MATH 2R03 (Linear Algebra II) -- 5 July 2018 -- Lecture 6

Last time: ORTHOGONAL/ORTHONORMAL SETS AND BASES & THE GRAM-SCHMIDT PROCESS

From a set of vectors $\{u_1, \dots, u_k\}$, we can produce an orthonormal set of vectors $\{v_1, \dots, v_k\}$ with $\text{span}\{u_1, \dots, u_k\} = \text{span}\{v_1, \dots, v_k\}$.

$$v_1 = \cancel{u_1} / \|u_1\|, \quad v_2 = \frac{\text{proj}_{\text{span}\{v_1\}^\perp} u_2}{\|\text{proj}_{\text{span}\{v_1\}^\perp} u_2\|} = \frac{u_2 - \langle u_2, v_1\rangle v_1}{\|u_2 - \langle u_2, v_1\rangle v_1\|},$$

$$v_3 = \frac{\text{proj}_{\text{span}\{v_1, v_2\}^\perp} u_3}{\|\text{proj}_{\text{span}\{v_1, v_2\}^\perp} u_3\|} = \frac{u_3 - (\langle u_3, v_1\rangle v_1 + \langle u_3, v_2\rangle v_2)}{\|u_3 - (\langle u_3, v_1\rangle v_1 + \langle u_3, v_2\rangle v_2)\|}, \dots$$

Example: Look at $P_2([0,1])$, space of polynomials on $[0,1]$ of degree ≤ 2, with the integral inner product. Find an orthonormal basis for $P_2([0,1])$ starting with the basis $\{1, x, x^2\}$.

Solution: Step 1: Set $u_1 = 1$, $u_2 = x$, $u_3 = x^2$.

Step 2: $v_1 = \dfrac{u_1}{\|u_1\|} = 1$.

Step 3: Find $\text{proj}_{\text{span}\{v_1\}^\perp} u_2$.

$$= u_2 - \overbrace{\langle u_2, v_1\rangle v_1}^{\text{proj}_{\text{span}\{v_1\}} u_2}$$

$$= x - \left(\int_0^1 x\,dx\right) \cdot 1$$

$$= x - \left[x^2/2\right]_0^1 = x - 1/2.$$

Then set $v_2 = \dfrac{x - 1/2}{\|x - 1/2\|}$ We already found $\|x-1/2\|^2 = 1/12$

So $\|x - 1/2\| = \dfrac{1}{2\sqrt{3}}$

So $v_2 = \dfrac{x - 1/2}{(1/2\sqrt{3})} = 2\sqrt{3}\,x - \sqrt{3}.$

Step 4 Find $u_3 - (\langle u_3, v_1\rangle v_1 + \langle u_3, v_2\rangle v_2)$

$$= x^2 - \left(\left(\int_0^1 x^2\,dx\right) + \left(\int_0^1 2\sqrt{3}\,x^3 - \sqrt{3}\,x^2\,dx\right)(2\sqrt{3}x - \sqrt{3})\right)$$

$$= x^2 - \left(\left[\frac{x^3}{3}\right]_0^1 + \left[\frac{\sqrt{3}}{2}x^4 - \frac{x^3}{\sqrt{3}}\right]_0^1 (2\sqrt{3}x - \sqrt{3})\right)$$

$$= x^2 - \left(\frac{1}{3} + \left(\frac{\sqrt{3}}{2} - \frac{1}{\sqrt{3}}\right)(2\sqrt{3}x - \sqrt{3})\right)$$

$$= x^2 - \frac{1}{3} - \left(3x - 2x - \frac{3}{2} + 1\right)$$

$$= x^2 - \frac{1}{3} - x + \frac{1}{2} = x^2 - x + 1/6.$$

Set $v_3 = \dfrac{x^2 - x + 1/6}{\| x^2 - x + 1/6 \|}$

We found $\| x^2 - x + \frac{1}{6} \|^2 = 1/180.$

$(1/6\sqrt{5})^2$

$$= \frac{x^2 - x + 1/6}{(1/6\sqrt{5})}$$

$$= 6\sqrt{5}\, x^2 - 6\sqrt{5}\, x + \sqrt{5}.$$

We ran out of u_i s. So the orthonormal basis we found is $v_1 = 1$, $v_2 = 2\sqrt{3}\, x - \sqrt{3}$

$$v_3 = 6\sqrt{5}\, x^2 - 6\sqrt{5}\, x + \sqrt{5}.$$

QR Decomposition : Extracting info. about certain matrices from a special case of the Gram-Schmidt Process

If A is an $m \times n$ matrix with n linearly independent column vectors in $\mathbb{R}^m$ (or $\mathbb{C}^m$)

We can apply Gram-Schmidt to these vectors to get an orthonormal set of n vectors in $\mathbb{R}^m$ (respectively $\mathbb{C}^m$).

