Lecture 20:

Recall:

Preliminary: QR factorization

Definition: $Q \in M_{n\times n}(\mathbb{R})$ is orthogonal if $Q^T Q = I_n$

Remark:
- $Q^{-1} = Q^T$
- Columns of Q forms orthonormal set.

Revisit: Gram-Schmidt orthogonalization

Let $A = \begin{pmatrix} | & | & & | \\ \vec{a}_1 & \vec{a}_2 & \dots, & \vec{a}_n \\ | & | & & | \end{pmatrix}$ are linearly independent.

G-S process converts $\{\vec{a}_1, \dots, \vec{a}_n\}$ to orthonormal set $\{\vec{q}_1, \vec{q}_2, \dots, \vec{q}_n\}$

G-S process:

<u>Step 1</u>: Let $\tilde{q}_1 = \vec{a}_1$. Normalize: $\vec{q}_1 = \dfrac{\tilde{q}_1}{\|\tilde{q}_1\|_2}$. Let $\alpha_{11} = \|\tilde{q}_1\|_2$

<u>Step 2</u>: Define: $\tilde{q}_2 = \vec{a}_2 - \alpha_{12}\vec{q}_1$. Choose α_{12} such that:

$\tilde{q}_2^{\,T}\vec{q}_1 = 0$. Then: $\alpha_{12} = \vec{q}_1^{\,T}\vec{a}_2$.

Normalize: $\vec{q}_2 = \dfrac{\tilde{q}_2}{\|\tilde{q}_2\|_2}$. Let $\alpha_{22} = \|\tilde{q}_2\|_2$.

<u>Step 3</u>: Suppose $\vec{q}_1, \vec{q}_2, \ldots, \vec{q}_{k-1}$ are constructed.

Let $\tilde{q}_k = \vec{a}_k - (\alpha_{1k}\vec{q}_1 + \alpha_{2k}\vec{q}_2 + \ldots + \alpha_{k-1,k}\vec{q}_{k-1})$ where $\alpha_{jk} = \vec{q}_j^{\,T}\vec{a}_k$.

Then: $\tilde{q}_k^{\,T}\vec{q}_i = 0$ for $i = 1, 2, \ldots, k-1$. Let $\alpha_{kk} = \|\tilde{q}_k\|_2$

Normalize: $\vec{q}_k = \tilde{q}_k / \|\tilde{q}_k\|_2$.

In summary,

$$\begin{cases} \vec{a}_1 = \alpha_{11}\vec{q}_1 \\ \vec{a}_2 = \alpha_{12}\vec{q}_1 + \alpha_{22}\vec{q}_2 \\ \vec{a}_3 = \alpha_{13}\vec{q}_1 + \alpha_{23}\vec{q}_2 + \alpha_{33}\vec{q}_3 \\ \quad \vdots \\ \vec{a}_k = \alpha_{1k}\vec{q}_1 + \alpha_{2k}\vec{q}_2 + \cdots + \alpha_{kk}\vec{q}_k \end{cases}$$

This is equivalent to:

$$A = \underset{M_{m\times n}}{\begin{pmatrix} \downarrow & \downarrow & & \downarrow \\ \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n \\ | & | & & | \end{pmatrix}} = \underbrace{\overset{M_{m\times n}}{\begin{pmatrix} \downarrow & & \downarrow \\ \vec{q}_1 & \cdots & \vec{q}_n \\ | & & | \end{pmatrix}}}_{Q} \underbrace{\overset{M_{n\times n}}{\begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ & \alpha_{22} & \cdots & \alpha_{2n} \\ & & \ddots & \vdots \\ & & & \alpha_{nn} \end{pmatrix}}}_{R}$$

Remark:

- Q = orthonormal; R = upper triangular
- Q may NOT be a square matrix ($Q \in M_{m\times n}$)
  R is a square matrix ($R \in M_{n\times n}$)
- Factorization of $A = QR$ is called the QR factorization.

(If all diagonal entries of R are positive, then QR factorization is unique)

<u>Detailed algorithm:</u> (QR factorization) Let $A = (\vec{a}_1, \ldots, \vec{a}_n)$ be full rank.

