Math 1600 Lecture 35, Section 2, 28 Nov 2014

Announcements:

Today we finish 5.3 and start 5.4. **Read** Section 5.4 for next class. Work through recommended homework questions.

Tutorials: None next week.

Help Centers: Monday-Friday 2:30-6:30 in MC 106, until Friday, Dec 5 2014.

Final exam: Monday, December 8, 9am to noon. All students in Section 001 write in NS1. For students in Section 002: A to TOPA write in NS1, and TU to Z write in NS7.

Review Sessions: Wednesday, in class, for section 002. Another being planned too. Bring questions!

Review of Section 5.2: Orthogonal Complements and Orthogonal Projections

Orthogonal Complements

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

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

Orthogonal projection

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 showed that $\text{proj}_W(\vec{v})$ is in W and $\text{perp}_W(\vec{v})$ is in $W^\perp$.

Here and in the rest of Section 5.2, we assume that every subspace has at least one orthogonal basis.

Theorem 5.11: Let W be a subspace of $\mathbb{R}^n$ and let $\vec{v}$ be a vector in $\mathbb{R}^n$. Then there are **unique** vectors $\vec{w}$ in W and $\vec{w}^\perp$ in $W^\perp$ such that $\vec{v} = \vec{w} + \vec{w}^\perp$.

Theorem 5.13: If W is a subspace of $\mathbb{R}^n$, then

$$\dim W + \dim W^\perp = n$$

The Rank Theorem follows if we take $W = \text{row}(A)$, since then $W^\perp = \text{null}(A)$.

Section 5.3: The Gram-Schmidt Process and the QR Factorization

The Gram-Schmidt Process

This is a fancy name for a way of converting a basis into an orthogonal or orthonormal basis. And it's pretty clear how to do it, given what we know.

Theorem 5.15 (The Gram-Schmidt Process): Let $\{\vec{x}_1, \ldots, \vec{x}_k\}$ be a basis for a subspace W of $\mathbb{R}^n$. Write $W_1 = \text{span}(\vec{x}_1)$, $W_2 = \text{span}(\vec{x}_1, \vec{x}_2)$, ..., $W_k = \text{span}(\vec{x}_1, \ldots, \vec{x}_k)$. Define:

$$\vec{v}_1 = \vec{x}_1$$

$$\vec{v}_2 = \text{perp}_{W_1}(\vec{x}_2) = \vec{x}_2 - \frac{\vec{v}_1 \cdot \vec{x}_2}{\vec{v}_1 \cdot \vec{v}_1} \vec{v}_1$$

$$\vec{v}_3 = \text{perp}_{W_2}(\vec{x}_3) = \vec{x}_3 - \frac{\vec{v}_1 \cdot \vec{x}_3}{\vec{v}_1 \cdot \vec{v}_1} \vec{v}_1 - \frac{\vec{v}_2 \cdot \vec{x}_3}{\vec{v}_2 \cdot \vec{v}_2} \vec{v}_2$$

$$\vdots$$

$$\vec{v}_k = \text{perp}_{W_{k-1}}(\vec{x}_k) = \vec{x}_k - \frac{\vec{v}_1 \cdot \vec{x}_k}{\vec{v}_1 \cdot \vec{v}_1} \vec{v}_1 - \cdots - \frac{\vec{v}_{k-1} \cdot \vec{x}_k}{\vec{v}_{k-1} \cdot \vec{v}_{k-1}} \vec{v}_{k-1}$$

Then for each i, $\{\vec{v}_1, \ldots, \vec{v}_i\}$ is an orthogonal basis for W_i. In particular, $\{\vec{v}_1, \ldots, \vec{v}_k\}$ is an orthogonal basis for $W = W_k$.

Notes: To compute perp_{W_i} you have to use the *orthogonal* basis of $\vec{v}_j$'s that you

have constructed already, not the original basis of $\vec{x}_j$'s.

The basis you get depends on the order of the vectors you start with. You should always do a question using the vectors in the order given, since that order will be chosen to minimize the arithmetic.

If you are asked to find an orthonormal basis, normalize each $\vec{v}_j$ at the end. (It is correct to normalize earlier, but can be messier.)

New: You don't need to check that the starting vectors are linearly independent. If they are dependent, then one or more of the $\vec{v}_j$'s will be $\vec{0}$, and you can just ignore it.

New: This theorem shows that every subspace has an orthogonal basis. We used the material from 5.2, but only for subspaces for which we had already computed an orthogonal basis, so the logic isn't circular.

Example 5.13: Apply Gram-Schmidt to construct an orthogonal basis for the subspace $W = \mathrm{span}(\vec{x}_1, \vec{x}_2, \vec{x}_3)$ of $\mathbb{R}^4$ where

$$\vec{x}_1 = \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}, \quad \vec{x}_2 = \begin{bmatrix} 2 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \quad \vec{x}_3 = \begin{bmatrix} 2 \\ 2 \\ 1 \\ 2 \end{bmatrix}$$

We get

$$\vec{v}_1 = \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}, \quad \vec{v}_2' = \begin{bmatrix} 3 \\ 3 \\ 1 \\ 1 \end{bmatrix}, \quad \vec{v}_3' = \begin{bmatrix} -1 \\ 0 \\ 1 \\ 2 \end{bmatrix}$$

If we want an orthonormal basis, we scale them:

$$\vec{q}_1 = \frac{1}{2}\begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}, \quad \vec{q}_2 = \frac{1}{\sqrt{20}}\begin{bmatrix} 3 \\ 3 \\ 1 \\ 1 \end{bmatrix}, \quad \vec{q}_3 = \frac{1}{\sqrt{6}}\begin{bmatrix} -1 \\ 0 \\ 1 \\ 2 \end{bmatrix}$$

New material

Example: Is $\vec{w} = \begin{bmatrix} 8 \\ -2 \\ 2 \\ 0 \end{bmatrix}$ in W?

