CISC 271   Class 6

Diagonalizable Matrices

Text Correspondence: §6.2

Main Concepts:

- *Diagonalizable: a matrix is similar to a diagonal matrix*
- *Similar matrices have the same eigenvalues*
- *Produces an especially useful decomposition*

Sample Problem, Data Analysis: When are unit eigenvectors a basis?

For many matrices that we will encounter, the eigenvectors $\vec{v}_j$ form a special basis. Let us try to understand why.

First, a useful convention in mathematics is that an eigenvector is often assumed to be of unit length. We can see immediately that any non-zero vector can be forced to be of unit length, just by dividing each entry by the norm of the vector. From now on, we will assume that an eigenvector $\vec{v}$ has the property

$$\|\vec{v}\| = 1$$

6.1 Similar Matrices

In linear algebra, when we say that two matrices are *similar*, there is a specific meaning.

Definition: similar matrices

For any matrix $A \in \mathbb{R}^{n\times n}$, and for matrix $C \in \mathbb{R}^{n\times n}$, A is *similar* to C, or $A \sim C$, is defined as:

There exists an invertible matrix $P \in \mathbb{R}^{n\times n}$ such that

$$A = P^{-1}CP \tag{6.1}$$

Because P in Definition 6.1 is invertible, we can also write

$$C = PAP^{-1}$$

6.2 Eigenvectors as a Basis

The idea of similarity is especially useful when a matrix is similar to a particular diagonal matrix. Consider any matrix $A \in \mathbb{R}^{n \times n}$ that has n linearly independent eigenvectors. For each eigenvector $\vec{v}_j$, we know that

$$A\vec{v}_j = \lambda_j \vec{v}_j \tag{6.2}$$

We can assemble these eigenvectors into a matrix E, which is

$$E = \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \end{bmatrix} \tag{6.3}$$

Applying Equation 6.2 to the columns of E, we get

$$AE = \begin{bmatrix} \lambda_1\vec{v}_1 & \lambda_2\vec{v}_2 & \cdots & \lambda_n\vec{v}_n \end{bmatrix} \tag{6.4}$$

Consider ways that we can re-write the right-hand side of Equation 6.4. One way is to factor it, which means expressing it as the product of two simpler matrices. We can decompose the matrix as $E\Lambda$, where E is the matrix of eigenvectors and Λ is a diagonal matrix of the eigenvalues, so that

$$E\Lambda = \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} \tag{6.5}$$

This means that AE and $E\Lambda$ are the same. Because the eigenvectors are linearly independent, the matrix E is full rank and invertible, so E^{-1} exists. Together, these results imply that

$$\begin{aligned} AE &= E\Lambda \\ \Rightarrow \quad AEE^{-1} &= E\Lambda E^{-1} \\ \Rightarrow \quad A &= E\Lambda E^{-1} \end{aligned} \tag{6.6}$$

Using a closely related line of reasoning, pre-multiplying by E^{-1}, implies that

$$\Lambda = E^{-1}AE \tag{6.7}$$

A matrix A that can be converted into a diagonal matrix is called *diagonalizable*. A necessary and sufficient condition for a matrix to be diagonalizable is that its eigenvectors are a basis, which means that there are n eigenvectors that are linearly independent.

A sufficient, but not necessary, condition for a matrix to be diagonalizable is that all of the eigenvalues are distinct. This condition is not necessary because some matrices have repeated eigenvalues and also have an eigenvector basis. A 3×3 example is

$$A = \begin{bmatrix} -1 & 3 & -1 \\ -3 & 5 & -1 \\ -3 & 3 & 1 \end{bmatrix} \quad \text{with} \quad \lambda_1 = 1 \;\; \lambda_2 = 2 \;\; \lambda_3 = 2 \quad \text{and} \quad E \sim \begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 3 \end{bmatrix}$$

where we have multiplied the columns of E by real numbers that give us a "human-readable" set of basis vectors.

