Matrices of linear transformations

In order to perform calculations about a linear transformation $L\colon V \to W$ one chooses bases $S = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ and $T = \{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n\}$ of the spaces V and W respectively. Then we have the isomorphism $I_S\colon V \to \mathbb{R}^n$ taking a vector $\mathbf{v}$ to the coordinate vector $[\mathbf{v}]_S$ with respect to basis S. Here n is the dimension of V. Similarly, one has the isomorphism $I_T\colon W \to \mathbb{R}^m$, where $m = \dim W$. Thus we have a diagram of transformations:

$$
\begin{array}{ccc}
V & \xrightarrow{\;L\;} & W \\
\Big\downarrow{\scriptstyle I_S} & & \Big\downarrow{\scriptstyle I_T} \\
\mathbb{R}^n & & \mathbb{R}^m
\end{array} \;.
$$

Since I_S is an isomorphism, there is an inverse isomorphism $I_S^{-1}\colon \mathbb{R}^n \to V$. Specifically, it assigns to an n-vector $\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ the vector $x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n$ in V. Then the composition of linear transformations I_S^{-1}, L and I_T defines the linear transformation $I_T \circ L \circ I_S^{-1}\colon \mathbb{R}^n \to \mathbb{R}^m$. [Why is the composition of linear transformations a linear transformation?] Since any transformation of $\mathbb{R}^n$ into $\mathbb{R}^m$ is a matrix one, we have $I_T \circ L \circ I_S^{-1}(\mathbf{x}) = A\mathbf{x}$ for some $m \times n$ matrix A. In other words, the diagram above can be completed:

$$
\begin{array}{ccc}
V & \xrightarrow{\;L\;} & W \\
\Big\downarrow{\scriptstyle I_S} & & \Big\downarrow{\scriptstyle I_T} \\
\mathbb{R}^n & \xrightarrow{\;A\;} & \mathbb{R}^m
\end{array} \;.
\tag{1}
$$

[Here we regard the matrix A as the transformation $\mathbf{x} \mapsto A\mathbf{x}$.] This diagram means that $I_T \circ L = A \circ I_S$, or equivalently, for any vector $\mathbf{v}$ in V:

$$
[L(\mathbf{v})]_T = A[\mathbf{v}]_S.
$$

Matrix A is called the *matrix associated with L and bases S and T*.

Example 1. Consider the derivation $L\colon \mathrm{Pol}(2) \to \mathrm{Pol}(1)$, $L(f) = f'$. Let's find the matrix A associated with L and bases $S = \{t^2, t, 1\}$, $T = \{t, 1\}$.

We have the isomorphisms $I_S\colon \mathrm{Pol}(2) \to \mathbb{R}^3$ and $I_T\colon \mathrm{Pol}(1) \to \mathbb{R}^2$. The first column of A is the 3-vector $A\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$. Since $\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = I_S(t^2)$, the first column of A is $AI_S(t^2)$. By (1), the latter equals $I_T L(t^2) = I_T((t^2)') = I_T(2t) = [2t]_T = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$. Similarly, the second column of A is $A\begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} = AI_S(t) = I_T L(t) = I_T(1) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, and the third column is zero because $I_T L(1) = I_T(0) = \mathbf{0}$. Thus

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

We can use A to find values of L. Indeed, by (1), $L(f) = I_T^{-1} A I_S(f)$. For example, by this formula $L(3t^2 - 2t + 1) = I_T^{-1} A [3t^2 - 2t + 1]_S = I_T^{-1} A \begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix} = I_T^{-1} \begin{bmatrix} 6 \\ -2 \end{bmatrix} = 6t - 2$, which agrees with the straightforward computation $L(3t^2 - 2t + 1) = (3t^2 - 2t + 1)' = 6t - 2$.

If we choose other bases S', T' in V, W, the matrix A' associated with L will be different, in general. Take $S' = \{t, 1, t^2\}$ and $T' = \{t + 1, t - 1\}$. As before, the the first column of A' is

$$I_T L(t) = I_T(1) = \begin{bmatrix} 0.5 \\ -0.5 \end{bmatrix},$$

because $1 = 0.5(t + 1) - 0.5(t - 1)$. The second column of A is

$$I_T L(1) = I_T(0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

and the third:

$$I_T L(t^2) = I_T(2t) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

as $2t = (t + 1) + (t - 1)$. Hence

$$A' = \begin{bmatrix} 0.5 & 0 & 1 \\ -0.5 & 0 & 1 \end{bmatrix}.$$

Using A', verify that $L(3t^2 - 2t + 1) = 6t - 2$.

