1212: Linear Algebra II

Dr. Vladimir Dotsenko (Vlad)

Lecture 3

Let us mention one consequence of the rank-nullity theorem from the previous class.

Proposition 1. *For any operator* $\varphi\colon V \to W$*, we have* $\operatorname{rk}(\varphi) \leqslant \min(\dim(V), \dim(W))$.

Proof. We have $\operatorname{rk}(\varphi) \leqslant \dim(W)$ because $\operatorname{Im}(\varphi) \subset W$, and the dimension of a subspace cannot exceed the dimension of the whole space. Also, $\operatorname{rk}(\varphi) = \dim(V) - \operatorname{null}(\varphi) \leqslant \dim(V)$.

(Alternatively, one can argue that being the number of pivots in the reduced row echelon form, the rank cannot exceed either the number of rows or the number of columns, but this proof shows some other useful techniques, so we mention it here). □

The rank-nullity theorem and its proofs actually tells us precisely how to simplify matrices of most general linear maps $\varphi\colon V \to W$. If we allowed to change bases of V and W independently, then rank is the only invariant: every $m \times n$-matrix A can be brought to the form $\begin{pmatrix} I_k & 0_{(n-k)\times k} \\ 0_{k\times(m-k)} & 0_{(n-k)\times(m-k)} \end{pmatrix}$, where $k = \operatorname{rk}(A)$. However, if we restrict ourselves to linear transformations $\varphi\colon V \to V$, then we can only change one basis, and under the changes we replace matrices A by $C^{-1}AC$, where C is the transition matrix. We know several things that remain the same under this change, e.g. the trace and the determinant, so the story gets much more subtle.

Reminder: eigenvalues and eigenvectors

Recall that if V is a vector space, $\varphi\colon V \to V$ is a linear transformation, and $\mathbf{v} \neq 0$ a vector for which $\varphi(\mathbf{v})$ is proportional to $\mathbf{v}$, that is $\varphi(\mathbf{v}) = c \cdot \mathbf{v}$ for some c, then $\mathbf{v}$ is called an *eigenvector* of φ, and c is the corresponding *eigenvalue*.

Usually, one first finds all eigenvalues, and then the corresponding eigenvectors. If φ is represented by an $n \times n$-matrix $A = A_{\varphi,\mathbf{e}}$, then eigenvalues are roots of the equation $\det(A - XI_n) = 0$. The left hand side of this equation is a degree n polynomial in X. Indeed, each matrix element is a polynomial in X of degree 0 or 1, and each term in the expansion of the determinant is a product of n elements, one from each row/column. Thus, every term is of degree at most n. The only term of degree n is the one where all elements are of degree one, that is taken from the diagonal. Therefore, X^n appears in our equation with coefficient $(-1)^n$.

Theorem 1. *If* φ *has* n *distinct eigenvalues, then the corresponding eigenvectors form a basis of* V.

Proof. We shall first prove the following statement: if $c_1, \ldots, c_k$ are pairwise distinct scalars, and $v_1, \ldots, v_k$ are eigenvectors of φ whose eigenvalues are $c_1, \ldots, c_k$, then $v_1, \ldots, v_k$ are linearly independent. Let us prove it by induction on k. For $k = 1$ there is nothing to prove: an eigenvector is a nonzero vector by definition. Suppose we know the statement for some k, and wish to prove it for $k+1$. Assume the contrary, let's say that $a_1v_1 + \cdots + a_kv_k + a_{k+1}v_{k+1} = 0$, and some of the coefficients are nonzero. We have

$$\varphi(a_1v_1 + \cdots + a_kv_k + a_{k+1}v_{k+1}) = a_1\varphi(v_1) + \cdots + a_k\varphi(v_k) + a_{k+1}\varphi(v_{k+1}) = 0,$$

so

$$a_1c_1v_1 + \cdots + a_kc_kv_k + a_{k+1}c_{k+1}v_{k+1} = 0.$$

Let us subtract from this equation c_{k+1} times the original one, obtaining

$$0 = a_1c_1v_1 + \cdots + a_kc_kv_k + a_{k+1}c_{k+1}v_{k+1} - c_{k+1}(a_1v_1 + \cdots + a_kv_k + a_{k+1}v_{k+1}),$$

which can be rewritten as

$$a_1(c_1 - c_{k+1})v_1 + a_2(c_2 - c_{k+1})v_2 + \cdots + a_k(c_k - c_{k+1})v_k = 0.$$

