A Note on Nearest Singular Polynomials *

Zhi Lihong
Institute of Systems Science, Academia Sinica
Beijing 100080, China
E-mail: lzhi@mmrc.iss.ac.cn

Wu Wenda
Beijing Municipal Computing Center
Beijing 100005,China
E-mail: wdwu@mmrc.iss.ac.cn

Abstract

The nearest singular polynomials to a given polynomial have been studied in [1], based on minimization of quadratic forms. An equivalent expression of the quadratic form is presented. It leads to a simple equation satisfied by the double zeros of the nearest singular polynomials.

§1. Introduction

An approach based on minimization of quadratic forms to study problems related to polynomials with inexact coefficients has been mentioned or proposed by several authors[1,2,3,4]. In [1], the nearest singular polynomials to a given polynomial have been defined, and by the above approach, a parametric quadratic form with the double root of the perturbed polynomial as parameter is obtained.

In this note, we will give an equivalent expression of the quadratic form, its first derivative can be factored into two nontrivial factors, and only one of them yields the local minimum. Therefore, for finding the double zeros

*Work supported in part by the national climbing project No.8505, and in part by Chinese-Austria Scientific-Technological Cooperation Project No, IV B 15, and Beijing Municipal committee of Science and technology

of the nearest singular polynomials, it is sufficient to use this factor with a simple expression.

In §2, we review briefly the approach given in [1]. The new expression of the quadratic form is proved in §3. The obvious consequences derived from it are given in §4, and numerical examples and discussions in §5.

§2. Karmarkar and Lakshman's Approach [1]

Given a monic polynomial f

$$f = x^m + \sum_{j=1}^{m} f_j x^{m-j} \tag{1}$$

with complex coefficients, a nearest singular polynomial to f is defined to be a monic polynomial h

$$h = (x-c)^2 (x^{m-2} + \sum_{j=1}^{m-2} \phi_j x^{m-2-j}) \tag{2}$$

such that

$$\mathcal{N} := \| f - h \|^2 = \sum_{j=1}^{m} |f_j - \phi_j + 2c\phi_{j-1} - c^2 \phi_{j-2}|^2$$

is minimized, where $\phi_{-1} = 0, \phi_0 = 1, \phi_{m-1} = 0, \phi_m = 0$, and our problem is to look for $c \in \mathcal{C}$ and $\phi_j's \in \mathcal{C}$ that minimize $\mathcal{N}$.

Let $\phi = [\phi_1 \quad \phi_2 \quad \cdots \quad \phi_{m-2}]^T$, $\mathcal{N}$ can be written as

$$\mathcal{N} = \phi^* Q_2^{(m-2)} \phi - \phi^* r - r^* \phi + s$$

where $*$ stands for conjugate transpose. Here $Q_2^{(m-2)} := (\ q_{ij}\ )$ is Hermitian, Toeplitz, positive definite and five-diagonal $(m-2) \times (m-2)$ matrix, with $q_{11} = 1 + 4c\overline{c} + (c\overline{c})^2, q_{21} = -2c(1 + c\overline{c}), q_{31} = c^2$, $r = [r_1 \quad r_2 \quad \cdots \quad r_{m-2}]^T$ with

$$\begin{aligned} r_1 &= f_1 - 2\overline{c} f_2 + \overline{c}^2 f_3 + 2c(1 + c\overline{c}), \\ r_2 &= f_2 - 2\overline{c} f_3 + \overline{c}^2 f_4 - c^2, \\ r_k &= f_k - 2\overline{c} f_{k+1} + \overline{c}^2 f_{k+2}, \quad k > 2 \end{aligned} \tag{3}$$

and

$$s = \sum_{j=1}^{m} |f_j|^2 + 2c\overline{f_1} + 2\overline{c} f_1 - c^2 \overline{f_2} - \overline{c}^2 f_2 + 4c\overline{c} + (c\overline{c})^2. \tag{4}$$

For fixed $c \in \mathcal{C}$, $\mathcal{N}$ attains a minimum at ϕ satisfying

$$Q_2^{(m-2)}\phi - r = 0,$$

and

$$\mathcal{N}_m := min\mathcal{N} = -r^*(Q_2^{(m-2)})^{-1}r + s, \tag{5}$$

which is a real valued function of the complex variable c.

