Eigenvectors: Similarity and Bases
Lay 5.4

November 11, 2013

1 Bases and coordinate vectors

Remember that if $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ is a basis for $\mathbb{R}^n$, then for each $\mathbf{x} \in \mathbb{R}^n$, we can write $\mathbf{x} = c_1\mathbf{b}_1 + \ldots + c_n\mathbf{b}_n$ for some unique set of scalars $c_1, \ldots, c_n$. We call the vector whose ith entry is c_i the $\mathcal{B}$-coordinate vector $[\mathbf{x}]_\mathcal{B}$.

Example 1.1. The set $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$, where

$$\mathbf{b}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad \mathbf{b}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

is a basis for $\mathbb{R}^2$. If

$$\mathbf{x} = \begin{bmatrix} 0 \\ 2 \end{bmatrix},$$

then $\mathbf{x} = 2\mathbf{b}_1 - 2\mathbf{b}_2$, so

$$[\mathbf{x}]_\mathcal{B} = \begin{bmatrix} 2 \\ -2 \end{bmatrix}.$$

2 Linear transformations in a basis

Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Suppose we have two bases $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ and $\mathcal{C} = \{\mathbf{c}_1, \ldots, \mathbf{c}_m\}$, for $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively. If $\mathbf{x}$ is any vector in $\mathbb{R}^n$, then $T(\mathbf{x}) \in \mathbb{R}^m$, and we can write $\mathbf{x}$ and its image in $\mathcal{B}$-coordinates and $\mathcal{C}$-coordinates respectively, with corresponding coordinate vectors $[\mathbf{x}]_\mathcal{B}$, $[T(\mathbf{x})]_\mathcal{C}$. For definiteness, let's say that $\mathbf{x} = r_1\mathbf{b}_1 + \ldots + r_n\mathbf{b}_n$. Then $T(\mathbf{x}) = r_1T(\mathbf{b}_1) + \ldots r_nT(\mathbf{b}_n)$, by linearity. Now, each $T(\mathbf{b}_i)$ appearing in this sum can be written as a linear combination of the $\mathbf{c}_i$'s. If you write each of these linear combinations and collect terms, you get a formula for the $\mathcal{C}$-coordinates of $T(\mathbf{x})$ in terms of the $\mathcal{B}$-coordinates of $\mathbf{x}$. It is

$$[T(\mathbf{x})]_\mathcal{C} = M[\mathbf{x}]_\mathcal{B}$$

where the $m \times n$ matrix M is

$$M = \begin{bmatrix} [T(\mathbf{b}_1)]_\mathcal{C} & \ldots & [T(\mathbf{b}_n)]_\mathcal{C} \end{bmatrix}.$$

We call M the matrix for T relative to the bases $\mathcal{B}$ and $\mathcal{C}$. We saw in a past lecture that every linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ can be written as multiplication by a matrix A. The matrix A represents "how the function T looks from the point of view of the standard bases". The matrix M we constructed above describes how the transformation looks from the point of view of $\mathcal{B}$ and $\mathcal{C}$.

Example 2.1. Let $T : \mathbb{R}^2 \to \mathbb{R}^3$ be the linear transformation with standard matrix

$$A = \begin{bmatrix} 1 & 4 \\ 1 & 4 \\ 0 & 1 \end{bmatrix}.$$

Let $\mathcal{B}$ be the basis for $\mathbb{R}^2$ from the previous example, and consider a basis for $\mathbb{R}^3$ denoted by $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3\}$, where

$$\mathbf{c}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{c}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{c}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Then

$$T(\mathbf{b}_1) = \begin{bmatrix} 9 \\ 9 \\ 2 \end{bmatrix}, \quad T(\mathbf{b}_2) = \begin{bmatrix} 5 \\ 5 \\ 1 \end{bmatrix}.$$

Thus,

$$[T(\mathbf{b}_1)]_{\mathcal{C}} = \begin{bmatrix} 9 \\ 0 \\ -7 \end{bmatrix}, \quad [T(\mathbf{b}_2)]_{\mathcal{C}} = \begin{bmatrix} 5 \\ 0 \\ -4 \end{bmatrix}.$$

Therefore, the matrix M is given by

$$\begin{bmatrix} 9 & 5 \\ 0 & 0 \\ -7 & -4 \end{bmatrix}.$$

2.1 On the same space $\mathbb{R}^n$

The above was introduced largely to consider what diagonalization actually means. Let's consider a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ (so the standard matrix A of T is square).

