KYBERNETIKA — VOLUME 11 (1975), NUMBER 6

New Algorithm for Polynomial Spectral Factorization with Quadratic Convergence I

ZDENĚK VOSTRÝ

The whole paper is divided into two parts appearing separately. In part I and II new efficient algorithms are derived for the numerical spectral factorization of polynomials arising in discrete and continuous optimality problems, respectively.

INTRODUCTION

If we are given a function $a(\zeta) = b(\zeta^{-1})\, b(\zeta)$, where $b(\zeta) = b_0 + b_1\zeta + \ldots + b_k\zeta^k$ and $b_0 b_k \neq 0$ without any loss of generality then the spectral factorization of $a(\zeta)$, $a(\zeta) = a_0 + a_1(\zeta + \zeta^{-1}) + \ldots + a_k(\zeta^k + \zeta^{-k})$, is defined as a polynomial $\varphi(\zeta)$ if

$$(1) \qquad \varphi(\zeta^{-1})\, \varphi(\zeta) = a(\zeta) \quad \text{and} \quad \varphi(\zeta) \neq 0 \quad \text{for} \quad |\zeta| < 1\,.$$

This problem is mentioned in [1–4].

There are several ways of numerical computation of the polynomial spectral factorization. For example

(i) computation of the roots of $a(\zeta)$ and their possible selection,

(ii) triangular factorization of semi-infinite matrix

(iii) Newton-Raphson method.

First method requires great amount of operations and its accuracy is given by multiplicity of the roots of $a(\zeta)$.

Second method is an iterative method with linear convergence. In [2] a very useful method for Cholesky triangular factorization of a positive definite Toeplitz matrix is described. By rearranging these expressions the next method is developed.

Consider two vector sequences

$$P^{(i)} = [p_0^{(i)}, p_1^{(i)}, \ldots, p_k^{(i)}]\,, \quad Q^{(i)} = [q_0^{(i)}, q_1^{(i)}, \ldots, q_k^{(i)}]\,, \quad i = 0, 1, 2 \ldots,$$

for which

$$P^{(0)} = [a_0, a_1, \ldots, a_k]\,, \quad Q^{(0)} = [a_1, a_2, \ldots, a_k, 0]$$

and

$$P_j^{(i+1)} = p_j^{(i)} - \frac{q_0^{(i)}}{p_0^{(i)}}\, q_j^{(i)}\,, \quad q_{j+1}^{(i+1)} = q_{j+1}^{(i)} - \frac{q_0^{(i)}}{p_0^{(i)}}\, p_{j+1}^{(i)}\,, \quad j = 0, 1, \ldots, (k-1)\,,$$

$$p_k^{(i+1)} = p_k^{(i)} = a_k\,, \quad q_k^{(i+1)} = q_k^{(i)} = 0\,,$$

then $\lim_{i\to\infty} Q^{(i)} = 0$, $\lim_{i\to\infty} P^{(i)} = P$ and the polynomial $P(\zeta) = p_0 + p_1\zeta + \ldots + p_k\zeta^k$ satisfies

$$\frac{1}{p_0}\, P(\zeta)\, P(\zeta^{-1}) = a(\zeta)\,.$$

Hence the $(1/p_0)\, P(\zeta)$ is the spectral factor of $a(\zeta)$.

This method requires only $2k + 1$ operations in one iteration and the error decreases as $|\lambda|^i$, where λ is the root of $a(\zeta)$ lying inside the unite circle and being maximal in modulus. If $|\lambda| = 1$ then for sufficiently large i the error decreases as $1/i$ only.

Third method is described separately and a new way of computation is developed.

NEWTON-RAPHSON METHOD

Consider a function $F(\zeta) = b(\zeta)\, b(\zeta^{-1}) - \varphi(\zeta)\, \varphi(\zeta^{-1})$. Denoting $\varphi(\zeta)$ as φ and $\varphi(\zeta^{-1})$ as $\bar{\varphi}$, the "differential" of F with respect to variations $\mathrm{d}\varphi$ in φ is

$$\mathrm{d}F = -\varphi\, \mathrm{d}\bar{\varphi} - \bar{\varphi}\, \mathrm{d}\varphi \tag{2}$$

To obtain a new value $\varphi + \mathrm{d}\varphi$, we solve the "linear" equation $F + \mathrm{d}F = 0$. The substitution $\varphi_{i+1} = \varphi_i + \mathrm{d}\varphi$ gives

$$\varphi^{(i)}\bar{\varphi}^{(i+1)} + \bar{\varphi}^{(i)}\varphi^{(i+1)} = b\bar{b} + \varphi^{(i)}\bar{\varphi}^{(i)}\,. \tag{3}$$

Properties of the sequence $\varphi^{(0)}, \varphi^{(1)}, \ldots$ are summarized in the folliwing theorems.

