SI231b: Matrix Computations

Lecture 12: Computations of QR Factorization

Yue Qiu

qiuyue@shanghaitech.edu.cn

School of Information Science and Technology

ShanghaiTech University

Oct. 19, 2022

One of the Top 10 Algorithms in the 20th Century[1]

Given a rectangular matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{A}$ can be factorized into the form

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

where

- $\mathbf{Q} \in \mathbb{R}^{m \times m}$ is an orthogonal matrix
- $\mathbf{R} \in \mathbb{R}^{m \times n}$ is upper-triangular

Reduced QR Factorization

For $m > n$, the reduced QR factorization given by

- $\mathbf{Q} \in \mathbb{R}^{m \times n}$ has orthonormal columns
- $\mathbf{R} \in \mathbb{R}^{n \times n}$ is upper-triangular
- also called 'economic' QR factorization in some cases

[1] https://doi.ieeecomputersociety.org/10.1109/MCISE.2000.814652

- a matrix $\mathbf{H} \in \mathbb{R}^{m \times m}$ is called a reflection matrix if

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

where $\mathbf{P}$ is an orthogonal projector.

- interpretation: denote $\mathbf{P}^{\perp} = \mathbf{I} - \mathbf{P}$, and observe

$$\mathbf{x} = \mathbf{P}\mathbf{x} + \mathbf{P}^{\perp}\mathbf{x}, \qquad \mathbf{H}\mathbf{x} = -\mathbf{P}\mathbf{x} + \mathbf{P}^{\perp}\mathbf{x}.$$

The vector $\mathbf{H}\mathbf{x}$ is a reflected version of $\mathbf{x}$, with $\mathcal{R}(\mathbf{P}^{\perp})$ being the "mirror"

- a reflection matrix is orthogonal:

$$\mathbf{H}^T\mathbf{H} = (\mathbf{I} - 2\mathbf{P})(\mathbf{I} - 2\mathbf{P}) = \mathbf{I} - 4\mathbf{P} + 4\mathbf{P}^2 = \mathbf{I} - 4\mathbf{P} + 4\mathbf{P} = \mathbf{I}$$

- Problem: given $\mathbf{x} \in \mathbb{R}^m$, find an orthogonal $\mathbf{H} \in \mathbb{R}^{m\times m}$ such that

$$\mathbf{H}\mathbf{x} = \begin{bmatrix} \beta \\ \mathbf{0} \end{bmatrix} = \beta \mathbf{e}_1, \qquad \text{for some } \beta \in \mathbb{R}.$$

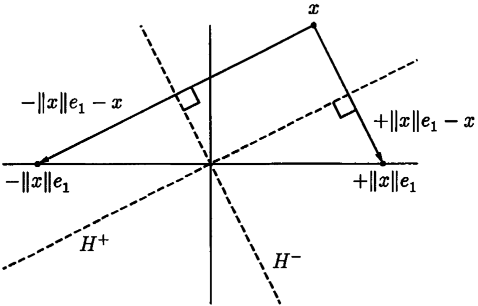

Figure 1: Householder reflection

- Householder reflection: let $\mathbf{v} \in \mathbb{R}^m$, $\mathbf{v} \neq \mathbf{0}$. Let
$$\mathbf{H} = \mathbf{I} - \frac{2}{\|\mathbf{v}\|_2^2}\mathbf{v}\mathbf{v}^T,$$
which is a reflection matrix with $\mathbf{P} = \mathbf{v}\mathbf{v}^T/\|\mathbf{v}\|_2^2$
- it can be verified that (try)
$$\mathbf{v} = \mathbf{x} \mp \|\mathbf{x}\|_2\mathbf{e}_1 \quad \Longrightarrow \quad \mathbf{H}\mathbf{x} = \pm\|\mathbf{x}\|_2\mathbf{e}_1;$$
the sign above may be determined to be the one that maximizes $\|\mathbf{v}\|_2$, for the sake of numerical stability (why?)
  - $\mathbf{v} = \mathbf{x} + \|\mathbf{x}\|_2\mathbf{e}_1$ if $x_1 > 0$
  - $\mathbf{v} = \mathbf{x} - \|\mathbf{x}\|_2\mathbf{e}_1$ if $x_1 < 0$

  Here, x_1 denotes the first entry of $\mathbf{x}$.