A general rule is that *distinct eigenvalues imply that the eigenvectors are a basis*.

6.3 Eigenvector Basis

If a matrix $A \in \mathbb{R}^{n \times n}$ is diagonalizable, then its eigenvectors are a basis. Using the convention that eigenvectors are of unit length, i.e., we require that $\|\vec{v}_j\| = 1$, we can represent any given vector $\vec{u}$ as

$$\begin{aligned} \vec{u} &= \alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2 + \cdots \alpha_n \vec{v}_n \\ &= \sum_{j=1}^{n} \alpha_j \vec{v}_j \end{aligned}$$

When we perform the multiplication $\vec{y} = A\vec{u}$ we get

$$\begin{aligned} A\vec{u} &= \alpha_1 A\vec{v}_1 + \alpha_2 A\vec{v}_2 + \cdots + \alpha_n A\vec{v}_n \\ &= \sum_{j=1}^{n} \alpha_j \lambda_j \vec{v}_j \\ &= \sum_{j=1}^{n} \lambda_j (\alpha_j \vec{v}_j) \end{aligned}$$

so each of the original terms $\alpha_j \vec{v}_j$ is multiplied by the eigenvalue λ_j. This is depicted in Figure 6.1 for $n = 2$ (2D, or the plane).

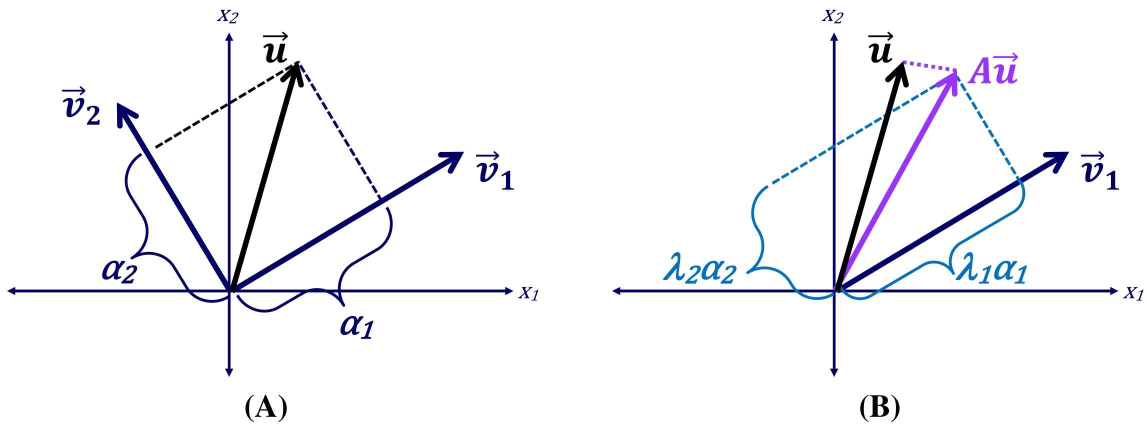

Figure 6.1: Distinct eigenvectors form a basis of a vector space, so any given vector in that space can be expressed as a weighted sum of the eigenvectors. The linear transformation of the matrix scales the eigenvectors and thus changes the given vector in a predictable manner. (A) The original vector $\vec{u}$ is a weighted sum of two eigenvectors, so $\vec{u} = \alpha_1\vec{v}_1 + \alpha_2\vec{v}_2$. (B) The eigenvectors are scaled by their corresponding eigenvalues so the original vector $\vec{u}$ becomes $\alpha_1\lambda_1\vec{v}_1 + \alpha_2\lambda_2\vec{v}_2$. In general, both the direction and magnitude of the original vector are changed by the linear transformation of the matrix A.

6.4 Nondiagonalizable Matrices

In linear algebra, a square matrix that is not diagonalizable is historically called a *defective matrix*. If we allow eigenvalues and eigenvectors to be complex – even when the entries of the original matrix are real – then, in a special technical sense that involves the Lebesgue measure, "almost no" matrix is defective.

