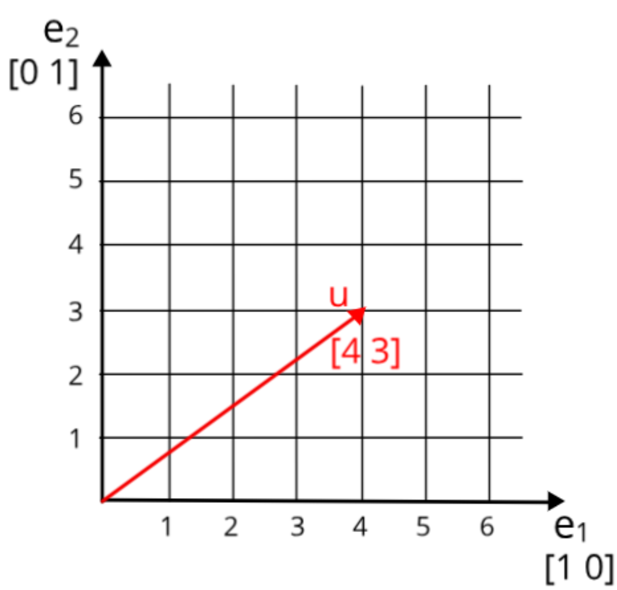
EECS 16A Designing Information Devices and Systems I
Spring 2019 Lecture Notes Note 10

10.1 Change of Basis for Vectors

Previously, we have seen that matrices can be interpreted as *linear transformations* between vector spaces. In particular, an $m \times n$ matrix A can be viewed as a function $A : U \to V$ mapping a vector $\vec{u}$ from vector space $U \in \mathbb{R}^n$ to a vector $A\vec{u}$ in vector space $V \in \mathbb{R}^m$. In this note, we explore a different interpretation of square, invertible matrices: as a *change of basis*.

Let's first start with an example. Consider the vector $\vec{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$. When we write a vector in this form, implicitly we are representing it in the *standard basis* for $\mathbb{R}^2$, $\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\vec{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. This means that we can write $\vec{u} = 4\vec{e}_1 + 3\vec{e}_2$. Geometrically, $\vec{e}_1$ and $\vec{e}_2$ defines a grid in $\mathbb{R}^2$, and $\vec{u}$ is represented by the coordinates in the grid, as shown in the figure below:

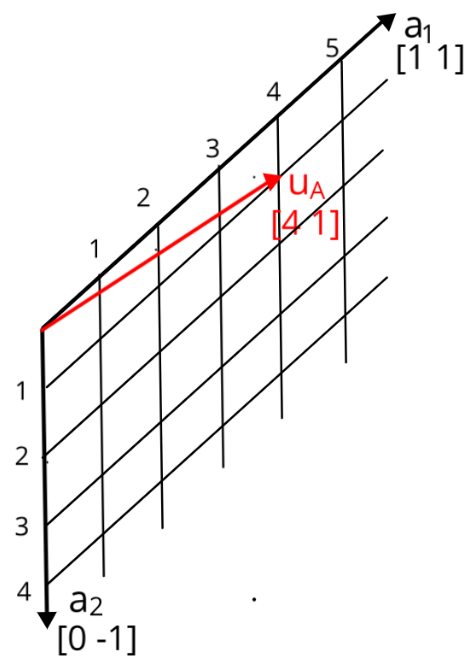


What if we want to represent $\vec{u}$ as a linear combination of another set of basis vectors, say $\vec{a}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\vec{a}_2 = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$? This means that we need to find scalars u_{a_1} and u_{a_2} such that $\vec{u} = u_{a_1}\vec{a}_1 + u_{a_2}\vec{a}_2$. We can write this equation in matrix form:

$$\begin{bmatrix} | & | \\ \vec{a}_1 & \vec{a}_2 \\ | & | \end{bmatrix} \begin{bmatrix} u_{a_1} \\ u_{a_2} \end{bmatrix} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} u_{a_1} \\ u_{a_2} \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$$

Thus we can find u_{a_1} and u_{a_2} by solving a system of linear equations. Since the inverse of $\begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix}$ is $\begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix}$, we get $u_{a_1} = 4$ and $u_{a_2} = 1$. Then we can write $\vec{u}$ as $4\vec{a}_1 + \vec{a}_2$. Geometrically, $\vec{a}_1$ and $\vec{a}_2$ defines a skewed grid from which the new coordinates are computed.



