Math 304 (Spring 2010) - Lecture 10

Computation of Eigenvalues and Eigenvectors

Emre Mengi
Department of Mathematics
Koç University

emengi@ku.edu.tr

Outline

Eigenvalues and Eigenvectors

- Algorithms to compute an extreme eigenvalue (*e.g.* with largest modulus)
  - Power Iteration (Watkins - 5.3, Fausett - 5.1)
  - Convergence Properties of Power Iteration (Watkins - 5.3, Fausett - 5.1)
  - Extensions of Power Iteration (Watkins - 5.3, Fausett 5.2)

Complex Vectors

Let $u, v \in \mathbf{C}^n$. ($\mathbf{C}^n$ - the set of complex vectors with n components)

Complex Vectors

Let $u, v \in \mathbf{C}^n$. ($\mathbf{C}^n$ - the set of complex vectors with n components)

- The standard inner (or dot) product of u and v is defined as

$$u^* v = \bar{u}^T v = \bar{u}_1 v_1 + \bar{u}_2 v_2 + \cdots + \bar{u}_n v_n$$

  $\bar{u}$: conjugate of u entrywise, u^* : complex conjugate transpose of u

Complex Vectors

Let $u, v \in \mathbf{C}^n$. ($\mathbf{C}^n$ - the set of complex vectors with n components)

- The standard inner (or dot) product of u and v is defined as

$$u^*v = \bar{u}^T v = \bar{u}_1 v_1 + \bar{u}_2 v_2 + \cdots + \bar{u}_n v_n$$

$\bar{u}$: conjugate of u entrywise, u^* : complex conjugate transpose of u

e.g. Let $u = \begin{bmatrix} 2+3i \\ 1-2i \end{bmatrix}, v = \begin{bmatrix} 1+i \\ -2+i \end{bmatrix}$.

Complex Vectors

Let $u, v \in \mathbf{C}^n$. ($\mathbf{C}^n$ - the set of complex vectors with n components)

- The standard inner (or dot) product of u and v is defined as

$$u^* v = \bar{u}^T v = \bar{u}_1 v_1 + \bar{u}_2 v_2 + \cdots + \bar{u}_n v_n$$

$\bar{u}$: conjugate of u entrywise, u^* : complex conjugate transpose of u

e.g. Let $u = \begin{bmatrix} 2+3i \\ 1-2i \end{bmatrix}, v = \begin{bmatrix} 1+i \\ -2+i \end{bmatrix}$.

$$u^* v = \begin{bmatrix} 2-3i & 1+2i \end{bmatrix} \begin{bmatrix} 1+i \\ -2+i \end{bmatrix} = (2-3i)(1+i) + (1+2i)(-2+i) = 1-4i$$

Complex Vectors

Let $u, v \in \mathbf{C}^n$. ($\mathbf{C}^n$ - the set of complex vectors with n components)

- The standard inner (or dot) product of u and v is defined as

$$u^* v = \bar{u}^T v = \bar{u}_1 v_1 + \bar{u}_2 v_2 + \cdots + \bar{u}_n v_n$$

$\bar{u}$: conjugate of u entrywise, u^* : complex conjugate transpose of u

e.g. Let $u = \begin{bmatrix} 2+3i \\ 1-2i \end{bmatrix}, v = \begin{bmatrix} 1+i \\ -2+i \end{bmatrix}$.

$$u^* v = \begin{bmatrix} 2-3i & 1+2i \end{bmatrix} \begin{bmatrix} 1+i \\ -2+i \end{bmatrix} = (2-3i)(1+i) + (1+2i)(-2+i) = 1-4i$$

- The 2-norm (or Euclidean length) of v is defined as

$$\|v\|_2 = \sqrt{v^* v} = \sqrt{|v_1|^2 + \cdots + |v_n|^2}.$$

Complex Vectors

Let $u, v \in \mathbf{C}^n$. ($\mathbf{C}^n$ - the set of complex vectors with n components)

- The standard inner (or dot) product of u and v is defined as

$$u^* v = \bar{u}^T v = \bar{u}_1 v_1 + \bar{u}_2 v_2 + \cdots + \bar{u}_n v_n$$

  $\bar{u}$: conjugate of u entrywise, $\quad u^*$: complex conjugate transpose of u

  e.g. Let $u = \begin{bmatrix} 2+3i \\ 1-2i \end{bmatrix}, v = \begin{bmatrix} 1+i \\ -2+i \end{bmatrix}$.

$$u^* v = \begin{bmatrix} 2-3i & 1+2i \end{bmatrix} \begin{bmatrix} 1+i \\ -2+i \end{bmatrix} = (2-3i)(1+i) + (1+2i)(-2+i) = 1-4i$$

- The 2-norm (or Euclidean length) of v is defined as

$$\|v\|_2 = \sqrt{v^* v} = \sqrt{|v_1|^2 + \cdots + |v_n|^2}.$$

- The vector u is said to be orthogonal to v if $u^* v = 0$.

