[25]

ALGEBRAICALLY SOLVABLE CHEBYSHEV APPROXIMATION PROBLEMS

Martin H. Gutknecht

Best rational Chebyshev approximations that can be computed algebraically, i.e. by solving an eigenvalue problem, have been a matter of interest for a long time [1,6,8-10]. On the basis of the Carathéodory-Fejér method we derive here a general and unified theory for such approximations on a disk (complex) and on an interval (real), respectively; the latter case extends to real rational trigonometric approximations. The given function must be a rational function of higher numerator or denominator degree. To our knowledge our theory covers all known cases of (nontrivially) algebraically computable best approximations and even most of those classes of Chebyshev approximations where the solution is known explicitly. (A historical survey will be published elsewhere.)

1. CF approximation: the special case of the smallest singular value

Let R_{mn} denote the space of complex rational functions with numerator degree at most m and denominator degree at most n, let $P_m := R_{m0}$, and let $R_{mn}(\overline{D})$ be the subset of those rational functions that have no poles on the closed unit disk $\overline{D}$. Moreover, let $\tilde{R}_{mn} \supset R_{mn}(\overline{D})$ consist of the functions of the form $\tilde{r} = \tilde{p}/q$ with $q \in P_n$ nonvanishing on $\overline{D}$ and $z \mapsto z^{-m}\tilde{p}(z)$ analytic and bounded in $1 < |z| \leq \infty$. Finally, $\|\cdot\|_S$ is the Chebyshev norm (essential supremum norm) on the unit circle S.

In [4] we proved results which on one hand enhance the Adamjan-Arov-Krein theory [2] in the finite rank case and on the other hand extend algorithms proposed by Kung [5] and Trefethen [11] (who introduced the name "CF approximation" and studied its deviation from the best approximation). Here we state part of these results in the special case where the weight function is the reciprocal of a polynomial and where

ISBN 0-12-174580-5

only the smallest singular value is considered:

THEOREM 1. *Let* $f \in \tilde{R}_{MN}$, $g \in P_L$, g *nonvanishing in* $\{z;\ 1 \leq |z| < \infty\} \cup \{0\}$, *and let either* $m = M - 1$, $n \geq N - 1$ *or* $m \geq M - 1$, $n = N - 1$. $(L, M, N \geq 0$ *are assumed to be the actual degrees.)*

Let $b \in P_N$ *(with all zeros outside* S*) denote the 'denominator' of* f*, so that* $a := fb \in \tilde{R}_{M0}$*, and let* $b^*(z) := z^N \overline{b}(1/z)$ *be the reciprocal polynomial of* b*. Denote the Fourier coefficients of* $t \mapsto (f/g)(e^{it})$ *and of* $t \mapsto (b^*/b)(e^{it})$ *by* ϕ_k *and* β_k*, respectively. Define the* $(n+1) \times (n+1)$ *Toeplitz matrices* $\underset{\sim}{T}(f/g;\ \ell) := (\phi_{\ell+j-k})_{j,k=0}^{n}$ *and* $\underset{\sim}{T}(b^*/b;\ \ell) := (\beta_{\ell+j-k})_{j,k=0}^{n}$ $(\ell \in \mathbb{Z})$*; let* $\underset{\sim}{R} := (\delta_{j+k,n})_{j,k=0}^{n}$ *(with Kronecker's delta) be the Hankel matrix with ones on the anti-diagonal and zeros elsewhere. Consider the smallest singular value* σ *and any corresponding singular vector* $\underset{\sim}{v} = (v_0, \ldots, v_n)^T$ *satisfying*

$$(1) \quad \underset{\sim}{T}\left(\frac{f}{g};\ m+1-L\right) \underset{\sim}{v} = \sigma\, \underset{\sim}{T}\left(\frac{b^*}{b};\ m+1-L\right) \underset{\sim}{R}\, \overline{\underset{\sim}{v}}.$$

