$v_1^{(1)}, \ldots, v_{n_1}^{(1)}, v_1^{(2)}, \ldots, v_{n_1}^{(2)}$$

are linearly independent. In fact, if

$$a_1 v_1^{(1)} + \ldots + a_{n_1} v_{n_1}^{(1)} + b_1 v_1^{(2)} + \ldots + b_{n_1} v_{n_1}^{(2)} = 0,$$

we apply U to obtain that

$$0 = a_1 U(v_1^{(1)}) + \ldots + a_{n_1} U(v_{n_1}^{(1)}) + b_1 U(v_1^{(2)}) + \ldots + b_{n_1} U(v_{n_1}^{(2)}) = b_1 v_1^{(1)} + \ldots + b_{n_1} v_{n_1}^{(1)}.$$

This gives $b_1 = \ldots = b_{n_1} = 0$, and hence $a_1 = \ldots = a_{n_1} = 0$. Notice that $v_1^{(1)}, \ldots, v_{n_1}^{(1)}$ belong to $\mathrm{N}(U)$, so we can find a basis of $\mathrm{N}(U) \cap \mathrm{R}(U^{m-2})$ of the form $v_1^{(1)}, \ldots, v_{n_1}^{(1)}, v_{n_1+1}^{(2)}, \ldots, v_{n_2}^{(2)}$. Together with the vectors $v_1^{(2)}, \ldots, v_{n_1}^{(2)}$ we get a basis of $\mathrm{R}(U^{m-2})$. In fact, by the formula for the dimension of the range space of a linear transformation, the dimensions of the subspaces $\mathrm{span}(v_1^{(2)}, \ldots, v_{n_1}^{(2)})$ and $\mathrm{N}(U) \cap \mathrm{R}(U^{m-2})$ add up to the dimension of $\mathrm{R}(U^{m-2})$. Also their intersection is the zero subspace $\{0\}$. In fact, if $\sum_{i=1}^{n_1} a_i v_i^{(2)} \in \mathrm{N}(U)$, applying U we get $\sum_{i=1}^{n_1} a_i U(v_i^{(2)}) = \sum_{i=1}^{n_1} a_i v_i^{(1)} = 0$, hence $a_1 = \ldots = a_{n_1} = 0$ because the vectors $v_1^{(1)}, \ldots, v_{n_1}^{(1)}$ are linearly independent. Next we find $v_1^{(3)}, \ldots, v_{n_2}^{(3)} \in \mathrm{R}(U^{m-3})$ which are mapped to $v_1^{(2)}, \ldots, v_{n_2}^{(2)}$, respectively. Then we find a basis of $\mathrm{N}(U) \cap \mathrm{R}(U^{m-3})$ which includes the previous basis $v_1^{(1)}, \ldots, v_{n_1}^{(1)}, v_{n_1+1}^{(2)}, \ldots, v_{n_2}^{(2)}$ of $\mathrm{N}(U) \cap \mathrm{R}(U^{m-2})$. The union of this basis and the set $v_1^{(2)}, \ldots, v_{n_2}^{(2)}, v_1^{(3)}, \ldots, v_{n_2}^{(3)}$ is a basis of $\mathrm{R}(U^{m-3})$. Proceeding in this way, we find a basis in V

$$\begin{array}{ccccccc}
v_1^{(1)} & \ldots & v_{n_1}^{(1)} & & & & \\
v_1^{(2)} & \ldots & v_{n_1}^{(2)} & \ldots & v_{n_2}^{(2)} & & \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \\
v_1^{(m)} & \ldots & v_{n_1}^{(m)} & \ldots & v_{n_2}^{(m)} & \ldots & v_{n_m}^{(m)}
\end{array} \tag{6}$$

satisfying the following property

(i) $n_1 \leq n_2 \leq \ldots \leq n_m$;

(ii) $n_1 + \ldots + n_i = \dim \mathrm{R}(U^{m-i})$, in particular, $n_1 + \ldots + n_m = \dim V(T, \lambda)$;

(iii) $(T - \lambda \mathrm{id}_V)(v_i^{(j+1)}) = v_i^{(j)}$ if $i = 1, \ldots, n_j$.