<u>Step 1:</u> Use G-S process to obtain orthonormal set $\{\vec{q}_1, \vec{q}_2, \ldots, \vec{q}_n\}$

<u>Step 2:</u> Compute $\alpha_{jk} = \vec{q}_j^{\,T} \vec{a}_k$

<u>Step 3:</u> Construct QR factorization:

$$A = QR = \begin{pmatrix} | & | & & | \\ \vec{q}_1 & \vec{q}_2 & \cdots & \vec{q}_n \\ | & | & & | \end{pmatrix} \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ & \alpha_{22} & & \alpha_{2n} \\ & & \ddots & \vdots \\ & & & \alpha_{nn} \end{pmatrix}$$

Example: Let $A=\begin{pmatrix} 12 & -51 & 4 \\ 6 & 167 & -68 \\ -4 & 24 & -41 \end{pmatrix}$. Find the QR factorization of A.

Solution: Use G-S process, we get:

$$\vec{q}_1 = \frac{(12,6,-4)^T}{\|(12,6,-4)^T\|_2} = \begin{pmatrix} 6/7 \\ 3/7 \\ -2/7 \end{pmatrix} \qquad \alpha_{11} = \|(12,6,-4)^T\|_2$$

Similarly, $\vec{q}_2 = \left(\frac{-69}{175}, \frac{158}{175}, \frac{6}{35}\right)^T$ and $\vec{q}_3 = \left(\frac{-58}{175}, \frac{6}{175}, \frac{-33}{35}\right)^T$

Step 2: Compute: $\alpha_{11} = \vec{q}_1^{\,T}\vec{a}_1 = 14$, $\alpha_{12} = \vec{q}_1^{\,T}\vec{a}_2 = 21, \ldots$

Alternatively, since $Q = (\vec{q}_1, \ldots, \vec{q}_n)$ is orthogonal,

$A = QR \Rightarrow R = Q^T A$.

We get: $R = \begin{pmatrix} 14 & 21 & -14 \\ 0 & 175 & -70 \\ 0 & 0 & 35 \end{pmatrix}$

Step 3: Construct the QR factorization of A:

$$A = \begin{pmatrix} 6/7 & -69/175 & -58/175 \\ 3/7 & 158/175 & 6/175 \\ -2/7 & 6/35 & -33/35 \end{pmatrix} \begin{pmatrix} 14 & 21 & -14 \\ 0 & 175 & -70 \\ 0 & 0 & 35 \end{pmatrix}.$$

$$\tilde{q}_2 = \vec{a}_2 - \alpha_{12}\vec{q}_1, \qquad \alpha_{12} = \vec{a}_2 \cdot \vec{q}_1$$

$$\vec{q}_2 = \frac{\tilde{q}_2}{\|\tilde{q}_2\|_2}$$

QR method to find eigenvalues

Algorithm: (QR algorithm)

Input : $A \in M_{n\times n}(\mathbb{R})$

Step 1: Let $A^{(0)} = A$. Compute QR factorization of $A^{(0)} = Q_0 R_0$.
Let $A^{(1)} = R_0 Q_0$.

Step 2: Assume $A^{(1)}, \dots, A^{(k)}$ are computed. Let $A^{(k)} = Q_k R_k$ be the QR factorization of $A^{(k)}$. Let $A^{(k+1)} = R_k Q_k$.

Observation: 1. QR method gives a sequence of matrices

$$\{ A^{(0)} = A,\ A^{(1)},\ A^{(2)},\ \ldots,\ A^{(k)},\ \ldots \}$$

2. Now, $A^{(1)} = R_0 Q_0 = Q_0^{-1} \underbrace{(Q_0 R_0)}_{\substack{\shortparallel \\ A = A^{(0)}}} Q_0$. $\therefore$ $A^{(1)}$ is similar to $A^{(0)} = A$

$\therefore$ $A^{(1)}$ has same sets of eigenvalues as $A^{(0)}$.

$$\big(\det(A^{(1)} - \lambda I) = \det(Q_0^{-1} A^{(0)} Q_0 - \lambda I)$$

$$= \det(Q_0^{-1} [A^{(0)} - \lambda I] Q_0) = \det(A^{(0)} - \lambda I) \big)$$

Similarly, $A^{(2)} = R_1 Q_1 = Q_1^{-1} \underbrace{Q_1 R_1}_{\substack{\shortparallel \\ A^{(1)}}} Q_1$.

$$\therefore\ A^{(0)} \underset{\text{similar}}{\sim} A^{(1)} \sim A^{(2)} \sim \ldots \sim A^{(k)} \sim \ldots$$

3. If $A^{(k)}$ converges to an upper triangular matrix, then the diagonal entries are the eigenvalues of A.

Example: Let $A = \begin{pmatrix} -149 & -50 & -154 \\ 537 & 180 & 546 \\ -27 & -9 & -25 \end{pmatrix}$