- let $\mathbf{H}_1 \in \mathbb{R}^{m\times m}$ be the Householder reflection w.r.t. $\mathbf{a}_1$. Transform $\mathbf{A}$ as

$$\mathbf{A}^{(1)} = \mathbf{H}_1\mathbf{A} = \begin{bmatrix} \times & \times & \dots & \times \\ 0 & \color{red}{\times} & \dots & \times \\ \vdots & \vdots & & \vdots \\ 0 & \color{red}{\times} & \dots & \times \end{bmatrix}$$

- let $\tilde{\mathbf{H}}_2 \in \mathbb{R}^{(m-1)\times(m-1)}$ be the Householder reflection w.r.t. $\mathbf{A}^{(1)}_{2:m,2}$ (marked red above). Transform $\mathbf{A}^{(1)}$ as

$$\mathbf{A}^{(2)} = \underbrace{\begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{H}}_2 \end{bmatrix}}_{=\mathbf{H}_2} \mathbf{A}^{(1)} = \begin{bmatrix} \times & \times \quad \dots \times \\ \mathbf{0} & \tilde{\mathbf{H}}_2\mathbf{A}^{(1)}_{2:m,2:n} \end{bmatrix} = \begin{bmatrix} \times & \times & \times & \dots & \times \\ 0 & \times & \times & \dots & \times \\ \vdots & 0 & \times & \dots & \times \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \times & \dots & \times \end{bmatrix}$$

- by repeatedly applying the trick above, we can transform $\mathbf{A}$ as the desired $\mathbf{R}$

$\mathbf{A}^{(0)} = \mathbf{A}$

for $k = 1, \ldots, n$

$\mathbf{A}^{(k)} = \mathbf{H}_k \mathbf{A}^{(k-1)}$, where

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

$\mathbf{I}_k$ is the $k \times k$ identity matrix; $\tilde{\mathbf{H}}_k$ is the Householder reflection of $\mathbf{A}^{(k-1)}_{k:m,k}$

end

- $\mathbf{H}_k$ introduces zeros under the diagonal of the k-th column
- the above procedure results in

$$\mathbf{A}^{(n)} = \mathbf{H}_n \cdots \mathbf{H}_2 \mathbf{H}_1 \mathbf{A}, \quad \mathbf{A}^{(n)} \text{ taking an upper triangular form}$$

- by letting $\mathbf{R} = \mathbf{A}^{(n)}$, $\mathbf{Q} = (\mathbf{H}_n \cdots \mathbf{H}_2 \mathbf{H}_1)^T$, we obtain the full QR
- a popularly used method for QR decomposition

Applying the Householder Matrix: HA

$$\mathbf{HA} = (\mathbf{I} - \beta \mathbf{v}\mathbf{v}^T)\mathbf{A} = \mathbf{A} - (\beta \mathbf{v})(\mathbf{v}^T \mathbf{A})$$

- takes $\mathcal{O}(4mn)$ flops, rather than $\mathcal{O}(m^2 n)$
- only acts on a submatrix of $\mathbf{A}$ as the process goes
- takes $\mathcal{O}(2mn^2 - \frac{2}{3}n^3)$ flops to obtain $\mathbf{R}$ $(m > n)$. What for $m < n$?

Computations of Q

Recall $\mathbf{Q} = (\mathbf{H}_n \cdots \mathbf{H}_2 \mathbf{H}_1)^T = \mathbf{H}_1 \mathbf{H}_2 \cdots \mathbf{H}_n$, with $\mathbf{H}_k = \mathbf{I} - \beta_k \mathbf{v}^{(k)} (\mathbf{v}^{(k)})^T$ and

$$\mathbf{v}^{(k)} = \begin{bmatrix} 0 & \cdots & 0 & v_k^{(k)} & v_{k+1}^{(k)} & \cdots & v_m^{(k)} \end{bmatrix}^T$$

By letting $\mathbf{Q}_{n+1} = \mathbf{I}$, and executing $\mathbf{Q}_k = \mathbf{H}_k \mathbf{Q}_{k+1}$ for $k = n : -1 : 1$, we obtain $\mathbf{Q} = \mathbf{Q}_1$

- efficiently computations by applying Householder matrix
- takes $\mathcal{O}(4m^2 n - 4mn^2 + \frac{4}{3}n^3)$ flops $(m > n)$, what for $m < n$?

