Palindromic cyclic reduction and matrix functions

Bruno Iannazzo and Beatrice Meini

Advanced Workshop on Matrix Geometries and Applications
ICTP, Trieste, July 9–12, 2013

Quadratic palindromic matrix polynomials

In the recent literature there are several definitions of palindromic matrix polynomials ($*$-palindromic, T-palindromic,...).

In our setting a quadratic palindromic matrix polynomial is

$$\varphi(z) = Pz^2 + Qz + P$$

where P and Q are $n \times n$ matrices.

Throughout we assume that Q is nonsingular.

Spectral properties

Observe that $\varphi(z^{-1}) = Pz^{-2} + Qz^{-1} + P = z^{-2}\varphi(z)$.

Therefore, if $\lambda \in \mathbb{C}$, $\lambda \neq 0$ and $x \in \mathbb{C}^n$, $x \neq 0$, are such that $\varphi(\lambda)x = 0$, then $\varphi(\lambda^{-1})x = 0$.

Spectral properties

Observe that $\varphi(z^{-1}) = Pz^{-2} + Qz^{-1} + P = z^{-2}\varphi(z)$.

Therefore, if $\lambda \in \mathbb{C}$, $\lambda \neq 0$ and $x \in \mathbb{C}^n$, $x \neq 0$, are such that $\varphi(\lambda)x = 0$, then $\varphi(\lambda^{-1})x = 0$.

Hence, with the convention that $1/0 = \infty$ and $1/\infty = 0$:

- the roots of $\varphi(z)$ come to pairs $(\lambda, 1/\lambda)$;
- if $\det(\varphi(z)) \neq 0$ for $|z| = 1$, then $\varphi(z)$ has exactly n roots in the open unit disk, and n roots outside the closed unit disk, i.e., the roots of $\varphi(z)$ have a splitting with respect to the unit circle.

Under which conditions do the roots of $\varphi(z)$ have a splitting with respect to the unit circle?

Spectral properties – cont.

- The roots of $\varphi(z)$ have a splitting with respect to the unit circle if and only if the real eigenvalues of $M = Q^{-1}P$ belong to $(-1/2, 1/2)$.

Spectral properties – cont.

- The roots of $\varphi(z)$ have a splitting with respect to the unit circle if and only if the real eigenvalues of $M = Q^{-1}P$ belong to $(-1/2, 1/2)$.
- To each real eigenvalue of M of modulus $1/2$, there corresponds a root of multiplicity 2 of $\varphi(z)$, equal to ± 1.
- To each real eigenvalue of M, lying outside the interval to $[-1/2, 1/2]$, there correspond two distinct roots on the unit circle of $\varphi(z)$.

Laurent matrix polynomials

Let $\mathcal{L}(z) = Pz^{-1} + Q + Pz$ and assume that the real eigenvalues of M belong to the interval $(-1/2, 1/2)$.
Then:

- $\mathcal{L}(z)$ is invertible in the annulus $\mathcal{A}_R = \{z \in \mathbb{C} : \; R < |z| < 1/R\}$, where $R = \rho(X_*) < 1$, and $X_* = -2M(I + (I - 4M^2)^{1/2})^{-1}$ solves $P + QX + PX^2 = 0$
- $\mathcal{L}(z)$ can be factorized as

$$\mathcal{L}(z) = (I - QX_* Q^{-1} z)(Q + PX_*)(I - X_* z^{-1}), \quad z \in \mathcal{A}_R$$

- if $\mathcal{H}(z) = \mathcal{L}(z)^{-1} = H_0 + \sum_{i=1}^{+\infty} H_i(z^i + z^{-i})$, we have

$$H_0 = (I - 4M^2)^{-1/2} Q^{-1}$$

Matrix square root

Let A be a square matrix without real nonpositive eigenvalues. The principal matrix square root $A^{1/2}$ is the unique solution of

$$X^2 - A = 0$$

having eigenvalues with positive real part.

The definition is extended to the case where A does not have negative real eigenvalues and the null eigenvalue is semisimple.

Matrix sign

Let A be a square matrix without pure imaginary eigenvalues, and let

$$J = V^{-1}AV = \begin{bmatrix} J_- & 0 \\ 0 & J_+ \end{bmatrix}$$

its Jordan normal form, where $\sigma(J_-) \in \mathbb{C}_-$, $\sigma(J_+) \in \mathbb{C}_+$.
The matrix sign of A is defined as

$$\mathrm{sign}(A) = V \begin{bmatrix} -I_p & 0 \\ 0 & I_q \end{bmatrix} V^{-1}$$

where $p = \mathrm{size} J_-$ and $q = \mathrm{size} J_+$.