The question arises: is there a formula relating A and A' (where A' is the matrix associated with L and bases S', T')? We have two diagrams glued:

$$
\begin{array}{ccc}
\mathbb{R}^n & \xrightarrow{A'} & \mathbb{R}^m \\
{\scriptstyle I_{S'}} \uparrow & & \uparrow {\scriptstyle I_{T'}} \\
V & \xrightarrow{L} & W \\
{\scriptstyle I_S} \downarrow & & \downarrow {\scriptstyle I_T} \\
\mathbb{R}^n & \xrightarrow{A} & \mathbb{R}^m
\end{array}
\quad .
$$

Recall that $[\mathbf{x}]_S = P_{SS'}[\mathbf{x}]_{S'}$, where $\mathbf{x}$ is a vector in $\mathbb{R}^n$, $P_{SS'}$ is the transition matrix from the basis S' to S. In other words, $I_S = P_{SS'} I_{S'}$. Similarly, $I_T = P_{TT'} I_{T'}$. Thus we can add additional arrows to our diagram:

$$
\begin{array}{ccc}
 & \mathbb{R}^n \xrightarrow{A'} \mathbb{R}^m & \\
P_{SS'} \Big(& {\scriptstyle I_{S'}}\uparrow \quad \uparrow{\scriptstyle I_{T'}} & \Big) P_{TT'} \\
 & V \xrightarrow{L} W & \\
 & {\scriptstyle I_S}\downarrow \quad \downarrow{\scriptstyle I_T} & \\
 & \mathbb{R}^n \xrightarrow{A} \mathbb{R}^m &
\end{array}
\quad . \tag{2}
$$

By the boundary diagram, we have $P_{TT'} A' = A P_{SS'}$, or equivalently,

$$A' = P_{TT'}^{-1} A P_{SS'}. \tag{3}$$

As an exercise, verify formula (2) for the example above. [First, show that

$$P_{SS'} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \quad P_{TT'} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.]$$

An important case is $V = W$. Then $n = m$ and A is a square matrix. If we choose $S = T$, then we simply say that A is the matrix associated with L and basis S. If $S' = T'$, then (3) takes form:

$$A' = P^{-1}AP, \tag{3}$$

where P is the transition matrix from S' to S. Square matrices A and A' related by formula (3) for some invertible matrix P are called *similar*. We've just seen that matrices associated with operator L with respect to different bases are similar. On the other hand, suppose $n \times n$ matrices A and A' are related by the formula (3) for some P. Consider the transformation $L \colon \mathbb{R}^n \to \mathbb{R}^n$, $L(\mathbf{x}) = A\mathbf{x}$. Then A is the matrix associated with L and the standard basis $S = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ of $\mathbb{R}^n$ and A' is the matrix associated with L and the basis $S' = \{P(\mathbf{e}_1), P(\mathbf{e}_2), \ldots, P(\mathbf{e}_n)\}$ (why). Thus we proved

Theorem 0.1. *Two $n \times n$ matrices are similar if and only if they are matrices associated with the same linear operator $L \colon V \to V$ (and some bases S and S' of V).*

Lec 36: Kernel and range of a linear transformation

Let $L \colon V \to W$ be a linear transformation. The *kernel* of L is the set of all vectors $\mathbf{v}$ in V such that $L(\mathbf{v}) = \mathbf{0}$. It is denoted by $\operatorname{Ker} L$. The *range* $\operatorname{Im} L$ of L is the set of all $\mathbf{w}$ in W such that $\mathbf{w} = L(\mathbf{v})$ for some $\mathbf{v}$ in V. Verify that $\operatorname{Ker} L$ and $\operatorname{Im} L$ are subspaces in V and W respectively.

Consider the derivation $L \colon \operatorname{Pol}(2) \to \operatorname{Pol}(2)$, $L(f) = f'$. Its kernel consists of all $ax^2 + bx + c$ with zero derivative, i. e. only of constant polynomials c. Hence $\operatorname{Ker} L = \operatorname{Span} 1$. We also have $\operatorname{Im} L = \operatorname{Pol}(1)$. Indeed, derivation lessens the degree of a polynomial, that is why the range can have the linear (and constant) polynomials only. On the other hand, every linear polynomial is a derivative of a quadratic one: $ax + b = (0.5ax^2 + bx)'$. Note that $\dim \operatorname{Ker} L = 1$ and $\dim \operatorname{Im} L = 2$.

