Nonlinear matrix equations and canonical factorizations

Beatrice Meini
joint work with Dario A. Bini

Dipartimento di Matematica, Università di Pisa, Italy

Structured numerical linear algebra problems
Cortona, Sept. 19–24, 2004

Università di Pisa

Outline

1. Some examples
   - Quadratic matrix equations
   - Matrix pth root: $X^p = A$
   - Power series matrix equations
2. Canonical factorization
3. Canonical factorization and matrix equations
   - Some questions
   - Existence of solutions
   - Shift technique

UNIVERSITÀ DI PISA

Outline

UNIVERSITÀ DI PISA

Outline

UNIVERSITÀ DI PISA

Outline

UNIVERSITÀ DI PISA

Quadratic matrix equations

Given the $m \times m$ matrix polynomial $A(z) = A_{-1} + zA_0 + z^2 A_1$ such that $\det A(z)$ has zeros

$$|\xi_1| \le \cdots \le |\xi_m| < |\xi_{m+1}| \le \cdots \le |\xi_{2m}|$$

compute the solution G of

$$A_{-1} + A_0 X + A_1 X^2 = 0$$

such that $\lambda(G) = \{\xi_1, \ldots, \xi_m\}$.
Such G is called the minimal solvent (Gohberg, Lancaster, Rodman '82)

Applications Quadratic eigenvalue problems (damped vibration problems), polynomial factorization, Markov chains, etc.

Functional interpretation (Gohberg, Lancaster, Rodman '82)

1. The matrix function $S(z) = z^{-1}A_{-1} + A_0 + zA_1$ can be factorized as

$$S(z) = (A_0 + zA_1 G)(I - z^{-1}G)$$

where

- $\det(A_0 + zA_1 G) \neq 0$ for $|z| \leq 1$;
- $\det(I - z^{-1}G) \neq 0$ for $|z| \geq 1$.

2. Conversely: if

$$S(z) = (U_0 + zU_1)(L_0 + z^{-1}L_{-1}) = U(z)L(z)$$

where $\det U(z) \neq 0$ for $|z| \leq 1$ and $\det L(z) \neq 0$ for $|z| \geq 1$, then $G = -L_0^{-1}L_{-1}$ is the minimal right solvent.

UNIVERSITÀ DI PISA

Functional interpretation (Gohberg, Lancaster, Rodman '82)

1. The matrix function $S(z) = z^{-1}A_{-1} + A_0 + zA_1$ can be factorized as

$$S(z) = (A_0 + zA_1 G)(I - z^{-1}G)$$

   where
   - $\det(A_0 + zA_1 G) \neq 0$ for $|z| \leq 1$;
   - $\det(I - z^{-1}G) \neq 0$ for $|z| \geq 1$.

2. Conversely: if

$$S(z) = (U_0 + zU_1)(L_0 + z^{-1}L_{-1}) = U(z)L(z)$$

   where $\det U(z) \neq 0$ for $|z| \leq 1$ and $\det L(z) \neq 0$ for $|z| \geq 1$, then $G = -L_0^{-1}L_{-1}$ is the minimal right solvent.

Università di Pisa

Matrix pth root

Assumptions $A \in \mathbb{C}^{m\times m}$ with no eigenvalues on the closed negative real axis.

Definition The principal matrix pth root of A, $A^{1/p}$, is the unique matrix X such that:

1. $X^p = A$.
2. The eigenvalues of X lie in the segment $\{\, z : -\pi/p < \arg(z) < \pi/p \,\}$.

Applications Computation of the matrix logarithm, computation of the matrix sector function (control theory).

UNIVERSITÀ DI PISA

Functional interpretation

Theorem (Bini, Higham, Meini 04)

Assume $p = 2q$, where q is odd. Let

$$S(z) = z^{-q} \sum_{j=0}^{p} z^j \binom{p}{j} \left(A + (-1)^{j+1} I\right).$$

If $U(z) = U_0 + zU_1 + \cdots + z^q U_q$ is such that $\det U(z) \neq 0$ for $|z| \leq 1$, and $S(z) = U(z)U(z^{-1})$ then

$$A^{1/p} = -\sigma^{-1}(qI + 2U'(-1)U(-1)^{-1})$$

where $\sigma = 1 + 2\sum_{j=1}^{\lfloor q/2 \rfloor} \cos(2\pi j/p)$.

