CS450: Numerical Analysis

Howard Hao Yang

Assistant Professor, ZJU-UIUC Institute

01/11/2023

QR Factorization

- Given a rectangular matrix $\boldsymbol{A} \in R^{m\times n}$ with $m > n$, it can be decomposed as

$$\boldsymbol{A} = \boldsymbol{Q}\begin{bmatrix}\boldsymbol{R}\\ \boldsymbol{O}\end{bmatrix}$$

where $\boldsymbol{Q}$ is an $m{\times}m$ orthogonal matrix and $\boldsymbol{R}$ is an upper triangular matrix

- This is known as the QR factorization
- Why we want this? (compared to the normal equation.
- Since it preserves the condition number, making the solution process more stable

Householder Transformations

- (Definition) Given a unit-length vector $\boldsymbol{x} \in R^n$, i.e., $\boldsymbol{x}^T\boldsymbol{x} = 1$. The following $n \times n$ matrix is called a Householder transformation

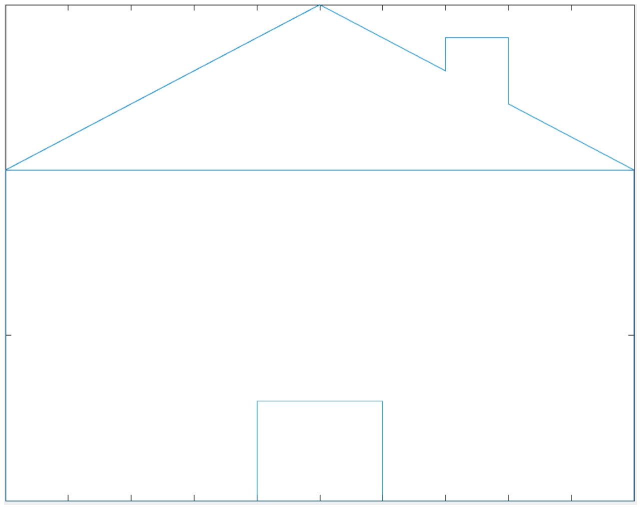
$$\boldsymbol{H} = \boldsymbol{I} - 2\boldsymbol{x}\boldsymbol{x}^T$$

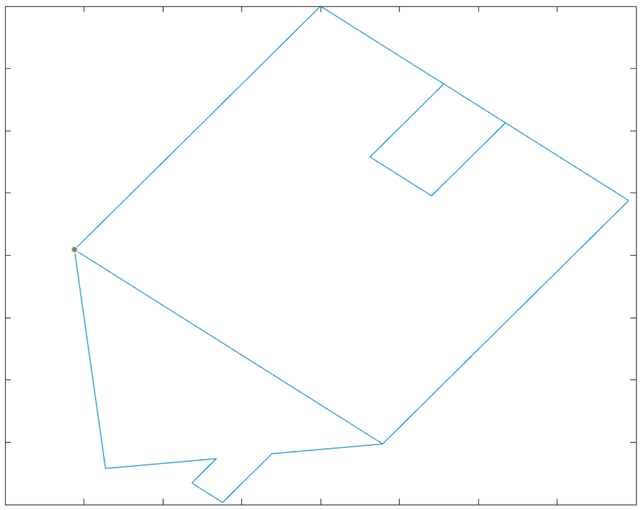
- The Householder transformation is symmetric and orthogonal
- Therefore $\boldsymbol{H}^{-1} = \boldsymbol{H}$

Illustrative Example

- Try the following Matlab code

```
>> A = [0 1 1 0 0 0.5 0.7 0.7 0.8 0.8 1 1 0.6 0.6 0.4 0.4;
1 1 0 0 1 1.5 1.3 1.4 1.4 1.2 1 0 0 0.3 0.3 0];
plot(A(1,:), A(2,:))
>> u = rand(2,1)

u =

    0.0975
    0.2785

>> u = u/norm(u)

u =

    0.3305
    0.9438

>> H = eye(2) - 2 * u * u'

H =

    0.7815   -0.6239
   -0.6239   -0.7815

>> A = H * A
```





Householder Transformations (cont'd)

- Given a vector $\boldsymbol{x} \in R^n$ with $\boldsymbol{x}^T\boldsymbol{x} = 1$, the Householder transformation is

$$\boldsymbol{H} = \boldsymbol{I} - 2\boldsymbol{x}\boldsymbol{x}^T$$

- Reflection property: for any vector $\boldsymbol{a} \in R^n$, $\boldsymbol{Ha}$ reflects $\boldsymbol{a}$ by the hyperplane perpendicular to $\boldsymbol{x}$

