Math 1600A Lecture 33, Section 2, 27 Nov 2013

Announcements:

Today we finish 5.1 and start 5.2. Continue **reading** Section 5.2 for Friday, and start reading 5.3. Work through recommended homework questions.

Tutorials: Quiz 6 covers Sections 4.4 and the Markov Chains part of 4.6.
Office hour: Wednesday, 12:30-1:30, MC103B.
Help Centers: Monday-Friday 2:30-6:30 in MC 106.

T/F: A matrix with orthogonal columns is called an orthogonal matrix.

T/F: An orthogonal matrix must be square.

Question: Why are orthonormal bases great? Are orthogonal bases great too?

Review of Section 5.1: Orthogonal and Orthonormal sets

Definition: A set of vectors $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$ in $\mathbb{R}^n$ is an **orthogonal set** if $\vec{v}_i \cdot \vec{v}_j = 0$ for $i \neq j$.

Theorem 5.1: An orthogonal set of nonzero vectors is always linearly independent.

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

You only need to check that the set spans W, since it is automatically linearly independent.

Fact: We'll show in Section 5.3 that every subspace *has* an orthogonal basis.

Recall that if $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$ is any basis of a subspace W, then any $\vec{w}$ in W can be written uniquely as a linearly combination of the vectors in the basis. In general, finding the coefficients involves solving a linear system. For an orthogonal basis, it is much easier:

Theorem 5.2: If $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$ is an orthogonal basis of a subspace W, and $\vec{w}$ is in W, then

$$\vec{w} = c_1 \vec{v}_1 + \cdots + c_k \vec{v}_k \quad \text{where} \quad c_i = \frac{\vec{w} \cdot \vec{v}_i}{\vec{v}_i \cdot \vec{v}_i}$$

Definition: An **orthonomal set** is an orthogonal set of unit vectors. An **orthonormal**

basis for a subspace W is a basis for W that is an orthonormal set.

The condition of being orthonormal can be expressed as

$$\vec{v}_i \cdot \vec{v}_j = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}$$

Question: How many orthonormal bases are there for $\mathbb{R}^3$?

Note that an orthogonal basis can be converted to an orthonormal basis by dividing each vector by its length.

Theorem 5.3: If $\{\, \vec{q}_1, \vec{q}_2, \ldots, \vec{q}_k \}$ is an orthonormal basis of a subspace W, and $\vec{w}$ is in W, then

$$\vec{w} = (\vec{w} \cdot \vec{q}_1)\, \vec{q}_1 + \cdots + (\vec{w} \cdot \vec{q}_k)\, \vec{q}_k$$

Orthogonal Matrices

Definition: A square matrix Q with real entries whose columns form an orthonormal set is called an **orthogonal** matrix!

Note: In $\mathbb{R}^2$ and $\mathbb{R}^3$, orthogonal matrices correspond exactly to the rotations and reflections. This is an important geometric reason to study them. Another reason is that we will see in Section 5.4 that they are related to diagonalization of symmetric matrices.

Theorems 5.4 and 5.5: Q is orthogonal if and only if $Q^T Q = I$, i.e. if and only if Q is invertible and $Q^{-1} = Q^T$.

Examples: $A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$.

Theorem 5.6: Let Q be an $n \times n$ matrix. Then the following statements are equivalent:
a. Q is orthogonal.
b. $\|Q\,\vec{x}\| = \|\,\vec{x}\|$ for every $\vec{x}$ in $\mathbb{R}^n$.
c. $Q\,\vec{x} \cdot Q\,\vec{y} = \vec{x} \cdot \vec{y}$ for every $\vec{x}$ and $\vec{y}$ in $\mathbb{R}^n$.

New material

Partial proof of 5.6:
(a) $\implies$ (c): Suppose Q is orthogonal, so $Q^T Q = I$. Then

$$Q\,\vec{x} \cdot Q\,\vec{y} = (Q\,\vec{x})^T (Q\,\vec{y}) = \vec{x}^T Q^T Q\,\vec{y} = \vec{x}^T\,\vec{y} = \vec{x} \cdot \vec{y}$$

(c) $\implies$ (a): Suppose (c) holds. Take $\vec{x} = \vec{e}_i$ and $\vec{y} = \vec{e}_j$. Then $Q\,\vec{x} = Q\,\vec{e}_i$ is the ith column of Q and $Q\,\vec{y} = Q\,\vec{e}_j$ is the jth column of Q. (c) says that the dot product of these columns is

$$Q\,\vec{e}_i \cdot Q\,\vec{e}_j = \vec{e}_i \cdot \vec{e}_j = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}$$

So the columns of Q are orthonomal, which means Q is orthogonal.

(c) $\implies$ (b) is clear, by taking $\vec{x} = \vec{y}$ in (c).

(b) $\implies$ (c): see text. $\square$

Theorem 5.7: If Q is orthogonal, then its **rows** form an orthonormal set too.

Proof: Since $Q^TQ = I$, we must also have $QQ^T = I$. But the last equation says exactly that the rows of Q are orthonormal. $\square$

Another way to put it is that Q^T is also an orthogonal matrix.

Theorem 5.8: Let Q be an orthogonal matrix. Then:
a. Q^{-1} is orthogonal.
b. $\det Q = \pm 1$
c. If λ is an eigenvalue of Q, then $|\lambda| = 1$.
d. If Q_1 and Q_2 are orthogonal matrices of the same size, then $Q_1 Q_2$ is orthogonal.

Proof:
(a) is Theorem 5.7, since $Q^{-1} = Q^T$.