From this example, we can see that the same vector $\vec{u}$ can be represented in multiple ways! In the standard basis $\vec{e}_1, \vec{e}_2$, the coordinates for $\vec{u}$ are $4, 3$. In the skewed basis $\vec{a}_1, \vec{a}_2$, the coordinates for $\vec{u}$ are $4, 1$. It's the same vector geometrically, but with different coordinates.

In general, suppose we are given a vector $\vec{u} \in \mathbb{R}^n$ in the standard basis and want to change to a different basis with linearly independent basis vectors $\vec{a}_1, \cdots, \vec{a}_n$. If we denote the vector in the new basis as $\vec{u}_a = \begin{bmatrix} u_{a_1} \\ \vdots \\ u_{a_n} \end{bmatrix}$, we solve the following equation $A\vec{u}_a = \vec{u}$, where A is the matrix $\begin{bmatrix} \vec{a}_1 & \cdots & \vec{a}_n \end{bmatrix}$. Therefore the change of basis is given by:

$$\vec{u}_a = A^{-1}\vec{u}$$

If we already have a vector $\vec{u}_a$ in the basis $\vec{a}_1, \cdots, \vec{a}_n$, how do we change it back to a vector $\vec{u}$ in the standard basis? We can reverse the change of basis transformation, thus $\vec{u} = A\vec{u}_a$.

Pictorially, the relationship between any two bases and the standard basis is given by:

$$\text{Basis } \vec{a}_1, \cdots, \vec{a}_n \underset{A^{-1}}{\overset{A}{\rightleftarrows}} \text{Standard Basis} \underset{B}{\overset{B^{-1}}{\rightleftarrows}} \text{Basis } \vec{b}_1, \cdots, \vec{b}_n$$

For example, consider a vector in the basis of $\vec{a}_1, \cdots, \vec{a}_n$:

$$\vec{u} = u_{a_1}\vec{a}_1 + \cdots + u_{a_n}\vec{a}_n = A\vec{u}_a$$

Suppose we would like to represent it as a linear combination of different basis vectors $\vec{b}_1, \cdots, \vec{b}_n$. In other words, we'd like to write it as:

$$\vec{u} = u_{b_1}\vec{b}_1 + \cdots + u_{b_n}\vec{b}_n = B\vec{u}_b$$

Setting these equal gives the following relationship:

$$\begin{aligned} B\vec{u}_b = u_{b_1}\vec{b}_1 + \cdots + u_{b_n}\vec{b}_n = \vec{u} &= u_{a_1}\vec{a}_1 + \cdots + u_{a_n}\vec{a}_n = A\vec{u}_a \\ B\vec{u}_b &= A\vec{u}_a \\ \vec{u}_b &= B^{-1}A\vec{u}_a \end{aligned}$$

Thus the change of basis transformation from basis $\vec{a}_1, \cdots, \vec{a}_n$ to basis $\vec{b}_1, \cdots, \vec{b}_n$ is given by $B^{-1}A$.

10.2 Change of Basis for Linear Transformations

Now that we know how to change the basis of vectors, let's shift our attention to linear transformations. We will answer these questions in this section: how do we change the basis of linear transformations and what does this mean? First let's review linear transformations. Suppose we have a linear transformation T represented by a $n \times n$ matrix that transforms $\vec{u} \in \mathbb{R}^n$ to $v \in \mathbb{R}^n$:

$$\vec{v} = T\vec{u}$$

Although T is represented by a matrix, we can think of it as a geometric transformation on vectors. Implicit in this representation is the choice of a coordinate system. Unless stated otherwise, we always assume that vectors lie in the coordinate system defined by the standard basis vectors $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$.

But what if our vectors represented in a different basis? Suppose we have basis vectors $\vec{a}_1, \cdots, \vec{a}_n \in \mathbb{R}^n$, and the vectors $\vec{u}$, $\vec{v}$ above are represented in this basis:

$$\begin{aligned} \vec{u} &= u_{a_1}\vec{a}_1 + \cdots + u_{a_n}\vec{a}_n \\ \vec{v} &= v_{a_1}\vec{a}_1 + \cdots + v_{a_n}\vec{a}_n \end{aligned}$$