Karmarkar proposed to write $c = a + ib$, where a and b are real, and to consider $\mathcal{N}_m$ as a real rational function of the real variables a and b. The problem then reduces to find real solutions (ξ, η) of the system

$$\frac{\partial \mathcal{N}_m}{\partial a} = 0, \frac{\partial \mathcal{N}_m}{\partial b} = 0 \tag{6}$$

for the determination $c = \xi + i\eta$ and the corresponding nearest singular polynomials.

§3. The Expression of the Quadratic Form

Let $Q_1^{(k)} = (\ q_{ij}\ )$ be a $k \times k$ Toeplitz, Hermitian, positive definite and tridiagonal matrix with $q_{11} = 1 + c\overline{c}, q_{21} = -c$ [1]. Denote the determinant of $Q_1^{(k)}$ by $q_1^{(k)} = \sum_{j=0}^{k} (c\overline{c})^j$ and the determinant of $Q_2^{(k)}$ in §2 for $m = k+2$ by $q_2^{(k)}$.

Lemma 1.

$$(q_1^{(k)})^2 = q_2^{(k)} - 2c\overline{c}q_2^{(k-1)} + (c\overline{c})^2 q_2^{(k-2)}.$$

Proof: It is easy to see that

$$(Q_1^{(k)})^2 = Q_2^{(k)} - \begin{bmatrix} c\overline{c} & & & & \\ & 0 & & & \\ & & \ddots & & \\ & & & 0 & \\ & & & & c\overline{c} \end{bmatrix}.$$

The lemma follows from the above identity. □

Lemma 1 imply the following recurrence relation.

Lemma 2.

$$q_2^{(k)} = q_1^{(k+1)} \frac{\partial^2 q_1^{(k+1)}}{\partial c \partial \overline{c}} - \frac{\partial q_1^{(k+1)}}{\partial c} \frac{\partial q_1^{(k+1)}}{\partial \overline{c}}. \tag{7}$$

$\square$

Now we are ready to prove

Theorem 1.

$$\mathcal{N}_m = \frac{f(c)\overline{f(c)}}{q_1^{(m-1)}} + \frac{P_2\overline{P_2}}{q_1^{(m-1)}q_2^{(m-2)}},$$

where

$$P_2 := q_1^{(m-1)} f'(c) - \frac{\partial q_1^{(m-1)}}{\partial c} f(c). \tag{8}$$

Proof: Let $L^{(m)} = (\quad l_{ij} \quad)$ be a $k \times k$ bi-diagonal Toeplitz matrix with $l_{11} = 1, l_{21} = -c$. Consider the auxiliary linear system:

$$((L^{(m)})^*)^2 (L^{(m)})^2 \widehat{\phi} = \widehat{r} \tag{9}$$

where

$$\begin{aligned} \widehat{\phi} &= [\phi \quad \phi_{m-1} \quad \phi_m]^T, \\ \widehat{r} &= [r \quad r_{m-1} \quad r_m]^T, \\ r_{m-1} &= f_{m-1} - 2\overline{c} f_m + \beta_1, \\ r_m &= f_m + \beta_2, \end{aligned}$$

β_1 and β_2 are parameters which will be specified later.