- Example: Let

$$\mathbf{J} = \begin{bmatrix} c & s \\ -s & c \end{bmatrix}$$

where $c = \cos(\theta)$, $s = \sin(\theta)$ for some θ. Consider $\mathbf{y} = \mathbf{J}\mathbf{x}$:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} c & s \\ -s & c \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} cx_1 + sx_2 \\ -sx_1 + cx_2 \end{bmatrix}.$$

It can be verified that

- $\mathbf{J}$ is orthogonal;
- $y_2 = 0$ if $\theta = \arctan(x_2/x_1)$, or if

$$c = \frac{x_1}{\sqrt{x_1^2 + x_2^2}}, \quad s = \frac{x_2}{\sqrt{x_1^2 + x_2^2}}.$$

- Givens rotations:

$$\mathbf{J}(i,k,\theta) = \begin{array}{c} \\ \\ i \\ \\ k \\ \\ \end{array} \begin{array}{c} \begin{array}{ccccc} \quad & i & & k & \quad \end{array} \\ \begin{bmatrix} \mathbf{I} & & & & \\ & c & & s & \\ & & \mathbf{I} & & \\ & -s & & c & \\ & & & & \mathbf{I} \end{bmatrix} \end{array}$$

where $c = \cos(\theta)$, $s = \sin(\theta)$.

- $\mathbf{J}(i,k,\theta)$ is orthogonal
- let $\mathbf{y} = \mathbf{J}(i,k,\theta)\mathbf{x}$. It holds that

$$y_j = \begin{cases} cx_i + sx_k, & j = i \\ -sx_i + cx_k, & j = k \\ x_j, & j \neq i,k \end{cases}$$

- y_k is forced to zero if we choose $\theta = \tan^{-1}(x_k/x_i)$.

Givens QR

- Example: consider a 4×3 matrix.

$$\mathbf{A} = \begin{bmatrix} \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \end{bmatrix} \xrightarrow{\mathbf{J}_{1,2}} \begin{bmatrix} \times & \times & \times \\ 0 & \times & \times \\ \times & \times & \times \\ \times & \times & \times \end{bmatrix} \xrightarrow{\mathbf{J}_{1,3}} \begin{bmatrix} \times & \times & \times \\ 0 & \times & \times \\ 0 & \times & \times \\ \times & \times & \times \end{bmatrix} \xrightarrow{\mathbf{J}_{1,4}} \begin{bmatrix} \times & \times & \times \\ 0 & \times & \times \\ 0 & \times & \times \\ 0 & \times & \times \end{bmatrix} \xrightarrow{\mathbf{J}_{2,3}}$$

$$\begin{bmatrix} \times & \times & \times \\ 0 & \times & \times \\ 0 & 0 & \times \\ 0 & \times & \times \end{bmatrix} \xrightarrow{\mathbf{J}_{2,4}} \begin{bmatrix} \times & \times & \times \\ 0 & \times & \times \\ 0 & 0 & \times \\ 0 & 0 & \times \end{bmatrix} \xrightarrow{\mathbf{J}_{3,4}} \begin{bmatrix} \times & \times & \times \\ 0 & \times & \times \\ 0 & 0 & \times \\ 0 & 0 & 0 \end{bmatrix} = \mathbf{R}$$

where $\mathbf{B} \xrightarrow{\mathbf{J}} \mathbf{C}$ means $\mathbf{B} = \mathbf{J}\mathbf{C}$; $\mathbf{J}_{i,k} = \mathbf{J}(i, k, \theta)$, with θ chosen to zero out the kth entry in the ith column vector of the matrix transformed by $\mathbf{J}_{i,k}$.