Theorem 1. [1, 3]. If $\varphi^{(0)} \neq 0$ for $|\zeta| \leqq 1$ i.e. $\varphi^{(0)}$ is a stable polynomial, the iterative solution of (3) has the properties bellow:

(i) $\varphi^{(i)} \neq 0$ for $|\zeta| \leqq 1$ implies that $\varphi^{(i+1)} \neq 0$ for $|\zeta| \leqq 1$

(ii) $\varphi^{(0)}, \varphi^{(1)}, \varphi^{(2)}, \ldots$ converges uniformly for $|\zeta| \leqq 1$ to the unique solution φ such that $\varphi\bar{\varphi} = b\bar{b}$ and $\varphi \neq 0$ for $|\zeta| < 1$,

(iii) for real $z \in \langle -1, 1\rangle$ and $\varphi^{(0)}(o) > 0$ the following inequality holds

$$\tfrac{1}{2}\varphi^{(i)}(z) < \varphi^{(i+1)}(z) \leqq \varphi^{(i)}(z)$$

(iv) the convergence of the sequence $\varphi^{(0)}, \varphi^{(1)}, \varphi^{(2)} \ldots$ is quadratic in nature.

Theorem 2. For any polynomial $\varphi^{(0)}$ and modulus of ζ equal to 1 holds

$$\varphi^{(i)}(\zeta)\,\varphi^{(i)}(\zeta^{-1}) \geqq b(\zeta)\,b(\zeta^{-1})\,, \quad i = 1, 2, \ldots,$$

where $\varphi^{(i)}$ is given by (3).

Proof. It is evident that for $|\zeta| = 1$

$$(\bar{\varphi}^{(i+1)} - \bar{\varphi}^{(i)})\,(\varphi^{(i+1)} - \varphi^{(i)}) \geqq 0\,.$$

Using (3),

$$\varphi^{i+1}\bar{\varphi}^{(i+1)} - b\bar{b} \geqq 0\,.$$

From the above theorem it follows

$$\sum_{l=0}^{k} (\varphi_l^{(i)})^2 \geqq \sum_{l=0}^{k} b_l^2 = a_0\,. \tag{4}$$

Now we introduce a new approach to solving the equation (3). Consider the substitution

$$\varphi^{(i+1)} = \tfrac{1}{2}(\varphi^{(i)} + x^{(i)}) \tag{5}$$

then

$$\bar{\varphi}^{(i)}x^{(i)} + \varphi^{(i)}\bar{x}^{(i)} = 2b\bar{b} \tag{6}$$

and for the sake of clarity the equation will be written as

$$\varphi\bar{x} + \bar{\varphi}x = c = 2b\bar{b}\,, \tag{7}$$

where

$$\varphi = \varphi_0 + \varphi_1\zeta + \ldots + \varphi_k\zeta^k\,;$$
$$x = x_0 + x_1\zeta + \ldots + x_k\zeta^k\,,$$
$$c = c_0 + c_1(\zeta + \zeta^{-1}) + \ldots + c_k(\zeta^k + \zeta^{-k})\,.$$

Equation (7) can be written in the following useful matrix form

$$\begin{bmatrix} \varphi_k & \varphi_{k-1} & \ldots & \varphi_1 & \varphi_0 & \varphi_0 & \varphi_1 & \ldots & \varphi_{k-1} & \varphi_k \\ 0 & \varphi_k & \ldots\ldots & \varphi_1 & 0 & \varphi_0 & \ldots & \ldots & & \varphi_{k-1} \\ \ldots & & & & & & & & & \\ \ldots & & & & & & & & & \\ 0 & 0 & \ldots & 0 & \varphi_k & 0 & 0 & \ldots & 0 & \varphi_0 \end{bmatrix} \begin{bmatrix} x_k \\ \cdot \\ \cdot \\ x_1 \\ x_0 \\ x_0 \\ x_1 \\ \vdots \\ x_k \end{bmatrix} = \begin{bmatrix} c_0 \\ c_1 \\ \cdot \\ \cdot \\ c_k \end{bmatrix}. \tag{8}$$

In a shorthand notation $\Phi X = C$.

Now we make use of the substitution $X = TY$, where Y has the same structure as X, such that the ΦT matrix will be triangular. Its elements are determined by the stability test of the polynomial φ as follows.