Let us find the matrix of $V(T,\lambda)$ of T with respect to this basis. We first reorder the vectors by taking the first m vectors from the first column in (6) starting from the top, then go to the second column and so on. Since $(T-\lambda \mathrm{id}_V)(v_1)=0, \quad (T-\lambda \mathrm{id}_V)(v_i)=v_{i-1}, i=2,\ldots,m$, we obtain

$$T(v_1)=\lambda v_1, \quad T(v_i)=\lambda v_i+v_{i-1}, \quad i=2,\ldots,m.$$

This shows that the first m columns of the matrix of T look like

$$\begin{pmatrix} \lambda & 1 & 0 & \ldots & \ldots & 0 \\ 0 & \lambda & 1 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \ldots & \ldots & 0 & \lambda & 1 \\ 0 & \ldots & \ldots & 0 & 0 & \lambda \end{pmatrix}$$

Continuing in this way we easily convince ourselves that the matrix of T in our basis is equal to the direct sum of n_1 Jordan blocks of size m, n_2-n_1 Jordan blocks of size $m-1$, and, finally, n_m-n_{m-1} Jordan blocks of size 1. All of them of course have λ at the diagonal.

Remark. *The proof shows for which eigenvectors C_1 with eigenvalue λ you can solve the equation $AC_2=\lambda C_2+C_1$. The vector C_1 must come from the range space $R(A-\lambda I_n)$ of the matrix $(A-\lambda I_n)^{m-1}$, where m is the nilpotency level of $A-\lambda I_n$.*

To finish the proof we use

Lemma 3. *Let λ be an eigenvalue of T. Then*

$$V=V(T,\lambda)\oplus W,,$$

where W is invariant with respect to T and $T-\lambda \mathrm{id}_V$ is invertible when restricted to W.

Proof. We know that $V(T,\lambda)=\mathrm{N}((T-\lambda \mathrm{id}_V)^m)$ for some $m>0$. Define $W=\mathrm{R}((T-\lambda \mathrm{id}_V)^m)$. Then the dimensions of the spaces $V(T,\lambda)$ and W add up to $\dim V$. It remains to show that $V(T,\lambda)\cap W=\{0\}$. If v is in the intersection we have $v=(T-\lambda \mathrm{id}_V)^m(w)$, for some $w\in V$, and hence $0=(T-\lambda \mathrm{id}_V)^m(v)=(T-\lambda \mathrm{id}_V)^{2m}(w)$. This implies that $w\in V(T,\lambda)$. But then $(T-\lambda \mathrm{id}_V)^m(w)=0$ and thus $v=0$.

Now we can finish the proof. Take a Jordan basis in $V(T,\lambda)$ and extend it to some basis of V. The matrix of T is the direct sum of a Jordan matrix and a matrix of T restricted to W. It is easy to see that the determinant of a block matrix is equal to the product of determinants of the blocks. This shows that the characteristic polynomial of an operator restricted to W divides the characteristic polynomial of T (this is true for any invariant subspace, see Theorem 5.21 from the book). By assumption it factors into the product of linear polynomials. By induction on dimension of the vector space, we may assume that the theorem is true for W. Since the restriction of $T-\lambda \mathrm{id}_V$ to W is inveritible, its eigenvalues are different from λ. Thus T restricted to W has a basis such that the matrix of T is the sum of Jordan blocks with no λ at the diagonal. Taking this basis and adding to this the basis for $V(T,\lambda)$ which we have just constructed we see that the matrix of T is the sum of Jordan blocks. The theorem is proven.

Problems

1. Find the Jordan form of the following matrices

$$a)\begin{pmatrix} 0 & 1 & 0 \\ -4 & 4 & 0 \\ -2 & 1 & 2 \end{pmatrix}, \quad b)\begin{pmatrix} 3 & -4 & 0 & 2 \\ 4 & -5 & -2 & 4 \\ 0 & 0 & 3 & -2 \\ 0 & 0 & 2 & -1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

2. For the matrix a) from Problem 1 compute A^{10}.

3. Prove that a matrix A is nilpotent if and only if its characteristic polynomial is equal to $(-\lambda)^n$.

4. Find the Jordan form of a matrix A satisfying $A^2 = A$.

5 Let $J = J_n(0)$. Find the Jordan form of J^2.

6. Count the number of different (up to permutation of blocks) Jordan matrices of size $n \leq 4$ with 0 at the diagonal.

7. Prove that two Jordan matrices are matrices of the same linear operator with respect to different bases if and only if one is obtained from another by permutation of its Jordan blocks.