Can we also represent the transformation T above, such that $\vec{v}_a = T_a\vec{u}_a$? Let's start by defining the matrix A:

$$A = \begin{bmatrix} | & & | \\ \vec{a}_1 & \cdots & \vec{a}_n \\ | & & | \end{bmatrix}$$

Since our vectors $\vec{u}$ and $\vec{v}$ are linear combinations of the basis vectors we can write them as $\vec{u} = A\vec{u}_a$ and $\vec{v} = A\vec{v}_a$. This is exactly the formula defined in the previous section to change a vector back to the standard basis! Now, starting with our transformation T in the standard basis, we can plug in these relationships:

$$\begin{aligned} T\vec{u} &= \vec{v} \\ TA\vec{u}_a &= A\vec{v}_a \\ A^{-1}TA\vec{u}_a &= \vec{v}_a \end{aligned}$$

By pattern matching, we see that if we set $T_a = A^{-1}TA$, we get our desired relationship, $T_a\vec{u}_a = \vec{v}_a$.

The correspondences stated above are all represented in the following diagram:

$$\begin{tikzcd}
\vec{u} \arrow[r, mapsto, "T"] \arrow[d, mapsto, shift right, "A^{-1}"'] & \vec{v} \arrow[d, mapsto, shift left, "A^{-1}"] \\
\vec{u}_a \arrow[r, mapsto, "T_a"'] \arrow[u, mapsto, shift right, "A"'] & \vec{v}_a \arrow[u, mapsto, shift left, "A"]
\end{tikzcd}$$

For example, there are two ways to get from $\vec{u}_a$ to $\vec{v}_a$. First is the transformation T_a. Second, we can trace out the longer path, applying transformations A, T and A^{-1} in order. This is represented in matrix form as $T_a = A^{-1}TA$ (Recall that matrices act from right to left: the closest matrix to the vector will be applied first.)

By the same logic, we can go in the reverse direction: $T = AT_aA^{-1}$. Let's look at whats happing at each step. First, A^{-1} transforms a vector in the standard basis into the A basis. Then, T_a then acts on it, returning a new vector which also in the A basis. Finally, multiplying by A returns a weighted sum of the columns of A, in other words, transforming the vector back into the original standard basis.

10.3 A Diagonalizing Basis

We just saw that transformations are different in different bases. Is there a special basis under which a transformation attains a particularly nice form? Making our transformation into a diagonal matrix is very useful since diagonal matrices are easily invertible, they can be raised to high powers without a lot of computation, and multiplication between diagonal matrices is commutative.

Suppose that we choose our basis vectors $\vec{a}_1, \cdots, \vec{a}_n$ to be the eigenvectors of the transformation matrix T, which have associated eigenvalues $\lambda_1, \cdots, \lambda_n$. What does $T\vec{u}$ look like now? Recall that $\vec{u}$ can be written in the new basis: $u_{a_1}\vec{a}_1 + \cdots + u_{a_n}\vec{a}_n$.

$$\begin{aligned} T\vec{u} &= T(u_{a_1}\vec{a}_1 + \cdots + u_{a_n}\vec{a}_n) \\ &= u_{a_1}T\vec{a}_1 + \cdots + u_{a_n}T\vec{a}_n \end{aligned}$$

By the definition of eigenvectors and eigenvalues:

$$
\begin{aligned}
&= u_{a_1}\lambda_1\vec{a}_1 + \cdots + u_{a_n}\lambda_n\vec{a}_n \\
&= \begin{bmatrix} | & \cdots & | \\ \vec{a}_1 & \cdots & \vec{a}_n \\ | & \cdots & | \end{bmatrix} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \begin{bmatrix} u_{a_1} \\ \vdots \\ u_{a_n} \end{bmatrix} \\
&= AD\vec{u}_a \\
&= ADA^{-1}\vec{u},
\end{aligned}
$$

where D is the diagonal matrix of eigenvalues and A is a matrix with the corresponding eigenvectors as its columns. Thus we have shown that in an eigenvector basis, $T = ADA^{-1}$. In particular, T_a, the counterpart of T in the eigenvector basis, is a diagonal matrix.

How is transforming T into a diagonal matrix useful? Say we want apply our transformation T many times. We might be interested in calculating T^k where k could be quite large. Although you could do this calculation with repeated matrix multiplication, it would quite tedious and even with a computer is slow for large matrices.