Stability test

$$
\begin{array}{lll}
\varphi_k\ \varphi_{k-1} \cdots \varphi_0 & & \\
\varphi_0\ \varphi_1\ \cdots \varphi_k & \Big|\ p_0 = -\dfrac{\varphi_k}{\varphi_0} & \\
\varphi^1_{k-1} \cdots\cdots \varphi^1_0 & & \varphi^1_j = \varphi_j + \varphi_{k-j} p_0,\ j = 0, 1 \ldots (k-1) \\
\varphi^1_0 \cdots\cdots \varphi^1_{k-1} & \Big|\ p_1 = -\dfrac{\varphi^1_{k-1}}{\varphi^1_0} & \\
\varphi^2_{k-2} \cdots \varphi^2_0 & & \varphi^2_j = \varphi^1_j + \varphi^1_{k-1-j} p_1,\ j = 0, 1 \ldots (k-2) \\
\cdots\cdots\cdots\cdots & & \\
\varphi^{k-1}_1\ \varphi^{k-1}_0 & & \\
\varphi^{k-1}_0\ \varphi^{k-1}_1 & \Big|\ p_{k-1} = -\dfrac{\varphi^{k-1}_1}{\varphi^{k-1}_0} & \\
\varphi^k_0 & & \varphi^k_0 = \varphi^{k-1}_0 + \varphi^{k-1}_1 p_{k-1}
\end{array}
$$

The equation $\Phi TY = C$ can be written as

(9)
$$
\begin{bmatrix}
0 \ldots 0 & 2\varphi^k_0 & 0 & \ldots & & 0 \\
0 \ldots 0 & 0 & \varphi^k_0 & \ldots & \varphi^2_{k-2} & \varphi^1_{k-1} \\
\cdots & & & & & \cdots \\
0 \ldots 0 & 0 & & \ldots 0 & \varphi^2_0 & \varphi^1_1 \\
0 \ldots 0 & 0 & & \ldots 0 & & \varphi^1_0
\end{bmatrix}
\begin{bmatrix} y_k \\ \cdot \\ \cdot \\ y_1 \\ y_0 \\ y_0 \\ y_1 \\ \vdots \\ y_k \end{bmatrix}
=
\begin{bmatrix} c_0 \\ c_1 \\ \cdot \\ \cdot \\ c_k \end{bmatrix}
$$

and gives the $y_0, y_1, \ldots, y_k$ after simple computation.

The substitution $X = TY$ can be computed as

$$
\begin{array}{lll}
y_k y_{k-1} \ \cdots & y_1 y_0 & \\
 & y_0 y_1 & \big|\ p_{k-1} \\
y^1_k y^1_{k-1} \ \cdots & y^1_2 y^1_1 y^1_0 & \\
 & y^1_0 y^1_1 y^1_2 & \big|\ p_{k-2} \\
y^2_k y^2_{k-1} \ \cdots & y^2_2 y^2_1 y^2_0 & \\
\cdots\cdots\cdots & & \\
y^{k-1}_k y^{k-1}_{k-1} \cdots & y_{k-1} y^{k-1}_0 & \\
y^{k-1}_0 y^{k-1}_1 \cdots & y^{k-1}_{k-1} y^{k-1}_k & \big|\ p_0 \\
x_k x_{k-1} \ \cdots & x_1 x_0 &
\end{array}
$$

where the third row of this scheme is given as the sum of the first row and the second row multiplied by p_{k-1}.

This algorithm for solving (7) requires only $\frac{3}{2}k^2 + \frac{9}{2}k + 1$ operations.

PROPERTIES OF THE EQUATION (7)

Denote ∂x the degree of a polynomial x. Premultiply (7) by $\zeta^{\partial b}$ and denote $\bar{b}\zeta^{\partial b} = \tilde{b}$ ($\tilde{b}$ is the reciprocal polynomial to b), then

$$\varphi\tilde{x} + \tilde{\varphi}x = 2b\tilde{b}\,. \tag{10}$$

For our purpose $\partial b = \partial\varphi$ and moreover we require $\partial x = \partial b$. If φ, b are stable polynomials (it particularly implies that the roots of φ and b do not lie on the unite circle) then (10) has only one solution and $\partial x = \partial b$. On the other hand, if $\varphi \neq 0$ and $b \neq 0$ for $|\zeta| < 1$ and $\varphi(\eta) = b(\eta) = 0$ for some η, $|\eta| = 1$, then many solutions of (10) with $\partial x = \partial b$ exist.