Householder Transformations (cont'd)

- Exercise: Let $n \geq 2$ and $\boldsymbol{u}, \boldsymbol{v} \in R^n$ be unit vectors (i.e., $\boldsymbol{u}^T\boldsymbol{u} = \boldsymbol{v}^T\boldsymbol{v} = 1$). Suppose $\boldsymbol{u} \neq \boldsymbol{v}$, let $\boldsymbol{x} = \frac{\boldsymbol{u}-\boldsymbol{v}}{\|\boldsymbol{u}-\boldsymbol{v}\|_2}$ and construct $\boldsymbol{H} = \boldsymbol{I} - 2\boldsymbol{x}\boldsymbol{x}^T$. Show that $\boldsymbol{H}\boldsymbol{u} = \boldsymbol{v}$.

Householder Transformations (cont'd)

- Exercise: Find a Householder transformation $\boldsymbol{H} = \boldsymbol{I} - 2\boldsymbol{x}\boldsymbol{x}^T$ such that $\boldsymbol{H}\boldsymbol{v} = \boldsymbol{w}$, where $\boldsymbol{v} = (2,1,2)^T$ and $\boldsymbol{w} = (3,0,0)^T$

Householder Transformations (Cont'd)

- Suppose we have $\boldsymbol{a} \in R^n$, and want to annihilate all elements below the first entry while preserving the norm
- Can we leverage some ideas from the Householder transformation?
- Problem: find vector $\boldsymbol{x} \in R^n$, such that $\boldsymbol{x}^T\boldsymbol{x} = 1$ and

$$\boldsymbol{H}\boldsymbol{a} = (\boldsymbol{I} - 2\boldsymbol{x}\boldsymbol{x}^T)\boldsymbol{a} = \alpha\boldsymbol{e}_1$$

where $\boldsymbol{e}_1 = (1,0,\cdots,0)^T$ and $\alpha = \|\boldsymbol{a}\|_2$

- Solution to this is

$$\boldsymbol{x} = \frac{\boldsymbol{a}}{\alpha} \pm \boldsymbol{e}_1$$

Householder Transformations (cont'd)

- Exercise: Find the Householder transformation matrix for a vector $\boldsymbol{a}$, if $\boldsymbol{a} = \boldsymbol{e}_1$

Householder Transformations (Cont'd)

- For a rectangular matrix $\boldsymbol{A} \in R^{m \times n}$ with $m > n$. Suppose $m = 6$, $\mathrm{n} = 5$ and we have computed the following

$$\boldsymbol{H_2 H_1 A} = \begin{bmatrix} \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & 0 & \mathbf{x} & \times & \times \\ 0 & 0 & \mathbf{x} & \times & \times \\ 0 & 0 & \mathbf{x} & \times & \times \\ 0 & 0 & \mathbf{x} & \times & \times \end{bmatrix}$$

- Concentrating on the highlighted entries, we determined a matrix $\widetilde{H}_3 \in R^{4 \times 4}$ such that

$$\tilde{H}_3 \begin{bmatrix} \mathbf{x} \\ \mathbf{x} \\ \mathbf{x} \\ \mathbf{x} \end{bmatrix} = \begin{bmatrix} \times \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Householder Transformations (Cont'd)

- Then, by forming $H_3 = \text{diag}(I_2, \widetilde{H}_3)$, we have

$$\boldsymbol{H_3 H_2 H_1 A} = \begin{bmatrix} \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & \times & \times \\ 0 & 0 & 0 & \times & \times \\ 0 & 0 & 0 & \times & \times \end{bmatrix}$$

- More generally, for a given vector $\boldsymbol{a} = \begin{bmatrix} \boldsymbol{a}_1 \\ \boldsymbol{a}_2 \end{bmatrix}$, where a_1 is a (k-1)-vector, with $1 \leq k < m$. If we take the Householder vector to be $\boldsymbol{v} = \begin{bmatrix} \mathbf{0} \\ \boldsymbol{a}_2 \end{bmatrix} - \alpha\, \boldsymbol{e}_k$, where $\alpha = \mp \|\boldsymbol{a_2}\|_2$, then the resulting Householder transformation annihilates the last $m - k$ components of $\boldsymbol{a}$.

