COM521500
Math. Methods for SP I
Lecture 9: QR Decomposition

QR Decomposition

Given a matrix $\mathbf{A} \in \mathbb{C}^{m \times n}$, we can have the following decomposition

$$\mathbf{A} = \mathbf{QR}$$

where $\mathbf{Q} \in \mathbb{C}^{m \times m}$ is unitary, and $\mathbf{R} \in \mathbb{C}^{m \times n}$ is upper triangular.

Classical Gram-Schmidt

For simplicity, assume for the time being that $\mathbf{A}$ has full column rank.

Set

$$\mathbf{q}_1 = \mathbf{a}_1/\|\mathbf{a}_1\|_2, \qquad r_{11} = \|\mathbf{a}_1\|_2$$

Consider finding $\mathbf{q}_2$. Since

$$\mathbf{a}_2 = r_{12}\mathbf{q}_1 + r_{22}\mathbf{q}_2,$$

and $\mathbf{q}_1^H\mathbf{q}_2 = 0$, we have that

$$r_{12} = \mathbf{q}_1^H\mathbf{a}_2.$$

Institute Comm. Eng. & Dept. Elect. Eng., National Tsing Hua University

3

Subsequently,

$$\mathbf{q}_2 = \mathbf{y}_2/\|\mathbf{y}_2\|_2, \qquad r_{22} = \|\mathbf{y}_2\|_2$$

where

$$\mathbf{y}_2 = \mathbf{a}_2 - r_{12}\mathbf{q}_1$$

Repeating the above steps, we have, for $k = 2, \ldots, n$,

$$r_{ik} = \mathbf{q}_i^H\mathbf{a}_k, \qquad i = 1, \ldots, k-1$$

$$\mathbf{y}_k = \mathbf{a}_k - \sum_{i=1}^{k-1} r_{ik}\mathbf{q}_i$$

$$\mathbf{q}_k = \mathbf{y}_k/\|\mathbf{y}_k\|_2, \qquad r_{kk} = \|\mathbf{y}_k\|_2$$

Institute Comm. Eng. & Dept. Elect. Eng., National Tsing Hua University

4

To complete the above Gram-Schmidt (GS) procedure, we need

$$\mathbf{y}_k \neq \mathbf{0}, \quad \forall k$$

for all k. This is true when $\mathbf{A}$ is of full column rank.

The GS procedure leads to a thin QR decomposition

$$\mathbf{A} = \mathbf{Q}_1 \mathbf{R}_1$$

where $\mathbf{Q}_1 \in \mathbb{C}^{m \times n}$ has orthonormal columns, and $\mathbf{R}_1 \in \mathbb{C}^{n \times n}$ is upper triangular.

We can always find a matrix $\mathbf{Q}_2 \in \mathbb{C}^{(m-n) \times n}$ so that

$$\mathbf{Q} = [\ \mathbf{Q}_1\ \mathbf{Q}_2\ ]$$

is unitary.

We can then form a complete QR decomposition

$$\mathbf{A} = [\ \mathbf{Q}_1\ \mathbf{Q}_2\ ] \begin{bmatrix} \mathbf{R}_1 \\ \mathbf{0} \end{bmatrix}$$

Theorem 9.1 If $\mathbf{A}$ has full column rank, then the decomposition

$$\mathbf{A} = \mathbf{Q}_1 \mathbf{R}_1$$

where $\mathbf{R}_1 \in \mathbb{C}^{n \times n}$ is upper triangular with *positive diagonal entries*, $\mathbf{Q}_1 \in \mathbb{C}^{m \times n}$ is semi-unitary, is unique.

Proof: Since

$$\mathbf{A}^H \mathbf{A} = \mathbf{R}_1^H \mathbf{R}_1,$$

$\mathbf{R}_1^H$ is also the Cholesky factor which is unique (cf., Theorem 7.3). It follows that $\mathbf{Q}_1 = \mathbf{A} \mathbf{R}_1^{-1}$ is also unique.