Solution: We can compute that

$$\text{proj}_W(\vec{w}) = 2\vec{v}_1 + \vec{v}_2' - \vec{v}_3' = \begin{bmatrix} 6 \\ 1 \\ -2 \\ 1 \end{bmatrix}$$

so the answer is no.

Example: Compute the projection of $\vec{w} = \begin{bmatrix} 8 \\ -2 \\ 2 \\ 0 \end{bmatrix}$ onto $\text{span}(\vec{x}_1, \vec{x}_2)$.

Solution: We use that $\text{span}(\vec{x}_1, \vec{x}_2) = \text{span}(\vec{v}_1, \vec{v}_2')$. So, by the work done for the previous example, we get $2\vec{v}_1 + \vec{v}_2'$. (Do *not* try to work directly with $\vec{x}_1$ and $\vec{x}_2$.)

Example 5.14: Find an orthogonal basis for $\mathbb{R}^3$ that contains the vector $\vec{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$.

Solution: Choose any two vectors $\vec{x}_2$ and $\vec{x}_3$ so that $\{\vec{v}_1, \vec{x}_2, \vec{x}_3\}$ is a basis for $\mathbb{R}^3$. For example, you can take

$$\vec{x}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \vec{x}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Then apply Gram-Schmidt, using the vectors in that order, so $\vec{v}_1$ doesn't change. (Details in text.)

QR Factorization

We can apply the above to obtain:

Theorem 5.16: Let A be an $m \times n$ matrix with **linearly independent columns**. Then A can be factored as $A = QR$, where Q is an $m \times n$ matrix with orthonormal columns and R is an invertible upper triangular $n \times n$ matrix.

Note that we must have $m \geq n$. (Why?)

This is useful for numerically approximating eigenvalues (see the Exploration after Section 5.3) and for least squares approximation (Chapter 7), but we won't cover these applications.

Explanation: Write $\vec{a}_1, \ldots, \vec{a}_n$ for the linearly independent columns of A. Apply Gram-Schmidt to produce orthonormal vectors $\vec{q}_1, \ldots, \vec{q}_n$ with $\text{span}(\vec{a}_1, \ldots, \vec{a}_i) = \text{span}(\vec{q}_1, \ldots, \vec{q}_i)$ for each i. Therefore we can find scalars r_{ij} such that:

$$\vec{a}_1 = r_{11}\vec{q}_1$$
$$\vec{a}_2 = r_{12}\vec{q}_1 + r_{22}\vec{q}_2$$
$$\vdots$$
$$\vec{a}_n = r_{1n}\vec{q}_1 + r_{2n}\vec{q}_2 + \cdots + r_{nn}\vec{q}_n$$

That is,

$$A = [\,\vec{a}_1 \cdots \vec{a}_n\,] = [\,\vec{q}_1 \cdots \vec{q}_n\,] \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_{nn} \end{bmatrix} = QR$$

One can also see that the diagonal entries r_{ii} are non-zero. (Explain.) Therefore, $\det R \neq 0$ and R is invertible.

Note that $r_{ij} = \vec{q}_i \cdot \vec{a}_j$.

Example 5.15: Find a QR factorization of $A = \begin{bmatrix} 1 & 2 & 2 \\ -1 & 1 & 2 \\ -1 & 0 & 1 \\ 1 & 1 & 2 \end{bmatrix}$.

Solution: The columns of A are the vectors from Example 5.13, so we get the matrix

$$Q = \begin{bmatrix} 1/2 & 3/\sqrt{20} & -1/\sqrt{6} \\ -1/2 & 3/\sqrt{20} & 0 \\ -1/2 & 1/\sqrt{20} & 1/\sqrt{6} \\ 1/2 & 1/\sqrt{20} & 2/\sqrt{6} \end{bmatrix}$$

We want to find R such that $A = QR$. Since the columns of Q are orthonormal, we have $Q^T Q = I$. So $Q^T A = Q^T Q R = R$ and one can compute R by matrix multiplication to find

$$R = Q^T A = \cdots = \begin{bmatrix} 2 & 1 & 1/2 \\ 0 & \sqrt{5} & 3\sqrt{5}/2 \\ 0 & 0 & \sqrt{6}/2 \end{bmatrix}$$

(See text for details.) Note that you can save some work, since you know that the entries below the diagonal must be zero.

Also note that this matrix multiplication is exactly working out the components of $\vec{a}_i$ with respect to the orthonormal basis of $\vec{q}_j$'s, using that $r_{ij} = \vec{q}_i \cdot \vec{a}_j$.

Section 5.4: Orthogonal Diagonalization of Symmetric Matrices

In Section 4.4 we learned all about diagonalizing a square matrix A. One of the difficulties that arose is that a matrix with real entries can have complex eigenvalues. In this section, we focus on the case where A is a symmetric matrix, and we will show that the eigenvalues of A are always real and that A is always diagonalizable!

Recall that a square matrix A is **symmetric** if $A^T = A$.

Examples: $\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$, $\begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix}$, $\begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$, $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix}$.

Non-examples: $\begin{bmatrix} 3 & -2 \\ 2 & 3 \end{bmatrix}$, $\begin{bmatrix} 1 & 2 & 3 \\ 5 & 4 & 5 \\ 3 & 2 & 6 \end{bmatrix}$.

Example 5.16: If possible, diagonalize $A = \begin{bmatrix} 1 & 2 \\ 2 & -2 \end{bmatrix}$. On board, if time.

Definition: A square matrix A is **orthogonally diagonalizable** if there exists an orthogonal matrix Q such that $Q^T A Q$ is a diagonal matrix D.