Example 2. Find the kernel and the range of the linear transformation $L \colon \mathbb{R}^3 \to \mathbb{R}^2$ given by $L(\begin{bmatrix} x \\ y \\ z \end{bmatrix}) = \begin{bmatrix} x+2z \\ 3z-y \end{bmatrix}$.

For $\mathbf{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ we have $L(\mathbf{x}) = A\mathbf{x}$, where

$$A = \begin{bmatrix} 1 & 0 & 2 \\ 0 & -1 & 3 \end{bmatrix}.$$

Hence the kernel of L is the nullspace of A, i. e. the space of solutions to $A\mathbf{x} = \mathbf{0}$. The RREF of A is

$$\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & -3 \end{bmatrix},$$

3

and we have z free and $x = -2z$, $y = 3z$. Then $\mathbf{x} = \begin{bmatrix} -2 \\ 3 \\ 1 \end{bmatrix} z$ and $\operatorname{Ker} L = \operatorname{Span}\{\begin{bmatrix} -2 \\ 3 \\ 1 \end{bmatrix}\}$.

The range of L is the column space of A (why?), i. e. $\operatorname{Im} L = \operatorname{Span}\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \end{bmatrix}\}$. Then one can easily see that $\operatorname{Im} L = \mathbb{R}^2$.

In these examples the following statement holds.

Theorem 0.2. $\dim \operatorname{Ker} L + \dim \operatorname{Im} L = \dim V$.

Let's prove this theorem. Choose bases S and T in V and W respectively. Then we have the diagram (1). Note that if $\mathbf{v}$ belongs to the kernel of L, then $\mathbf{x} = I_S(\mathbf{v}) = [v]_S$ belongs to the nullspace $\operatorname{Nullspace} A$ of A. Indeed, $A\mathbf{x} = AI_S(\mathbf{v}) = I_T L(\mathbf{v}) = I_T(\mathbf{0}) = \mathbf{0}$. Conversely, if $\mathbf{x}$ belongs to the nullspace of A, $\mathbf{v} = I_S^{-1}(\mathbf{x})$ is in $\operatorname{Ker} L$: $L(\mathbf{v}) = LI_S^{-1}(\mathbf{x}) = I_T^{-1}A(\mathbf{x}) = I_T^{-1}(\mathbf{0}) = \mathbf{0}$. This means that $\operatorname{Ker} L$ and $\operatorname{Null} A$ are isomorphic (by I_S). In particular, $\dim \operatorname{Ker} L = \dim \operatorname{Nullspace} A = \operatorname{Null} A$ (nullity of A). Verify, that the ranges of L and A are isomorphic (by I_T). Note that the range of A is the span of $A\mathbf{e}_1, \ldots, A\mathbf{e}_n$, where $\mathbf{e}_1, \ldots, \mathbf{e}_n$ is a basis in $\mathbb{R}^n$ (why?). Taking standard basis, this span is exactly the column space of A, and its dimension is, by definition, $\operatorname{rank} A$. Hence $\dim \operatorname{Im} L = \operatorname{rank} A$. Thus the formula in the theorem turns to

$$\operatorname{Null} A + \operatorname{rank} A = n,$$

which is known. The theorem is proved.

As a consequence of the correspondence between kernels of operators and nullspaces of their matrices, ranges and column spaces (see the reasoning above), and the theorem (1), we obtain:

Theorem 0.3. *Similar matrices have the same nullity and rank.*

Lec 37: Eigenvalues and eigenvectors of linear operators.

Let $L\colon V \to V$ be a linear operator. A real number λ is called an *eigenvalue* of L, if $L(\mathbf{v}) = \lambda\mathbf{v}$ for some nonzero vector $\mathbf{v}$. Then $\mathbf{v}$ is called an *eigenvector* associated with λ. [Note that all eigenvectors must be nonzero.] The set of all eigenvectors associated with an eigenvalue λ, plus the zero vector, is called the *eigenspace* associated with λ, and is denoted by $V(\lambda)$. One can show (do it!) that $V(\lambda)$ is a subspace of V. In particular one can find a basis in $V(\lambda)$ and $\dim V(\lambda)$. Eigenspaces corresponding to different eigenvalues intersect only by $\mathbf{0}$. This is because $\lambda\mathbf{v} = \mu\mathbf{v}$ implies $\lambda = \mu$ for a nonzero $\mathbf{v}$. Moreover, one can show that eigenvectors associated with different eigenvalues are linearly independent (try to do this).

