Abstract

This paper analyses the growth of the condition number of a class of modified moment matrices that arise when computing least squares polynomials in polygons of the complex plane. It is shown that if the polygon is inserted between two ellipses then the condition number of the (n+1) x (n+1) modified moment matrix is upper bounded by $2m(n+1)^2(\kappa)^{2n}$, where m is the number of edges of the polygon, and $\kappa \geq 1$ is a known ratio which is close to one if the two ellipses are close to each other.

On the condition number of modified moment matrices arising from least squares approximation in the complex plane

Youcef Saad

Technical Report YALEU/DCS/RR-347

January 3, 1985

This work was supported in part by the U.S. Office of Naval Research under grant N00014-82-K-0184, in part by Dept. Of Energy under Grant AC02-81ER10996, and in part by Army Research Office under contract DAAG-83-0177.

1. Introduction

A classical problem that arises when solving the linear system Ax=b, by iterative methods, is to find a polynomial $s(\lambda)$ of degree $\leq$n-1 which is such that the residual polynomial $R(\lambda)=1-\lambda s(\lambda)$ is small in a certain region H of the complex plane containing the sectrum of the matrix A. In [7, 8] the region H was chosen to be a polygon, and the problem of interest was to find a polynomial s_n of degree n-1 such that $||1-\lambda s(\lambda)||_\omega$ is minimum, where $||.||_\omega$ is the norm associated with the L_2 inner product $<.,.>$ with respect to the weight function $\omega(\lambda)$. In order to compute the least squares polynomial s_n in some basis $\{t_j\}_{j=0,n}$ one needs to compute the Gram matrix $M_n=\{<t_i,t_j>\}_{i,j=0,n}$ often referred to as the modified moment matrix, along with its Choleski factorization. Similar problems are encountered when one uses polynomial acceleration in eigenvalue algorithms [8] such as the subspace iteration algorithm, or when one is interested in computing a polynomial approximation to the exponential of a matrix. A crucial difficulty that arises then is to choose a good basis of polynomials, i.e. a basis for which the modified moment matrix is not too badly conditioned.

It is a well known fact that, in general, the use of the power basis $\{1,\lambda,\lambda^2,.....,\lambda^n\}$ is to be avoided for stability reasons [11, 12]. Instead, if for example one is interested in approximating a given function in the interval [-1,+1], a better alternative is to use of the basis of Chebyshev polynomials of the first kind $\{T_j(\lambda)\}_{j=0,n}$. More generally, if the interval is $[\alpha,\beta]$ then a good basis is $\{T_j[(\lambda-c)/d]\}_{j=0,n}$ where $c=(\beta+\alpha)/2$, $d=(\beta-\alpha)/2$. Assuming that the convex hull H can be enclosed in the ellipse centered at c, and having focal distance d and major semi axis a, a natural generalization is to use the following basis of polynomials:

$$t_j(\lambda) \;=\; T_j(\tfrac{\lambda-c}{d})/\;T_j(\tfrac{a}{d}) \quad k=0,1,..n \quad . \tag{1}$$

This basis has been successfully used in [7], for computing least-squares polynomials of much higher degrees than would have been allowed with the classical power basis. In this paper we will analyse this basis in more detail and establish an upper bound for the condition number of its Gram matrix.

In the real interval case, the idea of using modified moments instead of regular moments to compute orthogonal polynonmials is well established and was first proposed by Sack and Donovan [10]. Gautschi has analysed the condition numbers of the modified moment matrices in that context [2, 3], and has shown that the technique is reliable at least for finite intervals.

Although we will not show as a strong a resullt for the complex case, we will demonstrate that the condition number will not grow too rapidly if the polygon is well approximated by the enclosing ellipse.

