These are brief notes for the lecture on Monday October 25, 2010: they are not complete, but they are a guide to what I want to say today. They are not guaranteed to be correct.

4.7. Change of Basis

When we have a vector space V with a basis $\mathcal{B} = \{\underline{b}_1, \underline{b}_2, \ldots, \underline{b}_n\}$ specified, we know that every vector $\underline{x} \in V$ can be expressed uniquely as a linear combination of $\mathcal{B}$: that is, it is represented as the coordinate vector $[\underline{x}]_{\mathcal{B}}$. Recall that the entries in the coordinate vector are the coefficients of the linear combination.

Sometimes there is more than one natural basis to use for a vector space: if we have two bases $\mathcal{B}$ and $\mathcal{C}$, then each vector $\underline{x}$ can be expressed in terms of $\mathcal{B}$ as $[\underline{x}]_{\mathcal{B}}$, and in terms of $\mathcal{C}$ as $[\underline{x}]_{\mathcal{C}}$. Now, clearly these two representations are related, and either one determines the other: we now discuss how to transform from one to the other.

Example: Suppose that we have two bases $\mathcal{B} = \{\underline{b}_1, \underline{b}_2\}$ and $\mathcal{C} = \{\underline{c}_1, \underline{c}_2\}$. Since $\mathcal{C}$ is a basis, we can express V in terms of $\mathcal{C}$: suppose that we know

$$\underline{b}_1 = 4\underline{c}_1 + 3\underline{c}_2 \qquad \text{and} \qquad \underline{b}_1 = 2\underline{c}_1 + \underline{c}_2.$$

If we are given a vector $\underline{x}$ in terms of its $\mathcal{B}$-coordinates $[\underline{x}]_{\mathcal{B}}$, we can then determine $[\underline{x}]_{\mathcal{C}}$ as follows.

Suppose $[\underline{x}]_{\mathcal{B}} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$, that is,

$$\underline{x} = 3\underline{b}_1 + \underline{b}_2.$$

We can then substitute the expressions for $\underline{b}_1$ and $\underline{b}_2$ to obtain

$$\begin{aligned} \underline{x} &= 3(4\underline{c}_1 + 3\underline{c}_2) + (2\underline{c}_1 + \underline{c}_2) \\ &= 14\underline{c}_1 + 10\underline{c}_2 \end{aligned}$$

Hence we obtain $[\underline{x}]_{\mathcal{C}} = \begin{pmatrix} 14 \\ 10 \end{pmatrix}$.

Let's redo the problem from a slightly different perspective: we wish to obtain $[\underline{x}]_{\mathcal{C}}$, and we know $\underline{x} = 3\underline{b}_1 + \underline{b}_2$. Hence we know

$$[\underline{x}]_{\mathcal{C}} = [3\underline{b}_1]_{\mathcal{C}} + [\underline{b}_2]_{\mathcal{C}}$$

(since an equation about vectors remains true when we express it relative to coordinates!) We can express this in matrix form by forming a matrix whose columns are $[\underline{b}_1]_{\mathcal{C}}, [\underline{b}_1]_{\mathcal{C}}$:

$$[\underline{x}]_{\mathcal{C}} = [\ [\underline{b}_1]_{\mathcal{C}}\ [\underline{b}_2]_{\mathcal{C}}] \begin{pmatrix} 3 \\ 1 \end{pmatrix}.$$

But we know the columns of the matrix are

$$[\underline{b}_1]_{\mathcal{C}} = \begin{pmatrix} 4 \\ 3 \end{pmatrix} \qquad \text{and} \qquad [\underline{b}_2]_{\mathcal{C}} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

so we have

$$[\underline{x}]_{\mathcal{C}} = \begin{pmatrix} 4 & 2 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 14 \\ 10 \end{pmatrix}.$$

We can write this method up systematically: we obtain the following theorem.