Some properties resulting from the GS procedure:

Property 9.1 For full column rank $\mathbf{A}$,

$$\mathrm{span}\{\mathbf{a}_1, \ldots, \mathbf{a}_k\} = \mathrm{span}\{\mathbf{q}_1, \ldots, \mathbf{q}_k\}$$

for $k = 1, \ldots, n$.

Property 9.2 For full column rank $\mathbf{A}$,

$$R(\mathbf{A}) = R(\mathbf{Q}_1), \qquad R_\perp(\mathbf{A}) = R(\mathbf{Q}_2)$$

Property 9.3 For full column rank $\mathbf{A}$, $\mathbf{R}_1$ is nonsingular.

For rank deficient $\mathbf{A}$, a QR pair can be constructed.

In the GS procedure, do the following the modification

$$\mathbf{q}_k = \mathbf{0}, \quad r_{kk} = 0 \qquad \text{if } \mathbf{y}_k = \mathbf{0}.$$

for $k = 1, \ldots, n$.

Let $\mathbf{Q}_1' \in \mathbb{C}^{m \times r}$ be a matrix containing the nonzero $\mathbf{q}_k$'s.
Let $\mathbf{Q}_2' = [\, \mathbf{q}_{r+1}', \ldots, \mathbf{q}_n' \,]$ be a matrix so that $[\, \mathbf{Q}_1' \ \mathbf{Q}_2' \,]$ is unitary.

Replace the first zero column vector $\mathbf{q}_k$ by $\mathbf{q}_{r+1}'$, the 2nd $\mathbf{q}_k$ by $\mathbf{q}_{r+2}'$, and so on. A QR decomposition will then be obtained.

Application: Full column rank LS

Consider again the LS problem

$$\min_{\mathbf{x} \in \mathbb{C}^n} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2$$

where $\mathbf{A}$ is of full column rank.

Let

$$\mathbf{Q}^H \mathbf{b} = \begin{bmatrix} \mathbf{Q}_1^H \\ \mathbf{Q}_2^H \end{bmatrix} \mathbf{b} = \begin{bmatrix} \mathbf{c} \\ \mathbf{d} \end{bmatrix}$$

We have that

$$\|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2 = \|\mathbf{Q}^H(\mathbf{A}\mathbf{x} - \mathbf{b})\|_2^2$$

$$= \left\| \begin{bmatrix} \mathbf{R}_1 \\ \mathbf{0} \end{bmatrix} \mathbf{x} - \begin{bmatrix} \mathbf{c} \\ \mathbf{d} \end{bmatrix} \right\|_2^2$$

$$= \|\mathbf{R}_1\mathbf{x} - \mathbf{c}\|_2^2 + \|\mathbf{d}\|_2^2$$

Hence, the LS solution is

$$\mathbf{x}_{LS} = \mathbf{R}_1^{-1}\mathbf{c}$$

In practice, the classical GS procedure is numerically unstable.

There are several other ways of finding QR factors. In this course we consider

modified GS

Householder transform

Givens rotations.

Modified GS

Partition

$$\mathbf{R}_1 = \begin{bmatrix} \mathbf{r}_1^T \\ \vdots \\ \mathbf{r}_n^T \end{bmatrix}$$

The decomposition of $\mathbf{A}$ can be expressed as

$$\mathbf{A} = \mathbf{QR} = \sum_{i=1}^{n} \mathbf{q}_i \mathbf{r}_i^T$$

At stage k, consider

$$\mathbf{A} - \sum_{i=1}^{k-1} \mathbf{q}_i \mathbf{r}_i^T = \sum_{i=k}^{n} \mathbf{q}_i \mathbf{r}_i^T$$
$$= [\ \underbrace{\mathbf{0} \dots \mathbf{0}}_{k-1}\ r_{kk}\mathbf{q}_k,\ \ r_{k,k+1}\mathbf{q}_k + r_{kk}\mathbf{q}_k, \dots\ ]$$

Thus,

$$\mathbf{y}_k = (\mathbf{A} - \textstyle\sum_{i=1}^{k-1} \mathbf{q}_i \mathbf{r}_i^T)\mathbf{e}_k$$
$$\mathbf{q}_k = \mathbf{y}_k / \|\mathbf{y}_k\|_2,$$
$$\mathbf{r}_k^T = \mathbf{q}_k^H (\mathbf{A} - \textstyle\sum_{i=1}^{k-1} \mathbf{q}_i \mathbf{r}_i^T)$$

Unlike the classical GS, the modified GS has been found to be numerically very stable.