Power series matrix equations

An application M/G/1-type Markov chains, introduced by M. F. Neuts in the 80's, which model a large variety of queueing problems.

Problem Given nonnegative matrices $A_i \in \mathbb{R}^{m\times m}$, $i \geq -1$, such that $\sum_{i=-1}^{+\infty} A_i$ is stochastic, compute the minimal component-wise solution G, among the nonnegative solutions, of

$$X = A_{-1} + A_0 X + A_1 X^2 + \cdots$$

UNIVERSITÀ DI PISA

Some properties of G

Let $\phi(z) = zI - \sum_{i=-1}^{+\infty} z^{i+1} A_i$.
If the M/G/1-type Markov chain is positive recurrent, then:

- G is row stochastic.
- $\det \phi(z)$ has exactly m zeros in the closed unit disk.
- The eigenvalues of G are the zeros of $\det \phi(z)$ in the closed unit disk.

Therefore G is the spectral minimal solution, i.e., $\rho(G) \le \rho(X)$ for any other possible solution X.

UNIVERSITÀ DI PISA

The induced factorization

The function $S(z) = I - \sum_{i=-1}^{+\infty} z^i A_i$ can be factorized as

$$S(z) = \Big(I - \sum_{i=0}^{+\infty} z^i U_i\Big)(I - z^{-1} G), \quad |z| = 1,$$

where:

- $U(z) = I - \sum_{i=0}^{+\infty} z^i U_i$ is analytic for $|z| < 1$, convergent for $|z| \le 1$, and $\det U(z) \neq 0$ for $|z| \le 1$;
- $L(z) = I - z^{-1} G$ is analytic for $|z| > 1$, convergent for $|z| \ge 1$, and $\det L(z) \neq 0$ for $|z| > 1$, $\det L(1) = 0$.

UNIVERSITÀ DI PISA

The induced factorization

The function $S(z) = I - \sum_{i=-1}^{+\infty} z^i A_i$ can be factorized as

$$S(z) = \Big(I - \sum_{i=0}^{+\infty} z^i U_i\Big)(I - z^{-1}G), \quad |z| = 1,$$

where:

- $U(z) = I - \sum_{i=0}^{+\infty} z^i U_i$ is analytic for $|z| < 1$, convergent for $|z| \leq 1$, and $\det U(z) \neq 0$ for $|z| \leq 1$;
- $L(z) = I - z^{-1}G$ is analytic for $|z| > 1$, convergent for $|z| \geq 1$, and $\det L(z) \neq 0$ for $|z| > 1$, $\det L(1) = 0$.

UNIVERSITÀ DI PISA

The induced factorization

The function $S(z) = I - \sum_{i=-1}^{+\infty} z^i A_i$ can be factorized as

$$S(z) = \Big(I - \sum_{i=0}^{+\infty} z^i U_i\Big)(I - z^{-1} G), \quad |z| = 1,$$

where:

- $U(z) = I - \sum_{i=0}^{+\infty} z^i U_i$ is analytic for $|z| < 1$, convergent for $|z| \le 1$, and $\det U(z) \neq 0$ for $|z| \le 1$;
- $L(z) = I - z^{-1} G$ is analytic for $|z| > 1$, convergent for $|z| \ge 1$, and $\det L(z) \neq 0$ for $|z| > 1$, $\det L(1) = 0$.

Università di Pisa

Outline

Wiener algebra

Definition ($\mathcal{W}$)

The Wiener algebra $\mathcal{W}$ is the set of complex $m \times m$ matrix valued functions $A(z) = \sum_{i=-\infty}^{+\infty} z^i A_i$ such that $\sum_{i=-\infty}^{+\infty} |A_i|$ is finite.