We can write the original column vectors $u_1, \ldots, u_n$ in terms of the new vectors

$$\left[\begin{pmatrix} | & & | \\ u_1 & ,\ldots, & u_n \\ | & & | \end{pmatrix} = A\right]$$

$v_1, \ldots, v_n$ using the formulas from last time:

$$u_1 = \langle u_1, v_1 \rangle v_1 + \ldots + \langle u_1, v_n \rangle v_n$$
$$\vdots \qquad \vdots \qquad \vdots$$
$$u_n = \langle u_n, v_1 \rangle v_1 + \ldots + \langle u_n, v_n \rangle v_n$$

Rewrite this system as

$$\underbrace{\begin{pmatrix} | & & | \\ u_1 & \cdots & u_n \\ | & & | \end{pmatrix}}_{A} = \underbrace{\begin{pmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{pmatrix}}_{Q} \underbrace{\begin{pmatrix} \langle u_1, v_1 \rangle & \cdots & \langle u_n, v_1 \rangle \\ \vdots & \ddots & \vdots \\ \langle u_1, v_n \rangle & \cdots & \langle u_n, v_n \rangle \end{pmatrix}}_{R}$$

It has orthonormal column vectors

$$Q^T Q = \begin{pmatrix} - & v_1 & - \\ & \vdots & \\ - & v_n & - \end{pmatrix} \begin{pmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{pmatrix} = I$$

because $v_i \cdot v_j = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$

Look at R. At each stage of the G.S. Process, v_i is constructed to be orthogonal to $W_{i-1} = \text{span}\{v_1, \dots, v_{i-1}\} = \text{span}\{u_1, \dots, u_{i-1}\}$

In other words for $j < i$ we have

$\langle u_j, v_i \rangle = 0$. So R is upper triangular.

What about the diagonal entries $\langle u_i, v_i \rangle$?

If $\langle u_i, v_i \rangle = 0$, then we would have

$$u_i = \langle u_i, v_1 \rangle v_1 + \dots + \langle u_i, v_{i-1} \rangle v_{i-1} + \cancel{\langle u_i, v_i \rangle v_i}$$

i.e. u_i in $\text{span}\{v_1, \dots, v_{i-1}\}$
$= \text{span}\{u_1, \dots, u_{i-1}\}$

But the us, the columns of A, are linearly independent. So this cannot in fact happen.

So R is invertible.

To summarise:

If A is an $m \times n$ matrix with n linearly independent column vectors, then A can decomposed into a product $A = QR$, (Q $m \times n$, R $n \times n$), where the columns of Q are orthonormal and R is invertible and upper triangular.

QR DECOMPOSITION

Exercise Find a QR Decomposition of

$$A = \begin{pmatrix} 3 & 5 \\ 4 & 15 \\ 0 & 12 \end{pmatrix}.$$

Projection as Best Approximation

Theorem V – inner product space
W – finite-dim. subspace

For any u in V $\text{proj}_W u$ is the closest vector in W to u i.e. if w is another vector in W (other than $\text{proj}_W u$) then

$$\|u-w\| > \|u - \text{proj}_W u\|.$$

Application : Least squares Problem

A – $m \times n$ matrix over $\mathbb{R}$,
b in $\mathbb{R}^m$

The linear system $Ax=b$ may or may not have a solution.

Goal: Find x so that Ax is a close to b as possible i.e. $\|Ax-b\|$ is minimized.

i.e. if $Ax-b = \begin{pmatrix} e_1 \\ \vdots \\ e_m \end{pmatrix}$, then want to minimize $\sqrt{e_1^2+\ldots+e_m^2}$
hence "least squares"

Really what is varying is $Ax = x_1u_1+\ldots+x_nu_n$
where $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, $A = \begin{pmatrix} | & & | \\ u_1 & \cdots & u_n \\ | & & | \end{pmatrix}$

i.e.
Ax ranges over linear combinations of the columns of A i.e. over the column space, call it W.

So, by the Theorem, really what we're looking for is $\text{proj}_W b$ (element in the column space closest to b).

Since $b = \overbrace{\text{proj}_W b}^{Ax} + \text{proj}_{W^\perp} b$, we want x satisfying $Ax - b = -\text{proj}_{W^\perp} b$

i.e. with $Ax-b$ orthogonal to W.

The space orthogonal to the column space of A is the null space of A^T.

So in other words we want $Ax-b$ to satisfy $A^T(Ax-b) = 0$

i.e. $\boxed{A^TAx = A^Tb}$ ← Normal equation(s) of the system $Ax=b$.

When we have solutions to the normal equation, these are the "least squares solutions" and the "least squares error" is $\|Ax-b\|$ (where x is a least squares solution).

The solution to the normal equation(s) is unique when A^TA is invertible, which happens if A has linearly independent column vectors.

Connection to QR Decomposition

If A is an $m \times n$ matrix with linearly independent columnvectors, with QR decomposition $A = QR$ then for each b in $\mathbb{R}^m$ the system $Ax = b$ has a unique least squares solution:

Normal equation: $A^T A x = A^T b$

$$
\begin{aligned}
x &= (A^T A)^{-1} A^T b \\
&= ((QR)^T (QR))^{-1} (QR)^T b \\
&= (R^T Q^T Q R)^{-1} R^T Q^T b \\
&= (R^T R)^{-1} R^T Q^T b \quad \text{as } Q^T Q = I \\
&= R^{-1} (R^T)^{-1} R^T Q^T b
\end{aligned}
$$

$$\boxed{x = R^{-1} Q^T b}.$$