Matrix polar factor

Let A be a square nonsingular matrix.
The unitary polar factor, denoted by $\mathrm{polar}(A)$, is the unique matrix U such that

$$A = UH, \quad U^*U = I,$$

where H is Hermitian positive definite.

Geometric matrix mean

Let A and B be hermitian positive definite.
The matrix geometric mean between A and B is

$$A\#B = A^{1/2}(A^{-1/2}BA^{-1/2})^{1/2}A^{1/2}$$

Matrix functions and palindromic Laurent polynomials

Assume that

$$\mathcal{L}(z) = \frac{1}{4}(S - T)z^{-1} + \frac{1}{2}(S + T) + \frac{1}{4}(S - T)z$$

Matrix functions and palindromic Laurent polynomials

Assume that

$$\mathcal{L}(z) = \frac{1}{4}(S - T)z^{-1} + \frac{1}{2}(S + T) + \frac{1}{4}(S - T)z$$

The following properties hold:

1. If A does not have nonpositive real eigenvalues, $S = I$, $T = A^{-1}$, then $H_0 = A^{1/2}$.

Matrix functions and palindromic Laurent polynomials

Assume that

$$\mathcal{L}(z) = \frac{1}{4}(S - T)z^{-1} + \frac{1}{2}(S + T) + \frac{1}{4}(S - T)z$$

The following properties hold:

1. If A does not have nonpositive real eigenvalues, $S = I$, $T = A^{-1}$, then $H_0 = A^{1/2}$.
2. If A does not have imaginary eigenvalues, $S = A^{-1}$, $T = A$, then $H_0 = \mathrm{sign}(A)$.

Matrix functions and palindromic Laurent polynomials

Assume that

$$\mathcal{L}(z) = \frac{1}{4}(S - T)z^{-1} + \frac{1}{2}(S + T) + \frac{1}{4}(S - T)z$$

The following properties hold:

1. If A does not have nonpositive real eigenvalues, $S = I$, $T = A^{-1}$, then $H_0 = A^{1/2}$.
2. If A does not have imaginary eigenvalues, $S = A^{-1}$, $T = A$, then $H_0 = \mathrm{sign}(A)$.
3. If $\det A \neq 0$, $S = A^{-1}$, $T = A^*$, then $H_0 = \mathrm{polar}(A)$.

Matrix functions and palindromic Laurent polynomials

Assume that

$$\mathcal{L}(z) = \frac{1}{4}(S - T)z^{-1} + \frac{1}{2}(S + T) + \frac{1}{4}(S - T)z$$

The following properties hold:

1. If A does not have nonpositive real eigenvalues, $S = I$, $T = A^{-1}$, then $H_0 = A^{1/2}$.
2. If A does not have imaginary eigenvalues, $S = A^{-1}$, $T = A$, then $H_0 = \mathrm{sign}(A)$.
3. If $\det A \neq 0$, $S = A^{-1}$, $T = A^*$, then $H_0 = \mathrm{polar}(A)$.
4. If $A > 0$, $B > 0$, $S = A^{-1}$ and $T = B^{-1}$, then $H_0 = A \# B$.

Cyclic reduction (CR)

CR is a versatile algorithm invented by G. Golub [Buzbee, Golub, Nielson 1970] for the f.d. Poisson equation.
Given $A_0, A_1, A_2 \in \mathbb{R}^{N\times N}$ such that the roots of $\varphi(z) = \det(A_0 + A_1 z + A_2 z^2)$ are

$$|\xi_1| \le \cdots \le |\xi_N| < 1 < |\xi_{N+1}| \le \cdots \le |\xi_{2N}|,$$

including zeros at ∞ if $\deg\varphi(\lambda) < 2N$, CR generates the matrix sequences

$$\begin{aligned} A_0^{(k+1)} &= -A_0^{(k)} S^{(k)} A_0^{(k)}, \quad S^{(k)} = (A_1^{(k)})^{-1} \\ A_2^{(k+1)} &= -A_2^{(k)} S^{(k)} A_2^{(k)}, \\ A_1^{(k+1)} &= A_1^{(k)} - A_0^{(k)} S^{(k)} A_2^{(k)} - A_2^{(k)} S^{(k)} A_0^{(k)}, \quad k \ge 0 \end{aligned}$$

starting from $A_i^{(0)} = A_i$, $i = 1, 2, 3$.