Power Iteration

- Given a matrix $A \in \mathbf{R}^{n\times n}$ and an initial vector $q_0 \in \mathbf{C}^n$. The power iteration generates the sequence of vectors $\{q_k\}$ satisfying

$$q_k := \frac{Aq_{k-1}}{\|Aq_{k-1}\|}, \quad (k = 1, 2, \dots)$$

Power Iteration

- Given a matrix $A \in \mathbf{R}^{n \times n}$ and an initial vector $q_0 \in \mathbf{C}^n$. The power iteration generates the sequence of vectors $\{q_k\}$ satisfying

$$q_k := \frac{Aq_{k-1}}{\|Aq_{k-1}\|}, \quad (k = 1, 2, \dots)$$

- An explicit formula for the iterate q_k

$$q_1 = \frac{Aq_0}{c_1}$$

Power Iteration

- Given a matrix $A \in \mathbf{R}^{n \times n}$ and an initial vector $q_0 \in \mathbf{C}^n$. The power iteration generates the sequence of vectors $\{q_k\}$ satisfying

$$q_k := \frac{Aq_{k-1}}{\|Aq_{k-1}\|}, \qquad (k = 1, 2, \ldots)$$

- An explicit formula for the iterate q_k

$$q_1 = \frac{Aq_0}{c_1}, \quad q_2 = \frac{Aq_1}{c_2}$$

Power Iteration

- Given a matrix $A \in \mathbf{R}^{n\times n}$ and an initial vector $q_0 \in \mathbf{C}^n$. The power iteration generates the sequence of vectors $\{q_k\}$ satisfying

$$q_k := \frac{Aq_{k-1}}{\|Aq_{k-1}\|}, \quad (k = 1, 2, \dots)$$

- An explicit formula for the iterate q_k

$$q_1 = \frac{Aq_0}{c_1}, \; q_2 = \frac{Aq_1}{c_2} = \frac{A\left((Aq_0)/c_1\right)}{c_2}$$

Power Iteration

- Given a matrix $A \in \mathbf{R}^{n\times n}$ and an initial vector $q_0 \in \mathbf{C}^n$. The power iteration generates the sequence of vectors $\{q_k\}$ satisfying

$$q_k := \frac{Aq_{k-1}}{\|Aq_{k-1}\|}, \quad (k = 1, 2, \dots)$$

- An explicit formula for the iterate q_k

$$q_1 = \frac{Aq_0}{c_1}, \;\; q_2 = \frac{Aq_1}{c_2} = \frac{A\left((Aq_0)/c_1\right)}{c_2} = \frac{A^2 q_0}{c_1 c_2},$$

Power Iteration

- Given a matrix $A \in \mathbf{R}^{n\times n}$ and an initial vector $q_0 \in \mathbf{C}^n$. The power iteration generates the sequence of vectors $\{q_k\}$ satisfying

$$q_k := \frac{Aq_{k-1}}{\|Aq_{k-1}\|}, \qquad (k = 1, 2, \dots)$$

- An explicit formula for the iterate q_k

$$q_1 = \frac{Aq_0}{c_1}, \;\; q_2 = \frac{Aq_1}{c_2} = \frac{A\left((Aq_0)/c_1\right)}{c_2} = \frac{A^2 q_0}{c_1 c_2}, \quad \dots, \quad q_k = \frac{A^k q_0}{c_1 c_2 \dots c_k}$$

Power Iteration

- Given a matrix $A \in \mathbf{R}^{n \times n}$ and an initial vector $q_0 \in \mathbf{C}^n$. The power iteration generates the sequence of vectors $\{q_k\}$ satisfying

$$q_k := \frac{Aq_{k-1}}{\|Aq_{k-1}\|}, \qquad (k = 1, 2, \ldots)$$

- An explicit formula for the iterate q_k

$$q_1 = \frac{Aq_0}{c_1}, \quad q_2 = \frac{Aq_1}{c_2} = \frac{A\left((Aq_0)/c_1\right)}{c_2} = \frac{A^2 q_0}{c_1 c_2}, \quad \ldots, \quad q_k = \frac{A^k q_0}{c_1 c_2 \ldots c_k}$$

- Since $\|q_k\| = 1$, it follows that $(c_1 c_2 \ldots c_k) = \|A^k q_0\|$ and we have

$$q_k = \frac{A^k q_0}{\|A^k q_0\|}$$

Power Iteration

- Suppose A has n linearly independent vectors
  - the eigenvalues $\lambda_1, \ldots, \lambda_n$ (s.t. $|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_n|$)
  - and the associated eigenvectors $v_1, v_2, \ldots, v_n$.

