EECS 16A Designing Information Devices and Systems I
Fall 2016 Official Lecture Notes **Note 19**

Introduction

This note introduces some of the tools needed in the following notes when we analyze signals and linear time-invariant systems in EE16B. In particular, it introduces the critical idea behind the Fourier Transform, which is vital for understanding not just wireless communication, but the general operation of linear time-invariant systems.

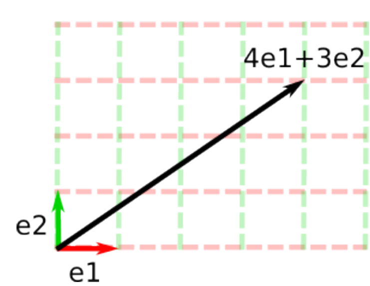
By the end of these notes, you should be able to:

1. Change a vector from one basis to another.
2. Take a matrix representation for a linear transformation in one basis and express that linear transformation in another basis.
3. Understand the importance of a diagonalizing basis and its properties.
4. Identify if a matrix is diagonalizable and if so, to diagonalize it.

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:

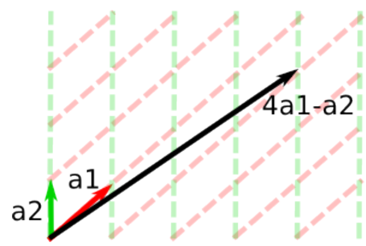


What if we want to represent $\vec{u}$ as a linear combination of another set of basis vectors, say $\vec{a}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\vec{a}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$? This means that we need to find scalars u_{a_1} and u_{a_2} such that $\vec{u} = u_{a_1}\vec{a}_1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



Notice the important thing here. 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 $4, -1$. Same vector geometrically, but 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 $\vec{u}_a = \begin{bmatrix} u_{a_1} \\ u_{a_2} \end{bmatrix}$, 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:

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

Now we know how to change a vector from the standard basis to any basis, how do we change a vector $\vec{u}_a$ in the basis $\vec{a}_1, \cdots, \vec{a}_n$ back to a vector $\vec{u}$ in the standard basis? We simply 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, given a vector $\vec{u} = u_{a_1}\vec{a}_1 + \cdots + u_{a_n}\vec{a}_n$ represented as a linear combination of the basis vectors $\vec{a}_1, \cdots, \vec{a}_n$, we can represent it as a different linear combination of basis vectors $\vec{b}_1, \cdots, \vec{b}_n$, $\vec{u} = u_{b_1}\vec{b}_1 + \cdots + u_{b_n}\vec{b}_n$ by writing:

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

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 $B^{-1}A$.

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 $v \in \mathbb{R}^n$:

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

The important thing to notice is that this linear transformation is a geometric thing: it maps vectors to vectors in a linear manner. It also happens to be represented as a matrix. Implicit in this representation is the choice of a coordinate system. Unless stated otherwise, we always assume that the coordinate system is that defined by the standard basis vectors $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$.

What is this transformation T in another basis? Suppose we have basis vectors $\vec{a}_1, \cdots, \vec{a}_n \in \mathbb{R}^n$. If $\vec{u}$ is represented by $\vec{u}_a = u_{a_1}\vec{a}_1 + \cdots + u_{a_n}\vec{a}_n$ and $\vec{v}$ is represented by $\vec{v}_a = v_{a_1}\vec{a}_1 + \cdots + v_{a_n}\vec{a}_n$ in this basis, what is the transformation T represented by in this basis, such that $T_a\vec{u}_a = \vec{v}_a$? If we define the matrix $A = \begin{bmatrix} \vec{a}_1 & \cdots & \vec{a}_n \end{bmatrix}$, we can write our vectors $\vec{u}$ and $\vec{v}$ as a linear combination of the basis vectors: $\vec{u} = A\vec{u}_a$ and $\vec{v} = A\vec{v}_a$. This is exactly the change of basis from the standard basis to the $\vec{a}_1, \cdots, \vec{a}_n$ basis.

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

Since we want $T_a\vec{u}_a = \vec{v}_a$, $T_a = A^{-1}TA$. In fact, 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 $T_a = A^{-1}TA$ (note that the order is reversed since this matrix acts on a vector $\vec{u}_a$ on the right).

By the same logic, we can go in the reverse direction. And then $T = AT_aA^{-1}$. Notice that this makes sense. A^{-1} takes a vector written in the standard basis and returns the coordinates in the A basis. T_a then acts on it and returns new coordinates in the A basis. Multiplying by A returns a weighted sum of the columns of A, in other words, giving us the vector back in the original standard basis.

A Diagonalizing Basis

Now we know what a linear transformation looks like under a different basis, is there a special basis under which the transformation attains a nice form? Let's suppose that the basis vectors $\vec{a}_1, \cdots, \vec{a}_n$ are eigenvectors of the transformation matrix T, with eigenvalues $\lambda_1, \cdots, \lambda_n$. What does $T\vec{u}$ looks like now? Recall that $\vec{u}$ can be written in the new basis: $u_{a_1}\vec{a}_1 + \cdots + u_{a_n}\vec{a}_n$.

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

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

This gives us many useful properties. Recall that matrix multiplication does not commute in general, but diagonal matrices commute. Can we construct other non-diagonal matrices that commute under multiplication? Suppose T_1 and T_2 can both be expressed in the same diagonal basis, that is they have the same eigenvectors but possibly different eigenvalues. Then $T_1 = AD_1A^{-1}$ and $T_2 = AD_2A^{-1}$. We show that T_1 and T_2 commute:

$$\begin{aligned}
T_1T_2 &= (AD_1A^{-1})(AD_2A^{-1}) \\
&= AD_1D_2A^{-1} \\
&= AD_2D_1A^{-1} \\
&= (AD_2)(A^{-1}A)(D_1A^{-1}) \\
&= (AD_2A^{-1})(AD_1A^{-1}) \\
&= T_2T_1
\end{aligned}$$

One of the other great properties has to do with raising a matrix to a power. If we can diagonalize a matrix, then this is really easy. Can you see for yourself why the argument above tells us that if $T = ADA^{-1}$, then $T^k = AD^kA^{-1}$? (Hint: all the terms in the middle cancel.) But raising a diagonal matrix to a power is just raising its individual elements to that same power.

Diagonalization

We saw from the previous section the usefulness of representing a matrix (i.e. a linear transformation) in a basis so that it is diagonal, so under what circumstances is a matrix diagonalizable? Recall from before that **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.

Example 19.1 (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, it is not diagonalizable.

Example 19.2 (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:

$$\begin{bmatrix} 4 & 2 & 4 \\ 2 & 1 & 2 \\ 4 & 2 & 4 \end{bmatrix} \vec{a} = 0 \quad \Longrightarrow \quad \begin{bmatrix} 2 & 1 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \vec{a} = 0$$

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:

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

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 $T = ADA^{-1}$.