Householder QR

It is easier to look at the real-value case; the idea in the complex case is more or less the same.

To see how the Householder QR algorithm works, it is necessary to understand the concepts of reflection matrices and Householder transformation.

A matrix is called a **reflection matrix** if it is given by

$$\mathbf{H} = \mathbf{I} - 2\mathbf{P}$$

where $\mathbf{P} \in \mathbb{R}^{m \times m}$ is an orthogonal projection matrix.

A reflection matrix is symmetric, and orthogonal.

Let $\mathbf{x} = \mathbf{P}\mathbf{x} + \mathbf{P}_\perp\mathbf{x}$. We have that

$$\mathbf{H}\mathbf{x} = -\mathbf{P}\mathbf{x} + \mathbf{P}_\perp\mathbf{x}$$

where the component $\mathbf{P}\mathbf{x}$ is reflected.

Householder transformation:

The Householder transformation is to find a matrix, denoted by $\mathbf{H}$ so that

$$\mathbf{H}\mathbf{x} = \beta\mathbf{e}_1,$$

i.e., the elements of $\mathbf{H}\mathbf{x}$ is eliminated except for the 1st.

Let

$$\mathbf{P} = \boldsymbol{\nu}(\boldsymbol{\nu}^T\boldsymbol{\nu})^{-1}\boldsymbol{\nu}^T$$

where $\boldsymbol{\nu}$ is to be determined.

Now,

$$\mathbf{H}\mathbf{x} = \mathbf{x} - \frac{2\boldsymbol{\nu}^T\mathbf{x}}{\boldsymbol{\nu}^T\boldsymbol{\nu}}\boldsymbol{\nu} \in \mathrm{span}\{\mathbf{x}, \boldsymbol{\nu}\}$$

If $\mathbf{H}\mathbf{x} = \beta\mathbf{e}_1$ is what we want, then

$$\boldsymbol{\nu} = \mathbf{x} + \alpha\mathbf{e}_1$$

for some coefficient α.

If we set $\alpha = \|\mathbf{x}\|_2$, then it can be verified that

$$\mathbf{H}\mathbf{x} = -\|\mathbf{x}\|_2\mathbf{e}_1$$

Note that a similar effect occurs when $\alpha = -\|\mathbf{x}\|_2$

Householder QR:

Let $\mathbf{H}_1 \in \mathbb{R}^{m \times m}$ be the Householder transformation for $\mathbf{a}_1$.

Then,

$$\mathbf{A}_1 = \mathbf{H}_1 \mathbf{A}$$

$$= \begin{bmatrix} \times & \times & \dots & \times \\ 0 & \times & \dots & \times \\ \vdots & \times & \dots & \times \\ 0 & \times & \dots & \times \end{bmatrix}$$

Let

$$\mathbf{H}_2 = \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{H}}_2 \end{bmatrix}$$

where $\tilde{\mathbf{H}}_2 \in \mathbb{R}^{m-1 \times m-1}$ is the Householder transform for $\mathbf{A}_1(2:m, 2)$. Then,

$$\mathbf{A}_2 = \mathbf{H}_2 \mathbf{A}_1$$

$$= \begin{bmatrix} \times & \times & \times & \dots & \times \\ 0 & \times & \times & \dots & \times \\ \vdots & 0 & \times & \dots & \times \\ \vdots & \vdots & \times & \dots & \times \\ 0 & 0 & \times & \dots & \times \end{bmatrix}$$

It follows that at stage k,

$$\mathbf{A}_k = \mathbf{H}_k \mathbf{A}_{k-1}$$

where

$$\mathbf{H}_k = \begin{bmatrix} \mathbf{I}_{k-1} & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{H}}_k \end{bmatrix},$$

and $\tilde{\mathbf{H}}_k \in \mathbb{R}^{(m-k+1)\times(m-k+1)}$ is the Householder transformation for $\mathbf{A}_{k-1}(k:m,k)$.