Then the unique best weighted approximation $\tilde{r}^*$ *of* f *from* $\tilde{R}_{mn}$ *for which* $||(f-r)/g||_S$ *becomes minimal among all* $\tilde{r} \in \tilde{R}_{mn}$ *is given by*

$$(2) \quad \tilde{r}^*(z) := f(z) - \sigma\, z^{J-L}\, \frac{b^*(z)}{b(z)}\, \frac{v^*(z)}{v(z)}\, g(z) ,$$

where $J := \max\{M-N,\ m-n\} = m - N + 1$, $v(z) := v_0 + \ldots + v_n z^n$, $v^*(z) = \overline{v}_n + \ldots + \overline{v}_0 z^n$. *By cancelling common factors and extracting powers of* z *the quotient* v^*/v *may be written*

$$(3) \quad \frac{v^*(z)}{v(z)} = z^{\gamma}\, \frac{q^*(z)}{q(z)}$$

with q *and* q^* *being mutually prime reciprocal polynomials of exact degree* $n' := n - \gamma - \mu$*, where* $\mu + 1$ *is the multiplicity of* σ *as a singular value of (1) and where* $\gamma \geq 0$ *if* $n = N - 1$*,* $\gamma = 0$ *otherwise.* q *has all its* n' *zeros outside* S *and is the* 'denominator' of $\tilde{r}^*$; $\tilde{p} := q\tilde{r}^* \in \tilde{R}_{m'0} \setminus \tilde{R}_{m'-1,0}$, $m' := m - \delta - \mu$ with $\delta \geq 0$ if $m = M - 1$, $\delta = 0$ otherwise, but always $\gamma\delta = 0$. Hence, $\tilde{r}^* \in \tilde{R}_{m',n'}$; moreover, $\tilde{r}^*$ is best out of $\tilde{R}_{m'+\nu,n'+\nu}$, where $\nu := \gamma + \delta + \mu$.

The special properties of the two cases $m = M - 1$, $n \geq N - 1$ and $m \geq M - 1$, $n = N - 1$, respectively, considered here are:

(i) It is always the smallest singular value that matters.

(ii) v^*/v has no poles on $\overline{D}$; therefore $\tilde{r}^* \in \tilde{R}_{m',n'}$ is analytic on $\overline{D}$, and hence $\tilde{r}^* \in \tilde{R}_{m',n'}(\overline{D})$ whenever $f \in R_{MN}(\overline{D})$ and $J - L + \gamma \geq 0$.

(iii) $\underset{\sim}{v} \in \mathbb{C}^{n+1}$ and the indices of the Fourier coefficients appearing in the Toeplitz matrices in (1) are particularly simple.

We might mention that in the nondegenerate case, where $\gamma = \delta = \mu = 0$ and $v = q$, the singular value problem (1) arises from comparing the coefficients of z^{m+1-L} through $z^{m+n+1-L}$ in the Laurent series with respect to S (the coefficients are well defined even if the series do not converge on S) of both sides of

$$(4) \quad \frac{\tilde{p}(z)}{g(z)} = \frac{f(z)}{g(z)} v(z) - \sigma z^{J-L} \frac{b^*(z)}{b(z)} v^*(z).$$

It can be shown that the coefficients of all higher powers are matched automatically as a consequence of (1).

A singular value problem equivalent to (1) may be obtained by comparing coefficients after multiplying (4) by $g(z)$. If both f and g have real coefficients both this singular value problem and (1) become generalized eigenvalue problems.

2. Complex rational approximation on a disk

The following is now an immediate consequence of Thm. 1, in particular of property (ii) mentioned above:

THEOREM 2. *In addition to the assumptions of Thm. 1 suppose that* $f \in R_{MN}(\bar{D})$ *and that either* $M \geq N$ *and* $L \leq m - N + \gamma + 1$ *or* $M < N = n + 1$, $m \geq n - \gamma$, *and* $L \leq m - n + \gamma$ $(= m-N+\gamma+1)$.

Then $\tilde{r}^*$ *constructed in Thm. 1 lies in* $R_{m',n'}(\bar{D})$, *and* $\tilde{r}^*$ *is the best approximation on* S *of* f *from* $R_{mn}(\bar{D})$ *with respect to the weight function* $1/g$. *If* $\tilde{r}^*$ *is degenerate, i.e. if* $m' < m$ *or* $n' < n$, *then* $\tilde{r}^*$ *is also best out of* $R_{m'+\ell,n'+\ell}(\bar{D})$ *for every* $\ell \leq \nu := \gamma + \delta + \mu$ *(defined in Thm. 1).*

Proof. We only have to make sure that $J - L + \gamma \geq 0$, where $J := \max\{M-N, m-n\} = m - N + 1$. If $M < N$ and $\gamma = 0$, we need $m \geq n$; so $n \geq N$ (which implies $\gamma = 0$, cf. Thm. 1) is impossible if $M < N$; hence $m \geq M - 1$, $n = N - 1$ in this case, and the condition reduces to $L \leq m - n + \gamma$. □

A not quite correct version of this theorem was given in [4, Thm. 5.2], where we missed the implications of the condition $L \leq J + \gamma$, since our derivation was based on the normalized case $M = N$, $L = 0$, where the condition is fulfilled.