(b): Since $I = Q^TQ$, we have

$$1 = \det I = \det(Q^TQ) = \det(Q^T)\det(Q) = \det(Q)^2.$$

Therefore $\det(Q) = \pm 1$.

(c) If $Q\,\vec{v} = \lambda\,\vec{v}$, then

$$\|\,\vec{v}\| = \|Q\,\vec{v}\| = \|\lambda\,\vec{v}\| = |\lambda|\|\,\vec{v}\|$$

so $|\lambda| = 1$, since $\|\,\vec{v}\| \neq 0$.

(d) Exericse, using properties of transpose. $\square$

Section 5.2: Orthogonal Complements and Orthogonal Projections

We saw in Section 5.1 that orthogonal and orthonormal bases are particularly easy to work with. In Section 5.3, we will learn how to find these kinds of bases. In this section, we learn the tools which will be needed in Section 5.3. We will also find a new way to understand the subspaces associated to a matrix.

Orthogonal Complements

If W is a plane through the origin, with normal vector $\vec{n}$, then the subspaces W and $\mathrm{span}(\vec{n})$ have the property that every vector in one is orthogonal to every vector in the other.

Definition: Let W be a subspace of $\mathbb{R}^n$. A vector $\vec{v}$ is **orthogonal** to W if $\vec{v}$ is orthogonal to every vector in W. The **orthogonal complement** of W is the set of all vectors orthogonal to W and is denoted $W^\perp$. So

$$W^\perp = \{\vec{v} \in \mathbb{R}^n : \vec{v} \cdot \vec{w} = 0 \text{ for all } \vec{w} \text{ in } W\}$$

In the example above, if we write $\ell = \mathrm{span}(\vec{n})$ for the line perpendicular to W, then $\ell = W^\perp$ and $W = \ell^\perp$.

Theorem 5.9: Let W be a subspace of $\mathbb{R}^n$. Then:
a. $W^\perp$ is a subspace of $\mathbb{R}^n$.
b. $(W^\perp)^\perp = W$
c. $W \cap W^\perp = \{\vec{0}\}$
d. If $W = \mathrm{span}(\vec{w}_1, \ldots, \vec{w}_k)$, then $\vec{v}$ is in $W^\perp$ if and only if $\vec{v} \cdot \vec{w}_i = 0$ for all i.

Explain (a), (c), (d) on whiteboard. (b) will be Corollary 5.12.

Theorem 5.10: Let A be an $m \times n$ matrix. Then

$$(\mathrm{row}(A))^\perp = \mathrm{null}(A) \quad \text{and} \quad (\mathrm{col}(A))^\perp = \mathrm{null}(A^T)$$

The first two are in $\mathbb{R}^n$ and the last two are in $\mathbb{R}^m$. These are the **four fundamental subspaces** of A.

Let's see why $(\mathrm{row}(A))^\perp = \mathrm{null}(A)$. A vector is in $\mathrm{null}(A)$ exactly when it is orthogonal to the rows of A. But the rows of A span $\mathrm{row}(A)$, so the vectors in $\mathrm{null}(A)$ are exactly those which are orthogonal to $\mathrm{row}(A)$, by 5.9(d).

The fact that $(\mathrm{col}(A))^\perp = \mathrm{null}(A^T)$ follows by replacing A with A^T.

Example: Let W be the subspace spanned by $\vec{v}_1 = [1, 2, 3]$ and $\vec{v}_2 = [2, 5, 7]$. Find a basis for $W^\perp$.

Solution: Let A be the matrix with $\vec{v}_1$ and $\vec{v}_2$ as rows. Then $W = \mathrm{row}(A)$, so

$W^\perp = \text{null}(A)$. Continue on whiteboard.

Orthogonal projection

Recall (from waaaay back in Section 1.2) that the formula for the projection of a vector $\vec{v}$ onto a nonzero vector $\vec{u}$ is:

$$\text{proj}_{\vec{u}}(\vec{v}) = \left(\frac{\vec{u}\cdot\vec{v}}{\vec{u}\cdot\vec{u}}\right)\vec{u}.$$

(Illustrated by this java applet, where red is $\vec{v}$, blue is $\vec{u}$ and yellow is the projection.)

We didn't name it then, but we also noticed that $\vec{v} - \text{proj}_{\vec{u}}(\vec{v})$ is orthogonal to $\vec{u}$. Let's call this $\text{perp}_{\vec{u}}(\vec{v})$.

So if we write $W = \text{span}(\vec{u})$, then $\vec{w} = \text{proj}_{\vec{u}}(\vec{v})$ is in W, $\vec{w}^\perp = \text{perp}_{\vec{u}}(\vec{v})$ is in $W^\perp$, and $\vec{v} = \vec{w} + \vec{w}^\perp$. We can do this more generally (sketch):

Definition: Let W be a subspace of $\mathbb{R}^n$ and let $\{\vec{u}_1, \ldots, \vec{u}_k\}$ be an orthogonal basis for W. For $\vec{v}$ in $\mathbb{R}^n$, the **orthogonal projection** of $\vec{v}$ onto W is the vector

$$\text{proj}_W(\vec{v}) = \text{proj}_{\vec{u}_1}(\vec{v}) + \cdots + \text{proj}_{\vec{u}_k}(\vec{v})$$

The **component of $\vec{v}$ orthogonal to W** is the vector

$$\text{perp}_W(\vec{v}) = \vec{v} - \text{proj}_W(\vec{v})$$

We will show soon that $\text{perp}_W(\vec{v})$ is in $W^\perp$.

Note that multiplying $\vec{u}$ by a scalar in the earlier example doesn't change W, $\vec{w}$ or $\vec{w}^\perp$. We'll see later that the general definition also doesn't depend on the choice of orthogonal basis.