Definition ($\mathcal{W}_+$ and $\mathcal{W}_-$)

The set $\mathcal{W}_+$ ($\mathcal{W}_-$) is the subalgebra of $\mathcal{W}$ made up by power series of the kind $\sum_{i=0}^{+\infty} z^i A_i$ ($\sum_{i=0}^{+\infty} z^{-i} A_i$).

Canonical factorization

Definition (Canonical factorization)

Let $A(z) = \sum_{i=-\infty}^{+\infty} z^i A_i \in \mathcal{W}$. A *canonical factorization* of $A(z)$ is a decomposition

$$A(z) = U(z)L(z), \qquad |z| = 1,$$

where $U(z) = \sum_{i=0}^{+\infty} z^i U_i \in \mathcal{W}_+$ and $L(z) = \sum_{i=0}^{+\infty} z^{-i} L_{-i} \in \mathcal{W}_-$ are invertible for $|z| \le 1$ and $1 \le |z| \le \infty$, respectively.

Definition (Weak canonical factorization)

The above decomposition is a *weak canonical factorization* if $U(z) = \sum_{i=0}^{+\infty} z^i U_i \in \mathcal{W}_+$ and $L(z) = \sum_{i=0}^{+\infty} z^{-i} L_{-i} \in \mathcal{W}_-$ are invertible for $|z| < 1$ and $1 < |z| \le \infty$, respectively.

Canonical factorization

Definition (Canonical factorization)

Let $A(z) = \sum_{i=-\infty}^{+\infty} z^i A_i \in \mathcal{W}$. A *canonical factorization* of $A(z)$ is a decomposition

$$A(z) = U(z)L(z), \qquad |z| = 1,$$

where $U(z) = \sum_{i=0}^{+\infty} z^i U_i \in \mathcal{W}_+$ and $L(z) = \sum_{i=0}^{+\infty} z^{-i} L_{-i} \in \mathcal{W}_-$ are invertible for $|z| \le 1$ and $1 \le |z| \le \infty$, respectively.

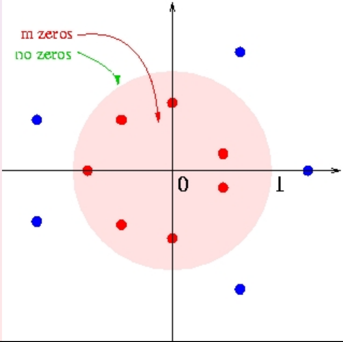
Definition (Weak canonical factorization)

The above decomposition is a *weak canonical factorization* if $U(z) = \sum_{i=0}^{+\infty} z^i U_i \in \mathcal{W}_+$ and $L(z) = \sum_{i=0}^{+\infty} z^{-i} L_{-i} \in \mathcal{W}_-$ are invertible for $|z| < 1$ and $1 < |z| \le \infty$, respectively.

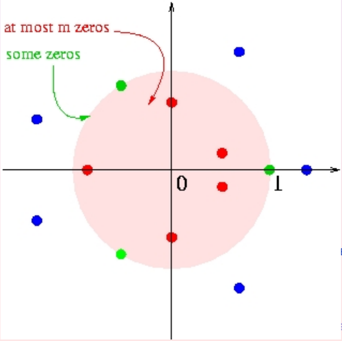
An example: $S(z)=\sum_{i=-1}^{+\infty} z^i A_i$

Location of the zeros of $\det(zS(z))$

Canonical factorization



Weak canonical factorization



Outline

1. Some examples
   - Quadratic matrix equations
   - Matrix pth root: $X^p = A$
   - Power series matrix equations
2. Canonical factorization
3. Canonical factorization and matrix equations
   - Some questions
   - Existence of solutions
   - Shift technique

Università di Pisa

Some questions

Let $S(z) = \sum_{i=-1}^{+\infty} z^i A_i \in \mathcal{W}$ and define $A(z) = zS(z)$. Consider

$$\sum_{i=-1}^{+\infty} A_i X^{i+1} = 0 \tag{1}$$

1. Existence of a canonical factorization $\Longrightarrow$ existence of a spectral minimal solution? Viceversa?
2. What can we say if the canonical factorization is weak?
3. Can we transform a weak canonical factorization into a canonical factorization?