Power Iteration

- Suppose A has n linearly independent vectors
  - the eigenvalues $\lambda_1, \ldots, \lambda_n$ (s.t. $|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_n|$)
  - and the associated eigenvectors $v_1, v_2, \ldots, v_n$.
- Since $\{v_1, v_2, \ldots, v_n\}$ is linearly independent, it is a basis for $\mathbf{C}^n$ and

$$q_0 = c_1 v_1 + c_2 v_2 + \cdots + c_n v_n \qquad \text{for some } c_1, \ldots, c_n \in \mathbf{C}$$

Power Iteration

- Suppose A has n linearly independent vectors
  - the eigenvalues $\lambda_1, \ldots, \lambda_n$ (s.t. $|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_n|$)
  - and the associated eigenvectors $v_1, v_2, \ldots, v_n$.
- Since $\{v_1, v_2, \ldots, v_n\}$ is linearly independent, it is a basis for $\mathbf{C}^n$ and

$$q_0 = c_1 v_1 + c_2 v_2 + \cdots + c_n v_n \qquad \text{for some } c_1, \ldots, c_n \in \mathbf{C}$$

- Reconsider the explicit formula for q_k in terms of eigenvectors

$$q_k = \frac{A^k q_0}{\|A^k q_0\|} = \frac{A^k(c_1 v_1 + c_2 v_2 + \cdots + c_n v_n)}{\|A^k(c_1 v_1 + c_2 v_2 + \cdots + c_n v_n)\|}$$

Power Iteration

- Suppose A has n linearly independent vectors
  - the eigenvalues $\lambda_1, \dots, \lambda_n$ (s.t. $|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_n|$)
  - and the associated eigenvectors $v_1, v_2, \dots, v_n$.
- Since $\{v_1, v_2, \dots, v_n\}$ is linearly independent, it is a basis for $\mathbf{C}^n$ and

$$q_0 = c_1 v_1 + c_2 v_2 + \cdots + c_n v_n \qquad \text{for some } c_1, \dots, c_n \in \mathbf{C}$$

- Reconsider the explicit formula for q_k in terms of eigenvectors

$$q_k = \frac{A^k q_0}{\|A^k q_0\|} = \frac{A^k(c_1 v_1 + c_2 v_2 + \cdots + c_n v_n)}{\|A^k(c_1 v_1 + c_2 v_2 + \cdots + c_n v_n)\|}$$

- Notice that $A^k v_j = \lambda_j^k v_j$ for all j, *e.g.*

$$A^2 v_j = A(A v_j) = A(\lambda_j v_j) = \lambda_j^2 v_j$$

Power Iteration

- Therefore the explicit formula simplifies as

$$q_k = \frac{c_1\lambda_1^k v_1 + c_2\lambda_2^k v_2 + \cdots + c_n\lambda_n^k v_n}{\|c_1\lambda_1^k v_1 + c_2\lambda_2^k v_2 + \cdots + c_n\lambda_n^k v_n\|}$$

Power Iteration

- Therefore the explicit formula simplifies as

$$
\begin{aligned}
q_k &= \frac{c_1\lambda_1^k v_1 + c_2\lambda_2^k v_2 + \ldots c_n\lambda_n^k v_n}{\|c_1\lambda_1^k v_1 + c_2\lambda_2^k v_2 + \ldots c_n\lambda_n^k v_n\|} \\
&= \frac{\lambda_1^k}{|\lambda_1^k|} \frac{c_1 v_1 + c_2\left(\frac{\lambda_2}{\lambda_1}\right)^k v_2 + \cdots + c_n\left(\frac{\lambda_n}{\lambda_1}\right)^k v_n}{\left\|c_1 v_1 + c_2\left(\frac{\lambda_2}{\lambda_1}\right)^k v_2 + \cdots + c_n\left(\frac{\lambda_n}{\lambda_1}\right)^k v_n\right\|}
\end{aligned}
$$

Power Iteration

- Therefore the explicit formula simplifies as

$$
\begin{aligned}
q_k &= \frac{c_1\lambda_1^k v_1 + c_2\lambda_2^k v_2 + \ldots c_n\lambda_n^k v_n}{\|c_1\lambda_1^k v_1 + c_2\lambda_2^k v_2 + \ldots c_n\lambda_n^k v_n\|} \\
&= \frac{\lambda_1^k}{|\lambda_1^k|}\ \frac{c_1 v_1 + c_2\left(\frac{\lambda_2}{\lambda_1}\right)^k v_2 + \cdots + c_n\left(\frac{\lambda_n}{\lambda_1}\right)^k v_n}{\left\|c_1 v_1 + c_2\left(\frac{\lambda_2}{\lambda_1}\right)^k v_2 + \cdots + c_n\left(\frac{\lambda_n}{\lambda_1}\right)^k v_n\right\|}
\end{aligned}
$$

- As $k \to \infty$ the terms $\left(\frac{\lambda_2}{\lambda_1}\right)^k, \ldots, \left(\frac{\lambda_n}{\lambda_1}\right)^k$ approach zero implying

