Math 415 - Lecture 26

Orthogonal Matrices and QR Decomposition

Monday October 26th 2015

Textbook reading: Chapter 3.4

Suggested practice exercises: 3.4: 13, 16, 17, 18. 13,

Khan Academy video: Gram-Schmidt Example

Strang lecture: Orthogonal Matrices and Gram-Schmidt Process.

1 Review

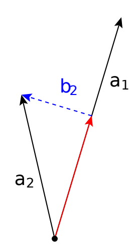
- Vectors $\mathbf{q}_1, \dots, \mathbf{q}_n$ are orthonormal if

$$\mathbf{q}_i^T \mathbf{q}_j = \begin{cases} 0, & \text{if } i \neq j, \\ 1, & \text{if } i = j. \end{cases}$$

- **Gram-Schmidt** orthonormalization: input: basis $\mathbf{a}_1, \dots, \mathbf{a}_n$ for V. output: orthonormal basis $\mathbf{q}_1, \dots, \mathbf{q}_n$ for V.

$$\begin{aligned} \mathbf{b}_1 &= \mathbf{a}_1, & \mathbf{q}_1 &= \tfrac{\mathbf{b}_1}{\|\mathbf{b}_1\|} \\ \mathbf{b}_2 &= \mathbf{a}_2 - \langle \mathbf{a}_2, \mathbf{q}_1 \rangle \mathbf{q}_1, & \mathbf{q}_2 &= \tfrac{\mathbf{b}_2}{\|\mathbf{b}_2\|} \\ \mathbf{b}_3 &= \mathbf{a}_3 - \langle \mathbf{a}_3, \mathbf{q}_1 \rangle \mathbf{q}_1 - \langle \mathbf{a}_3, \mathbf{q}_2 \rangle \mathbf{q}_2, & \mathbf{q}_3 &= \tfrac{\mathbf{b}_3}{\|\mathbf{b}_3\|} \\ & \dots & & \dots \end{aligned}$$

[-1cm]



Fact 1. *if A is any matrix A^TA is the matrix of dot products of the columns of A: Write $A = [\mathbf{a_1}, \ldots . \mathbf{a_n}]$ then*

$$A^TA = \begin{bmatrix} \mathbf{a_1}\cdot\mathbf{a_1} & \mathbf{a_1}\cdot\mathbf{a_2} & \mathbf{a_1}\cdot\mathbf{a_3} & \ldots \\ \mathbf{a_2}\cdot\mathbf{a_1} & \mathbf{a_2}\cdot\mathbf{a_2} & \mathbf{a_2}\cdot\mathbf{a_3} & \ldots \\ \mathbf{a_3}\cdot\mathbf{a_1} & \mathbf{a_3}\cdot\mathbf{a_2} & \mathbf{a_3}\cdot\mathbf{a_3} & \ldots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

Theorem 1. *The columns of Q are orthonormal $\iff Q^TQ = I$*

Definition. An **orthogonal matrix** is a square matrix Q with orthonormal columns.

2 The QR decomposition

In linear algebra "everything" is a matrix factorization.

- Gaussian elimination in terms of matrices: $A = LU$
- Gram-Schmidt in terms of matrices $A = QR$

Theorem 2 (QR decomposition)**.** *Let A be an $m\times n$ matrix of rank n. There is is a orthogonal matrix $m\times n$-matrix Q and an upper triangular $n\times n$ invertible matrix R such that*

$$A = QR.$$

Idea. Gram-Schmidt on the columns of A to get columns of Q.

Recipe
In general, to obtain $A = QR$:

- Gram-Schmidt on (columns of) A, to get (columns of) Q.
- Then $R = Q^TA$.

The resulting R is indeed upper triangular, and we get:

$$\begin{bmatrix} | & | & \\ \mathbf{a}_1 & \mathbf{a}_2 & \cdots \\ | & | & \end{bmatrix} = \begin{bmatrix} | & | & \\ \mathbf{q}_1 & \mathbf{q}_2 & \cdots \\ | & | & \end{bmatrix} \begin{bmatrix} \mathbf{q}_1^T\mathbf{a}_1 & \mathbf{q}_1^T\mathbf{a}_2 & \mathbf{q}_1^T\mathbf{a}_3 & \cdots \\ & \mathbf{q}_2^T\mathbf{a}_2 & \mathbf{q}_2^T\mathbf{a}_3 & \\ & & \mathbf{q}_3^T\mathbf{a}_3 & \\ & & & \ddots \end{bmatrix}$$

It should be noted that, no extra work is needed for computing R: all the inner products in R have been computed during Gram-Schmidt. (Just the LU decomposition encodes the steps of Gaussian elimination, the QR decomposition encodes the steps of Gram-Schmidt.)

Example 3. Find the QR decomposition of $A = \begin{bmatrix} 1 & 2 & 4 \\ 0 & 0 & 5 \\ 0 & 3 & 6 \end{bmatrix}$.