Università di Pisa

Existence of solutions and canonical factorization

Theorem

If there exists a c.f.

$$S(z) = U(z)L(z), \quad L(z) = L_0 + z^{-1}L_{-1}, \quad |z| = 1,$$

then $G = -L_0^{-1}L_{-1}$ *is the unique solution of* (1) *such that* $\rho(G) < 1$, *and it is the spectral minimal solution.*
Conversely, if there exists a solution G *of* (1) *such that* $\rho(G) < 1$ *and if* $A(z)$ *has exactly* m *roots in the open unit disk,* $\det A(z) \neq 0$ *for* $|z| = 1$, *then* $S(z)$ *has a c.f.*

$$S(z) = (U_0 + zU_1 + \cdots)(I - z^{-1}G), \quad |z| = 1.$$

Existence of solutions and canonical factorization

Theorem

If there exists a c.f.

$$S(z) = U(z)L(z), \quad L(z) = L_0 + z^{-1}L_{-1}, \quad |z| = 1,$$

then $G = -L_0^{-1}L_{-1}$ *is the unique solution of* (1) *such that* $\rho(G) < 1$*, and it is the spectral minimal solution.*
Conversely, if there exists a solution G *of* (1) *such that* $\rho(G) < 1$ *and if* $A(z)$ *has exactly* m *roots in the open unit disk,* $\det A(z) \neq 0$ *for* $|z| = 1$*, then* $S(z)$ *has a c.f.*

$$S(z) = (U_0 + zU_1 + \cdots)(I - z^{-1}G), \quad |z| = 1.$$

Existence of solutions and weak factorization

Theorem

If there exists a weak c.f.

$$S(z) = U(z)L(z), \quad L(z) = L_0 + z^{-1}L_{-1}, \quad |z| = 1,$$

such that $G = -L_0^{-1}L_{-1}$ is power bounded, then G is a spectral minimal solution of (1) such that $\rho(G) \le 1$.
Conversely, if $S'(z) \in \mathcal{W}$, if there exists a power bounded solution G of (1) such that $\rho(G) = 1$, and if all the zeros of $\det A(z)$ in the open unit disk are eigenvalues of G then there exists a weak c.f. of $S(z)$.

In general, weak c.f. $\not\Longrightarrow$ unique spectral minimal solution.
Can we transform a weak c.f. into a c.f?

Università di Pisa

Existence of solutions and weak factorization

Theorem

If there exists a weak c.f.

$$S(z) = U(z)L(z), \quad L(z) = L_0 + z^{-1}L_{-1}, \quad |z| = 1,$$

such that $G = -L_0^{-1}L_{-1}$ is power bounded, then G is a spectral minimal solution of (1) such that $\rho(G) \leq 1$.
Conversely, if $S'(z) \in \mathcal{W}$, if there exists a power bounded solution G of (1) such that $\rho(G) = 1$, and if all the zeros of $\det A(z)$ in the open unit disk are eigenvalues of G then there exists a weak c.f. of $S(z)$.

In general, weak c.f. $\not\Longrightarrow$ unique spectral minimal solution.
Can we transform a weak c.f. into a c.f?

Existence of solutions and weak factorization

Theorem

If there exists a weak c.f.

$$S(z) = U(z)L(z), \quad L(z) = L_0 + z^{-1}L_{-1}, \quad |z| = 1,$$

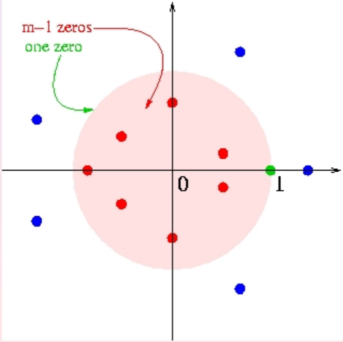
such that $G = -L_0^{-1}L_{-1}$ is power bounded, then G is a spectral minimal solution of (1) such that $\rho(G) \le 1$.
Conversely, if $S'(z) \in \mathcal{W}$, if there exists a power bounded solution G of (1) such that $\rho(G) = 1$, and if all the zeros of $\det A(z)$ in the open unit disk are eigenvalues of G then there exists a weak c.f. of $S(z)$.

