Computing Matrix Functions by Matrix Iterations

Bruno **Iannazzo**, University of **Insubria**, Italy

Manchester, May, 15-16, 2008

Matrix iterations and matrix functions

A function $f(A)$ can be solution of a matrix equation $F(X; A) = 0$; a natural way to compute it is to use iterative root-finding algorithms.

Numerical issues of an iteration:

- Competitivity with respect to the existing methods;
- Quick convergence (at least quadratic and in few steps!);
- Adequate accuracy.

Which functions can be computed by a given matrix iterations?

Example: square root

Newton's method is the most known root-finding algorithm. For the equation $x^2 - a = 0$ yields the iteration

$$x_{k+1} = \frac{1}{2}(x_k + a x_k^{-1}). \tag{1}$$

Newton's method for the matrix equation $X^2 - A = 0$ is

$$\begin{cases} H_k X_k + X_k H_k = -X_k^2 + A, \\ X_{k+1} = X_k + H_k. \end{cases}$$

Generalizing naively formula (1) yields

$$X_{k+1} = \frac{1}{2}(X_k + A X_k^{-1}),$$

which we call **simplified Newton's method**.

Example: square root

Two different methods which coincide if

- the former is well defined;
- X_0 commutes with A.

The simplified Newton method has advantages wrt the Newton method:

- Much less expensive;
- Easily proved convergence to the principal square root (if there exists) for $X_0 = A$;
- Wider applicability.

Problem: instability in certain cases.

Iterations which are functions of A

Consider the iterations $X_{k+1} = F(X_k)$ such that $X_k = s_k(A)$, for s_k function of A.

The iterates can be put together in a block upper triangular form by means of the same similarity.
If

$$MAM^{-1} = J_1 \oplus \cdots \oplus J_r,$$

then

$$Ms_k(A)M^{-1} = s_k(J_1) \oplus \cdots \oplus s_k(J_r).$$

The study of the convergence is reduced to the case in which A is a Jordan block.

Iterations which are functions of A

If A is diagonalizable then scalar convergence of $s_k(\lambda)$ for λ eigenvalue of A implies matrix convergence.

In general, "convergence on eigenvalues" does not implies convergence on Jordan blocks: if $x_{k+1} = x_k^2$,

$$X_0 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \qquad X_k = \begin{bmatrix} 1 & 2^k \\ 0 & 1 \end{bmatrix}.$$

Iterations which are functions of A

If the sequence s_k converges to a function f uniformly on a neighborhood K of the spectrum of A, then the matrix iteration converges, moreover

$$\|s_k(A) - f(A)\|_1 \leq c\|s_k - f\|_K,$$

where c depends on A.

If s_k is generated by a fixed point iteration usually **pointwise convergence to attractive fixed points implies uniform convergence in a neighborhood.**

Pure rational matrix iterations

A **pure rational matrix iteration** has the form

$$Z_{k+1} = \varphi(Z_k), \quad Z_0 \in \mathbb{C}^{n \times n},$$

where φ is a rational function (of degree at least 2).

Classification of a fixed point z^* for the scalar iteration:

- **attractive** if $|\varphi'(z^*)| < 1$;
- **rationally indifferent** if $\varphi'(z^*) = \exp(2i\pi\theta)$, $\theta \in \mathbb{Q}$;
- **irrationally indifferent** if $\varphi'(z^*) = \exp(2i\pi\theta)$, $\theta \notin \mathbb{Q}$;
- **repulsive** if $|\varphi'(z^*)| > 1$;

Theorem (I.)

Let $Z_{k+1} = \varphi(Z_k)$ be a pure rational matrix iteration. If for each eigenvalue λ of Z_0 the scalar sequence $z_{k+1} = \varphi(z_k)$, $z_0 = \lambda$, converges to a fixed point λ_*, attractive or rationally indifferent with infinite orbit, then there exists a **locally constant function** $f(z)$ such that for each initial value Z in a neighborhood of Z_0 **the matrix iteration converges** to $f(Z)$. Moreover, $f(Z)$ **is diagonalizable.**

The theorem can be extended:

- if λ_* is repulsive or $\lambda_* = \exp(2i\pi\theta)$ for almost each $\theta \in [0, 1]$ (λ_* not being a Cremer point or a root of unity) convergence can occurs only for diagonalizable matrices and in a finite number of steps;

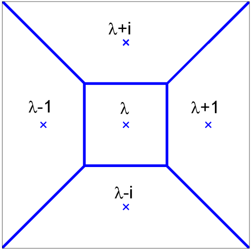
Matrix convergence

A pure rational matrix iteration converges to a **locally constant matrix function**.

