MA1112: Linear Algebra II

Dr. Vladimir Dotsenko (Vlad)

Lecture 6

Several commuting linear transformations have a common eigenvector

Let us illustrate how invariant subspaces can help to reduce a problem to a "smaller" one. Two linear transformations φ and ψ are said to *commute* if $\varphi \circ \psi = \psi \circ \varphi$, so that the result of consecutive application of φ and ψ does not depend on the order in which they are applied.

Theorem 1. *Any set of pairwise commuting transformations $\varphi_i \colon V \to V$ has a common eigenvector.*

Before we proceed to the proof, let us make a remark on the meaning of this result. The mathematical apparatus of Quantum Mechanics postulates that various physical quantities of particles (coordinates, momenta, etc.) are linear transformations of a certain space (Hilbert Space); vectors in that space correspond to different states of the given physical system. An eigenvector corresponds to a "pure" state where the quantity is exactly measurable, and any state is a mixture of pure states with various probabilities. Our result shows that commuting transformations correspond to quantities that can be exactly measured simultaneously. One of the famous examples in Quantum Mechanics is the Uncertainly Principle that claims that coordinates and momenta cannot be exactly measured simultaneously; this corresponds to the fact that certain linear transformations do not commute.

Proof. We shall prove it by induction on $\dim(V)$. If $\dim(V) = 1$, then any basis vector of V is a common eigenvector for these operators. Assume the statement is proved for $\dim(V) = k$. Let us prove it for $\dim(V) = k + 1$. If all the operators φ_i are scalar multiples of the identity map, that is there exist scalars c_i such that for all v and all i we have $\varphi_i(v) = c_i \cdot v$, then every non-zero vector is a common eigenvector of these transformations. Suppose that for some i the operator φ_i is not a scalar multiple of the identity map. Let us consider some eigenvalue λ of φ_i, and consider the solution space to the system of equations $\varphi_i(v) = \lambda \cdot v$. This solution space is a subspace W with $0 < \dim(W) < \dim(V)$. Let us note that W is an invariant subspace of all our transformations: if $w \in W$, and $w' = \varphi_k(w)$, we have

$$\varphi_i(w') = \varphi_i(\varphi_k(w)) = \varphi_k(\varphi_i(w)) = \varphi_k(\lambda \cdot w) = \lambda\varphi_k(w) = \lambda \cdot w',$$

so $w' \in W$. By induction, there is a common eigenvector in W, as required. □

Case $\varphi^2 = \varphi$

Let us assume that a linear transformation φ satisfies the equation $\varphi^2 = \varphi$, so that for every vector v, we have $\varphi(\varphi(v)) = \varphi(v)$. Note that this implies that all eigenvalues of φ are equal to 0 and 1: if $\varphi(v) = \lambda v$, then $\varphi(\varphi(v)) = \varphi(\lambda v) = \lambda\varphi(v) = \lambda^2 v$, and we see that $\lambda^2 = \lambda$. Therefore, such a linear transformation usually does not have distinct eigenvalues. Nevertheless, we shall see that it is always possible to find a basis where the matrix of φ is diagonal.

Lemma 1. *If* $\varphi^2 = \varphi$, *then* $\operatorname{Im}(\varphi) \cap \ker(\varphi) = \{0\}$.

Proof. Indeed, if $v \in \operatorname{Im}(\varphi) \cap \ker(\varphi)$, then $v = \varphi(w)$ for some w, and $0 = \varphi(w) = \varphi(\varphi(w)) = \varphi(w) = v$. □

Note that from this proof it is also clear that if $v \in \operatorname{Im}(\varphi)$, then $\varphi(v) = v$.

Lemma 2. *If* $\varphi^2 = \varphi$, *then* $V = \mathrm{Im}(\varphi) \oplus \ker(\varphi)$.

Proof. Indeed,

$$\dim(\mathrm{Im}(\varphi) + \ker(\varphi)) = \dim \mathrm{Im}(\varphi) + \dim \ker(\varphi) - \dim(\mathrm{Im}(\varphi) \cap \ker(\varphi)) = \mathrm{rk}(\varphi) + \mathrm{null}(\varphi) = \dim(V),$$

so the sum is a subspace of V of dimension equal to the dimension of V, that is V itself. Also, we already checked that the intersection is equal to 0, so the sum is direct. $\square$

Consequently, if we take a basis of $\ker(\varphi)$, and a basis of $\mathrm{Im}(\varphi)$, and join them together, we get a basis of V with respect to which the matrix of φ is the diagonal matrix $\begin{pmatrix} 0 & 0 \\ 0 & I_m \end{pmatrix}$, where $m = \mathrm{rk}(\varphi)$.

Case $\varphi^2 = 0$

However nice the approach from the previous section seems, sometimes it does not work that well. Though we always have

$$\dim \mathrm{Im}(\varphi) + \dim \ker(\varphi) = \dim(V),$$

the sum of these subspaces is not always direct, as the following example shows. If we know that $\varphi^2 = 0$, that is $\varphi(\varphi(v)) = 0$ for every $v \in V$, that implies $\mathrm{Im}(\varphi) \subset \ker(\varphi)$, so $\mathrm{Im}(\varphi) + \ker(\varphi) = \ker(\varphi)$. Let us discuss a way to handle this case, it will be very informative for our future results. We begin with a general definition which will be useful for packaging various constructions we shall use.

We start by picking a basis $e_1, \ldots, e_k$ of V relative to $\ker(\varphi)$. Note that $\varphi(e_1), \ldots, \varphi(e_k) \in \mathrm{Im}(\varphi) \subset \ker(\varphi)$. Let us pick a basis $f_1, \ldots, f_l$ of $\ker(\varphi)$ relative to $\mathrm{span}(\varphi(e_1), \ldots, \varphi(e_k))$. Let us show that the vectors

$$\varphi(e_1), \ldots, \varphi(e_k), f_1, \ldots, f_l$$

are linearly independent (and hence form a basis of $\ker(\varphi)$). Suppose that

$$b_1\varphi(e_1) + \cdots + b_k\varphi(e_k) + c_1 f_1 + \cdots + c_l f_l = 0.$$

Since $f_1, \ldots, f_l$ is a relative basis, we conclude that $c_1, \ldots, c_l$ are all equal to zero. Therefore,

$$b_1\varphi(e_1) + \cdots + b_k\varphi(e_k) = \varphi(b_1 e_1 + \cdots + b_k e_k) = 0,$$

so $b_1 e_1 + \cdots + b_k e_k \in \ker(\varphi)$. Since these vectors form a relative basis, we conclude that $b_1, \ldots, b_k$ are all equal to zero.

Thus

$$\varphi(e_1), \ldots, \varphi(e_k), f_1, \ldots, f_l$$

form a basis of $\ker(\varphi)$, which immediately implies that the vectors

$$e_1, \ldots, e_k, \varphi(e_1), \ldots, \varphi(e_k), f_1, \ldots, f_l$$

form a basis of V.

Reordering this basis, we obtain a basis

$$e_1, \varphi(e_1), \ldots, e_k, \varphi(e_k), f_1, \ldots, f_l,$$

relative to which the matrix of φ has a block diagonal form with k blocks $\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ on the diagonal, and all the other entries equal to zero.