These are brief notes for the lecture on Wednesday October 27, 2010: they are not complete, but they are a guide to what I want to say today. They are not guaranteed to be correct.

0.1. Eigenvectors and Eigenvalues

In this chapter we are going to focus on $n \times n$ matrices A. One very important aspect of square matrices is that when we interpret them as representing a linear transformation, $T : \mathbb{R}^n \to \mathbb{R}^n$ with $T : \underline{x} \mapsto A\underline{x}$, the transformation takes $\mathbb{R}^n$ to itself.

When we map a space to itself, we can ask some extra questions: for example, which points, if any, are fixed by T: that is, for which $\underline{x} \in \mathbb{R}^n$ is $T(\underline{x}) = \underline{x}$? Such values are called *fixed points.*

Another thing we can do with transformations from a space to itself is to iterate the transformation: that is compute

$$\underline{x},\ T(\underline{x}),\ T(T(\underline{x})),\ T(T(T(\underline{x}))), \ldots$$

We will examine what we say about these things for linear transformations.

Our first question might be: which fixed points does a matrix have? However, most linear transformations corresponding to matrices have none. One reason for this is that matrices such as, for example,

$$\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \text{ and } \begin{pmatrix} 1/2 & 0 \\ 0 & 1/3 \end{pmatrix}$$

have the effect of either stretching or shrinking vectors (and some matrices do both) and that some matrices, for example,

$$\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

have even more complicated behaviour. We will handle the stretching or shrinking question here (and if we work with complex numbers instead of real numbers) we can deal with the latter example too.

Consider the matrix

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$

and the corresponding transformation $T : \underline{x} \mapsto A\underline{x}$, for various vectors $\underline{x}$.

- $\underline{x} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$: $T(\underline{x}) =$

Of course, any linear transformation takes the zero vector to the zero vector!

- $\underline{x} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$: $T(\underline{x}) =$

- $\underline{x} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$: $T(\underline{x}) =$

Of course, for any linear transformation, if we know $T(\underline{x})$ then we know $T(c\underline{x}) = cT(\underline{x})$!

- $\underline{x} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$: $T(\underline{x}) =$

- $\underline{x} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$: $T(\underline{x}) =$

- $\underline{x} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$: $T(\underline{x}) =$

- $\underline{x} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$: $T(\underline{x}) =$

Of these, let's focus on the vectors $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$.

Each of these vectors has the property that $A\underline{x}$ is parallel to $\underline{x}$, that is, there is a real number λ (depending on which vector we pick) so that $A\underline{x} = \lambda\underline{x}$. (Aside: λ is the Greek letter lambda, corresponding to our letter l).

Definition. *Let A be an $n \times n$ matrix. A* non-zero *vector $\underline{x} \in \mathbb{R}^n$ is called an* eigenvector *of A if there exists some scalar $\lambda \in \mathbb{R}$ so that $A\underline{x} = \lambda\underline{x}$. If $\underline{x}$ is an eigenvector of A, the corresponding value λ is called an eigenvalue of A, and we say that λ is an eigenvalue of A with eigenvector $\underline{x}$.*

Note: while an eigenvector $\underline{x}$ must be non-zero (so that we are always excluding the trivial case $A\underline{0} = \underline{0}$) it is possible for the value λ to be zero. For example: $\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$ has eigenvector $\begin{pmatrix} 2 \\ -1 \end{pmatrix}$ with eigenvalue 0.

Note: If $\underline{x}$ is an eigenvector of A with eigenvalue λ, and if $c \in \mathbb{R}$ is a non-zero scalar, then $c\underline{x}$ is also an eigenvector with eigenvalue λ: however, since each is a scalar multiple of the other, we don't really get any new information from the two of them than from one.

Example:

$$\begin{pmatrix} 1 & 6 \\ 5 & 2 \end{pmatrix} \begin{pmatrix} 6 \\ -5 \end{pmatrix} = \begin{pmatrix} -24 \\ 20 \end{pmatrix} = -4 \begin{pmatrix} 6 \\ -5 \end{pmatrix}.$$

Thus we see that $\begin{pmatrix} 6 \\ -5 \end{pmatrix}$ is an eigenvector of this matrix, and -4 is the corresponding eigenvalue. This illustrates the following method: to show that a non-zero vector $\underline{x}$ is an eigenvector of A, compute $A\underline{x}$, and check that it is a scalar multiple of $\underline{x}$.