Example 1.

$$\varphi = 1 - \zeta\,, \quad b = 1 - \zeta\,, \quad \partial x = 1\,.$$

Denote $x = x_0 + x_1\zeta$, then from (10)

$$(1 - \zeta)\,\tilde{x} + (\zeta - 1)\,x = 2(1 - \zeta)(\zeta - 1)\,.$$

Hence

$$\tilde{x} - x = 2(\zeta - 1)$$

and the solution with $\partial x = 1$ can be written as $x = (1 - \zeta) + \tau(1 + \zeta)$, where $\tau \neq 1$ is a real number.

Our algorithm is based on the stability test and hence if any roots of the polynomial φ tends to 1 in modulus then at least one $\varphi_0^{(i)}$ approaches 0. Using a computer the computational errors increase as $\varphi_0^{(i)}$ approaches 0. In some cases the errors can be increased such that a new iteration of the algorithm is worse then the present one. Considering a finite number of decimal digits in a computer, this follows from Example 1.

These properties are not due to the choice of computational method of solving (6) but to the nature of this spectral factorization.

In many instances the troublesome roots may be isolated by a preliminary factorization.

There are two ways of numerical testing of the given algorithm: first, to compare the coefficients of $\varphi^{(i)}\bar{\varphi}^{(i)}$ and $b\bar{b}$, second, to compare the coefficients of $\varphi^{(i)}$ and b, where b is the accurate spectral factor of $a = b\bar{b}$. The proportionality between these errors depends on the roots of a and the modulus of the coefficients of b.

Example 2.

$$b = (\zeta^2 + 1)^3$$

a	20	0	15	0	6	0	1
$\varphi\bar{\varphi} - a$	10^{-6}	0	$2\cdot6 \,.\, 10^{-7}$	0	$-5\cdot2 \,.\, 10^{-7}$	0	$-2\cdot6 \,.\, 10^{-7}$
b	1	0	3	0	3	0	1
$\varphi - b$	$1\cdot24 \,.\, 10^{-2}$	0	$1\cdot23 \,.\, 10^{-2}$	0	$-1\cdot25 \,.\, 10^{-2}$	0	$-1\cdot23 \,.\, 10^{-2}$

$$b = (1 - 0{,}99\zeta)^2$$

a	5·880996	3·920598	0·9801
$\varphi\bar{\varphi} - a$	$1\cdot22 \,.\, 10^{-12}$	$5\cdot9 \,.\, 10^{-13}$	$-2\cdot01 \cdot 10^{-14}$
b	1	1·98	0·9801
$\varphi - b$	$3\cdot6 \,.\, 10^{-11}$	$-4 \,.\, ^{-13}$	$-3\cdot5 \,.\, 10^{-11}$

Choice of the starting polynomial

Theorem 3. Let b be a polynomial and $a = b(\zeta)\, b(\zeta^{-1}) = a_0 + a_1(\zeta + \zeta^{-1}) + \ldots$ $\ldots + a_k(\zeta^k + \zeta^{-k})$, then $\varphi^{(1)} = 1/\sqrt{(a_0)}\,(a_0 + a_1\zeta + \ldots + a_k\zeta^k)$ is a stable polynomial ($\varphi^{(1)} \neq 0$ for $|\zeta| \leqq 1$).

Proof. Consider the algorithm (3). Then for $\varphi^{(0)} = \sqrt{a_0}$ the $\varphi^{(1)}$ is a stable polynomial from Theorem 1.

As it follows from numerical examples this polynomial $\varphi^{(1)}$ is a suitable starting polynomial and corresponds to the starting polynomial in the second mentioned method.

Stop rule

One of the very important problem is to stop the iteration process in such a way that the result is stable in spite of computation errors and has a maximal reachable accuracy.

The first condition can be reached very simply because the computation of each iteration is based on the stability test of the previous iteration.

The second condition can be reached if the monotonicity of the iteration $\varphi^{(i)}(0)$ (see the condition (iii) of Theorem 1) and inequality (4) (see Theorem 2) are tested. The result of the iteration process is choosen in the following way.

(i) if $\varphi^{(n)}_{(0)} > \varphi^{(n-1)}_{(0)}$ or $\sum_{l=0}^{k} (\varphi^{(n)}_l)^2 < a_0$ then $\varphi = \varphi^{(n)}$ if $\varphi^{(n)}$ is stable, else $\varphi = \varphi^{(n-1)}$ (it follows from the algorithm that $\varphi^{(n-1)}$ is a stable polynomial);

(ii) if $\sum_{l=0}^{k} (\varphi^{(n)}_l)^2 - a_0 < a_0\, 10^{-14}$ or $n > 30$ then

$$\varphi = \varphi^{(n)} ;$$

(iii) if the stability test of $\varphi^{(n-1)}$ for computing $\varphi^{(n)}$ does not hold then

$$\varphi = \varphi^{(n-2)} .$$

NUMERICAL EXAMPLES

(Computer IBM 370, 16 decimal digits, program in the PL/I language.)
Consider two examples without troublesome roots reported in [3].