However, we just showed that we can write $T = ADA^{-1}$ where D is diagonal. Consider raising this to the power of k:

$$T^k = (ADA^{-1})^k$$

We can expand this out k times and then regroup the terms:

$$
\begin{aligned}
T^k =& (ADA^{-1})(ADA^{-1})\dots(ADA^{-1})(ADA^{-1}) \\
=& AD(A^{-1}A)D(A^{-1}\dots A)D(A^{-1}A)DA^{-1} \\
=& AD(I)D(I\dots)D(I)DA^{-1} \\
=& ADD\dots DDA^{-1}) \\
=& AD^kA^{-1}
\end{aligned}
$$

To raise a diagonal matrix to a power, one can raise each element to that power (try proving this to yourself by writing out the multiplication element by element). This is a much simpler calculation that repeatedly doing the full matrix multiplication! In addition, this formulation better allows us to write equations describing how the transformation changes with k – having equations will help us better understand and design systems.

10.4 Diagonalization

Under what circumstances is a matrix diagonalizable? **A $n \times n$ matrix T is diagonalizable if it has n linearly independent eigenvectors**. If it has n linearly independent eigenvectors $\vec{a}_1, \cdots, \vec{a}_n$ with eigenvalues $\lambda_1, \cdots \lambda_n$, then we can write:

$$T = ADA^{-1}$$

Where $A = \begin{bmatrix} \vec{a}_1 & \cdots & \vec{a}_n \end{bmatrix}$ and D is a diagonal matrix of eigenvalues. Note that the eigenvalues must be arranged to match the arrangement of eignevectors in A. For example, the eigenvalue associated with $\vec{a}_1$

should be in the first column on D, the eigenvalue associated with $\vec{a}_2$ should be in the second column on D, and so on.

Example 10.1 (A matrix that is not diagonalizable): Let $T = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$. To find the eigenvalues, we solve the equation:

$$\begin{aligned} \det(T - \lambda I) &= 0 \\ (1-\lambda)^2 &= 0 \\ \lambda &= 1 \end{aligned}$$

The eigenvector corresponding to $\lambda = 1$ is $\vec{a} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Since this matrix only has 1 eigenvector but exists in $\mathbb{R}^{2\times 2}$, it is not diagonalizable.

Example 10.2 (Diagonalization of a 3×3 matrix):

$$T = \begin{bmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{bmatrix}$$

To find the eigenvalues, we solve the equation:

$$\begin{aligned} \det(T - \lambda I) &= 0 \\ \lambda^3 - 6\lambda^2 - 15\lambda - 8 &= 0 \\ (\lambda - 8)(\lambda + 1)^2 &= 0 \\ \lambda &= -1, 8 \end{aligned}$$

If $\lambda = -1$, we need to find $\vec{a}$ such that:

$$(T - (-1)I)\vec{a} = 0 \quad \Longrightarrow \quad \begin{bmatrix} 4 & 2 & 4 \\ 2 & 1 & 2 \\ 4 & 2 & 4 \end{bmatrix} \vec{a} = 0 \quad \Longrightarrow \quad \begin{bmatrix} 2 & 1 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \vec{a} = 0$$

Thus the dimension of the nullspace of $T - (-1)I$ is 2, and we can find two linearly independent vectors in this basis:

$$\vec{a}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \quad \vec{a}_2 = \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix}$$

If $\lambda = 8$, we need to find $\vec{a}$ such that:

$$(T - (8)I)\vec{a} = 0 \quad \Longrightarrow \quad \begin{bmatrix} -5 & 2 & 4 \\ 2 & -8 & 2 \\ 4 & 2 & -5 \end{bmatrix} \vec{a} = 0 \quad \Longrightarrow \quad \begin{bmatrix} 1 & 0 & -1 \\ 0 & 2 & -1 \\ 0 & 0 & 0 \end{bmatrix} \vec{a} = 0$$

Thus the dimension of the nullspace of $T-(8)I$ is 1, and we find the vector in the nullspace:

$$\vec{a}_3 = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}$$

Now we define:

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 0 & -2 & 1 \\ -1 & 0 & 2 \end{bmatrix} \quad D = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 8 \end{bmatrix} \quad A^{-1} = \frac{1}{9}\begin{bmatrix} 4 & 2 & -5 \\ 1 & -4 & 1 \\ 2 & 1 & 2 \end{bmatrix}$$

Then T can be diagonalized as $T = ADA^{-1}$.