Householder Transformations (Cont'd)

- For a rectangular matrix $\boldsymbol{A} \in R^{m \times n}$ with $m > n$. The QR factorization to this matrix can be written as

$$\boldsymbol{H}_n \cdots \boldsymbol{H}_1 \boldsymbol{A} = \begin{bmatrix} \boldsymbol{R} \\ \boldsymbol{O} \end{bmatrix}$$

- The product of successive Householder transformations $\boldsymbol{H_n} \cdots \boldsymbol{H_1}$ is itself an orthogonal matrix. Thus, if we take $\boldsymbol{Q}^{\boldsymbol{T}} = \boldsymbol{H_n} \cdots \boldsymbol{H_1}$, or equivalently, $\boldsymbol{Q} = \boldsymbol{H_1} \cdots \boldsymbol{H_n}$, then

$$\boldsymbol{A} = \boldsymbol{Q} \begin{bmatrix} \boldsymbol{R} \\ \boldsymbol{O} \end{bmatrix}$$

A Complete Worked Example

- Do a Householder transform to obtain QR decomposition for the following

$$A = \begin{bmatrix} 1 & -1 & 4 \\ 1 & 4 & -2 \\ 1 & 4 & 2 \\ 1 & -1 & 0 \end{bmatrix}$$

- Step 1:

$$v_1 = x_1 - \mathrm{sign}(x_{11})\|x_1\|e_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} - 2\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \Rightarrow H_{v_1} = \begin{bmatrix} 1/2 & 1/2 & 1/2 & 1/2 \\ 1/2 & 1/2 & -1/2 & -1/2 \\ 1/2 & -1/2 & 1/2 & -1/2 \\ 1/2 & -1/2 & -1/2 & 1/2 \end{bmatrix} \Rightarrow H_{v_1}A = \begin{bmatrix} 2 & 3 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 4 \\ 0 & -5 & 2 \end{bmatrix}$$

Example (cont'd)

- Step 2:

$$v_2 = x_2'' - \text{sign}(x_{22}'')\|x_2''\|e_2 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ -5 \end{bmatrix} + (-1)\begin{bmatrix} 0 \\ 5 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ -5 \\ 0 \\ -5 \end{bmatrix} \Rightarrow H_{v_2} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}$$

$$\Rightarrow H_{v_2}(H_{v_1}A) = \begin{bmatrix} 2 & 3 & 2 \\ 0 & 5 & -2 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow Q = (H_{v_2}H_{v_1})^T = H_{v_1}^T H_{v_2}^T = \begin{bmatrix} 1/2 & -1/2 & 1/2 & -1/2 \\ 1/2 & 1/2 & -1/2 & -1/2 \\ 1/2 & 1/2 & 1/2 & 1/2 \\ 1/2 & -1/2 & -1/2 & 1/2 \end{bmatrix}$$

QR: Existence and Properties

- Question: Given a matrix $\boldsymbol{A} \in R^{m \times n}$, does the QR factorization always exist?
- Answer: Always exist.
- More formally, if $\boldsymbol{A} \in R^{m \times n}$, then there exists an orthogonal matrix $\boldsymbol{Q} \in R^{m \times m}$ and an upper triangular matrix $\boldsymbol{R} \in R^{m \times n}$ such that $\boldsymbol{A} = \boldsymbol{QR}$. Additionally, for $1 \leq k \leq n$, there is $\mathrm{span}(\boldsymbol{a_1}, \dots, \boldsymbol{a_k}) = \mathrm{span}(\boldsymbol{q_1}, \dots, \boldsymbol{q_k})$; and the thin QR factorization, $\boldsymbol{A} = \boldsymbol{Q_1 R_1}$ is unique

Householder Transformations (cont'd)

- Exercise: given a vector $\boldsymbol{a} = [2, 3, 4]^T$
  1. Specify an elementary elimination matrix that annihilates the third component of $\boldsymbol{a}$
  2. Specify a Householder transformation that annihilates the third component of $\boldsymbol{a}$

Givens Rotations

- Householder reflections are exceedingly useful for introducing zeros on a grand scale
- But sometimes we want to zero elements in a more selective way
- If reflections work, can we make rotations work, too?
- Given a two-dimension vector $\boldsymbol{a} = [a_1, a_2]^T$, we aim to form the following matrix (which is termed the Givens rotation)

$$\boldsymbol{G} = \begin{bmatrix} c & s \\ -s & c \end{bmatrix}$$

Givens Rotations (cont'd)

- Two epitome examples
  - A 2-by-2 orthogonal matrix is a rotation if it has the following form

$$Q = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix}$$

where $\boldsymbol{y} = \boldsymbol{Q}\boldsymbol{x}$ rotates vector $\boldsymbol{x}$ counterclock wise by an angle of θ

