Announcements

1. Quiz 6 covers section 4.4 and Markov chains part of 4.6
2. **Read** Section 5.0, 5.1,5.2 for Wednesday. Work through recommended homework questions.
3. Office hours today 11-12 MC 121.
4. Lecture notes (including this page) is available on my webpage:
   http://www.math.uwo.ca/~hbacard/teaching.html

Last time

Summary of Diagonalization

Objective: Given an $n \times n$ matrix A, we would like to determine whether A is diagonalizable, and if it is, find the invertible matrix P and the diagonal matrix D such that $P^{-1}AP = D$. The result may depend upon whether you are working over $\mathbb{R}$ or $\mathbb{C}$.

Steps

⇒ Compute the characteristic polynomial $det(A - \lambda I)$ of A.

⇒ Find the roots of the characteristic polynomial and their algebraic multiplicities by factoring.

⇒ If the algebraic multiplicities don't add up to n, then A is not diagonalizable, and you can stop. (If you are working over $\mathbb{C}$, this can't happen.)

⇒ For each eigenvalue λ, compute the dimension of the eigenspace E_λ. This is the geometric multiplicity of λ, and if it is less than the algebraic multiplicity, then A is not diagonalizable, and you can stop.

⇒ Compute a basis for the eigenspace E_λ.

⇒ If for each eigenvalue the geometric multiplicity equals the algebraic multiplicity, then you take the n eigenvectors you found and put them in the columns of a matrix P. Put the eigenvalues in the same order on the diagonal of a matrix D.

⇒ Check that $AP = PD$.

Computing powers If A is diagonalizable i.e, there is an invertible matrix P and a diagonal matrix D such that $P^{-1}AP = D$, then we can compute with less pain the powers of A.

One writes

$$P^{-1}AP = D \Leftrightarrow A = PDP^{-1},$$

and observe that

$$A^2 = (PDP^{-1})(PDP^{-1}) = PD(\underbrace{P^{-1}P}_{=I})DP^{-1} = PD^2P^{-1}$$

$\Rightarrow$ Pattern:

$$A^k = PD^kP^{-1}$$

Markov chains (4.6)

Recall: A **stochastic matrix** is a matrix $P = [P_{ij}]$ such that $P_{ij} \geq 0$ **and**

$$P_{1j} + P_{2j} + \cdots + P_{nj} = 1, \quad \text{for all } j.$$

$\Rightarrow$ Examples of such are **transition matrix** P **of a Markov chain**.

$\overrightarrow{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ is a **probability vector** if $x_i \geq 0$ **and**

$$x_1 + x_2 + \cdots x_n = 1$$

A probability vector $\overrightarrow{x}$ that satisfies $P\overrightarrow{x} = \overrightarrow{x}$ is called a **a steady state vector**. This is exactly the same thing as an eigenvector with eigenvalue 1

4 main Theorems

Theorem 4.30 If P is an $n \times n$ transition matrix of a Markov chain, then 1 is an eigenvalue of P

Theorem 4.31 Let P be an $n \times n$ transition matrix.

1. Then every eigenvalue λ has $|\lambda| \leq 1$
2. If P is regular (some power is positive), then all eigenvalues beside $\lambda = 1$ have $|\lambda| < 1$

Theorem 4.33 Let P be an $n \times n$ stochastic matrix all of whose entries are positive. Then as $k \longrightarrow \infty$, $P^k \longrightarrow L$, a matrix all of whose columns are equal to the same vector $\overrightarrow{x}$ which is a steady state probability vector for P

Example Let $P = \begin{bmatrix} 0.7 & 0.2 \\ 0.3 & 0.8 \end{bmatrix}$. An easy computation give the characteristic polynomial

$$\det(P - \lambda I) = \lambda^2 - 1.5\lambda + 0.5 = (\lambda - 1)(\lambda - 0.5)$$

$\Rightarrow \lambda_1 = 1$ and $\lambda_2 = 0.5$ are the eigenvalues. The eigenspaces are

$$E_1 = span(\begin{bmatrix} 2 \\ 3 \end{bmatrix}), \quad E_{0.5} = span(\begin{bmatrix} 1 \\ -1 \end{bmatrix}).$$

If take we $Q = \begin{bmatrix} 2 & 1 \\ 3 & -1 \end{bmatrix}$, we have $Q^{-1}PQ = \begin{bmatrix} 1 & 0 \\ 0 & 0.5 \end{bmatrix} = D$

$$\Rightarrow P^k = QD^kQ^{-1} = \begin{bmatrix} 2 & 1 \\ 3 & -1 \end{bmatrix} \begin{bmatrix} 1^k & 0 \\ 0 & (0.5)^k \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 3 & -1 \end{bmatrix}^{-1}$$

