MA 405, 4/5/17

John M. Lynch*

April 5, 2017

A=QR

<u>Recall</u>: Gram-Schmidt process for transforming a set of independent column vectors $\{v_1...v_n\}$ into a set of *orthonormal* vecs $\{q_1...q_n\}$

Works by successively projecting onto <u>lines</u> and subtracting.

<u>ex</u>

$$v_1 = \begin{bmatrix}1\\0\\1\end{bmatrix}, v_2 = v_3 = \tag{1}$$

which yields

$$q_1 = , q_2 = , q_3 = \tag{2}$$

check: these vecs are orthonormal.

Useful for <u>matrix factorization</u>!

example:

$$A = LU, \tag{3}$$

useful for repeatedly solving $A\vec{x} = \vec{b}$ for many $\vec{b}$

Different factorization:

$A = QR$, Q orthonormal columns, R upper triangular

Useful for solving least-squares problems for many data vectors.

How to get this factorization?

<u>ex</u>

$$A = \begin{bmatrix}v_1| & v_2| & v_3|\end{bmatrix} \in \Re^{3\times 3}, B = \begin{bmatrix}q_1| & q_2| & q_3|\end{bmatrix} \in \Re^{3\times 3} \tag{4}$$

What should R equal if we want A=QR?

$$v_1 \& q_1 : \tag{5}$$

They are already in the same direction:

$$\vec{v_1} = ()\vec{q_1} \tag{6}$$

is its projection onto $\vec{q_1}$!

Now, how to find the projection coefficient:

$$\vec{v_1} = (\frac{\vec{q_1}^T\vec{v_1}}{\vec{q_1}^T\vec{q_1}})\vec{q_1} \tag{7}$$

And since $\vec{q_1}$ has length 1,

$$\boxed{\vec{v_1} = (\vec{q_1}^T\vec{v_1})\vec{q_1}} \tag{8}$$

How are $\vec{v_2}$ and $\vec{q_2}$ related to each other?

Well, $\vec{v_2}$ has a part in the direction of $\vec{q_1}$, but all the rest of it is in the direction of $\vec{q_2}$ (because of the G-Sch process).

So,

$$\vec{v_2} = () \cdot \vec{q_1} + () \cdot \vec{q_2} \tag{9}$$

all of $\vec{v_2}$ is in one of these two directions...

Coefficients are found by *projections!*

$\vec{v_3}$ has parts in directions of q1 and q2, but all the rest is in direction of q3!

$$v_3 = (q1Tv3)q1 + (q2Tv3)q2 + (q3Tv3)q3 \tag{10}$$

Now we know how all vectors v are related to all vectors q. This should tell us what R needs to be.

Since q1 and v1 are related so simply,

$$R = \begin{bmatrix} q_1^Tv_1 & q_1^Tv_2 & q_1^Tv_3 \\ 0 & q_2^Tv_2 & q_2^Tv_3 \\ 0 & 0 & q_3^Tv_3 \end{bmatrix} \tag{11}$$

col R1 says you only first vector v1 to get to vector q1. 0 in R32 bc q2 doesn't depend on v3 at all.

Check: compare jth column of A to jth column of QR:

$$(QR)_j = Q \cdot (R)_{jthcol} = Q \begin{bmatrix} q1Tvj \\ q2Tvj \\ qjTvj \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

This is a linear combination of the different columns of Q, so it's also

$$(q1Tvj)q1 + (q2Tvj)q2 + ... + (qjTvj)qj + 0 \cdot q_{j+1} + 0 + 0 \cdot q_n \tag{12}$$

Now compare to v_j, the jth column of A, to the above. They're equal! So

<u>Theorem</u>: any A$\in \Re^{m\times n}$ having independent columns will have a "QR" factorization

$$A = Q \cdot R \tag{13}$$

where $Q \in \Re^{m\times n}$ has orthonormal columns and $R \in \Re^{n\times n}$ has upper-triangular structure and is <u>invertible</u>

Facts about orthonormal matrices, Q

1. $(Q^TQ)_{ij} = \sum_{k=1}^n Q^T_{ithrow} \cdot Q_{jthcol} = (Q_{ithcol})^T \cdot (Q_{jthcol}) = 0 i \neq j, 1 i = j$. i.e,
2. $Q^TQ = I$
3. If Q is square, Q is invertible, <u>and</u> $Q^{-1} = Q^T$
4. $||Q\vec{x}|| = ||\vec{x}||$ (Q doesn't change the length of things)
5. $(Q\vec{x})^T(Q\vec{y}) = \vec{x}^TQ^TQ\vec{y} = \vec{x}^T\vec{y}$ (Q preserves inner products)
6. <u>projections</u>: since q's are all orthogonal, they are a <u>basis</u> for Col(Q). Any vector $\vec{b}\in$Col(Q) has a linear combination $\vec{b} = c1q1 + c2q2 + ... + cnqn =$ a sum of <u>projections</u> of $\vec{b}$ onto 1D lines formed by the basis vectors q_j (where each of the coefficients can be found by <u>projection</u>: $c_j = (\vec{q_j}^T\vec{b})$)

*Undergraduate ECE/Physics, NCSU, Raleigh, NC 27705. E-Mail: jmlynch3@ncsu.edu

using QR to solve least squares

Suppose

$$A\vec{x}=\vec{b}$$

where $A \in \Re^{m\times n}$, overdetermined system. No exact solution, but the columns of A are independent.

So, as a least-square problem:

$$\hat{x}=\min||A\hat{x}-\vec{b}||^2$$

which can be found by solving the Normal equations:

$$A^TA\hat{x}=A^T\vec{b}$$

If I had a QR factorization,

$$(QR)^T(QR)\hat{x}=(QR)^T\vec{b}$$
$$R^TQ^TQR\hat{x}=R^TQ^T\vec{b}$$

but since Q has orthonormal columns, $Q^TQ=I$, and

$$R^TR\hat{x}=R^TQ^T\vec{b} \tag{14}$$

and, since R is invertible, so R^T is, too:

$$R^{-T}(R^TR\hat{x})=R^{-T}(R^TQ^T\vec{b})$$
$$\boxed{R\hat{x}=Q^T\vec{b}}$$

Now just solve *that* linear system for $\hat{x}$. But, since R is upper triangular, this is a triangular system, which is way easier to solve! Particularly for many $\vec{b}$ vectors.

End of Midterm Material

Next class: Chapter 4