  - A 2-by-2 orthogonal matrix is a reflection if it has the following form

$$Q = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{bmatrix}$$

where $\boldsymbol{y} = \boldsymbol{Q}\boldsymbol{x}$ reflects vector $\boldsymbol{x}$ across the line defined by $\mathrm{span}\{[\cos(\theta/2), \sin(\theta/2)]\}$

Givens Rotations (cont'd)

- If reflections work, can we make rotations work, too?
- Given a two-dimension vector $\boldsymbol{a} = [a_1, a_2]^T$, we aim to form the following matrix (which is termed the Givens rotation)

$$\boldsymbol{G} = \begin{bmatrix} c & s \\ -s & c \end{bmatrix}$$

- Here, c and s are basically cosine and sine, respectively, for some angle (which imposes $c^2 + s^2 = 1$). And we want

$$\boldsymbol{G}\boldsymbol{a} = \begin{bmatrix} c & s \\ -s & c \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} \alpha \\ 0 \end{bmatrix}$$

Givens Rotations (cont'd)

- Suppose we want to find c, s and α for the following equality

$$\boldsymbol{G}\boldsymbol{a} = \begin{bmatrix} c & s \\ -s & c \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} \alpha \\ 0 \end{bmatrix}$$

- Do elementary elimination

$$\begin{bmatrix} a_1 & a_2 \\ a_2 & -a_1 \end{bmatrix} \begin{bmatrix} c \\ s \end{bmatrix} = \begin{bmatrix} \alpha \\ 0 \end{bmatrix} \rightarrow \begin{bmatrix} a_1 & a_2 \\ 0 & -a_1 - a_2^2/a_1 \end{bmatrix} \begin{bmatrix} c \\ s \end{bmatrix} = \begin{bmatrix} \alpha \\ -\alpha a_2/a_1 \end{bmatrix}$$

- The above leads to solutions as

$$s = \frac{\alpha a_2}{a_1^2 + a_2^2}, \qquad c = \frac{\alpha a_1}{a_1^2 + a_2^2} \qquad \rightarrow \qquad c = \frac{a_1}{\sqrt{a_1^2 + a_2^2}}, \qquad s = \frac{a_2}{\sqrt{a_1^2 + a_2^2}}$$

$$\alpha = \sqrt{a_1^2 + a_2^2}$$

Example

- Develop a Givens rotation that annihilates the following vector

$$\boldsymbol{a} = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$$

- We can calculate c and s as the following

$$c = \frac{a_1}{\sqrt{a_1^2 + a_2^2}} = \frac{4}{5} = 0.8, \qquad s = \frac{a_2}{\sqrt{a_1^2 + a_2^2}} = \frac{3}{5} = 0.6,$$

- Hence,

$$\boldsymbol{G}\boldsymbol{a} = \begin{bmatrix} 0.8 & 0.6 \\ -0.6 & 0.8 \end{bmatrix} \begin{bmatrix} 4 \\ 3 \end{bmatrix} = \begin{bmatrix} 5 \\ 0 \end{bmatrix}$$

Givens Rotations (Cont'd)

- From a 2-dimension vector to a m-dimension vector: if I want to annihilate the j-th entry with the i-th entry via such a rotation, just identify the corresponding entries and use the rotation
- Example: m=5, i=2, j=4 in the following example

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & c & 0 & s & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & -s & 0 & c & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{bmatrix} = \begin{bmatrix} a_1 \\ \alpha \\ a_3 \\ 0 \\ a_5 \end{bmatrix}$$

Givens Rotations (Cont’d)

- More generally, for an n-by-n matrix, construct the rotation matrix by

$$G(i,k,\theta) = \begin{array}{cc} \begin{bmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & \cdots & c & \cdots & s & \cdots & 0 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & \cdots & -s & \cdots & c & \cdots & 0 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix} & \begin{array}{c} \\ \\ i \\ \\ k \\ \\ \\ \end{array} \\ \begin{array}{ccccccc} & & i & & k & & \end{array} & \end{array}$$

where $c = \cos(\theta)$ and $s = \sin(\theta)$

- For $\boldsymbol{x} \in R^n$, $\boldsymbol{y} = \boldsymbol{G}(i,k,\theta)^{\mathrm{T}}\boldsymbol{x}$ gives

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

Givens Rotations (Cont'd)

- Exercise: given a vector $\boldsymbol{a} = [2, 3, 4]^T$
  1. Specify an elementary elimination matrix that annihilates the third component of $\boldsymbol{a}$
  2. Specify a Householder transformation that annihilates the third component of $\boldsymbol{a}$
  3. Specify the Givens rotation that annihilates the third component of $\boldsymbol{a}$