Power Iteration

- Therefore the explicit formula simplifies as

$$
\begin{aligned}
q_k &= \frac{c_1\lambda_1^k v_1 + c_2\lambda_2^k v_2 + \ldots c_n\lambda_n^k v_n}{\|c_1\lambda_1^k v_1 + c_2\lambda_2^k v_2 + \ldots c_n\lambda_n^k v_n\|} \\
&= \frac{\lambda_1^k}{|\lambda_1^k|} \, \frac{c_1 v_1 + c_2\left(\frac{\lambda_2}{\lambda_1}\right)^k v_2 + \cdots + c_n\left(\frac{\lambda_n}{\lambda_1}\right)^k v_n}{\left\|c_1 v_1 + c_2\left(\frac{\lambda_2}{\lambda_1}\right)^k v_2 + \cdots + c_n\left(\frac{\lambda_n}{\lambda_1}\right)^k v_n\right\|}
\end{aligned}
$$

- As $k \to \infty$ the terms $\left(\frac{\lambda_2}{\lambda_1}\right)^k, \ldots, \left(\frac{\lambda_n}{\lambda_1}\right)^k$ approach zero implying

$$
\hat{v} := \lim_{k\to\infty} q_k = d\frac{c_1 v_1}{\|c_1 v_1\|} \qquad \text{where } \; d = \lim_{k\to\infty} \frac{\lambda_1^k}{|\lambda_1^k|} \; (|d| = 1)
$$

Power Iteration

- Therefore the explicit formula simplifies as

$$
\begin{aligned}
q_k &= \frac{c_1\lambda_1^k v_1 + c_2\lambda_2^k v_2 + \ldots c_n\lambda_n^k v_n}{\|c_1\lambda_1^k v_1 + c_2\lambda_2^k v_2 + \ldots c_n\lambda_n^k v_n\|} \\
&= \frac{\lambda_1^k}{|\lambda_1^k|} \frac{c_1 v_1 + c_2\left(\frac{\lambda_2}{\lambda_1}\right)^k v_2 + \cdots + c_n\left(\frac{\lambda_n}{\lambda_1}\right)^k v_n}{\left\|c_1 v_1 + c_2\left(\frac{\lambda_2}{\lambda_1}\right)^k v_2 + \cdots + c_n\left(\frac{\lambda_n}{\lambda_1}\right)^k v_n\right\|}
\end{aligned}
$$

- As $k \to \infty$ the terms $\left(\frac{\lambda_2}{\lambda_1}\right)^k, \ldots, \left(\frac{\lambda_n}{\lambda_1}\right)^k$ approach zero implying

$$\hat{v} := \lim_{k\to\infty} q_k = d\frac{c_1 v_1}{\|c_1 v_1\|} \qquad \text{where } \ d = \lim_{k\to\infty} \frac{\lambda_1^k}{|\lambda_1^k|} \ (|d| = 1)$$

- Assuming $c_1 \neq 0$ as $k \to \infty$ the sequence $\{q_k\}$ approaches an eigenvector associated with the eigenvalue with largest modulus.

Power Iteration

Summary

Power Iteration

Summary

- Convergence to dominant eigenvector:
  The sequence $\{q_k\}$ generated by the power iteration approaches a unit eigenvector $\hat{v}$ associated with the eigenvalue λ_1 with largest modulus (under reasonable assumptions).

Power Iteration

Summary

- Convergence to dominant eigenvector:
  The sequence $\{q_k\}$ generated by the power iteration approaches a unit eigenvector $\hat{v}$ associated with the eigenvalue λ_1 with largest modulus (under reasonable assumptions).

- Retrieval of dominant eigenvalue:
  Notice that $\hat{v}^* A \hat{v} = \hat{v}^* \lambda_1 \hat{v} = \lambda_1 \|\hat{v}\|^2 = \lambda_1$.

Power Iteration

<u>Pseudocode</u>

Given $A \in \mathbf{C}^{n \times n}$ and $q_0 \in \mathbf{C}^n$ s.t. $\|q_0\| = 1$.

for $k = 1, m$ **do**

$q_k \leftarrow Aq_{k-1}$

$q_k \leftarrow q_k / \|q_k\|$

end for

$v \leftarrow q_m$

$\lambda \leftarrow q_m^* A q_m$

Return (λ, v)

Power Iteration

Rate of Convergence : Suppose $\lim_{k\to\infty} v_k = \hat{v}$.

Power Iteration

Rate of Convergence : Suppose $\lim_{k\to\infty} v_k = \hat{v}$.

- Linear convergence : for some positive constant $c < 1$

$$\lim_{k\to\infty} \frac{\|v_{k+1} - \hat{v}\|}{\|v_k - \hat{v}\|} = c$$

Power Iteration

Rate of Convergence : Suppose $\lim_{k\to\infty} v_k = \hat{v}$.

- Linear convergence : for some positive constant $c < 1$

$$\lim_{k\to\infty} \frac{\|v_{k+1} - \hat{v}\|}{\|v_k - \hat{v}\|} = c$$

e.g. $\{10^{-k}\} = \{0.1, 0.01, 0.001, \ldots\}$ converges to 0 linearly.

Power Iteration

Rate of Convergence : Suppose $\lim_{k\to\infty} v_k = \hat{v}$.

- Linear convergence : for some positive constant $c < 1$

$$\lim_{k\to\infty} \frac{\|v_{k+1} - \hat{v}\|}{\|v_k - \hat{v}\|} = c$$

e.g. $\{10^{-k}\} = \{0.1, 0.01, 0.001, \ldots\}$ converges to 0 linearly.

$$\frac{10^{-(k+1)} - 0}{10^{-k} - 0} = 0.1$$

Power Iteration

Rate of Convergence : Suppose $\lim_{k\to\infty} v_k = \hat{v}$.

- Linear convergence : for some positive constant $c < 1$

$$\lim_{k\to\infty} \frac{\|v_{k+1} - \hat{v}\|}{\|v_k - \hat{v}\|} = c$$

e.g. $\{10^{-k}\} = \{0.1, 0.01, 0.001, \dots\}$ converges to 0 linearly.

$$\frac{10^{-(k+1)} - 0}{10^{-k} - 0} = 0.1$$

- Quadratic convergence : for some positive constant c

$$\lim_{k\to\infty} \frac{\|v_{k+1} - \hat{v}\|}{\|v_k - \hat{v}\|^2} = c.$$

Power Iteration

Rate of Convergence : Suppose $\lim_{k\to\infty} v_k = \hat{v}$.

- Linear convergence : for some positive constant $c < 1$

$$\lim_{k\to\infty} \frac{\|v_{k+1} - \hat{v}\|}{\|v_k - \hat{v}\|} = c$$

e.g. $\{10^{-k}\} = \{0.1, 0.01, 0.001, \ldots\}$ converges to 0 linearly.

$$\frac{10^{-(k+1)} - 0}{10^{-k} - 0} = 0.1$$

- Quadratic convergence : for some positive constant c

$$\lim_{k\to\infty} \frac{\|v_{k+1} - \hat{v}\|}{\|v_k - \hat{v}\|^2} = c.$$

e.g. $\{10^{-2^k}\} = \{10^{-2}, 10^{-4}, 10^{-8}, 10^{-16}, \ldots\}$ converges to 0 quadratically.

Power Iteration

Rate of Convergence : Suppose $\lim_{k\to\infty} v_k = \hat{v}$.

- Linear convergence : for some positive constant $c < 1$

$$\lim_{k\to\infty} \frac{\|v_{k+1} - \hat{v}\|}{\|v_k - \hat{v}\|} = c$$

e.g. $\{10^{-k}\} = \{0.1, 0.01, 0.001, \ldots\}$ converges to 0 linearly.

$$\frac{10^{-(k+1)} - 0}{10^{-k} - 0} = 0.1$$

- Quadratic convergence : for some positive constant c

$$\lim_{k\to\infty} \frac{\|v_{k+1} - \hat{v}\|}{\|v_k - \hat{v}\|^2} = c.$$

e.g. $\{10^{-2^k}\} = \{10^{-2}, 10^{-4}, 10^{-8}, 10^{-16}, \ldots\}$ converges to 0 quadratically.

$$\frac{10^{-2^{k+1}} - 0}{(10^{-2^k} - 0)^2} = \frac{10^{-2^{k+1}}}{10^{-2^k} 10^{-2^k}} = \frac{10^{-2^{k+1}}}{10^{-2^{k+1}}} = 1$$

Power Iteration

- Rate of Convergence: It can be shown that for some constant c

$$\lim_{k\to\infty} \frac{\|\hat{v} - q_{k+1}\|}{\|\hat{v} - q_k\|} = c\left|\frac{\lambda_2}{\lambda_1}\right|$$

Power Iteration

- Rate of Convergence: It can be shown that for some constant c

$$\lim_{k\to\infty} \frac{\|\hat{v} - q_{k+1}\|}{\|\hat{v} - q_k\|} = c \left|\frac{\lambda_2}{\lambda_1}\right|$$

  - When it is convergent, the power iteration converges only linearly.
  - The closer the moduli of the eigenvalues λ_2 and λ_1 are, the slower the convergence is.

Power Iteration

- Rate of Convergence: It can be shown that for some constant c

$$\lim_{k\to\infty} \frac{\|\hat{v} - q_{k+1}\|}{\|\hat{v} - q_k\|} = c\left|\frac{\lambda_2}{\lambda_1}\right|$$

  - When it is convergent, the power iteration converges only linearly.
  - The closer the moduli of the eigenvalues λ_2 and λ_1 are, the slower the convergence is.

- Dominant Eigenvalue:

Power Iteration

- Rate of Convergence: It can be shown that for some constant c

$$\lim_{k\to\infty} \frac{\|\hat{v} - q_{k+1}\|}{\|\hat{v} - q_k\|} = c \left| \frac{\lambda_2}{\lambda_1} \right|$$

  - When it is convergent, the power iteration converges only linearly.
  - The closer the moduli of the eigenvalues λ_2 and λ_1 are, the slower the convergence is.

- Dominant Eigenvalue: The eigenvalue with largest modulus is given by $q(\hat{v})$ where

$$r(x) = \frac{x^* A x}{x^* x}$$

  is called the *Rayleigh quotient* of $x \in \mathbf{C}^n$.

Power Iteration

- Rate of Convergence: It can be shown that for some constant c

$$\lim_{k\to\infty} \frac{\|\hat{v} - q_{k+1}\|}{\|\hat{v} - q_k\|} = c\left|\frac{\lambda_2}{\lambda_1}\right|$$

  - When it is convergent, the power iteration converges only linearly.
  - The closer the moduli of the eigenvalues λ_2 and λ_1 are, the slower the convergence is.

- Dominant Eigenvalue: The eigenvalue with largest modulus is given by $q(\hat{v})$ where

$$r(x) = \frac{x^* A x}{x^* x}$$

  is called the *Rayleigh quotient* of $x \in \mathbf{C}^n$.

  - Note that $r(\hat{v}) = \frac{\hat{v}^* A \hat{v}}{\hat{v}^* \hat{v}} = \frac{\hat{v}^* \lambda_1 \hat{v}}{\hat{v}^* \hat{v}} = \lambda_1$

Inverse Iteration

- Power iteration suffers from slow convergence, when $|\lambda_1| \approx |\lambda_2|$.

Inverse Iteration

- Power iteration suffers from slow convergence, when $|\lambda_1| \approx |\lambda_2|$.
- A key observation to speed-up power iteration

$$Av = \lambda v \quad \Longleftrightarrow \quad Av - \mu v = \lambda v - \mu v$$

Inverse Iteration

- Power iteration suffers from slow convergence, when $|\lambda_1| \approx |\lambda_2|$.
- A key observation to speed-up power iteration

$$
\begin{aligned}
Av = \lambda v \quad &\Longleftrightarrow \quad Av - \mu v = \lambda v - \mu v \\
&\Longleftrightarrow \quad (A - \mu I)v = (\lambda - \mu)v
\end{aligned}
$$

Inverse Iteration

- Power iteration suffers from slow convergence, when $|\lambda_1| \approx |\lambda_2|$.

- A key observation to speed-up power iteration

$$
\begin{aligned}
Av = \lambda v \quad &\Longleftrightarrow \quad Av - \mu v = \lambda v - \mu v \\
&\Longleftrightarrow \quad (A - \mu I)v = (\lambda - \mu)v \\
&\Longleftrightarrow \quad (\lambda - \mu)^{-1} v = (A - \mu I)^{-1} v
\end{aligned}
$$

Inverse Iteration

- Power iteration suffers from slow convergence, when $|\lambda_1| \approx |\lambda_2|$.
- A key observation to speed-up power iteration

$$
\begin{aligned}
Av = \lambda v \quad &\Longleftrightarrow \quad Av - \mu v = \lambda v - \mu v \\
&\Longleftrightarrow \quad (A - \mu I)v = (\lambda - \mu)v \\
&\Longleftrightarrow \quad (\lambda - \mu)^{-1} v = (A - \mu I)^{-1} v
\end{aligned}
$$

(λ, v) is an eigenpair of $A \Longleftrightarrow ((\lambda - \mu)^{-1}, v)$ is an eigenpair of $(A - \mu I)^{-1}$

Inverse Iteration

- Power iteration suffers from slow convergence, when $|\lambda_1| \approx |\lambda_2|$.

- A key observation to speed-up power iteration

$$\begin{aligned} Av = \lambda v \quad &\Longleftrightarrow \quad Av - \mu v = \lambda v - \mu v \\ &\Longleftrightarrow \quad (A - \mu I)v = (\lambda - \mu)v \\ &\Longleftrightarrow \quad (\lambda - \mu)^{-1}v = (A - \mu I)^{-1}v \end{aligned}$$

(λ, v) is an eigenpair of $A \Longleftrightarrow ((\lambda-\mu)^{-1}, v)$ is an eigenpair of $(A-\mu I)^{-1}$

- Suppose σ is a good estimate of an eigenvalue λ_l.

Inverse Iteration

- Power iteration suffers from slow convergence, when $|\lambda_1| \approx |\lambda_2|$.

- A key observation to speed-up power iteration

$$
\begin{aligned}
Av = \lambda v \quad &\Longleftrightarrow \quad Av - \mu v = \lambda v - \mu v \\
&\Longleftrightarrow \quad (A - \mu I)v = (\lambda - \mu)v \\
&\Longleftrightarrow \quad (\lambda - \mu)^{-1} v = (A - \mu I)^{-1} v
\end{aligned}
$$

(λ, v) is an eigenpair of $A \Longleftrightarrow ((\lambda-\mu)^{-1}, v)$ is an eigenpair of $(A-\mu I)^{-1}$

- Suppose σ is a good estimate of an eigenvalue λ_l.
  - That is $\frac{1}{|\lambda_l - \sigma|} \gg \frac{1}{|\lambda_j - \sigma|}$ for all $j \neq l$.

Inverse Iteration

- Power iteration suffers from slow convergence, when $|\lambda_1| \approx |\lambda_2|$.
- A key observation to speed-up power iteration

$$
\begin{aligned}
Av = \lambda v \quad &\Longleftrightarrow \quad Av - \mu v = \lambda v - \mu v \\
&\Longleftrightarrow \quad (A - \mu I)v = (\lambda - \mu)v \\
&\Longleftrightarrow \quad (\lambda - \mu)^{-1} v = (A - \mu I)^{-1} v
\end{aligned}
$$

(λ, v) is an eigenpair of $A \Longleftrightarrow ((\lambda-\mu)^{-1}, v)$ is an eigenpair of $(A-\mu I)^{-1}$

- Suppose σ is a good estimate of an eigenvalue λ_l.
  - That is $\frac{1}{|\lambda_l - \sigma|} \gg \frac{1}{|\lambda_j - \sigma|}$ for all $j \neq l$.
  - The eigenvalues of $(A - \sigma I)^{-1}$ are $\frac{1}{\lambda_1 - \sigma}, \quad \frac{1}{\lambda_2 - \sigma}, \quad \ldots, \quad \frac{1}{\lambda_n - \sigma}$

Inverse Iteration

- Power iteration suffers from slow convergence, when $|\lambda_1| \approx |\lambda_2|$.
- A key observation to speed-up power iteration

$$
\begin{aligned}
Av = \lambda v \quad & \iff \quad Av - \mu v = \lambda v - \mu v \\
& \iff \quad (A - \mu I)v = (\lambda - \mu)v \\
& \iff \quad (\lambda - \mu)^{-1} v = (A - \mu I)^{-1} v
\end{aligned}
$$

(λ, v) is an eigenpair of $A \iff ((\lambda - \mu)^{-1}, v)$ is an eigenpair of $(A - \mu I)^{-1}$

- Suppose σ is a good estimate of an eigenvalue λ_l.
  - That is $\frac{1}{|\lambda_l - \sigma|} \gg \frac{1}{|\lambda_j - \sigma|}$ for all $j \neq l$.
  - The eigenvalues of $(A - \sigma I)^{-1}$ are $\frac{1}{\lambda_1 - \sigma}, \quad \frac{1}{\lambda_2 - \sigma}, \quad \ldots, \quad \frac{1}{\lambda_n - \sigma}$
  - Power iteration applied to $(A - \sigma I)^{-1}$ must converge to v_l (associated with the eigenvalue $\frac{1}{|\lambda_l - \sigma|}$) quickly.

Inverse Iteration

- Rate of Convergence: Let λ_j be the eigenvalue second closest to σ.

$$\lim_{k\to\infty} \frac{\|\hat{v} - q_{k+1}\|}{\|\hat{v} - q_k\|} = c\left|\frac{1/(\lambda_j - \sigma)}{1/(\lambda_l - \sigma)}\right| = c\left|\frac{\lambda_l - \sigma}{\lambda_j - \sigma}\right|$$

Inverse Iteration

- Rate of Convergence: Let λ_j be the eigenvalue second closest to σ.

$$\lim_{k\to\infty} \frac{\|\hat{v} - q_{k+1}\|}{\|\hat{v} - q_k\|} = c\left|\frac{1/(\lambda_j - \sigma)}{1/(\lambda_l - \sigma)}\right| = c\left|\frac{\lambda_l - \sigma}{\lambda_j - \sigma}\right|$$

- Inverse iteration requires the product $(A - \sigma I)^{-1} q_k$, equivalently the solution of the linear system $(A - \sigma I)x = q_k$, at each iteration.

Inverse Iteration

- Rate of Convergence: Let λ_j be the eigenvalue second closest to σ.

$$\lim_{k\to\infty} \frac{\|\hat{v} - q_{k+1}\|}{\|\hat{v} - q_k\|} = c\left|\frac{1/(\lambda_j - \sigma)}{1/(\lambda_l - \sigma)}\right| = c\left|\frac{\lambda_l - \sigma}{\lambda_j - \sigma}\right|$$

- Inverse iteration requires the product $(A - \sigma I)^{-1} q_k$, equivalently the solution of the linear system $(A - \sigma I)x = q_k$, at each iteration.
  - In practice an LU factorization of $(A - \sigma I)$ is computed initially (at a cost of $2n^3/3$).

Inverse Iteration

- Rate of Convergence: Let λ_j be the eigenvalue second closest to σ.

$$\lim_{k\to\infty} \frac{\|\hat{v} - q_{k+1}\|}{\|\hat{v} - q_k\|} = c\left|\frac{1/(\lambda_j - \sigma)}{1/(\lambda_l - \sigma)}\right| = c\left|\frac{\lambda_l - \sigma}{\lambda_j - \sigma}\right|$$

- Inverse iteration requires the product $(A - \sigma I)^{-1} q_k$, equivalently the solution of the linear system $(A - \sigma I)x = q_k$, at each iteration.
  - In practice an LU factorization of $(A - \sigma I)$ is computed initially (at a cost of $2n^3/3$).
  - At each iteration the system

    $$(A - \sigma I)x = LUx = q_j$$

    is solved by forward and back substitutions (at a cost of $O(n^2)$).

Inverse Iteration

Pseudocode

Given $A \in \mathbf{C}^{n\times n}$, $q_0 \in \mathbf{C}^n$ s.t. $\|q_0\| = 1$ and $\sigma \in \mathbf{C}$.

Compute an LU factorization of $(A - \sigma I)$

for $k = 1, m$ **do**

  Solve $L\hat{x} = q_{k-1}$ by forward substitution.

  Solve $Ux = \hat{x}$ by back substitution.

  $q_k \leftarrow x/\|x\|$

end for

$v \leftarrow q_m$

$\lambda \leftarrow q_m^* A q_m$

Return (λ, v)

Rayleigh Iteration

- Rayleigh iteration is similar to the inverse iteration with the exception that the shifts σ are set to the Rayleigh quotient at every iteration, *i.e.*

$$q_k := \frac{(A - \sigma_{k-1} I)^{-1} q_{k-1}}{\|(A - \sigma_{k-1} I)^{-1} q_{k-1}\|} \qquad \text{where } \sigma_{k-1} := r(q_{k-1}) = \frac{q_{k-1}^* A q_{k-1}}{q_{k-1}^* q_{k-1}}$$

Rayleigh Iteration

- Rayleigh iteration is similar to the inverse iteration with the exception that the shifts σ are set to the Rayleigh quotient at every iteration, *i.e.*

$$q_k := \frac{(A - \sigma_{k-1} I)^{-1} q_{k-1}}{\|(A - \sigma_{k-1} I)^{-1} q_{k-1}\|} \qquad \text{where } \sigma_{k-1} := r(q_{k-1}) = \frac{q_{k-1}^* A q_{k-1}}{q_{k-1}^* q_{k-1}}$$

- Upside: Rayleigh iteration usually converges to an eigenvector v_l associated with an eigenvalue λ_l very quickly.

Rayleigh Iteration

- Rayleigh iteration is similar to the inverse iteration with the exception that the shifts σ are set to the Rayleigh quotient at every iteration, *i.e.*

$$q_k := \frac{(A - \sigma_{k-1} I)^{-1} q_{k-1}}{\|(A - \sigma_{k-1} I)^{-1} q_{k-1}\|} \qquad \text{where } \sigma_{k-1} := r(q_{k-1}) = \frac{q_{k-1}^* A q_{k-1}}{q_{k-1}^* q_{k-1}}$$

- Upside: Rayleigh iteration usually converges to an eigenvector v_l associated with an eigenvalue λ_l very quickly.
  - The quick convergence is due to the fact that $r(q_k)$ becomes an increasingly better estimate of $r(v_l) = \lambda_l$ as q_k approaches v_l.

Rayleigh Iteration

- Rate of Convergence: Suppose $\lim_{k\to\infty} q_k = \hat{v}$. Then

$$\lim_{k\to\infty} \frac{\|\hat{v} - q_{k+1}\|}{\|\hat{v} - q_k\|^2} = c$$

  - Rate of convergence is quadratic.

Rayleigh Iteration

- Rate of Convergence: Suppose $\lim_{k\to\infty} q_k = \hat{v}$. Then

$$\lim_{k\to\infty} \frac{\|\hat{v} - q_{k+1}\|}{\|\hat{v} - q_k\|^2} = c$$

  - Rate of convergence is quadratic.

- Downside: At each iteration an LU factorization of $(A - \sigma_k I)$ needs to be computed from scratch to solve $(A - \sigma_k I)x = q_k$ for x.
  - Each iteration costs $\frac{2n^3}{3}$ flops.

Rayleigh Iteration

Pseudocode

Given $A \in \mathbf{C}^{n \times n}$ and $q_0 \in \mathbf{C}^n$ s.t. $\|q_0\| = 1$.

for $k = 1, m$ **do**

$\sigma_{k-1} \leftarrow q_{k-1}^* A q_{k-1}$

Compute an LU factorization of $(A - \sigma_{k-1} I)$

Solve $L\hat{x} = q_{k-1}$ by forward substitution.

Solve $Ux = \hat{x}$ by back substitution.

$q_k \leftarrow x / \|x\|$

end for

$v \leftarrow q_m$

$\lambda \leftarrow q_m^* A q_m$

Return (λ, v)

Next Lecture

The QR Algorithm (Fausett - 5.3, Watkins - 5.6)