Functional formulation of CR

Let

$$\varphi_k(z) = A_0^{(k)} + A_1^{(k)} z + A_2^{(k)} z^2,$$

$$\psi_k(z) = z\varphi_k(z)^{-1}, \quad k = 0, 1, \ldots$$

Then, for any $k \geq 0$,

$$\varphi_{k+1}(z^2) = \varphi_k(z) \left(A_1^{(k)}\right)^{-1} \varphi_k(-z) \;\; \text{(Graeffe iteration)}$$

$$\psi_{k+1}(z^2) = \tfrac{1}{2}(\psi_k(z) + \psi_k(-z))$$

Properties of CR

Applicability: under mild conditions the matrices $A_1^{(k)}$ are invertible

Convergence property: the convergence is quadratic, more specifically:

$$||A_1^{(k)} - A_1^*|| = O(|\xi_N/\xi_{N+1}|^{2^k}),$$

$$||A_0^{(k)}|| = O(|\xi_N|^{2^k}), \quad ||A_2^{(k)}|| = O(|1/\xi_{N+1}|^{2^k})$$

Critical case: Under mild conditions, if $|\xi_N| = |\xi_{N+1}| = 1$ convergence turns to linear with rate $1/2$.

Convergence properties

If $|\xi_N| < 1 < |\xi_{N+1}|$ then:

- $\psi(z) = A_0 z^{-1} + A_1 + A_2 z$ is invertible in an annulus containing the unit circle and, setting $H(z) = \psi(z)^{-1} = \sum_{i=-\infty}^{+\infty} H_i z^i$, we have

$$\lim_k A_1^{(k)} = H_0^{-1}$$

- if the matrix equation $A_0 + A_1 X + A_2 X^2 = 0$ has a solution G with $\rho(G) < 1$, the sequence

$$\widehat{A}^{(k+1)} = \widehat{A}^{(k)} - A_0^{(k)} S^{(k)} A_2^{(k)}, \quad k \geq 0,$$

with $\widehat{A}^{(0)} = A_1$ is such that

$$\lim_k (-\widehat{A}^{(k)})^{-1} A_0 = G$$

Palindromic cyclic reduction (PCR)

Cyclic reduction applied to $\varphi(z) = Pz^2 + Qz + P$ generates the sequences

$$\begin{cases} P_0 = P, \quad Q_0 = Q, \\ P_{k+1} = -P_k Q_k^{-1} P_k, \\ Q_{k+1} = Q_k - 2P_k Q_k^{-1} P_k, \qquad k = 0, 1, 2, \ldots \end{cases}$$

Palindromic cyclic reduction (PCR)

Cyclic reduction applied to $\varphi(z) = Pz^2 + Qz + P$ generates the sequences

$$\begin{cases} P_0 = P, \quad Q_0 = Q, \\ P_{k+1} = -P_k Q_k^{-1} P_k, \\ Q_{k+1} = Q_k - 2P_k Q_k^{-1} P_k, \qquad k = 0, 1, 2, \ldots \end{cases}$$

By setting $X_k = Q_k$ and $H_k = 2P_{k+1}$, PCR can be rewritten as

$$\begin{cases} X_0 = Q, \quad H_0 = -2PQ^{-1}P \\ X_{k+1} = X_k + H_k, \\ H_{k+1} = -\frac{1}{2}(H_k X_{k+1}^{-1} H_k), \qquad k = 0, 1, 2, \ldots. \end{cases} \tag{1}$$

Remark: The recursive expression for X_k looks like Newton's method, where H_k is the Newton increment.

Newton's method

If A admits the principal square root, Newton's method applied to compute $A^{1/2}$ is:

$$Y_{k+1} = \frac{1}{2}(Y_k + AY_k^{-1}), \quad k = 0, 1, \ldots, \quad Y_0 = I$$

Newton's method

If A admits the principal square root, Newton's method applied to compute $A^{1/2}$ is:

$$Y_{k+1} = \frac{1}{2}(Y_k + AY_k^{-1}), \quad k = 0, 1, \ldots, \quad Y_0 = I$$

Connection with PCR: Set $M = Q^{-1}P$ and $A = I - 4M^2$.
Then

$$X_k = QY_k, \quad H_k = Q(Y_{k+1} - Y_k), \quad k = 0, 1, \ldots.$$

Therefore, up to the multiplication by Q, PCR and Newton's method generate the same matrix sequence

A three-term recurrence

Recurrences (1) can be expressed in terms of the matrices X_k, $k \geq 0$, only. Indeed, we find that

$$X_{k+1} = X_k + H_k = X_k - \frac{1}{2} H_{k-1} X_k^{-1} H_{k-1}$$

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

The resulting iteration is the three-term recurrence

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

with $X_0 = Q$, $X_1 = Q - 2PQ^{-1}P$.