THEOREM 1. *Let* $\mathcal{B} = \{\underline{b}_1, \ldots, \underline{b}_n\}$ *and* $\mathcal{C} = \{\underline{c}_1, \ldots, \underline{c}_n\}$ *be bases of a vector space* V*. Then there is a unique* $n \times n$ *matrix* $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ *so that*

$$[\underline{x}]_{\mathcal{C}} = \underset{\mathcal{C}\leftarrow\mathcal{B}}{P} \, [\underline{x}]_{\mathcal{B}}.$$

The columns of $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ *are the* $\mathcal{C}$*-coordinate vectors of the vectors in the basis* $\mathcal{B}$*. That is,*

$$\underset{\mathcal{C}\leftarrow\mathcal{B}}{P} = \left[\ [\underline{b}_1]_{\mathcal{C}}\ [\underline{b}_2]_{\mathcal{C}}\ \ldots\ [\underline{b}_n]_{\mathcal{C}}\ \right].$$

The matrix $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ is called the **change-of-coordinates matrix from $\mathcal{B}$ to $\mathcal{C}$**. Multiplication by $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ converts $\mathcal{B}$-coordinates to $\mathcal{C}$-coordinates.

The columns of $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ are linearly independent since they are the coordinates relative to $\mathcal{C}$ of the linearly independent set $\mathcal{B}$. Hence $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ is invertible. So, if we multiply both sides of the equation by $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}^{-1}$ we get a formula expressing $\mathcal{B}$-coordinates in terms of $\mathcal{C}$-coordinates.

$$\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}^{-1}[\underline{x}]_{\mathcal{C}} = [\underline{x}]_{\mathcal{B}}.$$

This means that the inverse of the change of basis matrix from $\mathcal{B}$ to $\mathcal{C}$ is the change of basis matrix from $\mathcal{C}$ to $\mathcal{B}$.

$$\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}^{-1} = \underset{\mathcal{B}\leftarrow\mathcal{C}}{P}$$

Working in $\mathbb{R}^n$

Suppose that $V = \mathbb{R}^n$, with bases $\mathcal{B}$ and $\mathcal{C}$, in addition to the standard basis $\mathcal{E} = \{\underline{e}_1, \ldots, \underline{e}_n\}$. Now, we have change of basis matrices

$$P_{\mathcal{B}} = \underset{\mathcal{E}\leftarrow\mathcal{B}}{P} = \left[\ [\underline{b}_1]_{\mathcal{E}}\ [\underline{b}_2]_{\mathcal{E}} \cdots\ [\underline{b}_n]_{\mathcal{E}}\ \right]$$

and

$$P_{\mathcal{C}} = \underset{\mathcal{E}\leftarrow\mathcal{C}}{P} = \left[\ [\underline{c}_1]_{\mathcal{E}}\ [\underline{c}_2]_{\mathcal{E}} \cdots\ [\underline{c}_n]_{\mathcal{E}}\ \right]$$

so that

$$[\underline{x}]_{\mathcal{E}} = P_{\mathcal{B}}\ [\underline{x}]_{\mathcal{B}} \qquad \text{and} \qquad [\underline{x}]_{\mathcal{E}} = P_{\mathcal{C}}\ [\underline{x}]_{\mathcal{C}}.$$

Hence,

$$[\underline{x}]_{\mathcal{C}} = P_{\mathcal{C}}^{-1}\ [\underline{x}]_{\mathcal{E}} = P_{\mathcal{C}}^{-1}\ P_{\mathcal{B}}[\underline{x}]_{\mathcal{B}}$$

which implies that

$$\underset{\mathcal{C}\leftarrow\mathcal{B}}{P} = P_{\mathcal{C}}^{-1}\ P_{\mathcal{B}}.$$

So, when we are working with different bases in $\mathbb{R}^n$, we can compute change of basis matrices, to change from coordinates with respect to $\mathcal{C}$ to coordinates with respect to $\mathcal{B}$, by working just with the matrices having columns coming from the representations with respect to the standard basis.

A computational note: to compute $P_{\mathcal{C}}^{-1}\ P_{\mathcal{B}}$, the best method is *not* to compute $P_{\mathcal{C}}^{-1}$ and then multiply: rather, it is quicker to append the matrices $P_{\mathcal{C}}P_{\mathcal{B}}$ in a $n \times 2n$ matrix, and row reduce until the first n columns give the identity. The remaining n columns will contain $P_{\mathcal{C}}^{-1}\ P_{\mathcal{B}}$.

CHAPTER 5

Eigenvalues and Eigenvectors

5.1. 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.