Consequence: only locally constant functions can be computed by pure iterations.

Example

Examples of locally constant functions: **the Voronoi maps** relative to a set of finite points.



In this class are the **matrix sign function** and the **matrix *p*-sector function.**

Rational matrix iterations involving A

A generalization of pure rational matrix iteration is obtained allowing A in the formula

$$Z_{k+1} = \varphi(Z_k, A), \quad Z_0 = g(A),$$

where $\varphi(z, t)$ is a rational function of two variable.

Theorem (Higham, 08)

If for each eigenvalue of A the scalar iteration converges to an attractive fixed point then the matrix iteration converges.

Rational matrix iterations involving A

The result can be improved:

(I.)

- the limit in a neighborhood of the spectrum is uniquely determined by the iteration (if convergence is at least linear);
- the limit of a rational matrix iterations (with convergence at least linear) is an analytic function;
- there can be given a bound for the convergence of nondiagonal elements.

Remark

- the limit of a rational iteration is an algebraic function.

Consequences

- **Nonanalytic functions** cannot be limit of rational iterations;
- **Transcendental functions** cannot be limit of rational iterations;

Iterations for the functions $\exp(A)$, $\log(A)$, etc... must be transcendental.

- Most of the theoretical results can be extended from rational functions to holomorphic functions;
- **Evident drawback**: computing the iteration step is usual as expensive as the computation of the function.

Summary

Which functions can be computed by a given matrix iterations?

class of iterations	functions	examples
pure rational iterations	locally constant functions	$\mathrm{sign}(A),\ \mathrm{sector}_p(A)$
rational iterations containing A	algebraic functions which are analytic	A^α
analytic iterations	analytic functions	$\exp(A),\ \log(A)$

Stability

What about the convergence in presence of small perturbations?

In general: local convergence $\Longrightarrow$ stability

But: scalar local convergence $\not\Longrightarrow$ matrix local convergence!

There can be **repulsive directions** in a neighborhood of the fixed point, moreover **fixed points can be non isolated**; it is pointless requiring local convergence

A model for the stability

It can be enough much less than local convergence.

Definition (Cheng, Higham, Kenney, Laub, 2001)

An iteration $X_{k+1} = \psi(X_k)$ is **stable in a neighborhood of a fixed point** X (N-stable) if the differential $d\psi$ at X is power bounded.

N-stable implies that a perturbation of some iterate cannot blow up.

An example: inverting a matrix by an iteration

Given

$$\varphi_1(X) = 2X - XAX, \qquad \varphi_2(X) = 2X - X^2 A,$$

the two iterations $X_{k+1} = \varphi_1(X_k)$ and $X_{k+1} = \varphi_2(X_k)$ are **mathematically equivalent** for $X_0 = A$, but

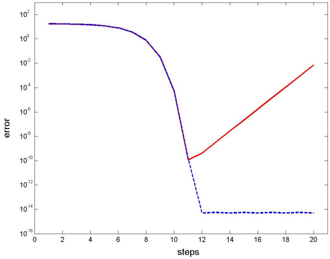
- $d\varphi_{1A^{-1}} = 0 \longrightarrow$ stable in a neigh'd of A^{-1}
- $d\varphi_{2A^{-1}} = I - A^T \otimes A^{-1} \longrightarrow$ stable if $\left|1 - \frac{\lambda_i}{\lambda_j}\right| < 1$, $\lambda_i \in \sigma(A)$

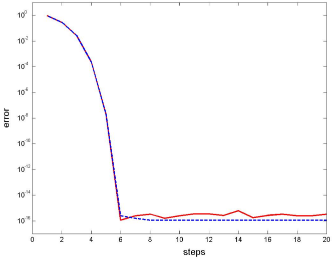
However $d\varphi_{2A^{-1}}|_{\text{matrices commuting with } A} = 0$

An example: inverting a matrix by an iteration

$$A = M^{-1} \begin{bmatrix} 1 & 0 \\ 0 & 2/3 \end{bmatrix} M \qquad A = M^{-1} \begin{bmatrix} 1 & 0 \\ 0 & 1/10 \end{bmatrix} M$$

$$\rho(d\varphi_{2A^{-1}}) = 3 \qquad \rho(d\varphi_{2A^{-1}}) = 1/10$$





General results on N-stability

Some general results are known for pure rational matrix iterations, i.e., iteration of the form $X_{k+1} = \varphi(X_k)$.