A formulation with commuting matrices

By setting $Y_k = X_{k-1}X_k^{-1}$, we find that

$$Y_{k+1} = \frac{1}{2}\left(I + 2Y_k^{-1} - (Y_k^{-1})^2\right), \quad k = 0, 1, 2, \ldots,$$

starting with $Y_0 = I - 2(PQ^{-1})^2$.
If we introduce also the matrix sequence

$$Z_{k+1} = Y_k Z_k, \quad k = 0, 1, 2, \ldots,$$

with $Z_0 = Q$, we find that $Z_k = X_k$ for any $k \geq 0$.

Therefore the sequences $\{Y_k\}_k$ and $\{Z_k\}_k$ represent another formulation of PCR.

The matrices Y_k, $k \geq 0$, commute each other and, in the case of convergence of PCR, converge to the identity matrix.

PCR and Gauss-Chebyshev quadrature

From the Cauchy integral formula one has

$$H_0 = \frac{1}{\pi}\int_{-1}^{1} \frac{(Q+2Pt)^{-1}}{\sqrt{1-t^2}}dt.$$

The Gauss-Chebyshev quadrature with k nodes is

$$C_k = \frac{1}{k}\sum_{j=0}^{k-1}\left(Q + 2P\cos\frac{(2j+1)\pi}{2k}\right)^{-1}.$$

and

$$C_{2^k} = Q_k^{-1}.$$

The k-th step of PCR is equivalent to G-C quadrature with 2^k knots

Trapezoidal rule

A change of variable leads to

$$H_0 = \frac{1}{\pi}\int_0^\pi (Q + 2P\cos\theta)^{-1}\, d\theta.$$

The trapezoidal rule with $k+1$ knots yields

$$T_k = \frac{1}{k}\left[\frac{(Q+2P)^{-1}}{2} + \sum_{j=1}^{k-1}\left(Q + 2P\cos\frac{\pi j}{k}\right)^{-1} + \frac{(Q-2P)^{-1}}{2}\right]$$

PCR and averaging technique

One has

$$T_{2^{k+1}} = \frac{1}{2}(T_{2^k} + C_{2^k}), \quad C_{2^{k+1}} = 2(C_{2^k}^{-1} + T_{2^k}^{-1})^{-1}$$

therefore the resulting iteration is

$$\begin{cases} A_1 = \frac{1}{2}\left((Q+2P)^{-1} + (Q-2P)^{-1}\right), \quad B_1 = Q^{-1}, \\ A_{k+1} = \frac{1}{2}(A_k + B_k), \\ B_{k+1} = 2A_k(A_k + B_k)^{-1}B_k, \quad k = 1, 2, \ldots, \end{cases}$$

where $B_k = Q_k^{-1}$ and $\lim_k A_k = \lim_k B_k = \lim_k Q_k^{-1} = H_0$.

Scaling technique

$$\begin{cases} X_0 = Q, \quad H_0 = -2PQ^{-1}P, \\ \gamma_k = \left| \dfrac{\det(X_k)^2}{\det(I - 4M^2)\det(Q)^2} \right|^{-1/(2n)}, \\ \widehat{X}_k = \gamma_k X_k, \qquad \widehat{H}_k = \gamma_k^{-1}(H_k + X_k/2) - \gamma_k X_k/2, \\ X_{k+1} = \widehat{X}_k + \widehat{H}_k, \qquad H_{k+1} = -\dfrac{1}{2}(\widehat{H}_k X_{k+1}^{-1} \widehat{H}_k), \qquad k = 0, 1, 2, \ldots. \end{cases}$$

The parameter γ_k is obtained at no additional cost during the inversion of X_k.

In certain cases the convergence is dramatically accelerated

PCR: convergence properties

The following properties hold:

1. PCR is well defined and convergent if and only if the real eigenvalues of M belong to the interval $[-1/2, 1/2]$, and the real eigenvalues of modulus $1/2$ (if any) are semisimple.