Now, as $k \longrightarrow \infty$ we have $(0.5)^k \longrightarrow 0$, so

$$P^k \longrightarrow \begin{bmatrix} 2 & 1 \\ 3 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 3 & -1 \end{bmatrix}^{-1} = \begin{bmatrix} 0.4 & 0.4 \\ 0.6 & 0.6 \end{bmatrix}$$

$\Rightarrow$ check that $\begin{bmatrix} 0.4 \\ 0.6 \end{bmatrix}$ is **the** steady state vector of P

Theorem 4.34 Let P be a regular $n \times n$ matrix $\overrightarrow{x}$ the steady state probability vector of P. The for any initial probability vector $\overrightarrow{x}_0$ the sequence of iterates $\overrightarrow{x}_k = P^k \overrightarrow{x}_0 \longrightarrow \overrightarrow{x}$ as $k \longrightarrow \infty$.

New material

Orthogonality in $\mathbb{R}^n$

Goal: Generalize what we did in $\mathbb{R}^2$ and $\mathbb{R}^3$.

Remember: We use the dot product and say that $\overrightarrow{u}$ and $\overrightarrow{v}$ are orthogonal if $\overrightarrow{u} \cdot \overrightarrow{v} = 0$

Definition A set of vectors $\{\overrightarrow{v_1}, ..., \overrightarrow{v_k}\}$ in $\mathbb{R}^n$ is called an **orthogonal set if** all pairs of distinct vectors in the set are orthogonal i.e

$$\overrightarrow{v_i} \cdot \overrightarrow{v_j} = 0, \quad \text{whenever } i \neq j.$$

Example Show that $\{\overrightarrow{v_1}, \overrightarrow{v_2}, \overrightarrow{v_3}\}$ is an orthogonal set in $\mathbb{R}^3$ if

$$\overrightarrow{v_1} = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}, \quad \overrightarrow{v_2} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \quad \overrightarrow{v_3} = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}.$$

Solution Just check that $\overrightarrow{v_1} \cdot \overrightarrow{v_2} = 0$, $\overrightarrow{v_1} \cdot \overrightarrow{v_3} = 0$ and $\overrightarrow{v_2} \cdot \overrightarrow{v_3} = 0$.

$$\overrightarrow{v_1} \cdot \overrightarrow{v_2} = 2(0) + 1(1) + (-1)(1) = 0, \quad \text{and same for the other ones.}$$

Theorem 5.1 If $\{\overrightarrow{v_1}, ..., \overrightarrow{v_k}\}$ is an orthogonal set of **nonzero vectors** in $\mathbb{R}^n$ then **necessarily** $\overrightarrow{v_1}, ..., \overrightarrow{v_k}$ are linearly independent.

Quick Proof Suppose there are scalars $c_1, ..., c_k$ such that

$$c_1\overrightarrow{v_1} + \cdots + c_k\overrightarrow{v_k} = \overrightarrow{0}.$$

We wish to show that these scalars are all equal to 0. Let's start with c_1.
Key idea: dot by $\overrightarrow{v_1}$ on both sides in the above equality.

$$(c_1\overrightarrow{v_1} + \cdots + c_k\overrightarrow{v_k}) \cdot \overrightarrow{v_1} = \overrightarrow{0} \cdot \overrightarrow{v_1} = 0$$

But $(c_1\overrightarrow{v_1} + \cdots + c_k\overrightarrow{v_k}) \cdot \overrightarrow{v_1} = c_1\overrightarrow{v_1} \cdot \overrightarrow{v_1} = c_1|\overrightarrow{v_1}|^2$, so we get

$$c_1|\overrightarrow{v_1}|^2 = 0 \quad \Rightarrow \quad c_1 = 0, \quad (\text{because } |\overrightarrow{v_1}| \neq 0).$$

$\Rightarrow$ Proceeding this way one shows that all c_i are 0. ■

Definition An **Orthogonal basis** for a subspace W of $\mathbb{R}^n$ is a basis that is also orthogonal.

Example The previous vectors

$$\overrightarrow{v_1} = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}, \quad \overrightarrow{v_2} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \quad \overrightarrow{v_3} = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

form an orthogonal basis for $\mathbb{R}^3$. Indeed they are linearly independend because orthogonal, and there are 3 vectors.