The scaling operator $L\colon V \to V$, $L(\mathbf{v}) = 2\mathbf{v}$, has the only eigenvalue, which is 2. Indeed, $V(2) = V$. For more elaborated examples we need to develop some machinery. First, consider the case of matrix operators $L\colon \mathbb{R}^n \to \mathbb{R}^n$, $L(\mathbf{x}) = A\mathbf{x}$ for an $n \times n$ matrix A. We define the *eigenvalues and eigenvectors of the matrix A* to be those of the operator L.

So, λ is an eigenvalue of A if

$$A\mathbf{x} = \lambda\mathbf{x}$$

for some nonzero n-vector $\mathbf{x}$. We can rewrite this as

$$(A - \lambda I_n)\mathbf{x} = \mathbf{0}.$$

This is equivalent to singularity of $A - \lambda I_n$, i. e.

$$\det(A - \lambda I_n) = 0.$$

If we regard λ as variable, the left-hand side of the equation above is a polynomial $p(\lambda)$ of degree n. This is called the *characteristic polynomial* of A. We proved

Theorem 0.4. *The eigenvalues of a matrix A are exactly the roots of its characteristic polynomial $p(\lambda)$. The eigenspace $\mathbb{R}^n(\lambda)$ is the nullspace of the matrix $A - \lambda I_n$.*

Example 3. If A is a diagonal matrix with the diagonal entries $d_1, d_2, \ldots, d_n$, then $p(\lambda) = (d_1 - \lambda)(d_2 - \lambda) \cdots (d_n - \lambda)$. Hence the eigenvalues of A are exactly the diagonal entries of A. The eigenspace $\mathbb{R}^n(d_i)$ is spanned by the ith standard vector $\mathbf{e}_i$.

Example 4. Let

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Then

$$A - \lambda I_2 = \begin{bmatrix} -\lambda & 1 \\ 1 & -\lambda \end{bmatrix}$$

and $p(\lambda) = \det(A - \lambda I_2) = \lambda^2 - 1 = (\lambda - 1)(\lambda + 1)$. Then the eigenvalues of A are 1 and -1. Find a basis in $\mathbb{R}^2(1)$. According to the theorem above, it must be a basis of the nullspace of

$$A - 1 \cdot I_2 = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}.$$

Verify that the vector $\mathbf{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ represents a basis of the nullspace. Similarly, check that $\mathbb{R}^2(-1)$ is spanned by $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

Example 5. Let

$$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

We have $p(\lambda) = \lambda^2 + 1$. Since p does not have real roots, there are no eigenvalues (and eigenvectors) of A.

Example 6. Let

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

We have $p(\lambda) = (1 - \lambda)^2$. Then the only eigenvalue is 1. Verify that $\mathbb{R}^2(1)$ is one-dimensional with a basis $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

Most examples are 2×2 matrices since otherwise $p(\lambda)$ has degree 3 or more, and the roots of such a polynomial may not be easy to find. In some some special cases it is not hard though. Try to find eigenvalues and bases in eigenspaces for the matrix

$$\begin{bmatrix} 0 & 0 & 9 \\ 0 & 2 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

Example 7. Let

$$A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}.$$

We have $p(\lambda) = -\lambda(1 - \lambda)^2$, and the eigenvalues are 0 and 1. Check that $\mathbb{R}^3(0) = \text{Span}\{\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}\}$ and $\mathbb{R}^3(1) = \text{Span}\{\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\}$.

