SI231b: Matrix Computations

Lecture 16: Eigenvalue Computations

Yue Qiu

qiuyue@shanghaitech.edu.cn

School of Information Science and Technology

ShanghaiTech University

Nov. 09, 2021

Power Iterations for a Set of Vectors

From the Power Iteration, we know that

- $A^k q_0$ converges to the eigenvector associated with the largest eigenvalue in magnititude.
- if we start with a set of linearly independent vectors $\{q_1, \ q_2, \ \cdots, q_r\}$, then $A^k\{q_1, \ q_2, \ \cdots, q_r\}$ should converge (under suitable assumptions) to a subspace spanned by eigenvectors of A associated with r largest eigenvalues in magnititude.

Simultaneous Iteration: applying power iteration to several vectors at once. Sometimes it is called **block power iteration**.

Define $\mathrm{V}^{(0)}$ to be the $n \times r$ matrix,

$$\mathrm{V}^{(0)} = \begin{bmatrix} v_1^{(0)} & v_2^{(0)} & \cdots & v_r^{(0)} \end{bmatrix}.$$

After k steps of applying A , we obtain

$$\mathrm{V}^{(k)} = \mathrm{A}^k \mathrm{V}^{(0)} = \begin{bmatrix} v_1^{(k)} & v_2^{(k)} & \cdots & v_r^{(k)} \end{bmatrix}.$$

Assume

1. The leading $r+1$ eigenvalues are distinct in absolute value;

$$|\lambda_1| > |\lambda_2| > \cdots > |\lambda_r| > |\lambda_{r+1}| \geq |\lambda_{r+2}| \geq \cdots |\lambda_n|$$

2. All the leading principle sub-matrices $\mathrm{Q}^T \mathrm{V}^{(0)}$ are nonsingular.
   - Q is the matrix with $\mathrm{q}_1, \mathrm{q}_2, \cdots, \mathrm{q}_r$ as columns;
   - $\mathrm{q}_1, \mathrm{q}_2, \cdots, \mathrm{q}_r$ are eigenvectors associated with eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_r$.

```
choose V^(0) with r linear independent columns
for k = 1, 2, ···
      V^(k) = AV^(k-1)
      Q^(k)R^(k) = V^(k)   reduced QR factorization
end
```

Under the assumptions, we have as $k \to \infty$,

- For real symmetric matrix A (Q has orthonormal columns)

$$\|\mathrm{q}_j^{(k)} - (\pm q_j)\| = \mathcal{O}(C^k),$$

  for $1 \le j \le r$, where $C < 1$ is the constant

$$C = \max_{1 \le k \le r} \frac{|\lambda_{k+1}|}{|\lambda_k|}$$

- For unsymmetric matrix A (Q does not have orthonormal columns)

$$\mathcal{R}(\mathrm{Q}^{(k)}) \to \mathcal{R}(\mathrm{Q})$$

For Unnormalized Simultaneous Iteration, as $k \to \infty$, the vectors $q^{(1)}, q^{(2)}, \cdots, q^{(r)}$ all converge to multiples of the same dominant eigenvector q_1. Therefore, they form an ill-conditioned basis of span $\left\{q^{(1)}, q^{(2)}, \cdots, q^{(r)}\right\}$.

The remedy is simple, we should build orthonormal basis at each iteration $\rightsquigarrow$ **Simultaneous Iteration/Subspace Iteration**

Subspace Iteration:

```
random selection Q^(0) with orthonormal columns
for k = 1, 2, ...
      Z_k = A Q^(k-1)
      Z_k = Q^(k) R^(k)   reduced QR factorization
end
```

- Z_k and $Q^{(k)}$ has the same column space
- equal to the column space of $A^k Q^{(0)}$

- $\mathcal{R}(Q^{(k)})$ converge to subspace associated with r largest eigenvalues in magnititude (dominant invariant subspace).
- $\lambda\left(\left(Q^{(k)}\right)^H AQ^{(k)}\right) \to \{\lambda_1,\ \lambda_2,\ \cdots,\lambda_r\}$
- $\left|\lambda_i^{(k)} - \lambda_i\right| = \mathcal{O}\left(\left|\frac{\lambda_{r+1}}{\lambda_i}\right|^k\right),\ i = 1,\ 2,\ \cdots,\ r$
- also called simultaneously iteration or orthogonal iteration
- when $r = n$, it coincides with QR iteration

QR Iteration:

$$
\begin{aligned}
&\mathrm{A}^{(0)} = \mathrm{A}\\
&\texttt{for}\ \ k = 1,\ 2,\ \cdots\\
&\qquad \mathrm{Q}^{(k)}\mathrm{R}^{(k)} = \mathrm{A}^{(k-1)} \quad \texttt{QR factorization of}\ \ \mathrm{A}^{(k-1)}\\
&\qquad \mathrm{A}^{(k)} = \mathrm{R}^{(k)}\mathrm{Q}^{(k)}\\
&\texttt{end}
\end{aligned}
$$

Facts:

- $\mathrm{A}^{(k)}$ is similar to A
- Eigenvalues of $\mathrm{A}^{(k)}$ should be easier to compute than that of A.
- $\mathrm{A}^{(k)}$ should converge fast (expected) to a form whose eigenvalues are easily computed.
    - upper triangular form

For an $n \times n$ matrix A, each iteration requires $\mathcal{O}(n^3)$ flops to compute the QR factorization.