In general, weak c.f. $\not\Longrightarrow$ unique spectral minimal solution.
Can we transform a weak c.f. into a c.f?

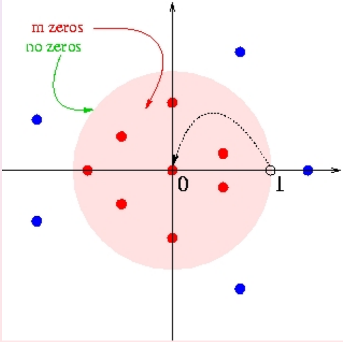
UNIVERSITÀ DI PISA

Shift technique: removing zeros of modulus 1

Before shifting



After shifting



Assumptions

- $S(z) = \sum_{i=-N}^{+\infty} z^i A_i \in \mathcal{W}$ and $S'(z) \in \mathcal{W}$, where $N \geq 1$.
- There is only one simple zero λ of $\det S(\lambda)$ on the unit circle.
- $\mathbf{v}$ is a vector such that $S(\lambda)\mathbf{v} = 0$, $\mathbf{v} \neq 0$.

In problems arising in Markov chains these assumptions are satisfied, moreover $\lambda = 1$ and $\mathbf{v} = (1, 1, \ldots, 1)^{\mathrm{T}}$.

Università di Pisa

Shift technique

Define

$$\widetilde{S}(z) = S(z)(I - z^{-1}\lambda Q)^{-1}, \quad Q = \mathbf{v}\mathbf{u}^{\mathrm{T}}$$

where $\mathbf{u}$ is any fixed vector such that $\mathbf{v}^{\mathrm{T}}\mathbf{u} = 1$. Let $\widetilde{A}(z) = z^N \widetilde{S}(z)$.
Then:

- $\widetilde{S}(z) = \sum_{i=-N}^{+\infty} z^i \widetilde{A}_i \in \mathcal{W}$.
- if $z \notin \{0, \lambda\}$, then $\det \widetilde{A}(z) = 0 \iff \det A(z) = 0$;
- $\det \widetilde{A}(0) = 0$ and $\widetilde{A}(0)\mathbf{v} = 0$;
- $\det \widetilde{A}(z) \neq 0$ if $|z| = 1$.

Università di Pisa

Shift technique

Define

$$\widetilde{S}(z) = S(z)(I - z^{-1}\lambda Q)^{-1}, \quad Q = \mathbf{v}\mathbf{u}^T$$

where $\mathbf{u}$ is any fixed vector such that $\mathbf{v}^T\mathbf{u} = 1$. Let $\widetilde{A}(z) = z^N \widetilde{S}(z)$. Then:

- $\widetilde{S}(z) = \sum_{i=-N}^{+\infty} z^i \widetilde{A}_i \in \mathcal{W}$.
- if $z \notin \{0, \lambda\}$, then $\det \widetilde{A}(z) = 0 \iff \det A(z) = 0$;
- $\det \widetilde{A}(0) = 0$ and $\widetilde{A}(0)\mathbf{v} = 0$;
- $\det \widetilde{A}(z) \neq 0$ if $|z| = 1$.

UNIVERSITÀ DI PISA

Shift technique

Define

$$\widetilde{S}(z) = S(z)(I - z^{-1}\lambda Q)^{-1}, \quad Q = \mathbf{v}\mathbf{u}^{\mathrm{T}}$$

where $\mathbf{u}$ is any fixed vector such that $\mathbf{v}^{\mathrm{T}}\mathbf{u} = 1$. Let $\widetilde{A}(z) = z^N \widetilde{S}(z)$. Then:

- $\widetilde{S}(z) = \sum_{i=-N}^{+\infty} z^i \widetilde{A}_i \in \mathcal{W}$.
- if $z \notin \{0, \lambda\}$, then $\det \widetilde{A}(z) = 0 \iff \det A(z) = 0$;
- $\det \widetilde{A}(0) = 0$ and $\widetilde{A}(0)\mathbf{v} = 0$;
- $\det \widetilde{A}(z) \neq 0$ if $|z| = 1$.