PCR: convergence properties

The following properties hold:

1. PCR is well defined and convergent if and only if the real eigenvalues of M belong to the interval $[-1/2, 1/2]$, and the real eigenvalues of modulus $1/2$ (if any) are semisimple.
2. If PCR is convergent then
   $P_k \to 0, \quad Q_k \to Q(1 - 4M^2)^{1/2} = H_0^{-1}$

PCR: convergence properties

The following properties hold:

1. PCR is well defined and convergent if and only if the real eigenvalues of M belong to the interval $[-1/2, 1/2]$, and the real eigenvalues of modulus $1/2$ (if any) are semisimple.
2. If PCR is convergent then
   $P_k \to 0, \quad Q_k \to Q(1 - 4M^2)^{1/2} = H_0^{-1}$
3. The convergence is:
   - quadratic if the real eigenvalues of M belong to the interval $(-1/2, 1/2)$;
   - linear if M has at least one semisimple eigenvalue equal to $\pm 1/2$.

PCR and $A^{1/2}$, $\mathrm{sign}(A)$, $\mathrm{polar}(A)$, $A\#B$

Let $P = \frac{1}{4}(B - A)$ and $Q = \frac{1}{2}(B + A)$. Then, given the matrices A and B, chosen according to the table below, the sequence Q_k converges to the corresponding matrix function.

A	B	$\lim_k Q_k$
$\sigma(A) \cap \mathbb{R}_- = \emptyset$	I	$A^{1/2}$
$\sigma(A) \cap \mathbb{C}_0 = \emptyset$	A^{-1}	$\mathrm{sign}(A)$
$\det(A) \neq 0$	A^{-*}	$\mathrm{polar}(A)$
$A > 0$	$B > 0$	$A\#B$

Numerical experiments

Set $P = \frac{1}{4}(B_\epsilon - A_\epsilon)$, $Q = \frac{1}{2}(A_\epsilon + B_\epsilon)$, where A_ϵ, B_ϵ are $n \times n$ symmetric positive definite, depending on $0 < \epsilon < \frac{1}{2}$, such that $M = Q^{-1}P$ has eigenvalues

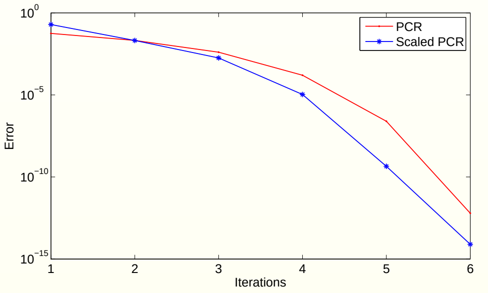
$$\frac{1}{2} - \epsilon, \frac{1}{3}, \ldots, \frac{1}{n+1}$$

Under these assumptions

$$\lim_k Q_k = A_\epsilon \# B_\epsilon$$

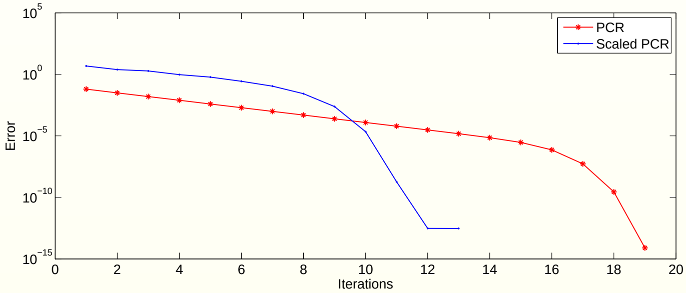
$n = 10$, $\epsilon =$ 1.e-5

	Absolute err.	Relative err.
PCR	2.0e-14	5.4e-15
Commuting mat.	2.3e-14	6.2e-15
3-term iter	9.6e-13	1.7e-14
Scaled PCR	1.7e-14	3.0e-15



$n = 10$, $\epsilon =$1.e-10

	Absolute err.	Relative err.
PCR	6.4e-12	1.7e-12
Commuting mat.	8.3e-12	2.2e-12
3-term iter	8.1e-10	1.9e-10
Scaled PCR	4.1e-12	7.3e-13



Open issues

- We have related palindromic quadratic matrix polynomials with $A^{1/2}$, $\mathrm{sign}(A)$, $\mathrm{polar}(A)$, $A\#B$. Are there any other matrix functions related to these polynomials?
- Can the theory developed for palindromic CR be extended to general CR? In particular, necessary and sufficient conditions for convergence, scaling technique, can be given for general CR?