A necessary condition for a matrix to be nondiagonalizable is that it must have repeated eigenvalues. For example, suppose that a matrix A is $n \times n$ and has k distinct eigenvalues; if $k < n$, then at least one eigenvalue is repeated and we say that such an eigenvalue has an *algebraic multiplicity* greater than one. This is another way of saying that there is a "multiple root" of the characteristic equation $F(\lambda) = 0$ for the matrix A.

We must be careful here, because nontrivial algebraic multiplicity is not a *sufficient* condition for a matrix to be nondiagonalizable. We must also examine the eigenvectors of the repeated eigenvalue to determine whether or not the matrix has *geometric multiplicity*, which is the dimension of the nullspace of $[A - \lambda I]$.

A specific example of a nondiagonalizable matrix is the 2×2 matrix

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \tag{6.8}$$

The matrix A in Equation 6.8 has a single eigenvalue $\lambda = 1$ and a single eigenvector $\vec{v} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

A general example of a nondiagonalizable matrix, of which Equation 6.8 is an instance, is any matrix with a nontrivial Jordan block of size 2×2. This is a *bidiagonal* matrix of the form

$$A = \begin{bmatrix} a & 1 & & & \\ & a & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & a \end{bmatrix} \tag{6.9}$$

The matrix A in Equation 6.9 has a single eigenvalue $\lambda = a$ and a single eigenvector $\vec{v} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$

6.5 Matrix Powers

Diagonalizable matrices, especially ones with distinct eigenvalues, are common in practice and have easily described matrix powers. We can raise a square matrix A to an integer power A^k using a simple recursive rule based on the identity matrix I:

$$\begin{aligned} A^0 &= I \\ A^{k+1} &= A^k I \end{aligned}$$

Because A is diagonalizable, we can see that

$$\begin{aligned} A^2 &= AA = E\Lambda E^{-1} E\Lambda E^{-1} \\ &= E\Lambda^2 E^{-1} \\ A^3 &= AA^2 = E\Lambda E^{-1} E\Lambda^2 E^{-1} \\ &= E\Lambda^3 E^{-1} \end{aligned}$$

and so on for A^k.

Consider: a diagonalizable matrix A for which each eigenvalue is non-negative. For such a matrix, we can write the eigenvalue matrix Λ in terms of a diagonal matrix D so that

$$\Lambda = DD$$

Each entry of the diagonal matrix D is, by definition, $d_{jj} = \sqrt{\lambda_j}$ and is a real number. so it is easy to compute a new matrix C that is

$$C = EDE^{-1} \tag{6.10}$$

The matrix C in Equation 6.10 is constructed so that

$$A = CC = C^2$$

This implies that, for a diagonalizable matrix A that has non-negative eigenvalues, there is a *square root matrix* C:

$$C = A^{1/2} \tag{6.11}$$

The method of solving Equation 6.11 by using the decomposition of Equation 6.10 is, numerically, not the way that `MATLAB` currently performs the computation of

```
C=sqrtm(A)
```

In practice there are multiple solutions for a matrix C such that $A = CC$ and `MATLAB` finds a solution that is numerically reasonable.

Extra Notes

6.6 Extra Notes: Small Perturbations

The eigenvalue/eigenvector decomposition of a matrix can be used to analyze the numerical stability of the matrix. We can consider three examples, with extensive analysis being beyond the scope of this course.