2. Modified moment matrices and their condition numbers

2.1. Background

Consider a polygon H in the complex plane whose boundary ∂H consists of m edges E_ν, $\nu=1,m$, which join the vertices $h_{\nu-1}$ and h_ν of H. Note that in [7] H is convex, and symmetric with respect to the real axis and is build with a particular procedure so that m is even. We will not make such restrictions here. We denote by c_ν the center of the ν-th edge and by d_ν its half-width, i.e.

$$c_\nu = \tfrac{1}{2}(h_\nu + h_{\nu-1}) \;,\quad d_\nu = \tfrac{1}{2}(h_\nu - h_{\nu-1}) \tag{2}$$

On each edge we define the generalized Chebyshev weight

$$\omega_\nu(\lambda) = \frac{2}{\pi}\,|d_\nu^{\,2} - (\lambda - c_\nu)^2|^{-1/2} \tag{3}$$

A weight function on the boundary of the polygon is then defined as the function whose restriction to each edge E_ν is $\omega_\nu(\lambda)$. An inner product on the space of complex polynomials $\mathbf{P}_n$ of degree not exceeding n is therefore defined by

$$<p,q> = \int_{\partial H} p(\lambda)\overline{q(\lambda)}\,\omega(\lambda)\,|d\lambda| \;=\; \sum_{\nu=1}^{m} \int_{E_\nu} p(\lambda)\overline{q(\lambda)}\,\omega_\nu(\lambda)\,|d\lambda| \tag{4}$$

Assume that the convex hull H can be enclosed in the ellipse $\mathcal{E}(c,d,a)$ having center c, focal distance d and major semi axis a. Consider the basis of shifted and scaled Chebyshev polynomials (1). We will assume that c is real while d is either real or purely imaginary. To the basis (1) is associated the (n+1)x(n+1) Gram matrix M_n whose elements $m_{i,j}$ are defined by

$$m_{i,j} = < t_j,\, t_i > \quad i,j=0,1,2..,n. \tag{5}$$

Note that M_n is a hermitian positive definite matrix. In case H is symmetric with respect to the real axis, then M_n is even *real* symmetric positive definite [9]. We will denote by $\tau(M_n)$ the spectral condition number of the matrix M_n, i.e. the ratio of its largest eigenvalue to its smallest one. The important question which we address here is to find an upper bound for the condition number $\tau(M_n)$.

The motivation for such a question is that any least squares problem using polynomials of degree $\leq n$, will require the Choleski factorization of M_n. This is the case when one computes the orthogonal polynomials with respect to $<.,.>$, or when one solves the least squares problem that arises from some acceleration techniques for solving linear systems [7, 8, 11, 12]. If the Gram matrix is highly ill-conditioned it is difficult to compute the least squares polynomials in the corresponding basis. This is to compare with near linear dependance of a system of vectors which is often measured by the ratio of the largest to the smallest singular values of the system, i.e. by the square root of the condition number of its Gram matrix.

2.2. Basic properties

We start by recalling some well known properties of the Chebyshev polynomials, see [6]. First notice that the mapping

$$z = (\lambda - c)/d \tag{6}$$

transforms the ellipse $\mathcal{E}(c,d,a)$ in the λ-plane into the ellipse $\mathcal{E}(0,1,\alpha)$ in the z-plane, where $\alpha = a/d$. Then the polynomials (1) can be written as

$$t_j(\lambda) = T_j(z)/T_j(\alpha), \tag{7}$$

where T_j denotes the Chebyshev polynomial of the first kind of degree j and where λ and z are related by (6). Furthermore, the mapping

$$z = (w + w^{-1})/2 \tag{8}$$

maps the circle $C(0,\rho)$ of center the origin and radius $\rho \geq 1$ in the w-plane into the ellipse $\mathcal{E}(0,1,\alpha)$ of the z-plane where α and ρ are related by

$$\alpha = (\rho + \rho^{-1})/2 \tag{9}$$

or, equivalently

$$\rho = \alpha + \sqrt{\alpha^2 - 1} \tag{10}$$

An important property is that the Chebyshev polynomial T_j can be expressed as

$$T_j(z) = (w^j + w^{-j})/2 \tag{11}$$

where w and z are related by the transform (8). As a consequence, when z belongs to the ellipse $\mathcal{E}(0,1,\alpha)$, we have $T_j(z) = (\rho^j e^{j\theta} + \rho^{-j} e^{-j\theta})/2$ where $0 \leq \theta < 2\pi$, so $T_j(z)$ reaches its maximum modulus when $\theta = 0$, i.e. when $z = \alpha$. By (7), this shows that the maximum modulus of the polynomials $t_j(\lambda)$ on the ellipse $\mathcal{E}(c,d,a)$ is one.