The inverses of $(L^{(m)})^*$ and $L^{(m)}$ are obvious. Thus

$$\widehat{\phi} = (L^{(m)})^{-2}((L^{(m)})^*)^{-2}\widehat{r},$$

$$((L^{(m)})^*)^{-2}\widehat{r} = \begin{bmatrix} f_1 + (m-1)\overline{c}^{m-2}\beta_1 + m\overline{c}^{m-1}\beta_2 + 2c \\ f_2 + (m-2)\overline{c}^{m-3}\beta_1 + (m-1)\overline{c}^{m-2}\beta_2 - c^2 \\ f_3 + (m-3)\overline{c}^{m-4}\beta_1 + (m-2)\overline{c}^{m-3}\beta_2 \\ \cdots \\ f_{m-2} + 2\overline{c}\beta_1 + 3\overline{c}^2\beta_2 \\ f_{m-1} + \beta_1 + 2\overline{c}\beta_2 \\ f_m + \beta_2 \end{bmatrix}. \tag{10}$$

The last two components ϕ_{m-1} and ϕ_m of $\widehat{\phi}$ are:

$$\begin{aligned}
\phi_{m-1} &= \sum_{i=1}^{m-1}(m-i)c^{m-i-1}f_i + \beta_1 \sum_{i=1}^{m-1}(m-i)^2 c^{m-i-1}\overline{c}^{m-i-1} \\
&\quad +\beta_2 \sum_{i=1}^{m-1}(m-i)(m-i+1)c^{m-i-1}\overline{c}^{m-i} + mc^{m-1} \\
&= f'(c) + \beta_1 \frac{\partial^2 q_1^{(m-1)}}{\partial c \partial \overline{c}} + \beta_2 \left(\overline{c}\frac{\partial^2 q_1^{(m-1)}}{\partial c \partial \overline{c}} + \frac{\partial q_1^{(m-1)}}{\partial c} \right), \\
\phi_m &= \sum_{i=1}^{m}(m-i+1)c^{m-i}f_i + \beta_1 \sum_{i=1}^{m-1}(m-i+1)(m-i)c^{m-i}\overline{c}^{m-i-1} \\
&\quad +\beta_2 \sum_{i=1}^{m}(m-i+1)^2 c^{m-i}\overline{c}^{m-i} + (m+1)c^m \\
&= (cf(c))' + \beta_1 \left(c\frac{\partial^2 q_1^{(m-1)}}{\partial c \partial \overline{c}} + \frac{\partial q_1^{(m-1)}}{\partial \overline{c}} \right) + \beta_2 \left(c\overline{c}\frac{\partial^2 q_1^{(m-1)}}{\partial c \partial \overline{c}} + 2c\frac{\partial q_1^{(m-1)}}{\partial c} + q_1^{(m-1)} \right),
\end{aligned}$$

and $\phi_{m-1} = \phi_m = 0 \quad \text{iff}$

$$\begin{aligned}
\beta_1 &= \frac{1}{q_2^{(m-2)}} \left(\frac{\partial^2 q_1^{(m-1)}}{\partial c \partial \overline{c}}\overline{c}f(c) + \frac{\partial q_1^{(m-1)}}{\partial c}f(c) - c\frac{\partial q_1^{(m-1)}}{\partial c}f'(c) - q_1^{(m-1)}f'(c) \right), \\
\beta_2 &= \frac{1}{q_2^{(m-2)}} \left(\frac{\partial q_1^{(m-1)}}{\partial \overline{c}}f'(c) - \frac{\partial^2 q_1^{(m-1)}}{\partial c \partial \overline{c}}f(c) \right).
\end{aligned} \tag{11}$$