Theorem 2.2. *Let the $n \times n$ matrix A be diagonalizable. If $A = PDP^{-1}$ (with D diagonal), and we denote by $\mathcal{B}$ the basis of $\mathbb{R}^n$ formed from the columns of P, then D is the $\mathcal{B}$-matrix for the transformation $\mathbf{x} \mapsto A\mathbf{x}$.*

When we write a matrix product $A\mathbf{x}$ in its diagonalized form $PDP^{-1}\mathbf{x}$, what we are actually doing in the computation of $PDP^{-1}\mathbf{x}$ is the following (where $\mathbf{b}_1, \ldots, \mathbf{b}_n$ are the basis of eigenvectors of A):

- Mapping $\mathbf{x}$ to $[\mathbf{x}]_{\mathcal{B}}$
- Multiplying the entries of $[\mathbf{x}]_{\mathcal{B}}$ by the corresponding eigenvalues;
- Mapping back to the standard basis $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$.

Example 2.3. We will illustrate the above list with a specific example. Let

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

A has eigenvalues -1 and 3, with corresponding eigenvectors $\mathbf{b}_1 = (-1, 1)$ and $\mathbf{b}_2 = (1, 1)$.

This means that $A = PDP^{-1}$, where

$$P = \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}, \quad D = \begin{bmatrix} -1 & 0 \\ 0 & 3 \end{bmatrix}, \quad P^{-1} = \begin{bmatrix} -1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}.$$

Let's take $\mathbf{x} = (2, -1)$. Then $A\mathbf{x} = (0, 3)$ by explicit calculation. We will show how this arises from the form $A = PDP^{-1}$.

- $P^{-1}\mathbf{x} = (-3/2, 1/2)$. Notice that $\mathbf{x} = (-3/2)\mathbf{b}_1 + (1/2)\mathbf{b}_2$, so $P^{-1}\mathbf{x} = [\mathbf{x}]_{\mathcal{B}}$, as described above.
- D acts diagonally on $[\mathbf{x}]_{\mathcal{B}}$ to give $(3/2, 3/2)$. This agrees with the fact that $\mathbf{x} = (-3/2)\mathbf{b}_1 + (1/2)\mathbf{b}_2$ and the $\mathbf{b}_i$'s are eigenvectors.
- P now changes basis back to the standard basis, and we see $P \begin{bmatrix} 3/2 \\ 3/2 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$, as we calculated earlier.

3 Similarity of Matrix Representations

For linear transformations mapping $\mathbb{R}^n$ to $\mathbb{R}^n$, the characterization of similar matrices from the last section holds whether or not the transformation is diagonalizable. That is:

Theorem 3.1. *If we have an $n \times n$ matrix A such that $A = PCP^{-1}$, then C is the $\mathcal{B}$-matrix for the transformation $\mathbf{x} \mapsto A\mathbf{x}$, where $\mathcal{B}$ is the basis made up of the columns of P. Similarly, if $\mathcal{B}$ is a basis for $\mathbb{R}^n$, then the $\mathcal{B}$-matrix for A is given by $P^{-1}AP$, where P is the matrix whose columns are the vectors in $\mathcal{B}$.*

Example 3.2. Consider the matrix A and the basis $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ for $\mathbb{R}^2$ given by

$$A = \begin{bmatrix} 1 & 3 \\ 0 & 0 \end{bmatrix}, \quad \mathbf{b}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \mathbf{b}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

Then if T is the linear transformation defined by $T(\mathbf{x}) = A\mathbf{x}$, then the $\mathcal{B}$-matrix for T (denoted by $[T]_\mathcal{B}$) is given by $[T]_\mathcal{B} = P^{-1}AP$, where

$$P = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}, \quad P^{-1} = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}.$$

Doing the calculation of the product $P^{-1}AP$ gives

$$[T]_\mathcal{B} = \begin{bmatrix} 2 & 1 \\ -2 & -1 \end{bmatrix}.$$