In addition to the inner product $<.,.>$ associated with the weight function ω defined in Section 2.1, we will sometimes also use the following inner product defined on $\mathcal{E}(0,1,\alpha)$

$$<p,q>_v = \int_{\mathcal{E}(0,1,\alpha)} p(z)\, \overline{q(z)}\, v(z)|dz|,$$

with

$$v(z) = \frac{1}{2\pi} |1 - z^2|^{-1/2}$$

to which we associate the norm $||.||_v$. It is easy to verify that the polynomials

$$T_j(z)/T_j(\alpha) \tag{12}$$

are orthogonal with respect to the above inner product [1] and that

$$||T_j(z)/T_j(\alpha)||_v^2 = 1-\epsilon_j \text{ with}$$

$$\epsilon_0 = 0 \text{ and } \epsilon_j = 2/[\rho^{2j}+\rho^{-2j}], \text{ for } j\neq 0.$$

In what follows we denote by (.,.) the hermitian inner product in $\mathbf{C}^{n+1}$, and by $||.||$ its induced norm. Next we recall two useful properties shown in [7], concerning the Gram matrices M_n.

Proposition 1: *Let*

$$p(\lambda) = \sum_{i=0}^{n} \eta_i t_i(\lambda) \qquad \text{and} \qquad q(\lambda) = \sum_{i=0}^{n} \theta_i t_i(\lambda)$$

be any two polynomials of degree not exceeding n. Then the following expression for the inner product <p,q> holds

$$< p , q > = (M_n\, \boldsymbol{\eta}, \boldsymbol{\theta}) \tag{13}$$

where $\boldsymbol{\eta} = (\eta_0, \eta_1,..\eta_n)^T$ *and* $\boldsymbol{\theta} = (\theta_0, \theta_1,..\theta_n)^T$.

The second property is a formula used in [7] for computing the matrix M_n.

Proposition 2: *Let each of the polynomial* $t_j(\lambda)$ *be expressed on each edge* E_ν *as*

$$t_j(\lambda) = \sum_{i=0}^{j} \gamma_{i,j}^{(\nu)} T_i(\xi) , \tag{14}$$

$$\text{where} \quad \xi = (\lambda - c_\nu) / d_\nu . \tag{15}$$

Then the coefficients of the modified moment matrix M_n *are given by*

$$m_{i,j} = \sum_{\nu=1}^{m} \sum_{k=0}^{i}{}' \gamma_{k,j}^{(\nu)} \overline{\gamma}_{k,i}^{(\nu)} \qquad i=0,1,...j. \tag{16}$$

where $\sum_{k=0}^{i}{}'$ *is defined by* $\sum_{k=0}^{i}{}' \alpha_k \equiv 2\, \alpha_0 + \sum_{k=1}^{i} \alpha_k$

2.3. An upper bound for $\tau(M_n)$

Our main result will be proved with the help of three lemmas. In all of this section we assume that the polygon H is enclosed in the ellipse $\mathcal{E}(c,d,a)$.

Lemma 3: *Let M_n be the modified moment matrix of size $(n+1) \times (n+1)$ and $x=(\xi_j)_{j=0,n}$ any complex vector of length n+1. Then*

$$(M_n x,x) \leq 2\ m\ (n+1)\ ||x||^2. \tag{17}$$

Proof: By (13), we have

$$(M_n x,x) = \int_{\partial H} |\ p(\lambda)\ |^2\ \omega(\lambda)\ |d\lambda| = ||\ p\ ||_\omega^2 \tag{18}$$

where $p(\lambda) = \sum_{j=0}^{n} \xi_j t_j(\lambda)$

By the Cauchy-Schwartz inequality

$$|p(\lambda)\ | = |\ \sum_{j=0}^{n} \xi_j t_j(\lambda)| \leq [\ \sum_{j=0}^{n} |\xi_j\ |^2\ ]^{1/2}\ [\ \sum_{j=0}^{n} |t_j(\lambda)|^2\ ]^{1/2}. \tag{19}$$

