On Newton's method and Halley's method for *p*th roots of matrices

Chun-Hua Guo
University of Regina
Canada

May 15, 2008

Principal pth root

Let $p \geq 2$ be an integer. Suppose that $A \in \mathbb{C}^{n\times n}$ has no negative real eigenvalues and all zero eigenvalues are semisimple. Let the Jordan canonical form of A be

$$Z^{-1}AZ = \mathrm{diag}(J_1, J_2, \ldots, J_p).$$

Then the principal pth root of A is

$$A^{1/p} = Z\mathrm{diag}(J_1^{1/p}, J_2^{1/p}, \ldots, J_p^{1/p})Z^{-1},$$

where for $m_k \times m_k$ Jordan block $J_k = J_k(\lambda_k)$

$$J_k^{1/p} = \begin{bmatrix} f(\lambda_k) & f'(\lambda_k) & \cdots & \frac{f^{(m_k-1)}(\lambda_k)}{(m_k-1)!} \\ & f(\lambda_k) & \ddots & \vdots \\ & & \ddots & f'(\lambda_k) \\ & & & f(\lambda_k) \end{bmatrix} \text{ with } f(z) = z^{1/p}.$$

Newton's method and Halley's method

We can find $A^{1/p}$ without JCF.
Newton's method:

$$\begin{aligned} X_0 &= I, \\ X_{k+1} &= \frac{1}{p}\left((p-1)X_k + AX_k^{1-p}\right). \end{aligned}$$

Halley's method:

$$\begin{aligned} X_0 &= I, \\ X_{k+1} &= X_k\left((p+1)X_k^p + (p-1)A\right)^{-1}\left((p-1)X_k^p + (p+1)A\right). \end{aligned}$$

Stable versions of these methods are given in [Iannazzo, 2006] and [Iannazzo, 2007].

Residual relation for Newton's method

Let the residual be defined by $R(X_k) = I - AX_k^{-p}$.
(The usual definition $R(X_k) = X_k^p - A$ does not work well.)
Assume that $\rho(I - A) \le 1$. So $\rho(R(X_0)) \le 1$.
Assume that X_k is defined and nonsingular, with $\rho(R(X_k)) \le 1$.
Then $X_{k+1} = X_k(I - \frac{1}{p}R(X_k))$ is defined and nonsingular, and

$$R(X_{k+1}) = I - (I - \frac{1}{p}R(X_k))^{-p}(I - R(X_k)) = \sum_{i=2}^{\infty} c_i (R(X_k))^i,$$

where $c_i > 0$ for $i \ge 2$ and $\sum_{i=2}^{\infty} c_i = 1$. So $\rho(R(X_{k+1})) \le 1$.
If $\|R(X_0)\| \le 1$, then $\|R(X_k)\| \le \|(R(X_0))^{2^k}\|$.

The scalar case

Theorem

Let λ be any complex number in $E = \{z : |z-1| \leq 1\}$. Then Newton's method with $x_0 = 1$ converges to $\lambda^{1/p}$.

Proof.

If $\lambda = 0$, then $x_k = ((p-1)/p)^k$, converging to 0 linearly. If $\lambda \neq 0$ and $|\lambda - 1| \leq 1$, then $\{x_k\}$ is a Cauchy sequence (by the residual relation) and hence converges to a pth root of λ. However, the set E is connected and for $\lambda = 1 \in E$ we know that $\{x_k\}$ converges to 1, the principal pth root of 1. It follows that for each $\lambda \in E$, $\{x_k\}$ converges to the principal pth root of λ. □

Return to the matrix case

Scalar convergence implies matrix convergence [Higham, 2008, Theorem 4.15].

Theorem

If all eigenvalues of A are in $\{z : |z-1| \leq 1\}$ and all zero eigenvalues of A (if any) are semisimple, then Newton's method converges to $A^{1/p}$.