- **Givens QR:** assume $m \geq n$. Perform a sequence of Givens rotations to annihilate the lower triangular parts of $\mathbf{A}$ to obtain

$$\underbrace{(\mathbf{J}_{m,n}\ldots\mathbf{J}_{n+2,n}\mathbf{J}_{n+1,n})\ldots(\mathbf{J}_{2m}\ldots\mathbf{J}_{24}\mathbf{J}_{23})(\mathbf{J}_{1m}\ldots\mathbf{J}_{13}\mathbf{J}_{12})}_{\mathbf{Q}^T}\mathbf{A} = \mathbf{R}$$

  where $\mathbf{R}$ takes the upper triangular form, and $\mathbf{Q}$ is orthogonal.

- applying Givens rotations $\mathbf{J}_{i,k}\mathbf{A}$ only updates the i, k row of $\mathbf{A}$, i.e.,

$$\mathbf{A}([i,j],:) = \begin{bmatrix} c & s \\ -s & c \end{bmatrix} \mathbf{A}([i,j],:)$$

- takes $\mathcal{O}(3mn^2 - n^3)$ flops to get $\mathbf{R}$, what for $\mathbf{Q}$?
- can be faster than Householder QR if $\mathbf{A}$ has certain sparse structures and we exploit them

$$\mathbf{x}_{LS} = \arg\min \|\mathbf{b} - \mathbf{A}\mathbf{x}\|_2^2$$

Using orthogonal projection

- solving $\mathbf{A}\mathbf{x} = \mathbf{P}\mathbf{b}$ to obtain $\mathbf{x}_{LS}$
  - $\mathbf{A}$ has orthonormal basis $\{\mathbf{q}_1, \mathbf{q}_2, \cdots, \mathbf{q}_n\}$ (can be computed using QR factorization),

$$\mathbf{x}_{LS} = \mathbf{R}^{-1}\mathbf{Q}^T\mathbf{b} \quad \text{(reduced QR)}$$

  - using $\mathbf{P} = \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T$,

$$(\mathbf{A}^T\mathbf{A})\mathbf{x}_{LS} = \mathbf{A}^T\mathbf{b} \quad \text{(normal equation)}$$

Using optimality condition

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

$$\nabla f(\mathbf{x}) = 0 \Longrightarrow \mathbf{x}_{LS} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b},$$

Rank-deficient LS, cf. [Golub-van Loan 13]

In the real field $\mathbb{R}$

For $\mathbf{A} \in \mathbb{R}^{m \times n}$, the pseudoinverse of $\mathbf{A}$ denoted by $\mathbf{A}^{+} \in \mathbb{R}^{n \times m}$ satisfying the Moore–Penrose conditions[2]

1. $\mathbf{A}\mathbf{A}^{\dagger}\mathbf{A} = \mathbf{A}$
2. $\mathbf{A}^{\dagger}\mathbf{A}\mathbf{A}^{\dagger} = \mathbf{A}^{\dagger}$
3. $(\mathbf{A}\mathbf{A}^{\dagger})^{T} = \mathbf{A}\mathbf{A}^{\dagger}$
4. $(\mathbf{A}^{\dagger}\mathbf{A})^{T} = \mathbf{A}^{\dagger}\mathbf{A}$

When $\mathbf{A}$ has full rank and $m > n$

- $\mathbf{A}^{\dagger} = (\mathbf{A}^{T}\mathbf{A})^{-1}\mathbf{A}^{T}$
  - In terms of reduced QR factorization of $\mathbf{A}$

$$\mathbf{A}^{\dagger} = (\mathbf{A}^{T}\mathbf{A})^{-1}\mathbf{A}^{T} = \mathbf{R}^{-1}\mathbf{Q}^{T}$$

[2] R. Penrose, A Generalized Inverse for Matrices. *Mathematical Proceedings of the Cambridge Philosophical Society*, 51(3), 1955

Readings

You are supposed to read

- Lloyd N. Trefethen and David Bau III. *Numerical Linear Algebra*, SIAM, 1997.

  Lecture 6, 8, 11

- Gene H. Golub and Charles F. Van Loan. *Matrix Computations*, Johns Hopkins University Press, 2013.

  Chapter 5.1 – 5.3