> **Additional Resources** For more on diagonalization, read *Strang* pages 304 - 309. In *Schuam's*, read pages 299 - 301 and try Problems 9.9 to 9.21, and 9.45 to 9.51.

10.5 A Proof with Diagonalization

Not only can diagonalization help us raise matrices to powers, it can also help us understand some underlying properties of matrices. In this example, we use diagonalization to understand when matrices are commutative.

Show that if two $n \times n$ diagonalizable matrices A and B have the same eigenvectors, then their matrix multiplication is commutative (i.e. $AB = BA$)

Let $\vec{v}_1, \ldots, \vec{v}_n$ be the eigenvectors of A and B, and $P = [\vec{v}_1, \ldots, \vec{v}_n]$ be a matrix constructed from the eigenvectors of A and B. Also let $\lambda_{a1}, \ldots, \lambda_{an}$ be the eigenvalues of A, and $\lambda_{b1}, \ldots, \lambda_{bn}$ be the eigenvalues of B.

Using diagonalization, we can decompose A and B such that $A = PD_AP^{-1}$ and $B = PD_BP^{-1}$, where D_A is the diagonal matrix of A, and D_B is the diagonal matrix of B:

$$D_A = \begin{bmatrix} \lambda_{a1} & & \\ & \ddots & \\ & & \lambda_{an} \end{bmatrix} \qquad D_B = \begin{bmatrix} \lambda_{b1} & & \\ & \ddots & \\ & & \lambda_{bn} \end{bmatrix}$$

Now we can write

$$\begin{aligned} AB &= (PD_AP^{-1})(PD_BP^{-1}) \\ &= PD_A(P^{-1}P)D_BP^{-1} \\ &= PD_A(I)D_BP^{-1} \\ &= PD_AD_BP^{-1} \end{aligned}$$

Because D_A and D_B are both diagonal matrices, $D_AD_B = D_BD_A$. To see why this is the case, let's look at it element by element:

$$D_A D_B = \begin{bmatrix} \lambda_{a1} & & \\ & \ddots & \\ & & \lambda_{an} \end{bmatrix} \begin{bmatrix} \lambda_{b1} & & \\ & \ddots & \\ & & \lambda_{bn} \end{bmatrix} \tag{1}$$

$$= \begin{bmatrix} \lambda_{a1}\lambda_{b1} & & \\ & \ddots & \\ & & \lambda_{an}\lambda_{bn} \end{bmatrix} \tag{2}$$

$$= \begin{bmatrix} \lambda_{b1}\lambda_{a1} & & \\ & \ddots & \\ & & \lambda_{bn}\lambda_{an} \end{bmatrix} \tag{3}$$

$$= \begin{bmatrix} \lambda_{b1} & & \\ & \ddots & \\ & & \lambda_{bn} \end{bmatrix} \begin{bmatrix} \lambda_{a1} & & \\ & \ddots & \\ & & \lambda_{an} \end{bmatrix} = D_B D_A \tag{4}$$

Thus:

$$\begin{aligned} PD_A D_B P^{-1} &= PD_B D_A P^{-1} \\ &= PD_B I D_A P^{-1} \\ &= (PD_B P^{-1})(PD_A P^{-1}) \\ &= BA \end{aligned}$$

We've now shown that A and B are commutative if they have the same eigenvectors. We can use this to design matrices that commute!

10.6 Practice Problems

These practice problems are also available in an interactive form on the course website (http://inst.eecs.berkeley.edu/~ee16a/sp19/hw-practice/).

1. If $\vec{v} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ is in the standard basis, what is $\vec{v}$ in the basis of $\vec{a}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \vec{a}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \vec{a}_3 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$?

2. True or False: A 3×3 matrix with the following eigenvectors is diagonalizable: $\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ 4 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$.

3. Which of the following matrices is equivalent to $\begin{bmatrix} 1 & 0 \\ 0.5 & 0.5 \end{bmatrix}^k$?

(a) $\begin{bmatrix} 1 & 0 \\ 0 & 0.5^k \end{bmatrix}$

(b) $\begin{bmatrix} 1 & 0 \\ 1 - 0.5^k & 0.5^k \end{bmatrix}$

(c) $\begin{bmatrix} 0.5^k & 0 \\ 0.5^k - 1 & 1 \end{bmatrix}$

(d) $\begin{bmatrix} 1 & 0 \\ 1 + 0.5^k & 0.5^k \end{bmatrix}$