Shift technique

Define

$$\widetilde{S}(z) = S(z)(I - z^{-1}\lambda Q)^{-1}, \quad Q = \mathbf{v}\mathbf{u}^{\mathrm{T}}$$

where $\mathbf{u}$ is any fixed vector such that $\mathbf{v}^{\mathrm{T}}\mathbf{u} = 1$. Let $\widetilde{A}(z) = z^N \widetilde{S}(z)$. Then:

- $\widetilde{S}(z) = \sum_{i=-N}^{+\infty} z^i \widetilde{A}_i \in \mathcal{W}$.
- if $z \notin \{0, \lambda\}$, then $\det \widetilde{A}(z) = 0 \iff \det A(z) = 0$;
- $\det \widetilde{A}(0) = 0$ and $\widetilde{A}(0)\mathbf{v} = 0$;
- $\det \widetilde{A}(z) \neq 0$ if $|z| = 1$.

Università di Pisa

Shift technique

Define

$$\widetilde{S}(z) = S(z)(I - z^{-1}\lambda Q)^{-1}, \quad Q = \mathbf{v}\mathbf{u}^T$$

where $\mathbf{u}$ is any fixed vector such that $\mathbf{v}^T\mathbf{u} = 1$. Let $\widetilde{A}(z) = z^N\widetilde{S}(z)$. Then:

- $\widetilde{S}(z) = \sum_{i=-N}^{+\infty} z^i \widetilde{A}_i \in \mathcal{W}$.
- if $z \notin \{0, \lambda\}$, then $\det \widetilde{A}(z) = 0 \iff \det A(z) = 0$;
- $\det \widetilde{A}(0) = 0$ and $\widetilde{A}(0)\mathbf{v} = 0$;
- $\det \widetilde{A}(z) \neq 0$ if $|z| = 1$.

Weak $\longrightarrow$ canonical factorization

If $S(z)$ has a weak canonical factorization

$$S(z) = U(z)L(z)$$

where $\det U(z) \neq 0$ if $|z| = 1$, then $\widetilde{S}(z)$ has a canonical factorization

$$\widetilde{S}(z) = \widetilde{U}(z)\widetilde{L}(z),$$

where

$$\begin{aligned} \widetilde{U}(z) &= U(z), \\ \widetilde{L}(z) &= L(z)(I - z^{-1}\lambda Q)^{-1} \end{aligned}$$

Back to matrix equations

Let $S(z) = \sum_{i=-1}^{+\infty} z^i A_i$ and let G, with $\rho(G) = |\lambda|$, be the spectral minimal solution of $\sum_{i=-1}^{+\infty} A_i X^{i+1} = 0$.
Then the matrix equation

$$\sum_{i=-1}^{+\infty} \widetilde{A}_i X^{i+1} = 0$$

has one minimal spectral solution

$$\widetilde{G} = G - \lambda Q.$$

Moreover $\rho(\widetilde{G}) = \rho_2(G) < 1$.

Computational issues

- Shift technique $\Longrightarrow$ larger isolation ratio of the roots of $S(z)$ with respect to the unit circle.
- Experimentally, larger isolatio ratio $\Longrightarrow$ faster speed of convergence of functional iterations, cyclic reduction.
- Experimentally, larger isolatio ratio $\Longrightarrow$ better numerical stability

A theorethical proof of the latter experimental observations is still missing

UNIVERSITÀ DI PISA

New book

Numerical Methods for Structured Markov Chains

D.A. Bini (University of Pisa)
G. Latouche (Université Libre de Bruxelles)
B. Meini (University of Pisa)

Oxford University Press, 2005

UNIVERSITÀ DI PISA