Theorem

[Iannazzo, 2006] *If all eigenvalues of A are in $\{z : \mathrm{Re} z > 0, |z| \leq 1\}$, then Newton's method converges to $A^{1/p}$.*

Residual relation for Halley's method

Let $R(X_k) = I - AX_k^{-p}$. When $\rho((R(X_k)) < \frac{2p}{p+1}$,

$$X_{k+1} = X_k \frac{I - \frac{p+1}{2p} R(X_k)}{I - \frac{p-1}{2p} R(X_k)}$$

is defined and nonsingular.

$$R(X_{k+1}) = I - \left(\frac{I - \frac{p-1}{2p} R(X_k)}{I - \frac{p+1}{2p} R(X_k)} \right)^p (I - R(X_k)).$$

Let

$$f(t) = 1 - \left(\frac{1 - \frac{p-1}{2p} t}{I - \frac{p+1}{2p} t} \right)^p (1 - t).$$

Properties of $f(t)$

For $|t| < \frac{2p}{p+1}$, $f(t) = \sum_{i=3}^{\infty} c_i t^i$, where $\sum_{i=3}^{\infty} c_i = 1$ and $c_3 = (p^2 - 1)/(12p^2)$.

Conjecture

$c_i > 0$ for all $i \geq 3$.

If this conjecture is proved, then:
If $\|R(X_0)\| \leq 1$ then $\|R(X_k)\| \leq \|R(X_0)^{3^k}\|$.
This would prove the convergence of Halley's method when $\sigma(A) \subset \{z : |z - 1| < 1\}$. [Iannazzo, 2007] proved the convergence of Halley's method when $\sigma(A) \subset \mathbb{C}_+$.
Without proving this conjecture, we have:
If $\|R(X_0)\| \leq q \leq 1$ for a sufficiently small q then $\|R(X_k)\| \leq \|R(X_0)^{3^k}\|$.

Newton, Halley, and binomial expansion

Theorem

Suppose that all eigenvalues of A are in $\{z : |z-1| < 1\}$ and write $A = I - B$ (so $\rho(B) < 1$). Let $(I-B)^{1/p} = \sum_{i=0}^{\infty} c_i B^i$ be the binomial expansion (so $c_i < 0$ for $i \geq 1$). Then the sequence X_k generated by Newton's method or by Halley's method has the Taylor expansion $X_k = \sum_{i=0}^{\infty} c_{k,i} B^i$. For Newton's method we have $c_{k,i} = c_i$ for $i = 0, 1, \ldots, 2^k - 1$, and for Halley's method we have $c_{k,i} = c_i$ for $i = 0, 1, \ldots, 3^k - 1$.

Proof.

For Newton's method, take $B = J(0)_{2^k \times 2^k}$. Then $\|R(X_k)\| \leq \|R(X_0)^{2^k}\| = \|B^{2^k}\| = 0$ so $X_k = (I-B)^{1/p}$, which implies $c_{k,i} = c_i$ for $i = 0, 1, \ldots, 2^k - 1$. □

Conjecture about $c_{k,i}$ for Newton's method

For Newton's method we know

- $c_{k,0} = 1$ for $k \geq 0$.
- $c_{0,i} = 0$ for $i \geq 1$.
- $c_{1,1} = -1/p$, $c_{1,i} = 0$ for $i \geq 2$.
- $c_{2,i} < 0$ for $i \geq 1$.
- $c_{k,i} < 0$ for $k \geq 3$ and $i = 1 : 2^k - 1$.
- $\sum_{i=0}^{\infty} c_{k,i} = ((p-1)/p)^k$ for $k \geq 0$.

Conjecture

For Newton's method, $c_{k,i} < 0$ for $k \geq 3$ and $i \geq 2^k$.

Conjecture about $c_{k,i}$ for Halley's method

For Halley's method we know

- $c_{k,0} = 1$ for $k \geq 0$.
- $c_{0,i} = 0$ for $i \geq 1$.
- $c_{1,i} < 0$ for $i \geq 1$.
- $c_{k,i} < 0$ for $k \geq 2$ and $i = 1 : 3^k - 1$.
- $\sum_{i=0}^{\infty} c_{k,i} = ((p-1)/(p+1))^k$ for $k \geq 0$.

Conjecture

For Halley's method, $c_{k,i} < 0$ for $k \geq 2$ and $i \geq 3^k$.

Scaling

If all eigenvalues of A are in $E = \{z : |z-1| < 1\}$, faster convergence can be achieved by a proper scaling of A. We write $A = c(I - B)$ with $c > 0$ and $B = I - \frac{1}{c}A$. Then $A^{1/p} = c^{1/p}(I-B)^{1/p}$ and $(I-B)^{1/p}$ is computed by Newton's method or Halley's method. The best constant c minimizes $\rho(B)$. The eigenvalues of A are known when A is obtained from a preprocessing procedure [Guo, Higham, 2006] based on the Schur decomposition and the computation of matrix square roots. If A has real eigenvalues $\lambda_1 \geq \ldots \geq \lambda_n$, the optimal c is $(\lambda_1 + \lambda_n)/2$. If A has complex eigenvalues, a near-optimal c can be obtained by using the idea in the proof of Proposition 4.5 in [Guo, Higham, 2007] and a bisection procedure on the parameter c.