In the following, we assume β_1 and β_2 to take the above values so that $\widehat{\phi} = [\phi \;\; 0 \;\; 0]^T$. Since

$$((L^{(m)})^*)^2(L^{(m)})^2 = \begin{bmatrix} Q_2^{(m-2)} & * \\ * & * \end{bmatrix},$$

we have

$$\begin{aligned}
\widehat{r}^*\widehat{\phi} &= r^*\phi, \\
r^*(Q_2^{(m-2)})^{-1}r &= r^*\phi = \widehat{r}^*\widehat{\phi} = \widehat{r}^*(L^{(m)})^{-2}((L^{(m)})^*)^{-2}\widehat{r} \\
&= (((L^{(m)})^*)^{-2}\widehat{r})^*(((L^{(m)})^*)^{-2}\widehat{r}) \\
&= s - \frac{f(c)\overline{f(c)}}{q_1^{(m-1)}} - \frac{P_2\overline{P_2}}{q_1^{(m-1)}q_2^{(m-2)}},
\end{aligned}$$

by (10) and (11). □

§4. The Equation satisfied by c

Instead of the system (5), we consider

$$\frac{\partial \mathcal{N}_m}{\partial c} = 0, \frac{\partial \mathcal{N}_m}{\partial \overline{c}} = 0. \tag{12}$$

Obviously, $\frac{\partial \mathcal{N}_m}{\partial c}, \frac{\partial \mathcal{N}_m}{\partial \overline{c}}$ are complex conjugate to each other and one of them is sufficient to determine c. The trace T and determinant D of the Hessian matrix

$$\begin{bmatrix} \frac{\partial^2 \mathcal{N}_m}{\partial a^2} & \frac{\partial^2 \mathcal{N}_m}{\partial a \partial b} \\ \frac{\partial^2 \mathcal{N}_m}{\partial a \partial b} & \frac{\partial^2 \mathcal{N}_m}{\partial b^2} \end{bmatrix}$$

are

$$T = 4\frac{\partial^2 \mathcal{N}_m}{\partial c \partial \overline{c}},$$

and

$$D = 4\left(\left(\frac{\partial^2 \mathcal{N}_m}{\partial c \partial \overline{c}}\right)^2 - \frac{\partial^2 \mathcal{N}_m}{\partial c^2}\frac{\partial^2 \mathcal{N}_m}{\partial \overline{c}^2}\right).$$

By direct computation, we have

$$\frac{\partial \mathcal{N}_m}{\partial c} = \frac{1}{q_1^{(m-1)}(q_2^{(m-2)})^2}\overline{P_2}P_3 \tag{13}$$

where

$$P_3 := q_2^{(m-2)}P_2' - \frac{\partial q_2^{(m-2)}}{\partial c}P_2. \tag{14}$$

Furthermore

$$\frac{\partial^2 \mathcal{N}_m}{\partial c \partial \overline{c}} = \frac{1}{q_1^{(m-1)}(q_2^{(m-2)})^3}\left(P_3\overline{P_3} - \left(q_2^{(m-2)}\frac{\partial^2 q_2^{(m-2)}}{\partial c \partial \overline{c}} - \frac{\partial q_2^{(m-2)}}{\partial c}\frac{\partial q_2^{(m-2)}}{\partial \overline{c}}\right)P_2\overline{P_2}\right), \tag{15}$$

and, for $P_2 = 0$,

$$D = \frac{4P_2'\overline{P_2'}}{(q_1^{(m-1)})^4(q_2^{(m-2)})^2}((q_1^{(m-1)})^2P_2'\overline{P_2'} - (q_2^{(m-2)})^2 f(c)\overline{f(c)}). \tag{16}$$

Theorem 2. $\mathcal{N}_m$ attains its local minimum at c satisfying $P_2 = 0$ if

$$(q_1^{(m-1)})^2P_2'\overline{P_2'} - (q_2^{(m-2)})^2 f(c)\overline{f(c)} > 0, \tag{17}$$

the minimum value of $\mathcal{N}_m$ is

$$\frac{f(c)\overline{f(c)}}{q_1^{(m-1)}},$$

and the perturbed polynomial is

$$h = f(x) - \frac{f(c)}{q_1^{(m-1)}} \sum_{j=0}^{m-1} (\overline{c}x)^j.$$

Proof: The last part comes from the uniqueness of ϕ for fixed c, and

$$\|f-h\|^2 = \left\| \frac{f(c)}{q_1^{(m-1)}} \sum_{j=0}^{m-1} (\overline{c}x)^j \right\|^2 = \frac{f(c)\overline{f(c)}}{(q_1^{(m-1)})^2} \sum_{j=0}^{m-1} (c\overline{c})^j = \frac{f(c)\overline{f(c)}}{q_1^{(m-1)}}.$$

The other parts are obvious. □

Theorem 3 The zero set of $P_2(P_3)$ contains all zeros of f with multiplicity $\geq 2(\geq 3)$.