Theorem 5.2 Let $\{\overrightarrow{v_1}, ..., \overrightarrow{v_k}\}$ be an orthogonal basis of a subspace $W \subset \mathbb{R}^n$ and let $\overrightarrow{w}$ be any vector in W. Then the unique scalars $c_1, ..., c_k$ such that

$$\overrightarrow{w} = c_1\overrightarrow{v_1} + \cdots + c_k\overrightarrow{v_k},$$

are given by the formula

$$c_i = \frac{\overrightarrow{w} \cdot \overrightarrow{v_i}}{\overrightarrow{v_i} \cdot \overrightarrow{v_i}}, \quad \text{for all } i$$

Example Find the coordinates of $\overrightarrow{w} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ in the orthogonal basis

$$\mathcal{B} = \{\overrightarrow{v_1} = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}, \overrightarrow{v_2} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \overrightarrow{v_3} = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}\}$$

Solution: The coordinates are:

$$c_1 = \frac{\overrightarrow{w} \cdot \overrightarrow{v_1}}{\overrightarrow{v_1} \cdot \overrightarrow{v_1}} = \frac{2+2-3}{2+1+1} = \frac{1}{6}$$

Similarly one finds

$$c_2 = \frac{5}{2}, \quad c_3 = \frac{2}{3}$$

So we can write

$$[\overrightarrow{w}]_{\mathcal{B}} = \begin{bmatrix} 1/6 \\ 5/2 \\ 2/3 \end{bmatrix}$$

Definition A set $\{\overrightarrow{v_1}, ..., \overrightarrow{v_k}\}$ of vectors is called an **orthonormal set** if

1. it's an orthogonal set **and**
2. all vectors are unit vectors i.e $|\overrightarrow{v_i}| = 1$ for every i.

An **orthonormal basis** for a subspace W of $\mathbb{R}^n$ is a basis of W that is an orthonormal set.

Example

1. The standard basis $\{\overrightarrow{e_1}, ..., \overrightarrow{e_n}\}$ is an orthonormal basis for $\mathbb{R}^n$.
2. We can normalize any orthogonal basis to an orthonormal basis:
   If $\{\overrightarrow{v_1}, ..., \overrightarrow{v_k}\}$ is an orthogonal basis then we can get an orthonormal basis by divind by the length i.e
   $$\overrightarrow{u_i} = \frac{1}{|\overrightarrow{v_i}|}\overrightarrow{v_i}.$$
   $\Rightarrow$ What do you get with the previous orthogonal basis $\mathcal{B}$ of $\mathbb{R}^3$?

Thereom 5.3 Let $\{\overrightarrow{q_1}, ..., \overrightarrow{q_k}\}$ be an orthonormal basis for a subspace W of $\mathbb{R}^n$ and let $\overrightarrow{w}$ be any vector in W. Then $\overrightarrow{w}$ can is uniquely written as

$$\overrightarrow{w} = (\overrightarrow{w} \cdot \overrightarrow{q_1})\overrightarrow{q_1} + \cdots + (\overrightarrow{w} \cdot \overrightarrow{q_k})\overrightarrow{q_k}.$$

$\Rightarrow$ This is just Theorem 5.2 with the fact that $(\overrightarrow{q_i} \cdot \overrightarrow{q_i}) = 1$.

Orthogonal Matrices

Main Idea: The columns form an **orthonormal set** (absurd terminology...)

Theorem 5.4 The columns of an $m \times n$ matrix Q form an orthonormal set if and only if

$$Q^T Q = I_n.$$

We are mainly interested to the case when $m = n$

Definition An $n \times n$ matrix Q whose columns form an orthonormal set is called an **orthogonal matrix**.

Theorem 5.5 An square matrix Q is orthogonal if and only if $Q^{-1} = Q^T$.

Remark Thanks to this theorem it's **very easy** to compute the inverse of an orthogonal matrix ! :D

Example Show that the following matrices are orthogonal and find their inverses.

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & -1 \\ -1 & 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

Theorem 5.6 For an $n \times n$ matrix Q, the following statements are equivalent

1. Q is orthogonal.
2. $|Q\overrightarrow{x}| = |\overrightarrow{x}|$ for every $\overrightarrow{x} \in \mathbb{R}^n$. "$Q$ preserves the length"
3. $Q\overrightarrow{x} \cdot Q\overrightarrow{y} = \overrightarrow{x} \cdot \overrightarrow{y}$, for every $\overrightarrow{x}$ and $\overrightarrow{y}$ in $\mathbb{R}^n$. "Q preserves the dot product"

Switching to the board from here.