Remembering that the maximum modulus of each t_j on $\mathcal{E}(c,d,a)$ is one and since the convex hull H is inside the ellipse $\mathcal{E}(c,d,a)$, the maximum principle implies that each $t_j(\lambda)$ has modulus not exceeding one for λ belonging to ∂H. Hence from (19)

$$|p(\lambda)\ |^2 \leq\ (n+1)\ ||\ x\ ||^2\ , \text{ for } \lambda \in\ \partial H.$$

Integrating both sides according to (18) and observing that the ω-norm of the constant function one is 2m (one can use the expression of $m_{0,0}$ from Proposition 2 to show it) we obtain (17)□

Our next lemma for establishing the main result is the following.

Lemma 4: *Let $x=(\xi_j)_{j=0,n}$ be any nonzero vector of $\mathbf{C}^{n+1}$, $p(\lambda) = \sum_{j=0}^{n} \xi_j t_j(\lambda)$ and let $\kappa_n(p)$ be the constant defined by*

$$\max_{\lambda \in \mathcal{E}(c,d,a)} |p(\lambda)| = \kappa_n(p) \max_{\lambda \in \partial H} |p(\lambda)| \tag{20}$$

Then

$$||\ x\ ||^2 \leq (\kappa_n(p))^2\ (n+1)\ \ (M_n x,x) \tag{21}$$

Proof: The proof will be split into 4 steps.

1) Writing everything in terms of the z variable we have

$$p(\lambda) = \sum_{j=0}^{n} \xi_j\ T_j\ (z)/T_j\ (\alpha)$$

where z and λ are related by the linear mapping (6). We will denote by q the resulting polynomial in the variable z, i.e. we have $q(z) \equiv p(\lambda) \equiv p(d\ z+c)$. From Section 2.2 we have

$$||\ q(z)\ ||_v^{\ 2} = \sum_{j=0}^{n} \xi_j^{\ 2}\ (1-\epsilon_j) \tag{22}$$

Moreover,

$$\| q(z) \|_v^2 = \int_{\mathcal{E}(0,1,\alpha)} |q(z)|^2 v(z) |dz| \le \max_{\mathcal{E}(0,1,\alpha)} |q(z)|^2 \| \mathbf{1}(z) \|_v^2 ,$$

where $\mathbf{1}(z)$ represents the constant function unity. Since $\| \mathbf{1}(z) \|_v = \| T_0(z)/T_0(\alpha) \|_v = 1$, the above inequality yields

$$\| q \|_v^2 \le \max_{\mathcal{E}(0,1,\alpha)} |q(z)|^2 = \max_{\mathcal{E}(c,d,a)} |p(\lambda)|^2$$

Hence

$$\| q \|_v^2 \le (\kappa_n(p))^2 \max_{\partial H} |p(\lambda)|^2. \tag{23}$$

2) Consider

$$p(\lambda) = \sum_{j=0}^{n} \xi_j t_j(\lambda).$$

Using the expressions (14) this expression becomes

$$p(\lambda) = \sum_{j=0}^{n} \xi_j \sum_{i=0}^{j} \gamma_{i,j}^{(\nu)} T_i(\xi) = \sum_{i=0}^{n} [\sum_{j=i}^{n} \gamma_{i,j}^{(\nu)} \xi_j] \ T_i(\xi) , \ \text{for } \nu=1,2,..m, \tag{24}$$

or in condensed form

$$p(\lambda) = (G^{(\nu)}x, t(\xi)) , \qquad \nu=1,m ,$$

where $G^{(\nu)}$ is the upper triangular matrix whose nonzero entries are the $\gamma_{i,j}^{(\nu)}$ and $t(\xi)$ is the (n+1)-vector whose i^{th} component is $T_i(\xi)$, i=0,n. We then have

$$|p(\lambda)|^2 \le \| G^{(\nu)} x \|^2 \| t(\xi) \|^2 \le (n+1) \| G^{(\nu)} x \|^2.$$