The process continues for $n-1$ stages. We then obtain

$$\mathbf{A}_{n-1} = \mathbf{H}_{n-1}\ldots\mathbf{H}_2\mathbf{H}_1\mathbf{A}$$

The matrix $\mathbf{A}_{n-1} \triangleq \mathbf{R}$ is upper triangular.

Since each $\mathbf{H}_k$ is orthogonal, $\prod_{i=1}^{n-1} \mathbf{H}_i \triangleq \mathbf{Q}$ is also orthogonal.

Givens Rotations

Again, in the Givens QR decomposition, it is instructive to look at the real case.

Let

$$
\mathbf{J}_{ik}(\theta) = \begin{bmatrix} \mathbf{I} & & & & \\ & c(\theta) & & s(\theta) & \\ & & \mathbf{I} & & \\ & -s(\theta) & & c(\theta) & \\ & & & & \mathbf{I} \end{bmatrix}
$$

where $c(\theta) = \cos(\theta)$, $s(\theta) = \sin(\theta)$, the (i, k) entry is $-s(\theta)$, and so on.

A Givens rotation $\mathbf{J}_{ik}(\theta)$ is orthogonal.

For example,

$$
\mathbf{J}_{21}(\theta)\mathbf{J}_{21}^T(\theta) = \begin{bmatrix} c(\theta) & s(\theta) & \\ -s(\theta) & c(\theta) & \\ & & \mathbf{I} \end{bmatrix} \begin{bmatrix} c(\theta) & -s(\theta) & \\ s(\theta) & c(\theta) & \\ & & \mathbf{I} \end{bmatrix}
$$

$$
= \begin{bmatrix} c^2(\theta) + s^2(\theta) & -c(\theta)s(\theta) + c(\theta)s(\theta) & \\ -c(\theta)s(\theta) + c(\theta)s(\theta) & c^2(\theta) + s^2(\theta) & \\ & & \mathbf{I} \end{bmatrix}
$$

$$
= \begin{bmatrix} \mathbf{I} & \\ & \mathbf{I} \end{bmatrix}
$$

Consider $\mathbf{J}_{21}(\theta)\mathbf{A}$:

$$
\mathbf{J}_{21}(\theta)\mathbf{A} = \begin{bmatrix} c(\theta) & s(\theta) & \\ -s(\theta) & c(\theta) & \\ & & \mathbf{I} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \dots \\ a_{21} & a_{22} & \dots \\ \vdots & \vdots & \end{bmatrix}
$$

$$
= \begin{bmatrix} \times & \times & \dots \\ -s(\theta)a_{11} + c(\theta)a_{21} & \times & \dots \\ \vdots & \vdots & \end{bmatrix}
$$

Set $\theta = \tan^{-1}(a_{21}/a_{11})$. Then

$$
-s(\theta)a_{11} + c(\theta)a_{21} = 0.
$$

For notational simplicity, let $\mathbf{J}_{ik} = \mathbf{J}_{ik}(\theta)$ where θ is chosen to annihilate the (i,k) entry of the transformed matrix.

By performing a sequence of Givens rotations to annihilate the lower triangular parts of $\mathbf{A}$, we obtain

$$
\underbrace{\mathbf{J}_{n,n-1} \dots (\mathbf{J}_{n,2} \dots \mathbf{J}_{32})(\mathbf{J}_{n1} \dots \mathbf{J}_{21})}_{\mathbf{Q}^T} \mathbf{A} = \mathbf{R}
$$

which is upper triangular.