Example: stable 2×2 ***matrix; consider***

$$A = \begin{bmatrix} +101 & -90 \\ +110 & -98 \end{bmatrix}$$

The eigenvalues of this matrix are $\{+1, +2\}$ and the respective eigenvectors are, approximately,

$$\vec{v}_1 = \begin{bmatrix} 0.6727 \\ 0.7399 \end{bmatrix} \qquad \vec{v}_2 = \begin{bmatrix} 0.6690 \\ 0.7433 \end{bmatrix}$$

Any given vector $\vec{u}$ will be transformed to $\vec{y} = A\vec{u}$ as

$$\begin{aligned} \vec{y} &= \lambda_1\alpha_1\vec{v}_1 + \lambda_2\alpha_2\vec{v}_2 \\ &= \alpha_1\vec{v}_1 + 2\alpha_2\vec{v}_2 \end{aligned}$$

so the transformed $\vec{y}$ is relatively insensitive to small changes in $\vec{u}$.

Example: unstable** 2×2 **matrix; consider

$$A = \begin{bmatrix} +100.999 & +90.001 \\ +110 & -98 \end{bmatrix}$$

For this matrix, the eigenvalues are approximately $\{-139.2134, +142.2134\}$ and the respective unit eigenvectors are, approximately,

$$\vec{v}_1 = \begin{bmatrix} -0.3508 \\ +0.9364 \end{bmatrix} \qquad \vec{v}_2 = \begin{bmatrix} +0.9092 \\ +0.4163 \end{bmatrix}$$

Any given vector $\vec{u}$ will be transformed to $\vec{y} = A\vec{u}$ as, approximately,

$$\begin{aligned} \vec{y} &= \lambda_1\alpha_1\vec{v}_1 + \lambda_2\alpha_2\vec{v}_2 \\ &\approx -139\alpha_1\vec{v}_1 + 142\alpha_2\vec{v}_2 \end{aligned}$$

so the transformed $\vec{w}$ is *highly sensitive* to small changes in $\vec{u}$.

Example: nearly singular** 2×2 **matrix; consider

$$A = \begin{bmatrix} +101.00 & +89.99 \\ +109.99 & +98.00 \end{bmatrix}$$

For this matrix, the eigenvalues are approximately $\{-5.0251 \times 10^{-7}, +199\}$ and the respective unit eigenvectors are, approximately,

$$\vec{v}_1 = \begin{bmatrix} -0.6652 \\ +0.7466 \end{bmatrix} \qquad \vec{v}_2 = \begin{bmatrix} +0.6764 \\ +0.7366 \end{bmatrix}$$

Any given vector $\vec{u}$ will be transformed to $\vec{y} = A\vec{u}$ as, approximately,

$$\begin{aligned} \vec{y} &= \lambda_1\alpha_1\vec{v}_1 + \lambda_2\alpha_2\vec{v}_2 \\ &\approx 199\alpha_1\vec{v}_1 \end{aligned}$$

so the transformed $\vec{y}$ is *highly sensitive* to small changes in the direction of $\vec{v}_1$ and *insensitive* to changes in the direction of $\vec{v}_2$.

Example: stability analysis of a symmetric linear system; consider

Suppose that the matrix A is symmetric and that we seek a numerical solution to the matrix-vector equation $A\vec{x} = \vec{b}$. How stable is the solution, that is, how sensitive is the solution to small changes in the entries of $\vec{b}$?

We know that the analytic solution to the system is $\vec{x} = A^{-1}\vec{b}$. Because A is symmetric, so is A^{-1}; finding the eigenvector decomposition of the matrix A gives

$$\vec{x} = \left(Q\Lambda Q^T\right)\vec{b}$$

where Q is orthogonal and Λ is diagonal. Expanding the factorization produces

$$\begin{aligned}\vec{x} &= Q\Lambda\left(Q^T\vec{b}\right)\\ &= Q\Lambda\vec{c}\\ &= Q\begin{bmatrix}\lambda_1 c_1\\ \lambda_2 c_2\\ \lambda_3 c_3\\ \vdots\\ \lambda_n c_n\end{bmatrix}\end{aligned}$$

If $|\lambda_n|$ is large, then small perturbations in c_n will result in large perturbations in $\vec{x}$.

End of Extra Notes