Thus, for $\lambda \in \partial H$,

$$|p(\lambda)|^2 \le (n+1) \max_{\nu=1,m} \| G^{(\nu)} x \|^2. \tag{25}$$

3) By using Propositions 1 and 2 we get

$$\| p \|_\omega^2 = \sum_{\nu=1}^{m} \sum_{i=0}^{n}{}' | \sum_{j=i}^{n} \gamma_{i,j}^{(\nu)} \xi_j |^2$$

where Σ' is as defined in Proposition 2. Therefore,

$$\| p \|_\omega^2 \ge \sum_{\nu=1}^{m} \| G^{(\nu)} x \|^2 \tag{26}$$

4) Using (22), (23), (25), (26) and (18) in this order we obtain

$$\|x\|^2 \le \| q(z) \|_v^2 \le (\kappa_n(p))^2 \max_{\partial H} |p(\lambda)|^2 \le (n+1) (\kappa_n(p))^2 \max_{\nu=1,m} \| G^{(\nu)} x \|^2$$
$$\le (n+1) (\kappa_n(p))^2 \sum_{\nu=1}^{m} \| G^{(\nu)} x \|^2 \le (n+1) (\kappa_n(p))^2 \| p \|_\omega^2 = (n+1) (\kappa_n(p))^2 (Mx,x) \ \square$$

In the above lemma, $\kappa_n(p)$ depends on the polynomial p. Let us define κ_n as the maximum of all $\kappa_n(p)$ for p belonging to the space of polynomials of degree $\leq n$. As will be seen shortly κ_n is not infinite. With this definition we have the following corollary.

Corollary 5: The condition number of the modified moment matrix M_n satisfies the inequality

$$\tau(M_n) \leq 2\ m\ (n+1)^2\ \kappa_n^2$$

Proof: Lemma 3 yields an upper bound for the largest eigenvalue of M_n, while Lemma 4 yields a lower bound when we replace $\kappa_n(p)$ by κ_n. Taking the ratio gives the above upper bound for $\tau(M_n)$□

Thus, we have transformed the initial problem into that of finding an upper bound for κ_n. It is to be expected that κ_n will increase to infinity as n goes to infinity but our goal is only to show that it does not increase too fast in certain favorable circumstances. We now prove our third and final lemma.

Lemma 6: *Assume that the polygonal boundary ∂H can be inserted between an inner ellipse $\mathcal{E}(c,d_I,a_I)$ and an outer ellipse $\mathcal{E}(c,d,a)$, with $d_I \leq d$. Then we have*

$$\kappa_n \leq \left(\frac{a+\sqrt{a^2-d_I^2}}{a_I+\sqrt{a_I^2-d_I^2}} \right)^n \tag{27}$$

Proof:

$$[\kappa_n(p)]^{-1} = [\max_{\partial H} |p(\lambda)|\ ] \ / \ [\max_{\mathcal{E}(c,d,a)} |p(\lambda)|\ ]$$

By the maximum principle we have

$$[\kappa_n(p)]^{-1} \geq [\max_{\mathcal{E}(c,d_I,a_I)} |p(\lambda)|\ ] \ / \ [\max_{\mathcal{E}(c,d,a)} |p(\lambda)|\ ]$$

We will now work with the z variable *associated with the inner ellipse,* i.e.

$$z=(\lambda-c)/d_I.$$

which maps the ellipses $\mathcal{E}(c,d_I,a_I)$ and $\mathcal{E}(c,d,a)$ into the ellipses $\mathcal{E}(0,1,a_I/d_I)$ and $\mathcal{E}(0,d/d_I,a/d_I)$ respectively. We will set $\alpha_I=a_I/d_I$. As before we will denote by q(z) the polynomial in the variable z which results from this mapping, i.e. we set $q(z)\equiv p(\lambda)\equiv p(d_I z+c)$.