Now, let's see what happens for $L: V \to V$. We can choose a basis S in V and consider the matrix A associated with L and S. Then the eigenvalues of L coincide with those of A. Indeed, by the diagram

$$\begin{array}{ccc} V & \xrightarrow{L} & V \\ {\scriptstyle I_S}\downarrow & & \downarrow{\scriptstyle I_S} \\ \mathbb{R}^n & \xrightarrow{A} & \mathbb{R}^m \end{array} \quad,$$

if $L(\mathbf{v}) = \lambda\mathbf{v}$, then for $\mathbf{x} = I_S(\mathbf{v}) = [\mathbf{v}]_S$ one has: $A(\mathbf{x}) = AI_S(\mathbf{v}) = I_S L(\mathbf{v}) = I_S(\lambda\mathbf{v}) = \lambda I_S(\mathbf{v}) = \lambda\mathbf{x}$. So, if λ is an eigenvalue of L, then it is an eigenvalue of A. And the converse is true (why?). The correspondence $\mathbf{v} \leftrightarrow \mathbf{x}$ above is actually an isomorphism between the eigenspaces $V(\lambda)$ and $\mathbb{R}^n(\lambda)$ of L and A respectively.

Example 8. Let $L: \text{Pol}(2) \to \text{Pol}(2)$ be given by $L(at^2 + bt + c) = ct^2 + (b + 2c)t + c$. Find the eigenvalues of L and the corresponding eigenspaces. In the basis $S = \{t^2, t, 1\}$ the associated matrix is

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

We have $p(\lambda) = -\lambda(1 - \lambda)^2$. Then the eigenvalues of A are 0 and 1. Solving the corresponding homogeneous systems, we get $\mathbb{R}^3(0) = \text{Span}\{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\}$ and $\mathbb{R}^3(0) = \text{Span}\{\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\}$. Applying the isomorphism I_S^{-1}, we obtain that $V(0)$ is spanned by t^2 and $V(1)$ — by t, where $V = \text{Pol}(2)$.

It follows from Theorem 1 and the discussion above that

Theorem 0.5. *Similar matrices have the same eigenvalues.*

Another way to prove this is to show that the characteristic polynomials $p(\lambda)$ and $q(\lambda)$ of similar matrices A and B coincide. In fact, if $A = PBP^{-1}$, then $p(\lambda) = \det(A - \lambda I_n) = \det(PBP^{-1} - \lambda I_n) = \det(PBP^{-1} - P(\lambda I_n)P^{-1}) = \det(P(B - \lambda I_n)P^{-1}) = \det(P)\det(B - \lambda I_n)\det(P)^{-1} = \det(B - \lambda I_n) = q(\lambda)$.

Lec 38: Diagonalizable operators.

An important class of linear operators is that of diagonalizable ones (in fact, they are a majority of all operators). A linear operator $L\colon V \to V$ is called *diagonalizable*, if it's matrix in some basis is a diagonal one. A square matrix A is called *diagonalizable*, if it is similar to a diagonal matrix.

Theorem 0.6. *Let $L\colon V \to V$ be a linear operator and A its associated matrix with respect to a basis S of V. The following statements are equivalent:*

(1) L is diagonalizable;

(2) A is diagonalizable;

(3) there is a basis of V consisting of eigenvectors for A;

(4) there is a basis of V consisting of eigenvectors for L.

Proof. If L is diagonalizable, there is a basis E of V for which the associated matrix D is diagonal. Then if $P = P_{SE}$ is the transition matrix from E to S, we have by formula (3): $D = P^{-1}AP$ (or $A = PDP^{-1}$). Hence A and D are similar, which means that A is diagonalizable.

Now let A be diagonalizable. Then for some P one has $P^{-1}AP = D$ — a diagonal matrix. By example 3, the standard basis $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ of $\mathbb{R}^n$ consists of eigenvectors for D. Then $\{P\mathbf{e}_1, \ldots, P\mathbf{e}_n\}$ is a basis (since P is invertible). Moreover, its vectors are eigenvectors for A. Indeed, if $D\mathbf{e}_i = \lambda_i \varepsilon_i$, then $A(P\mathbf{e}_i) = (AP)\mathbf{e}_i = (PD)\mathbf{e}_i = P(\lambda_i \mathbf{e}_i) = \lambda_i(P\mathbf{e}_i)$.

The rest implications (3) $\Rightarrow$ (4) and (4) $\Rightarrow$ (1) are left as exercises. $\qquad\square$

Using statement (3) of the theorem, one can show that in examples 3,4 and 7 the matrices are diagonalizable. Say, in example 4, A is similar to the diagonal matrix $D = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$: $A = P^{-1}DP$, where the transition matrix P has corresponding eigenvectors as columns: $P = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$. One could also take $P = \begin{bmatrix} 2 & 2 \\ 3 & -3 \end{bmatrix}$. In examples 5,6,8 the matrices (and linear operators) are not diagonalizable (why?).