The singular case

If A has semisimple zero eigenvalues, then the convergence of Newton iteration is linear with rate $(p-1)/p$, but accurate approximations to $A^{1/p}$ can be obtained by computing $Z_k = pX_k - (p-1)X_{k-1}$. For Halley's method one would use $Z_k = \frac{1}{2}\left((p+1)X_k - (p-1)X_{k-1}\right)$ for the improvement.

Proposition

Suppose that A has semisimple zero eigenvalues and let $R(Y) = Y^p - A$. Then for Newton's method or Halley's method

$$\|R(X_k)\| = O(\|X_k - A^{1/p}\|^p), \quad \|R(Z_k)\| = O(\|Z_k - A^{1/p}\|).$$

pth root of nonsingular H-matrix

It is shown in [Johnson, 1982] that $A^{1/p}$ is a nonsingular M-matrix for every nonsingular M-matrix A.

Theorem

Let A be a nonsingular H-matrix with positive diagonal entries. Then the principal pth root of A exists and is a nonsingular H-matrix whose diagonal entries have positive real parts.

Corollary

Let A be a real nonsingular H-matrix with positive diagonal entries. Then the principal pth root of A exists and is also a real nonsingular H-matrix with positive diagonal entries.

pth root of singular M-matrix

Theorem

Let A be a singular M-matrix with semisimple zero eigenvalues. Then $A^{1/p}$ is also a singular M-matrix with semisimple zero eigenvalues.

Corollary

Let A be an irreducible singular M-matrix. Then $A^{1/p}$ is also an irreducible singular M-matrix.

Computing pth root of nonsingular H-matrix

Let A be a nonsingular H-matrix with positive diagonal entries. Then the large convergence region [Iannazzo, 2007] for Halley's method allows one to compute $A^{1/p}$ directly by Halley's method (with $X_0 = I$). However, a better strategy is as follows.
Let s be the largest diagonal entry of A. Then $A = s(I - B)$ with $\rho(B) < 1$. (If A is a nonsingular M-matrix, we also have $B \geq 0$.)
We compute $A^{1/p}$ through $A^{1/p} = s^{1/p}(I - B)^{1/p}$. To find $(I - B)^{1/p}$ we generate a sequence X_k by Newton's method or Halley's method, with $X_0 = I$ in each case.

Structure preserving for M-matrices

We would like to know whether the approximations X_k are nonsingular M-matrices.

Proposition

For Newton's method or Halley's method, the matrix X_k is a nonsingular M-matrix for all nonsingular M-matrices (of all sizes) $A = I - B$ with $B \geq 0$ if and only if $c_{k,i} \leq 0$ for all $i \geq 1$.

Thus, when A is a nonsingular M-matrix with $a_{ii} \leq 1$, X_1 and X_2 from Newton's method are always nonsingular M-matrices (diagonal entries also ≤ 1); X_1 from Halley's method is always a nonsingular M-matrix.

Square roots of M-matrices

When $p = 2$, it is shown in [Meini, 2004] that if A is a nonsingular M-matrix with all diagonal entries ≤ 1, then the matrices X_k generated by Newton's method are all nonsingular M-matrices.

Conjecture

The matrices X_k generated by Newton's method $(with\ X_0 = I)$ are nonsingular M-matrices for every nonsingular M-matrix A.

This conjecture is of purely theoretical interest, since it is more appropriate to compute $A^{1/2}$ though $A^{1/2} = s^{1/2}(I - B)^{1/2}$ by applying Newton's method (with $X_0 = I$) to compute $(I - B)^{1/2}$, or equivalently, to compute $A^{1/2}$ directly by applying Newton's method with $X_0 = s^{1/2}I$.

Structure preserving for H-matrices

Proposition

Let A be a real nonsingular H-matrix with $0 < a_{ii} \leq 1$ for all i. If $c_{k,i} \leq 0$ for all $i \geq 1$ for Newton's method or Halley's method, the matrix X_k is a real nonsingular H-matrix with $0 < (X_k)_{ii} \leq 1$.

Thus, for Newton's method or Halley's method structure preserving for nonsingular M-matrices implies structure preserving for real nonsingualr H-matrices with positive diagonal entries.