Proof: By the expressions of P_2 and P_3. □

Thus when f has a double root, then $h = f$.

Theorem 4. For c satisfying $P_3 = 0$, the trace is non-positive.

Proof: By (15), if $P_3 = 0$ then

$$T = -\frac{4}{q_1^{(m-1)}(q_2^{(m-2)})^3} \left(q_2^{(m-2)} \frac{\partial^2 q_2^{(m-2)}}{\partial c \partial \overline{c}} - \frac{\partial q_2^{(m-2)}}{\partial c} \frac{\partial q_2^{(m-2)}}{\partial \overline{c}} \right) P_2 \overline{P_2},$$

We also have:

$$\begin{aligned}
& q_2^{(m-2)} \tfrac{\partial^2}{\partial c \partial \overline{c}} q_2^{(m-2)} - \tfrac{\partial}{\partial c} q_2^{(m-2)} \tfrac{\partial}{\partial \overline{c}} q_2^{(m-2)} \\
&= \textstyle\sum_{j=0}^{2m-4} \alpha_j (c\overline{c})^j \sum_{j=1}^{2m-4} j^2 \alpha_j (c\overline{c})^{j-1} - c\overline{c} (\sum_{j=1}^{2m-4} j \alpha_j (c\overline{c})^{j-1})^2 \\
&= \textstyle\sum_{j=1}^{2m-4} j^2 \alpha_j (c\overline{c})^{j-1} + c\overline{c} (\sum_{j=1}^{2m-4} \alpha_j (c\overline{c})^{j-1} \sum_{j=1}^{2m-4} j^2 \alpha_j (c\overline{c})^{j-1} \\
&\quad \textstyle - (\sum_{j=1}^{2m-4} j \alpha_j (c\overline{c})^{j-1})^2),
\end{aligned}$$

where $\alpha_j = \alpha_{2m-4-j} = \frac{1}{6}(j+1)(j+2)(j+3) > 0, \forall j \leq m-2$, and the first term is obviously positive, and the second term is positive by the Minkowski inequality. It imply the trace is non-positive. □

It is easy to see that the solution of $P_3 = 0$ is useless for our present purpose. But we would like to note that $P_3 = 0$ is useful when we consider nearest singular polynomial with triple root. See example 3.

In summary, it is sufficient to solve $P_2 = 0$ and to compare the values $f(c)\overline{f(c)}/q_1^{(m-1)}$ among the solutions satisfying (17). Note that the nearest singular polynomial may not be unique; see the examples below.

§5. Numerical Examples

Example 1: $\quad f = x^5 - x.$

There are 4 nearest singular polynomials due to the geometry of the zeros of f, one of them is given below:

$$\begin{aligned} h \;\approx\; & x^5 + 0.03895547966x^4 + 0.06708530296x^3 + 0.1155277233x^2 \\ & -0.8010494959x + 0.3426130279. \end{aligned}$$

The zeros of h are:

zeros of h	zeros of f
0.5806857529 (double)	0, 1
-1.050883646	-1
-0.07472166958+0.9804509313 i	i
-0.07472166958-0.9804509313 i	-i

$$\mathcal{N}_m = 0.1763296120.$$

The other 3 can be obtained by rotation with an angle $\pi/2$, π and $3\pi/2$ respectively.