By induction hypothesis, this implies

$$a_1(c_1 - c_{k+1}) = \cdots = a_k(c_k - c_{k+1}) = 0,$$

and since all c_i are distinct, we have $a_1 = \cdots = a_k = 0$, and consequently $a_{k+1}v_{k+1} = 0$, implying $a_{k+1} = 0$.

To complete the proof, we observe that because of what we just proved, a linear transformation with n distinct eigenvalue has n linearly independent eigenvectors, and they must form a basis (else they can be extended to a basis consisting of more than $\dim(V)$ vectors). □

If eigenvalues are not distinct, everything can become more complicated.

Example 1. Let $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. We have $\det(A - XI_2) = (1 - X)^2$, so the only eigenvalue is 1. We have $Av = v$ for every vector v, so any basis is a basis of eigenvectors.

Example 2. Let $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. We have $\det(A - XI_2) = X^2$, so the only eigenvalue is 0. Moreover, all eigenvectors are scalar multiples of $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, so in this case there is no basis of eigenvectors.

Example 3. Let $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. We have $\det(A - XI_2) = X^2 + 1$, so if we use real scalars, then there are no eigenvalues, and if scalars are complex, the eigenvalues are $\pm i$, and there is a basis consisting of eigenvectors.

If $v_1, \ldots, v_n$ is a basis of eigenvectors for φ, then by inspection we conclude that φ is represented by a diagonal matrix with respect to that basis (all off-diagonal entries are zero). For this reason, a linear operator is said to be *diagonalisable* if it has a basis of eigenvectors.

One example of diagonalisation was discussed in the first half of the module in the context of Fibonacci numbers, Lecture 22. There have been further applications outlined in Lectures 23 and 24. Our goal in this semester is to investigate what happens in the cases when there is no basis of eigenvectors.

Unless otherwise specified, we shall now use complex numbers as scalars, to at least ensure that each linear transformation has eigenvalues. Indeed, eigenvalues, as we discussed before, are roots of the *characteristic polynomial* of A, $\chi_A(X) = \det(A - XI_n)$, and over complex numbers, every polynomial equation has a root.

Let us discuss a few more examples of diagonalisability.

Example 4. Consider $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$. Then $A - XI_3 = \begin{pmatrix} -X & 1 & 0 \\ 0 & -X & 1 \\ 1 & 0 & -X \end{pmatrix}$, and $\chi_A(X) = 1 - X^3$. The complex roots of this polynomial are complex roots of unity, 1 and $\frac{-1 \pm i\sqrt{3}}{2}$. Since they are distinct, there is a basis of eigenvectors for A.

Example 5. Let us consider the matrix $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 14 & -23 & 10 \end{pmatrix}$. The characteristic polynomial of this matrix is $14 - 23X + 10X^2 - X^3$. We note that $c = 1$ is a root of this polynomial, so it is divisible by $X - 1$. Doing the long division

$$\begin{array}{r} -X^2 \quad +9X-14 \\ X-1\overline{)\,-X^3+10X^2-23X+14} \\ \underline{X^3 \quad -X^2} \qquad\qquad \\ 9X^2-23X \quad\;\; \\ \underline{-9X^2 \quad +9X} \quad\;\; \\ -14X+14 \\ \underline{14X-14} \\ 0 \end{array}$$

we observe that the other roots are roots of $-X^2+9X-14=-(X-2)(X-7)$. Therefore, the eigenvalues of this matrix are 1, 2, and 7. By the theorem we proved, there exists a basis of eigenvectors.