Solution. We apply Gram-Schmidt to the columns of A:

$$\begin{bmatrix}1\\0\\0\end{bmatrix} = \mathbf{q}_1$$
$$\begin{bmatrix}2\\0\\3\end{bmatrix} - \langle \begin{bmatrix}2\\0\\3\end{bmatrix}, \mathbf{q}_1\rangle \mathbf{q}_1 = \begin{bmatrix}0\\0\\3\end{bmatrix}, \quad \begin{bmatrix}0\\0\\1\end{bmatrix} = \mathbf{q}_2$$
$$\begin{bmatrix}4\\5\\6\end{bmatrix} - \langle \begin{bmatrix}4\\5\\6\end{bmatrix}, \mathbf{q}_1\rangle \mathbf{q}_1 - \langle \begin{bmatrix}4\\5\\6\end{bmatrix}, \mathbf{q}_2\rangle \mathbf{q}_2 = \begin{bmatrix}0\\5\\0\end{bmatrix}, \quad \begin{bmatrix}0\\1\\0\end{bmatrix} = \mathbf{q}_3$$

Solution (continued)**.** Hence: $Q = [\mathbf{q}_1 \ \mathbf{q}_2 \ \mathbf{q}_3] = \begin{bmatrix}1&0&0\\0&0&1\\0&1&0\end{bmatrix}$. Note Q is a permutation matrix and so orthogonal. Why? Q has orthonormal columns so $Q^TQ = 1$! To find R in $A = QR$, note that $Q^TA = Q^TQR = R$.

$$R = Q^TA = \begin{bmatrix}1&0&0\\0&0&1\\0&1&0\end{bmatrix}\begin{bmatrix}1&2&4\\0&0&5\\0&3&6\end{bmatrix} = \begin{bmatrix}1&2&4\\0&3&6\\0&0&5\end{bmatrix}$$

Note R is upper triangular. Summarizing, we have

$$\begin{bmatrix}1&2&4\\0&0&5\\0&3&6\end{bmatrix} = \begin{bmatrix}1&0&0\\0&0&1\\0&1&0\end{bmatrix}\begin{bmatrix}1&2&4\\0&3&6\\0&0&5\end{bmatrix}$$

Example 4. Find the QR decomposition of $A = \begin{bmatrix}1&1&2\\0&0&1\\1&0&0\end{bmatrix}$.

3 Applications of $A = QR$

3.1 Using QR to solve systems of equations

QR decomposition can be used to solve systems of linear equations.

$$\begin{aligned} A\mathbf{x} = \mathbf{b} \quad &\Longleftrightarrow \quad QR\mathbf{x} = \mathbf{b} \\ &\Longleftrightarrow \quad R\mathbf{x} = Q^T\mathbf{b} \end{aligned}$$

$R\mathbf{x} = Q^T\mathbf{b}$ is triangular, so solve it by back substitution. QR is a little slower than LU, but makes up in numerical stability.

Theorem 2. *Let A be matrix with linear independent columns. Suppose $A\mathbf{x} = \mathbf{b}$ has no solution. Then the solution of $R\mathbf{x} = Q^T\mathbf{b}$ is the least square solution of $A\mathbf{x} = \mathbf{b}$.*

Proof.

$$\begin{aligned} A^T A\hat{\mathbf{x}} = A^T\mathbf{b} \quad &\iff \quad \underbrace{(QR)^T QR}_{=R^T Q^T QR}\,\hat{\mathbf{x}} = (QR)^T\mathbf{b} \\ &\iff \quad R^T R\hat{\mathbf{x}} = R^T Q^T\mathbf{b} \\ &\iff \quad R\hat{\mathbf{x}} = Q^T\mathbf{b} \end{aligned}$$

Again, this is triangular, solved by back substitution.

$\hat{\mathbf{x}}$ is a least square solution of $A\mathbf{x} = \mathbf{b} \iff R\hat{\mathbf{x}} = Q^T\mathbf{b}$ (where $A = QR$)

□

Remark. $R\mathbf{x} = Q^T\mathbf{b}$ always gives the best possible solution to $A\mathbf{x} = \mathbf{b}$.

Example 5. Let $A = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 2 & 1 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$.Find the least square solution of $A\mathbf{x} = \mathbf{b}$ using QR-decomposition.

Solution. Let us first apply Gram-Schmidt to the columns of A. (We are going to work first with unnormalized vectors, and normalize at the end. Check that this also works!) We have $\mathbf{b_1} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$ and $\mathbf{b_2} = \mathbf{a_2} - \frac{\langle \mathbf{a_2}, \mathbf{b_1} \rangle}{\langle \mathbf{b_1}, \mathbf{b_1} \rangle}\mathbf{b_1} = \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix} - \frac{9}{9}\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$.

Solution (continued)**.** Normalizing we get

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

We have $A = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 2 & 1 \end{bmatrix}$, and $Q = \begin{bmatrix} 1/3 & 0 \\ 2/3 & 1/\sqrt{2} \\ 2/3 & -1/\sqrt{2} \end{bmatrix}$. Then

$$R = \begin{bmatrix} \mathbf{q_1} \cdot \mathbf{a_1} & \mathbf{q_1} \cdot \mathbf{a_2} \\ 0 & \mathbf{q_2} \cdot \mathbf{a_2} \end{bmatrix} = \begin{bmatrix} 3 & 3 \\ 0 & \sqrt{2} \end{bmatrix}.$$

Now $A\mathbf{x} = \mathbf{b}$ is not consistent.

Solution. So we do least squares, but in this case ($A = QR$) we know the normal equations are

$$R\hat{\mathbf{x}} = Q^T\mathbf{b}, \qquad \begin{bmatrix} 3 & 3 \\ 0 & \sqrt{2} \end{bmatrix}\hat{\mathbf{x}} = \begin{bmatrix} 1/3 & 2/3 & 2/3 \\ 0 & 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/3 \\ 0 \end{bmatrix}$$

So $\hat{\mathbf{x}} = \begin{bmatrix} 1/9 \\ 0 \end{bmatrix}$, and $\hat{\mathbf{b}} = A\hat{\mathbf{x}} = 1/9 \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$.