Let $A^{(0)} = A$. Compute the QR factorization of $A^{(0)}$:

$$A^{(0)} = \begin{pmatrix} -0.27 & -0.71 & 0.65 \\ 0.96 & -0.16 & 0.22 \\ -0.05 & 0.69 & 0.73 \end{pmatrix} \begin{pmatrix} 558 & 187 & 568 \\ 0 & 0.07 & 3.46 \\ 0 & 0 & 0.105 \end{pmatrix} = Q_0 R_0$$

$$A^{(1)} = R_0 Q_0 = \begin{pmatrix} 3.53 & * & * \\ -0.096 & 2.36 & * \\ -0.007 & 0.0996 & 0.1053 \end{pmatrix}$$ (Quite close to upper triangular)

$$A^{(14)} = \begin{pmatrix} 3.0716 & * & * \\ 0.0193 & 0.9284 & * \\ 0 & 0 & 2 \end{pmatrix}$$ (Very close to upper triangular)

Diagonal entries very close to eigenvalues 1, 2, 3.

Convergence of QR method

We will state it without proof (Out of scope)

Theorem: Let A be a real symmetric non-singular matrix. The sequence $\{A^{(k)}\}$ generated by the QR method converges to an upper triangular matrix and the diagonal entries of $A^{(k)}$ converges to eigenvalues of A.

Remark:
1. Power method computes ONE eigenvalue
2. QR method computes ALL eigenvalues

Relationship between Power's method and QR method

Motivation: Consider $A \in M_{n\times n}(\mathbb{R})$ with eigenvalues:
symmetric

$$|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n|$$

Say $\{\vec{q}_1, \vec{q}_2, \ldots, \vec{q}_n\}$ are eigenvectors of $\lambda_1, \lambda_2, \ldots, \lambda_n$ respectively.

Properties of $\{\vec{q}_i\}_{i=1}^n$:

$$\underbrace{A\vec{q}_i}_{=\lambda_i\vec{q}_i} \cdot \vec{q}_j = (A\vec{q}_i)^T\vec{q}_j = q_i^T \overbrace{\overbrace{A^T}^{A}\vec{q}_j}^{=\lambda_j\vec{q}_j}$$

$$\iff \lambda_i\, \vec{q}_i \cdot \vec{q}_j = \lambda_j\, \vec{q}_i \cdot \vec{q}_j$$

$$\iff (\lambda_i - \lambda_j)\, \vec{q}_i \cdot \vec{q}_j = 0 \iff \vec{q}_i \cdot \vec{q}_j = 0$$

$\therefore$ $\{\vec{q}_j\}_{j=1}^n$ is orthogonal.

Observation:

If we choose $\vec{x}^{(0)} = c_i \vec{q}_i + c_{i+1} \vec{q}_{i+1} + \dots + c_n \vec{q}_n \quad (c_i \neq 0)$

$$\vec{x}^{(k)} = \frac{A^k\left(\vec{x}^{(0)}\right)}{\| \cdots \|} = \frac{c_i \lambda_i^k \vec{q}_i + \dots + c_n \lambda_n^k \vec{q}_n}{\| \cdots \|}$$

$$= \frac{c_i \lambda_i^k \left(\vec{q}_i + \cancelto{0}{\left(\frac{\lambda_{i+1}}{\lambda_i}\right)^k c_{i+1} \vec{q}_{i+1}} + \cancelto{0}{\dots} \right)}{\| \cdots \|}$$

then the power method will converge to $|\lambda_i|$

Remark: If we are "smart" (?),

$$\vec{x}_i^{(0)} = c_i \vec{q}_i + \dots + c_n \vec{q}_n$$

then we can define:

$$M_{n\times n} \ni X^{(0)} = \begin{pmatrix} | & | & & | \\ \vec{x}_1^{(0)} & \vec{x}_2^{(0)} & \cdots & \vec{x}_n^{(0)} \\ | & | & & | \end{pmatrix}$$

Apply Power method on $X^{(0)}$:

$$A^k X^{(0)} = \begin{pmatrix} | & | & & | \\ A^k\vec{x}_1^{(0)} & A^k\vec{x}_2^{(0)} & \cdots & A^k\vec{x}_n^{(0)} \\ | & | & & | \\ \downarrow & \downarrow & & \downarrow \\ |\lambda_1| & |\lambda_2| & & |\lambda_n| \end{pmatrix}$$