Math 2270-6

Notes of 11/15/2019

Least Squares with $A = QR$

- Suppose we write

$$\boxed{A = QR} \tag{1}$$

where

$$Q = m \begin{array}{c} \\ \end{array} \begin{array}{cc} n & m-n \\ \left(Q_1 \right. & \left. Q_2 \right) \end{array} \tag{2}$$

is *orthogonal* and

$$R = \begin{array}{c} n \\ m-n \end{array} \begin{array}{c} n \\ \begin{pmatrix} R_1 \\ 0 \end{pmatrix} \end{array} \tag{3}$$

with R_1 being upper triangular.

- Earlier we discussed how to obtain

$$A = Q_1 R_1,$$

for example by the Gram-Schmidt Process.

- To get Q from Q_1 we simply add vectors to the orthonormal basis of the column space of A to get an orthonormal bass of $\mathbb{R}^m$.

- We won't actually need Q_2, but it's useful to describe the idea.
- An orthogonal matrix Q is one that satisfies

$$Q^{-1} = Q^T \quad \text{or} \quad Q^T Q = Q Q^T = I. \tag{4}$$

Obviously, if Q is orthogonal, so is Q^T.

- A significant property of an orthogonal matrix is that multiplying with it does not alter the norm of a vector:

$$\|Qx\|^2 = (Qx)^T(Qx) = x^T Q^T Q x = x^T x = \|x\|^2. \tag{5}$$

- Thus the first n columns of Q form an orthonormal basis of the column space of A. We obtain

$$\begin{aligned} \|Ax - b\|^2 &= \|Q^T(Ax - b)\|^2 \\ &= \|Q^T Ax - Q^T b\|^2 \\ &= \left\| \begin{pmatrix} R_1 x \\ 0 \end{pmatrix} - \begin{pmatrix} Q_1^T b \\ Q_2^T b \end{pmatrix} \right\|^2 \\ &= \left\| R_1 x - Q_1^T b \right\|^2 + \left\| Q_2^T b \right\|^2 . \end{aligned} \tag{6}$$

- Of the two terms on the right we have no control over the second, and we can render the first one zero by solving (the square triangular $n \times n$ linear system)

$$R_1 x = Q_1^T b. \tag{7}$$

6.7 Inner Product Space

- Today we will generalize the notion of inner products.
- The purpose is to get more flexibility in applications.
- The general process of generalizing a mathematical concept consists of two steps:

1. Decide what are the key properties of that concept.

2. Figure out what else has those properties, and what is implied by these properties.

- We actually followed this procedure when generalizing from $\mathbb{R}^n$ to general vector spaces.
- Recall our **Definition:** A **vector space**[-1-] is a nonempty set V of objects, called vectors, on which are defined two operations, called **addition** and **multiplication by scalars (real numbers)**, subject to the ten axioms (or rules) listed below. The axioms must hold for all vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ in V, and for all scalars c and d.

1. The sum of $\mathbf{u}$ and $\mathbf{v}$, denoted by $\mathbf{u}+\mathbf{v}$, is in V.

2. $\mathbf{u}+\mathbf{v}=\mathbf{v}+\mathbf{u}$.

3. $(\mathbf{u}+\mathbf{v})+\mathbf{w}=\mathbf{u}+(\mathbf{v}+\mathbf{w})$.

[-1-] Also called a **linear space**

4. There is a zero vector $\mathbf{0}$ in V such that

$$\mathbf{u} + \mathbf{0} = \mathbf{u}.$$

5. For each $\mathbf{u}$ in V, there is a vector $-\mathbf{u}$ in V such that $u + (-\mathbf{u}) = \mathbf{0}$.

6. The scalar multiple of $\mathbf{u}$ by c, denoted by $c\mathbf{u}$, is in V.

7. $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$.

8. $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$.

9. $c(d\mathbf{u}) = (cd)\mathbf{u}$.

10. $1\mathbf{u} = \mathbf{u}$.

- After we made that definition we found out that the standard concepts (linear combination, span, spanning set, linear independence, basis, dimension) all made sense in the new context.

- Then we came up with all kinds of space we did not think of before: polynomials, sequences, symmetric matrices, triangular matrices, solution sets of differential equations or difference equations, etc.

- We'll do the same for inner products.

- **Definition** (p. 378, textbook): An **inner product** on a vector space V is a function that, to each pair of vectors $\mathbf{u}$ and $\mathbf{v}$ in V, associates a real number $< \mathbf{u}, \mathbf{v} >$ and satisfies the following axioms, for all vectors $\mathbf{u}$ and $\mathbf{v}$ in V and all scalars c:

1. $< \mathbf{u}, \mathbf{v} > = < \mathbf{v}, \mathbf{u} >$.

2. $< \mathbf{u} + \mathbf{v}, \mathbf{w} > = < \mathbf{u}, \mathbf{w} > + < \mathbf{v}, \mathbf{w} >$

3. $< c\mathbf{u}, \mathbf{v} > = c < \mathbf{u}, \mathbf{v} >$

4. $< \mathbf{u}, \mathbf{u} > \geq 0$ and $< \mathbf{u}, \mathbf{u} > = 0$ if and only if $\mathbf{u} = \mathbf{0}$.

A Vector space with an inner product is called an **inner product space**.

- Clearly, $\mathbb{R}^n$ with our previously defined

$$< \mathbf{u}, \mathbf{v} > = (\mathbf{u}, \mathbf{v}) = \mathbf{v}^T \mathbf{u}$$

  is an inner product space. This is our first example.

- Clearly, we could now do everything we did in this chapter again, except that we use angular brackets instead of parentheses. We would find that almost everything we did, except the computation of actual numerical values, depended only on the properties listed in the above definition.

- So the main task is to identify interesting examples.

- Example 1. Suppose $\mathbb{R} = \mathbb{R}^2$ and

$$< \mathbf{u}, \mathbf{v} > = 4u_1v_1 + 5u_2v_2.$$

- This definition satisfies our definition.

- Example 2, textbook. Let $V = P_n$, the space of polynomials of degree n. Let t_0, t_1, ..., t_n be a set of distinct real numbers. Let p and q be two polynomials in V. Then define

$$< p, q >= \sum_{i=0}^{n} p(t_i) q(t_i).$$

- As before, we define in general

$$\|p\| = \sqrt{< p, p >}.$$

- Compute the norm of some polynomials in example 2.

- Two vectors $\mathbf{u}$ and $\mathbf{v}$ in an inner product space are **orthogonal** with respect to the given inner product if $< \mathbf{u}, \mathbf{v} >= 0$ as before.
- Orthogonal projection and best approximation work as before.
- The Gram-Schmidt process works exactly as before.
- Best approximation works as before.
- The Pythagorean Theorem works as before.

The Cauchy-Schwarz Inequality

$$| < \mathbf{u}, \mathbf{v} > | \leq \|\mathbf{u}\|\|\mathbf{v}\|$$

The Triangle Inequality

- We get, as before

$$\|\mathbf{u}+\mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|.$$

- Example: For continuous functions f and g define

$$< f, g >= \int_0^1 f(x)g(x)\mathrm{d}x.$$

- Start the Gram-Schmidt Process to get orthogonal polynomials.

- Example: Approximate e^x by a linear function.