To show that λ is an eigenvalue of A we have to work a little harder. We have to find a solution $\underline{x}$ to the equation $A\underline{x} = \lambda\underline{x}$. Let's do an example: let $\lambda = 7$ and

$$A = \begin{pmatrix} 2 & 4 \\ 5 & 3 \end{pmatrix}.$$

To show that λ is an eigenvalue of A we need to find $\underline{x}$ so that $A\underline{x} = \lambda\underline{x}$. This is equivalent to solving

$$\begin{pmatrix} 2 & 4 \\ 5 & 3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 7\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

or

$$\begin{pmatrix} 2-7 & 4 \\ 5 & 3-7 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

that is to say,

$$\begin{pmatrix} -5 & 4 \\ 5 & -4 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Clearly $x_1 = 4, x_2 = 5$ is a solution to this, with $\underline{x} \neq \underline{0}$. Hence $\lambda = 7$ *is* an eigenvalue of A.

Note: For any scalar λ:

$$A\underline{x} = \lambda\underline{x} \quad \iff \quad A\underline{x} - \lambda I\underline{x} = \underline{0} \quad \iff \quad \underline{x} \in \text{ Nul}(A - \lambda I)$$

DEFINITION. *If* $\dim(Nul(A - \lambda I)) > 0$ *(that is, if there are non-zero vectors in the null space) then* $Nul(A - \lambda I)$ *is called the* eigenspace *for* A *corresponding to the eigenvalue* λ*, since any* $\underline{x} \in Nul(A - \lambda I)$ *satisfies* $A\underline{x} = \lambda\underline{x}$*.*

Note: This means that all non-zero vectors in the eigenspace are eigenvectors of A with eigenvalue λ.

Example: Find a basis for the eigenspace corresponding to the eigenvector 2 for the matrix

$$A = \begin{pmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{pmatrix}.$$

THEOREM 1. *The eigenvalues of an upper triangular matrix (or of a lower triangular matrix) are its diagonal entries.*

Proof:

Note:

$$
\begin{aligned}
&\text{The number } 0 \text{ is an eigenvalue of } A \\
\iff &\text{ there exists } \underline{x} \neq \underline{0} \text{ so that } A\underline{x} = 0\underline{x} \\
\iff &\text{ there exists } \underline{x} \neq \underline{0} \text{ so that } A\underline{x} - 0\underline{x} = \underline{0}. \\
\iff &\text{ there exists } \underline{x} \neq \underline{0} \in \text{Nul}(A) \\
\iff &\ A \text{ is not invertible.}
\end{aligned}
$$

Eigenvectors with the same eigenvalue (together with the zero vector) form a subspace. In some sense, they behave similarly to each other. To continue the analogy, eigenvectors having different eigenvalues must be somehow very different from each other.

For example, if $\underline{v}_1, \underline{v}_2$ are eigenvectors of A with eigenvalues λ_1, λ_2 respectively, and $\lambda_1 \neq \lambda_2$, then $\underline{v}_1, \underline{v}_2$ cannot be scalar multiples of each other. Indeed if

$$\underline{v}_1 = c\underline{v}_2$$

then

$$A\underline{v}_1 = Ac\underline{v}_2 \implies \lambda_1\underline{v}_1 = c\lambda_2\underline{v}_2$$

so $c\lambda_1 = c\lambda_2$, so $c = 0$. But then $\underline{v}_1 = \underline{0}$, so it wasn't an eigenvector in the first place!

The following (often very useful) theorem is a typical rephrasing of that idea:

THEOREM 2. *If $\underline{v}_1, \ldots, \underline{v}_r$ are eigenvectors corresponding to distinct eigenvalues $\lambda_1, \ldots, \lambda_r$ of an $n \times n$ matrix A, then the set $\{\underline{v}_1, \ldots, \underline{v}_r\}$*

Proof: Suppose that the vectors $\underline{v}_1, \ldots, \underline{v}_r$ are eigenvectors with distinct eigenvalues $\lambda_1, \ldots, \lambda_r$, and that they are linearly dependent. Then we know that there is a p so that

$$\underline{v}_{p+1} = c_1\underline{v}_1 + \cdots + c_p\underline{v}_p$$

and $\underline{v}_1, \ldots, \underline{v}_p$ are linearly independent. Now multiply both sides of this equation by A: