The Matrix Eigenvalue Problem

Niels Henrik Abel
1802-1829

Niels Henrik Abel
1802-1829

Abel's Theorem: There are no formulas for finding the roots of generic polynomial of degree greater than 4.

Power Method and its Variations

Power Method

Let $A \in \mathbb{C}^{n\times n}$ be diagonalizable with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{C}$ satisfying

$$|\lambda_1| \gneqq |\lambda_2| \geq \cdots \geq |\lambda_n|$$

and let $v_1, \ldots, v_n \in \mathbb{C}^n \setminus \{0\}$ such that $Av_i = \lambda_i v_i, i = 1, 2, \ldots n$. λ_1 is called the dominant eigenvalue of A and v_1 a corresponding dominant eigenvector.

Let $x \in \mathbb{C}^n$ such that $x = c_1 v_1 + \cdots + c_n v_n$ with $c_1 \neq 0$. Then,

$$\left\| A^j(x)/\lambda_1^j - c_1 v_1 \right\| \to 0 \text{ as } j \to \infty.$$

Power Method

Let $A \in \mathbb{C}^{n\times n}$ be diagonalizable with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{C}$ satisfying

$$|\lambda_1| \gneq |\lambda_2| \geq \cdots \geq |\lambda_n|$$

and let $v_1, \ldots, v_n \in \mathbb{C}^n \setminus \{0\}$ such that $Av_i = \lambda_i v_i, i = 1, 2, \ldots n$. λ_1 is called the dominant eigenvalue of A and v_1 a corresponding dominant eigenvector.

Let $x \in \mathbb{C}^n$ such that $x = c_1 v_1 + \cdots + c_n v_n$ with $c_1 \neq 0$. Then,

$$\left\| A^j(x)/\lambda_1^j - c_1 v_1 \right\| \to 0 \text{ as } j \to \infty.$$

Moreover, if $c_2 \neq 0$ and $|\lambda_2| > |\lambda_3|$, then the convergence is linear at the rate $|\lambda_2|/|\lambda_1|$, i.e.,

$$\lim_{j\to\infty} \frac{\left\| A^{(j+1)}(x)/\lambda_1^{(j+1)} - c_1 v_1 \right\|}{\left\| A^j(x)/\lambda_1^j - c_1 v_1 \right\|} = \frac{|\lambda_2|}{|\lambda_1|}$$

(Ex: Prove the above limit!)

Power Method

Let $x = [x_1, x_2, \cdots x_n]^T \in \mathbb{C}^n \setminus \{0\}$ be arbitrarily chosen. Set $q_0 = x/s_0$ where $s_0 = x_i$ such that $|x_i| = \|x\|_\infty$.

for $j = 1, 2, \ldots$

Set $\hat{q}_j = A(q_{j-1})$

Find $s_j = \hat{q}_j(i)$ *such that* $|\hat{q}_j(i)| = \|\hat{q}_j\|_\infty$.

Set $q_j = \hat{q}_j / s_j$.

Power Method

Let $x = [x_1, x_2, \cdots x_n]^T \in \mathbb{C}^n \setminus \{0\}$ be arbitrarily chosen. Set $q_0 = x/s_0$ where $s_0 = x_i$ such that $|x_i| = \|x\|_\infty$.

for $j = 1, 2, \ldots$

Set $\hat{q}_j = A(q_{j-1})$

Find $s_j = \hat{q}_j(i)$ *such that* $|\hat{q}_j(i)| = \|\hat{q}_j\|_\infty$.

Set $q_j = \hat{q}_j/s_j$.

If $x = c_1 v_1 + \cdots + c_n v_n$ with with $c_1 \neq 0$, then

(i) $\lim_{j\to\infty} q_j = \hat{v}_1$, where $A\hat{v}_1 = \lambda_1 \hat{v}_1$, with $\|\hat{v}_1\|_\infty = 1$, and $\hat{v}_1(j) = 1$ for *some* $1 \leq j \leq n$.

(ii) $\lim_{j\to\infty} s_j = \lambda_1$.

Power Method

Let $x = [x_1, x_2, \cdots x_n]^T \in \mathbb{C}^n \setminus \{0\}$ be arbitrarily chosen. Set $q_0 = x/s_0$ where $s_0 = x_i$ such that $|x_i| = \|x\|_\infty$.

for $j = 1, 2, \ldots$
  Set $\hat{q}_j = A(q_{j-1})$
  Find $s_j = \hat{q}_j(i)$ *such that* $|\hat{q}_j(i)| = \|\hat{q}_j\|_\infty$.
  Set $q_j = \hat{q}_j / s_j$.

If $x = c_1 v_1 + \cdots + c_n v_n$ with with $c_1 \neq 0$, then

(i) $\lim_{j\to\infty} q_j = \hat{v}_1$, where $A\hat{v}_1 = \lambda_1 \hat{v}_1$, with $\|\hat{v}_1\|_\infty = 1$, and $\hat{v}_1(j) = 1$ for *some* $1 \leq j \leq n$.

(ii) $\lim_{j\to\infty} s_j = \lambda_1$.

(Ex: Prove these!)

Power Method

Further, if $c_1, c_2 \neq 0$ and $|\lambda_1| > |\lambda_2| > |\lambda_3|$, then $\{q_j\}$ converges to $\hat{v}_1$ linearly at the rate $\frac{|\lambda_2|}{|\lambda_1|}$, that is,

$$\lim_{j \to \infty} \frac{\|q_{j+1} - \hat{v}_1\|}{\|q_j - \hat{v}_1\|} = \frac{|\lambda_2|}{|\lambda_1|},$$

Power Method

Further, if $c_1, c_2 \neq 0$ and $|\lambda_1| > |\lambda_2| > |\lambda_3|$, then $\{q_j\}$ converges to $\hat{v}_1$ linearly at the rate $\frac{|\lambda_2|}{|\lambda_1|}$, that is,

$$\lim_{j\to\infty} \frac{\|q_{j+1} - \hat{v}_1\|}{\|q_j - \hat{v}_1\|} = \frac{|\lambda_2|}{|\lambda_1|},$$

(Ex: Prove this!)

Power Method

Further, if $c_1, c_2 \neq 0$ and $|\lambda_1| > |\lambda_2| > |\lambda_3|$, then $\{q_j\}$ converges to $\hat{v}_1$ linearly at the rate $\frac{|\lambda_2|}{|\lambda_1|}$, that is,

$$\lim_{j\to\infty} \frac{\|q_{j+1} - \hat{v}_1\|}{\|q_j - \hat{v}_1\|} = \frac{|\lambda_2|}{|\lambda_1|},$$

(Ex: Prove this!)

The Power Method is used to compute a dominant eigenvector of the massive non-negative Google Matrix in Google's PageRank Algorithm. For details see:

K. Bryan and T. Leise. *The $25,000,000,000 Eigenvector: The Linear Algebra behind Google.* SIAM Rev., 48(3), 569-581.

Shift and Invert Method

For $\rho \in \mathbb{C}$, the eigenvalues of $A - \rho I$ are $\lambda_i - \rho$, $i = 1, \dots n$ where the listing $\lambda_1, \dots, \lambda_n$ is determined by

$$|\lambda_1 - \rho| \geq \cdots \geq |\lambda_{n-1} - \rho| \geq |\lambda_n - \rho|.$$

Suppose that $|\lambda_{n-1} - \rho| \gneq |\lambda_n - \rho|$.

Shift and Invert Method

For $\rho \in \mathbb{C}$, the eigenvalues of $A - \rho I$ are $\lambda_i - \rho$, $i = 1, \ldots n$ where the listing $\lambda_1, \ldots, \lambda_n$ is determined by

$$|\lambda_1 - \rho| \geq \cdots \geq |\lambda_{n-1} - \rho| \geq |\lambda_n - \rho|.$$

Suppose that $|\lambda_{n-1} - \rho| \gneq |\lambda_n - \rho|$. Then $1/(\lambda_n - \rho)$ is a dominant eigenvalue of $(A - \rho I)^{-1}$.

Shift and Invert Method

For $\rho \in \mathbb{C}$, the eigenvalues of $A - \rho I$ are $\lambda_i - \rho$, $i = 1, \ldots n$ where the listing $\lambda_1, \ldots, \lambda_n$ is determined by

$$|\lambda_1 - \rho| \geq \cdots \geq |\lambda_{n-1} - \rho| \geq |\lambda_n - \rho|.$$

Suppose that $|\lambda_{n-1} - \rho| \gneq |\lambda_n - \rho|$. Then $1/(\lambda_n - \rho)$ is a dominant eigenvalue of $(A - \rho I)^{-1}$.

Let $x = [x_1, x_2, \cdots x_n]^T \in \mathbb{C}^n \setminus \{0\}$ be arbitrarily chosen. Set $q_0 = x/s_0$ where $s_0 = x_i$ such that $|x_i| = \|x\|_\infty$.

for $j = 1, 2, \ldots$
  Set $\hat{q}_j = (A - \rho I)^{-1}(q_{j-1})$
  Find $s_j = \hat{q}_j(i)$ *such that* $|\hat{q}_j(i)| = \|\hat{q}_j\|_\infty$.
  Set $q_j = \hat{q}_j / s_j$.

Shift and Invert Method

For $\rho \in \mathbb{C}$, the eigenvalues of $A - \rho I$ are $\lambda_i - \rho$, $i = 1, \ldots n$ where the listing $\lambda_1, \ldots, \lambda_n$ is determined by

$$|\lambda_1 - \rho| \geq \cdots \geq |\lambda_{n-1} - \rho| \geq |\lambda_n - \rho|.$$

Suppose that $|\lambda_{n-1} - \rho| \gneq |\lambda_n - \rho|$. Then $1/(\lambda_n - \rho)$ is a dominant eigenvalue of $(A - \rho I)^{-1}$.

Let $x = [x_1, x_2, \cdots x_n]^T \in \mathbb{C}^n \setminus \{0\}$ be arbitrarily chosen. Set $q_0 = x/s_0$ where $s_0 = x_i$ such that $|x_i| = \|x\|_\infty$.

for $j = 1, 2, \ldots$
  Set $\hat{q}_j = (A - \rho I)^{-1}(q_{j-1})$ *(Explicit inverse computation is a bad idea!)*
  Find $s_j = \hat{q}_j(i)$ *such that* $|\hat{q}_j(i)| = \|\hat{q}_j\|_\infty$.
  Set $q_j = \hat{q}_j/s_j$.

Shift and Invert Method

for $j = 1, 2, \ldots$

  Solve $(A - \rho I)\hat{q}_j = q_{j-1}$ *for* $\hat{q}_j$.

  Find $s_j = \hat{q}_j(i)$ *such that* $|\hat{q}_j(i)| = \|\hat{q}_j\|_\infty$.

  Set $q_j = \hat{q}_j / s_j$.

Shift and Invert Method

for $j = 1, 2, \ldots$

  Solve $(A - \rho I)\hat{q}_j = q_{j-1}$ *for* $\hat{q}_j$.

  (Costs $2n^3/3 + O(n^2)$ flops. Still not good enough!)

  Find $s_j = \hat{q}_j(i)$ *such that* $|\hat{q}_j(i)| = \|\hat{q}_j\|_\infty$.

  Set $q_j = \hat{q}_j / s_j$.

Shift and Invert Method

Best idea: Find a permutation matrix P, a unit lower triangular matrix L and upper triangular matrix U such that

$$P(A - \rho I) = LU$$

for $j = 1, 2, \ldots$
  Set $b = Pq_{j-1}$
  Solve $Ly = b$ *for* y *(Costs n^2 flops)*
  Solve $U\hat{q}_j = y$ *for* $\hat{q}_j$ *(Costs n^2 flops)*
  Find $s_j = \hat{q}_j(i)$ *such that* $|\hat{q}_j(i)| = \|\hat{q}_j\|_\infty$.
  Set $q_j = \hat{q}_j / s_j$.

Shift and Invert Method

Best idea: Find a permutation matrix P, a unit lower triangular matrix L and upper triangular matrix U such that

$$P(A - \rho I) = LU$$

for $j = 1, 2, \ldots$
  Set $b = Pq_{j-1}$
  Solve $Ly = b$ *for* y (Costs n^2 flops)
  Solve $U\hat{q}_j = y$ *for* $\hat{q}_j$ (Costs n^2 flops)
  Find $s_j = \hat{q}_j(i)$ *such that* $|\hat{q}_j(i)| = \|\hat{q}_j\|_\infty$.
  Set $q_j = \hat{q}_j / s_j$.

When the additional conditions for the Power Method to converge for $(A - \rho I)^{-1}$ are satisfied, the sequence $\{q_j\}$ converges to an eigenvector of A linearly at the rate

$$\frac{|\lambda_n - \rho|}{|\lambda_{n-1} - \rho|}.$$

Rayleigh Quotient Method

Let $q \in \mathbb{C}^n \setminus \{0\}$ and $A \in \mathbb{C}^{n \times n}$. Then $\rho := \frac{q^* A q}{q^* q}$ is called the Rayleigh Quotient associated with A and q.

If q is an eigenvector of A, then ρ is a corresponding eigenvalue of A. Else, ρ is the unique scalar that solves $\min\limits_{\mu \in \mathbb{C}} \|Aq - \mu q\|_2$.

Rayleigh Quotient Method

Let $q \in \mathbb{C}^n \setminus \{0\}$ and $A \in \mathbb{C}^{n\times n}$. Then $\rho := \frac{q^*Aq}{q^*q}$ is called the Rayleigh Quotient associated with A and q.

If q is an eigenvector of A, then ρ is a corresponding eigenvalue of A. Else, ρ is the unique scalar that solves $\min_{\mu\in\mathbb{C}} \|Aq - \mu q\|_2$.

Theorem Let $A \in \mathbb{C}^{n\times n}$ and $q, v \in \mathbb{C}^n$ with $\|q\|_2 = \|v\|_2 = 1$ and $Av = \lambda v$ for some scalar λ. Then $\rho := \frac{q^*Aq}{q^*q}$ satisfies

$$|\lambda - \rho| \leq 2\|A\|_2\|v - q\|_2.$$

Rayleigh Quotient Method

Set $q_0 = x/s_0$ where $x \in \mathbb{C}^n \setminus \{0\}$ and $s_0 = x_i$ satisfies $|s_0| = \|x\|_\infty$. Also set $\rho_0 = \frac{q_0^* A q_0}{q_0^* q_0}$.

for $j = 1, 2, \ldots$

Solve $(A - \rho_{j-1} I)\hat{q}_j = q_{j-1}$ *for* $\hat{q}_j$

Find $s_j = \hat{q}_j(i)$ *such that* $|\hat{q}_j(i)| = \|\hat{q}_j\|_\infty$.

Set $q_j = \hat{q}_j / s_j$ *and* $\rho_j = \frac{q_j^* A q_j}{q_j^* q_j}$.

Rayleigh Quotient Method

Set $q_0 = x/s_0$ where $x \in \mathbb{C}^n \setminus \{0\}$ and $s_0 = x_i$ satisfies $|s_0| = \|x\|_\infty$. Also set $\rho_0 = \frac{q_0^* A q_0}{q_0^* q_0}$.

for $j = 1, 2, \ldots$

Solve $(A - \rho_{j-1} I)\hat{q}_j = q_{j-1}$ *for* $\hat{q}_j$

(Costs $2n^3/3 + O(n^2)$ flops)

Find $s_j = \hat{q}_j(i)$ *such that* $|\hat{q}_j(i)| = \|\hat{q}_j\|_\infty$.

Set $q_j = \hat{q}_j/s_j$ *and* $\rho_j = \frac{q_j^* A q_j}{q_j^* q_j}$.

Upper Hessenberg Matrices

A matrix $A \in \mathbb{F}^{n\times n}$ is said to be upper Hessenberg if $a_{ij} = 0$ for $i > j + 1$. Thus A is of the form

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & \ddots & \ddots & \vdots \\ & & & \\ & & a_{n,n-1} & a_{nn} \end{bmatrix}.$$

A is said to be properly upper Hessenberg or an irreducible upper Hessenberg matrix if $a_{i+1,i} \neq 0$ for every $i = 1, 2, \ldots, n-1$.

Exercise: Finding a QR decomposition of an $n \times n$ upper Hessenberg matrix costs $O(n^2)$ flops. What is the special form of Q in this case?
Further, GEPP on an $n \times n$ upper Hessenberg matrix costs $O(n^2)$ flops.

Transformation to Upper Hessenberg form

Theorem 3 Given any matrix $A \in \mathbb{R}^{n \times n}$, there exists an orthogonal matrix Q and an upper Hessenberg matrix H such that $Q^T A Q = H$. If $A^T = A$, then, H is a symmetric tridiagonal matrix.
If $A \in \mathbb{C}^{n \times n}$, then there exists a unitary matrix Q such that $Q^* A Q = H$. In such a case if $A^* = A$, then H is a Hermitian tridiagonal matrix.

Rayleigh Quotient Method

Find a unitary matrix Q and upper-Hessenberg matrix H such that $Q^*AQ = H$ and perform Rayleigh Quotient iterations on H!

[Costs $O(n^3)$ flops]

Rayleigh Quotient Method

Find a unitary matrix Q and upper-Hessenberg matrix H such that $Q^*AQ = H$ and perform Rayleigh Quotient iterations on H!

[Costs $O(n^3)$ flops]

Set $q_0 = x/s_0$ where $x \in \mathbb{C}^n \setminus \{0\}$ and $s_0 = x_i$ satisfies $|s_0| = \|x\|_\infty$.
Also set $\rho_0 = \frac{q_0^* H q_0}{q_0^* q_0}$.

Rayleigh Quotient Method

Find a unitary matrix Q and upper-Hessenberg matrix H such that $Q^*AQ = H$ and perform Rayleigh Quotient iterations on H!

[Costs $O(n^3)$ flops]

Set $q_0 = x/s_0$ where $x \in \mathbb{C}^n \setminus \{0\}$ and $s_0 = x_i$ satisfies $|s_0| = \|x\|_\infty$.
Also set $\rho_0 = \frac{q_0^* H q_0}{q_0^* q_0}$.
for $j = 1, 2, \ldots$
  Solve $(H - \rho_{j-1} I)\hat{q}_j = q_{j-1}$ *for* $\hat{q}_j$
  (Costs $O(n^2)$ flops)
  Find $s_j = \hat{q}_j(i)$ *such that* $|\hat{q}_j(i)| = \|\hat{q}_j\|_\infty$.
  Set $q_j = \hat{q}_j/s_j$ *and* $\rho_j = \frac{q_j^* H q_j}{q_j^* q_j}$.

Rayleigh Quotient Method

Find a unitary matrix Q and upper-Hessenberg matrix H such that $Q^* AQ = H$ and perform Rayleigh Quotient iterations on H!

[Costs $O(n^3)$ flops]

Set $q_0 = x/s_0$ where $x \in \mathbb{C}^n \setminus \{0\}$ and $s_0 = x_i$ satisfies $|s_0| = \|x\|_\infty$.
Also set $\rho_0 = \frac{q_0^* H q_0}{q_0^* q_0}$.
for $j = 1, 2, \ldots$

Solve $(H - \rho_{j-1} I)\hat{q}_j = q_{j-1}$ *for* $\hat{q}_j$
(Costs $O(n^2)$ flops)
Find $s_j = \hat{q}_j(i)$ *such that* $|\hat{q}_j(i)| = \|\hat{q}_j\|_\infty$.
Set $q_j = \hat{q}_j / s_j$ *and* $\rho_j = \frac{q_j^* H q_j}{q_j^* q_j}$.

Rayleigh Quotient iterations may not converge!

Rayleigh Quotient Method

Find a unitary matrix Q and upper-Hessenberg matrix H such that $Q^*AQ = H$ and perform Rayleigh Quotient iterations on H!

[Costs $O(n^3)$ flops]

Set $q_0 = x/s_0$ where $x \in \mathbb{C}^n \setminus \{0\}$ and $s_0 = x_i$ satisfies $|s_0| = \|x\|_\infty$.
Also set $\rho_0 = \frac{q_0^* H q_0}{q_0^* q_0}$.
for $j = 1, 2, \ldots$
  Solve $(H - \rho_{j-1} I)\hat{q}_j = q_{j-1}$ *for* $\hat{q}_j$
  (Costs $O(n^2)$ flops)
  Find $s_j = \hat{q}_j(i)$ *such that* $|\hat{q}_j(i)| = \|\hat{q}_j\|_\infty$.
  Set $q_j = \hat{q}_j/s_j$ *and* $\rho_j = \frac{q_j^* H q_j}{q_j^* q_j}$.

Rayleigh Quotient iterations may not converge!

But when they do the convergence rate is usually quadratic. For Hermitian matrices, they convergence for *almost* all choices of starting vectors and when it happens, the convergence is cubic.