1) $\partial b = 3$

b	89	27	7	1
a	8004	2491	622	85

after CPU time 0.07s and $n = 4$ (n-number of iterations)

$\varphi - b$	$-7{\cdot}1 \,.\, 10^{-15}$	$3{\cdot}5 \,.\, 10^{-15}$	$2{\cdot}2 \,.\, 10^{-16}$	$-1{\cdot}4 \,.\, 10^{-17}$
$\varphi\bar{\varphi} - a$	$-2{\cdot}7 \,.\, 10^{-12}$	$5{\cdot}7 \,.\, 10^{-14}$	$-5{\cdot}7 \,.\, 10^{-14}$	$-1{\cdot}1 \,.\, 10^{-14}$

2) $\partial b = 8$

b	1 3,01	3,7488 2,2309	0·1704 −1·7945666
	−3·19483744	−2·594828029	−0·858277728

$$a = b\bar{b}$$

after CPU time 0.70s and $n = 16$

$$\|\varphi - b\| = 2{\cdot}4 \,.\, 10^{-8} , \quad \|\varphi\bar{\varphi} - a\| = 1{\cdot}4 \,.\, 10^{-13}$$

$$\|b\| = \max_{0 \leq i \leq \partial b} |b_i|) .$$

3) Consider a polynomial b with single roots on the unite circle $b = 1 + \zeta + \zeta^2 + \ldots + \zeta^{10}$ then after CPU time 1·32s, $n = 22$

$$\|\varphi - b\| = 3{\cdot}6 \,.\, 10^{-7}\,, \quad \|\varphi\bar{\varphi} - a\| = 8{\cdot}1 \,.\, 10^{-12}\,.$$

4) Consider a polynomial b no root of which lies on or closed to the unite circle

$$b = 6 + 5\zeta + 4\zeta^2 + 3\zeta^3 + 2\zeta^4 + \zeta^5$$

then after CPU time 0·13s and $n = 5$

$$\|\varphi - b\| = 1{\cdot}1 \,.\, 10^{-15}\,, \quad \|\varphi\bar{\varphi} - a\| = 1{\cdot}8 \,.\, 10^{-14}$$

(note that $a_0 = 91$).

5) Consider polynomials roots of which tends to 1 in modulus

b	$\|\varphi - b\|$	$\|\varphi\varphi - a\|$	CPU time	number of iteration
$(1 + 0{\cdot}99\xi)^2$	$3{\cdot}6 \,.\, 10^{-11}$	$1{\cdot}2 \,.\, 10^{-12}$	0·20	17
$(1 + 0{\cdot}999\xi)^2$	$1{\cdot}18 \,.\, 10^{-8}$	$4{\cdot}5 \,.\, 10^{-10}$	0·27	23
$(1 + 0{\cdot}9999\xi)^2$	$2 \,.\, 10^{-6}$	$9 \,.\, 10^{-9}$	0·33	28
$(1 + \xi)^2$	$6{\cdot}9 \,.\, 10^{-5}$	$1{\cdot}18 \,.\, 10^{-8}$	0·33	28

CONCLUSION

It follows from other examples that the accuracy of results depends only on the roots nearest to the unite circle and on its multiplicity. In cases where b has not troublesame roots the polynomial φ is correct to fifteen decimal digits.

(Received March 20, 1975.)

REFERENCES

[1] G. T. Wilson: Factorization of the covariance generating function of a pure moving average process. SIAM J. Numer. Anal. *6* (1969), 1–7.

[2] J. Rissanen: Algorithm for triangular decomposition of block Hankel and Toeplity matrices with application to factoring positive matrix polynomial. Mathematics of computation *27* (1973), 147–154.

[3] G. T. Wilson: The factorization of matricial spectral densities. SIAM J. Appl. Math. *23*, (1972), 420–426.

[4] W. G. Tuel, Jr.: Computer algorithm for spectral factorization of rational matrices. IBM J. (1968), 163–170.

Ing. Zdeněk Vostrý, CSc.: Ústav teorie informace a automatizace ČSAV (Institute of Information Theory and Automation — Czechoslovak Academy of Sciences), Pod vodárenskou věží 4, 180 76 Praha 8. Czechoslovakia.