Let z_* be the point of $\mathcal{E}(0,1,\alpha_I)$ where the maximum of $|q(z)|$ is reached. Then

$$[\kappa_n(p)]^{-1} \geq \max_{\mathcal{E}(0,1,\alpha_I)} |q(z)|/|q(z_*)|$$

$$\geq \min_{q \in \mathbf{P}_n} \ \max_{z \in \mathcal{E}(0,1,\alpha_I)} |q(z)/q(z_*)|$$

In which $\mathbf{P}_n$ represents the space of polynonials of degree $\leq$n. Let q(z) be expressed in the basis of Chebyshev polynomials as

$$q(z) = \sum_{j=0}^{n} \eta_j T_j(z)$$

and denote by η the vector with components η_i, i=0,n. Using (11) we find that

$$[\kappa_n(p)]^{-1} \geq \min_{\eta} \max_{w \in C(0,\rho_I)} [\sum_{j=0}^{n} \eta_j (w^j+w^{-j})/2] / [\sum_{j=0}^{n} \eta_j (w_*^j+w_*^{-j})/2]$$

where w_* and z_* are related by the transform (8) and ρ_I and α_I by the relation (10). We can rewrite the above inequality as

$$\begin{aligned}[\kappa_n(p)]^{-1} &\geq \min_{\eta} \max_{w \in C(0,\rho_I)} |w/w_*|^{-n} [\sum_{j=0}^{n} \eta_j (w^{n+j}+w^{n-j})/2] / [\sum_{j=0}^{n} \eta_j (w_*^{n+j}+w_*^{n-j})/2] \\ &= [\rho_I / |w_*|]^{-n} \min_{\eta} \max_{w \in C(0,\rho_I)} [\sum_{j=0}^{n} \eta_j (w^{n+j}+w^{n-j})] / [\sum_{j=0}^{n} \eta_j (w_*^{n+j}+w_*^{n-j})] \\ &\geq [\rho_I / |w_*|]^{-n} \min_{Q \in \mathbf{P}_{2n}} \max_{w \in C(0,\rho_I)} |Q(w)/Q(w_*)|\end{aligned}$$

It is known that among all polynomials of degree 2n, such that $Q(w_*)=1$, the one that minimizes the infinity norm on the circle is given by $(w/w_*)^{2n}$, see [4]. By replacing we finally get that

$$[\kappa_n(p)]^{-1} \geq \left(\frac{\rho_I}{|w_*|} \right)^n$$

In order to lower bound the term of the right hand side, we must now find the maximum of $|w_*|$ when z_* belongs to the outer ellipse $\mathcal{E}(0,d/d_I,a/d_I)$. Note that $|w_*|$ is the radius of that circle which is mapped by (8) into the ellipse of center the origin and focal distance 1, passing through the point z_*. For a larger radius $|w_*|$, the ellipse is larger. The largest possible $|w_*|$ corresponds to the outmost confocal ellipse that intersects the outer ellipse $\mathcal{E}(0,d/d_I,a/d_I)$. For $d_I \leq d$, this is realized when z_* is on the major principal axis, i.e. when $z_*=a/d_I$. Hence the result by using (10)□

As a result of corollary 5 and the above lemma we have the following final theorem.

Theorem 7: *Assume that the polygonal boundary ∂H can be inserted between an inner ellipse $\mathcal{E}(c,d_I,a_I)$ and an outer ellipse $\mathcal{E}(c,d,a)$, with $d_I \leq d$. Then the condition number of the modified moment matrix M_n satisfies the inequality*

$$\tau(M_n) \leq 2\, m\, (n+1)^2 \left(\frac{a+\sqrt{a^2-d_I^2}}{a_I+\sqrt{a_I^2-d_I^2}} \right)^{2n}$$