- too computationally expensive!

Improvement:

Perform a similarity transform A to obtain a form $A^{(0)} = (Q^{(0)})^H A Q^{(0)}$

- the QR decomposition of $A^{(0)}$ should be computationally cheap
- $A^{(k)}$ $(k = 1, 2, \cdots)$ should have similar structure with $A^{(0)}$ so that the QR decomposition at each iteration is computationally cheap

Motivation: perform similarity transform A to an upper Hessenberg form (zeros below the first subdiagonal), i.e., $Q^H AQ = H$ where

$$H = \begin{bmatrix} \times & \times & \times & \times \\ \times & \times & \times & \times \\ & \times & \times & \times \\ & & \times & \times \end{bmatrix}$$

Advantage: QR factorization of an upper Hessenberg matrix requires $\mathcal{O}(n^2)$ flops (how?).

- by using Givens rotations

QR Iteration with Hessenberg Reduction:

```
A = Q^H H Q,  A^(0) = H,   H is upper Hessenberg
for k = 1, 2, ...
      Q^(k) R^(k) = A^(k-1)   QR factorization of A^(k-1)
      A^(k) = R^(k) Q^(k)
end
```

$A = Q^H H Q$, $A^{(0)} = H$, H is upper Hessenberg
for $k = 1, 2, \cdots$
$\quad Q^{(k)} R^{(k)} = A^{(k-1)}$ QR factorization of $A^{(k-1)}$
$\quad A^{(k)} = R^{(k)} Q^{(k)}$
end

Key: $A^{(k)}$ is of upper Hessenberg form (how to preserve?)

- by using Givens rotations to compute the QR factorization (how to prove?)

Benefit: $\mathcal{O}(n^2)$ flops for QR factorization.

For an $n \times n$ matrix $\mathrm{A} = \begin{bmatrix} \mathrm{a}_1 & \mathrm{a}_2 & \cdots & \mathrm{a}_n \end{bmatrix}$.

A Naive Try

Let Q_1 be the Householder reflection matrix that reflects a_1 to $-\mathrm{sign}(\mathrm{a}_1(1))\|\mathrm{a}_1\|_2\mathrm{e}_1$,

$$\mathrm{A} = \begin{bmatrix} \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \end{bmatrix} \rightarrow \underbrace{\begin{bmatrix} \times & \times & \times & \times \\ & \times & \times & \times \\ & \times & \times & \times \\ & \times & \times & \times \end{bmatrix}}_{\mathrm{Q}_1\mathrm{A}} \rightarrow \underbrace{\begin{bmatrix} \times & \times & \times & \times \\ {\color{red}\times} & \times & \times & \times \\ {\color{red}\times} & \times & \times & \times \\ {\color{red}\times} & \times & \times & \times \end{bmatrix}}_{\mathrm{Q}_1\mathrm{A}\mathrm{Q}_1^H}$$

Mission failed!

Less Ambitious Try

Let $\tilde{a}_1 = A(2:n, 1)$ and Q_1 be the Householder reflection matrix that reflects $\tilde{a}_1$ to $-\text{sign}(\tilde{a}_1(1))\|\tilde{a}_1\|_2 e_1$,

$$A = \begin{bmatrix} \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \end{bmatrix} \rightarrow \underbrace{\begin{bmatrix} \times & \times & \times & \times \\ \times & \times & \times & \times \\ & \times & \times & \times \\ & \times & \times & \times \end{bmatrix}}_{Q_1 A} \rightarrow \underbrace{\begin{bmatrix} \times & \times & \times & \times \\ \times & \times & \times & \times \\ & \times & \times & \times \\ & \times & \times & \times \end{bmatrix}}_{Q_1 A Q_1^H}$$

Repeat the above procedure to the 2nd column of $Q_1 A Q_1^H \cdots$

Given an $n \times n$ matrix A, the following algorithm reduces A to an upper Hessenberg form.

Hessenberg Reduction:

```
for k = 1 : n − 2
    x= A(k+1:n, k)
    v_k = sign(x(1))||x||_2 e_1 + x
    v_k = v_k / ||v_k||_2
    A(k+1 : n, k : n) = A(k+1 : n, k : n) − 2v_k(v_k^H A(k+1 : n, k : n))
    A(1 : n, k+1 : n) = A(1 : n, k+1 : n) − 2(A(1 : n, k+1 : n)v_k)v_k^H
end
```

Readings

You are supposed to read

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

  Lecture 26, 28

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

  Chapter 7.3 – 7.4