Example 2: $\quad f = x^5 + 1.$

There are 5 nearest singular polynomials, one of them is:

$$\begin{aligned} h \;\approx\; & x^5 - 0.1944450407x^4 - 0.2533066502x^3 - 0.3299866060x^2 \\ & -0.4298788054x + 0.4399900361, \end{aligned}$$

with zeros:

zeros of h	zeros of f
0.7676270658(double)	$e^{\pm\pi i/5}$
-0.907566588	-1
$-0.2166162160 \pm 0.8808010592i$	$e^{\pm 3\pi i/5}$

$$\mathcal{N}_m = 0.7092712403.$$

The other four can be obtained by rotation. Note that in this example $c = 0$ is a solution, which is a saddle point.

Example 3:

$$\begin{aligned} f &= (x - 0.89 - 0.03i)(x - 0.88 + 0.02i)(x - 0.87)(x - 1) \\ &= x^4 + (-3.64 - 0.01i)x^3 + (4.9637 + 0.0273i)x^2 + (-3.005606 - 0.024782i)x \\ &\quad +0.681906 + 0.007482i. \end{aligned}$$

The nearest singular polynomial of f is unique:

$$\begin{aligned} h \approx\ & x^4 + (-3.639999897 - 0.01000012076i)x^3 + (4.963700115 + 0.02729986094i)x^2 \\ & +(-3.005605870 - 0.02478216008i)x + 0.6819061456 + 0.007481815756i, \end{aligned}$$

with the roots given below:

zeros of h	zeros of f
$0.8768135619 - 0.01006779565i$ (double)	$0.88 - 0.02i$, 0.87
$0.8866786823 + 0.02982563187i$	$0.89 + 0.03i$
$0.9996940915 + 0.0003100788900i$	1

$$\mathcal{N}_m = 0.1552760144 \times 10^{-12}.$$

For c satisfying $P_3 = 0$, we can get an unique nearest singular polynomial with a triple root:

$$\begin{aligned} h \approx\ & x^4 + (-3.639968566 - 0.01002119406i)x^3 + (4.963698969 + 0.02730077333i)x^2 \\ & +(-3.005622319 - 0.02477096306i)x + 0.6819004541 + 0.007485892787i. \end{aligned}$$

The zeros of h are:

zeros of h	zeros of f
$0.8817735725 + 0.002337412033i$ (triple)	$0.88 - 0.02i$, 0.87
$0.9946478484 + 0.003008957959i$,	1

$$\mathcal{N}_m = 0.3311925673 \times 10^{-8}.$$

Remark: We have simplified the equation satisfied by the double zeros c, but how to find all c efficiently is another problem. Fortunately the following bound is helpful.Let c_{opt} denote an optimal c. We have

$$|c_{opt}| \leq 1 + \max_j |f_j| + (|f_{m-1}|^2 + |f_m|^2)^{\frac{1}{2}}$$

The techniques used here can be applied to the nearest singular polynomial $h(x)$ with a zero of multiplicity > 2, it will be reported later.

Acknowledgment The authors thank Prof. H. J. Stetter for bringing related works to attention.

References

[1] N.Karmarkar, Lakshman Y.N. Approximate Polynomial Greatest Common Divisors and Nearest Singular Polynomials. In Proceedings of the International Symposium on Symbolic and Algebraic Computation, Zurich, Switzerland, July 1996, 35-39.

[2] N.Karmarkar, Lakshman Y.N. On Approximate Polynomial Greatest Common Divisors. In Proceedings of Workshop on Symbolic-Numeric Algebra for Polynomial,France, July 1996, 25.

[3] R.M. Corless, P.M.Gianni, B.M.Trager and S.M.Watt The Singular Value Decomposition for Polynomial Systems. In Proceedings of the International Symposium on Symbolic and Algebraic Computation, Montreal, Canada, July 1995, 195-207.

[4] N.Karmarkar, Lakshman Y.N. to appear.