2.4. Example

To illustrate the above theorem, let us assume that H is simply a rectangle centered at c and having half-length L along the real axis and half-width l. It is most natural to take as inner ellipse the ellipse centered at c with major semi axis a_I=L and minor semi axis l. This ellipse has focal distance $d_I=(L^2-l^2)^{1/2}$. To simplify, consider as outer ellipse the ellipse centered at c, having focal distance d_I and passing through the points $(c\pm L)\pm il$. We find the major semi axis of this ellipse by using the following relation for an ellipse passing through $c+L+il$ and $c-L+il$

$$d2 = a2 \left[1 - \frac{l^2}{a2-L^2} \right]$$

where $d2=d_I^2$, $a2=a^2$, see [5]. Solving for a2 we find that $a2= L(L+l)$ or $a=[L(L+l)]^{1/2}$. This gives the growth ratio

$$\kappa = \frac{\sqrt{L}+\sqrt{l}}{\sqrt{L+l}}. \tag{28}$$

When $l<<L$ this ratio is close to 1. For example if l=0.01 and L=1, then κ=1.0945. In this case the bound of the theorem yields $\tau(M_n)\leq 2.9\text{x}10^3$ for n=10 and $\tau(M_n)\leq 6.5\text{x}10^4$ for n=20. When l=0.1 and L=1, then κ=1.2549 and the bound of the theorem yields $\tau(M_n)\leq 9.1\text{x}10^4$ for n=10 and $\tau(M_n)\leq 3.1\text{x}10^7$ for n=20. As a final example, let us take L=l which is in some sense the worst case e example. Then the two ellipses become circles and as is easily seen the basis (1) reduces to the basis of scaled monic polynomials $(\lambda/R)^j$ where $R=\sqrt{2L}$ is the radius of the enclosing circle. From (28) we get $\kappa=\sqrt{2}$ and Theorem 7 yields $\tau(M_n)\leq 8(n+1)^2 2^n$. Thus $\tau(M_{10})\leq 9.9\text{x}10^5$ and $\tau(M_{20})\leq 3.7\text{x}10^9$. Note that these are only upper bounds and one can expect the actual condition numbers to be much smaller in general.

On the practical side, badly conditioned Gram matrices can easily be detected by computing the Choleski factorization in a progressive way, i.e. by updating it as each column appears. By doing so, one can stop just before the condition number becomes too large, and thus work with the largest 'numerically allowable' degree n. In most cases, the Chebyshev basis enables us to compute least-squares polynomials of orders as high as 50 with no major difficulty. However, as n increases, another limitation which appears is computational cost since the number of required arithmetic operations grows like $O(m\ n^3)$.

References

[1] P.H. Davis. *Interpolation and Approximation.* Blaisdell, Waltham, Mass., 1963.

[2] W. Gautschi. On generating orthogonal polynomials. *SIAM J. Sci. and Stat. Computing* 3:289-317, 1982.

[3] W. Gautschi. Construction of the Gauss-Christoffel quadrature formula. *Math. Comp.* 22:251-270, 1968.

[4] G.G. Lorentz. *Approximation of functions.* Holt, Rinehart & Winston, New York, 1966.

[5] T.A. Manteuffel. *An iterative method for solving nonsymmetric linear systems with dynamic estimation of parameters.* Technical Report UIUCDCS-75-758, University of Illinois at Urbana-Champaign, 1975. Ph.D. dissertation.

[6] T.J. Rivlin. *The Chebyshev Polynomials.* J.Wiley and Sons Inc., New York, 1976.

[7] Y. Saad. Least squares polynomials in the complex plane and their use for solving sparse nonsymmetric linear systems. *SIAM J. Stat. Sci. Comput.* : , . submitted.

[8] Y. Saad. *Least squares polynomials in the complex plane with applications to solving sparse nonsymmetric matrix problems.* Technical Report 276, Yale University, 1983.

[9] Y. Saad. *Practical use of polynomial preconditionings for the Conjugate Gradient Method.* Technical Report YALEU/DCS/RR-282, Yale University, 1983. to appear in SSISC.

[10] R.A. Sack, A.F. Donovan. An algorithm for Gaussian quadrature given modified moments. *Numer. Math.* 18:465-478, 1971.

[11] D. C. Smolarski. *Optimum semi-iterative methods for the solution of any linear algebraic system with a square matrix.* Technical Report UIUCDCS-R-81-1077, University of Illinois at Urbana-Champaign, 1981. PhD Thesis.

[12] D.C. Smolarski and P.E. Saylor. *Optimum paramaters for the solution of linear equations by Richardson' iteration.* 1982. Unpublished paper.