Theorem (Higham, 08)

Let f be an idempotent function that is Frèchet differentiable at $X = f(X)$. If X_k obtained by a pure rational matrix iteration which converges superlinearly to $f(X_0)$ for all X_0 close to X then $df_X = d\varphi_X$.

Consequence: an iteration converging superlinearly to a function such that df_X is idempotent is N-stable at X.

General results on N-stability

The previous result can be generalized.

Theorem (I.)

Let S be a fixed point of a pure rational matrix iteration, if the eigenvalues of S are attractive fixed points, then S is diagonalizable and the iteration is N-stable at S.

If S is a Jordan block relative to a rationally indifferent fixed point, then any sequence X_k converging to S must be finite (the same holds for repulsive and rationally indifferent points which are not Cremer points).

Pure rational iterations are stable!

Possible approaches to remove instability

- Projection of each iterate into a stable subspace;
- Preprocessing the matrix to get stability;
- Modifying the iterations trying to reach a pure iteration.

The first two approaches seem too expensive

For the third approach there is no general recipe anyway...

Example: pth root of a matrix

For the equation $X^p = A$, the family

$$X_{k+1} = \frac{(p-1)X_k + AX^{1-p}}{p}, \quad X_0 = I,$$

is not stable, but the equivalent

$$\begin{cases} X_{k+1} = X_k \left(\frac{(p-1)I + M_k}{p}\right), & X_0 = I, \\ M_{k+1} = \left(\frac{(p-1)I + M_k}{p}\right)^{-p} M_k, & M_0 = A. \end{cases}$$

is stable [I., SIMAX, 2006]

Systems of rational iterations

A relatively unexplored topic is the study of **systems of (pure) rational iterations**.

Example. The Cyclic Reduction is a system of pure rational iterations,

$$\begin{cases} A_{k+1} = -A_k B_k^{-1} A_k, \\ B_{k+1} = B_k - A_k B_k^{-1} C_k - C_k B_k^{-1} A_k, \\ C_{k+1} = -C_k B_k^{-1} C_k, \end{cases}$$

for $k = 0, 1, \ldots$ and $A_0 = A$, $B_0 = B$, $C_0 = C$ initial values. The iteration is N-stable.

- Is it possible to see the CR as a single block iteration?
- Can the limit described as a function of A, B and C?
- Is there a general convergence result?

Palindromic cyclic reduction: all purpose iteration

If $A = C$ can be rewritten as the three-terms (pure) rational iteration

$$X_{k+1} = \frac{1}{2}(X_k + 2X_{k-1} - X_{k-1}X_k^{-1}X_{k-1}),$$

this recurrence is related to many "matrix functions":

- $X_0 = A$, $X_1 = \frac{1}{2}(A + I)$, $X_k \to A^{1/2}$;
- $X_0 = A$, $X_1 = \frac{1}{2}(A + A^{-1})$, $X_k \to \text{sign}(A)$;
- $X_0 = A$, $X_1 = \frac{1}{2}(A + A^{-*})$, $X_k \to \text{polar}(A)$;
- $X_0 = A$, $X_1 = \frac{1}{2}(A + B)$, A, B, def. pos. $X_k \to A\#B$.

- C. Kenney, A.J. Laub, *Rational iterative methods for the matrix sign function*, SIMAX, 1991.
- N. Higham, *Functions of Matrices: Theory and Computation*, 2008.
- B. Iannazzo, *On the Newton method for the matrix pth root*, SIMAX, 2006.
- B. Iannazzo, *Notes on rational matrix iterations*, preprint, 2007.
- B. Iannazzo, *Numerical Solution of Certain Nonlinear Matrix Equations*, PhD Thesis, 2007.