Notice that $\tilde{r}^*$ *is also the best approximation of* f *from* R_{mn} *with respect to the weighted norm* $||(f-r)/g^*||_D$ on the *disk* D (or on $\bar{D}$) since for all $r \in R_{mn}$ with finite error, $(f-r)/g^*$ is analytic on $\bar{D}$, so that the maximum error is taken on S, where $|g(z)| = |g^*(z)|$.

3. Real rational approximation on an interval and real rational trigonometric approximation

For $m, n \geq 0$ let

$$T_m^R := \{z \mapsto \bar{a}_m z^{-m} + \ldots + \bar{a}_1 z^{-1} + a_0 + a_1 z + \ldots + a_m z^m ;\ a_0 \in \mathbb{R},\ a_1,\ldots,a_m \in \mathbb{C}\},$$

$$T_{mn}^R := \{\hat{p}/\hat{q} ;\ \hat{p} \in T_m^R,\ \hat{q} \in T_n^R,\ \hat{q}(z) > 0 \text{ on } S\},$$

and let T_m^{RS}, T_m^{RS} denote the corresponding subsets of functions $\hat{r}$ satisfying $\hat{r}(z) = \overline{\hat{r}}(1/z)$. (Hence, e.g., $\hat{p} \in T_m^{RS}$ has real coefficients.) Upon the substitution $z = \exp(it)$ approximation on S from T_{mn}^R is seen to be equivalent to real rational trigonometric approximation on $[0,2\pi]$, and upon the substitution $\xi = \cos t = (z+1/z)/2$ approximation on S from T_{mn}^{RS} is seen to be equivalent to (ordinary) real rational approximation on the ξ-interval $I := [-1,1]$.

The Chebyshev-CF method for approximation from T_{mn}^{RS} and the Fourier-CF method for approximation from T_{mn}^R have been defined in [3,4,12]. In analogy to §2 these methods allow to compute all known cases of algebraically computable best approximations on an interval, including those presented by Achieser [1], Mirakyan [8], Talbot [9,10], and Lam [6].

Both for the denominator of the given function and for the weight function we make use of the face that $\hat{b} \in T_N^R$ satisfying $\hat{b}(z) > 0$ on S can be factored in the form $\hat{b}(z) = b(z)\bar{b}(1/z)$ with $b \in P_N$, $b(z) \neq 0$ on $\bar{D}$.

THEOREM 3. Let $f \in T_{MN}^R$, $\hat{g} \in T_L^R$, $\hat{g}(z) > 0$ on S, and assume that either $m = M - 1$, $n \geq N - 1$ or $m \geq M - 1$, $n = N - 1$. ($L, M, N \geq 0$ are assumed to be the actual degrees.)

f and $\hat{g}$ can be factored in the form

$$(5)\quad f(z) = \frac{\hat{a}(z)}{b(z)\bar{b}(1/z)},\quad \hat{g}(z) = g(z)\bar{g}(1/z),$$

respectively, with $\hat{a} \in T_M^R$, $b \in P_N$, $b(z) \neq 0$ on $\bar{D}$, $g \in P_L$, $g(z) \neq 0$ on $\{0\} \cup \mathbb{C}\backslash D$. Now define $\tilde{r}^*$ by (2) as the solution of the approximation problem (for f and g) of Thm. 1. (Note that $a(z) := \hat{a}(z)/\overline{b}(1/z)$.) Finally, let $\tilde{e} := f - \tilde{r}^*$,

(6) $\hat{r}(z) := f(z) - \tilde{e}(z) - \tilde{e}(1/z)$.

In the notation of Thm. 1 define

$$m'' := \max\{m',\ N + n' - m - \gamma - 1 + L\},$$

and assume that $L \leq 2m - N - n' + \gamma + 1$ holds, so that $m'' \leq m$.

Then $\hat{r} \in T^R_{m'',n'}$, and $\hat{r}$ is the best weighted approximation of f from $T^R_{m''+\ell,n'+\ell}$ for every $\ell \leq \nu + m' - m'' = m - m'' + \gamma$ in the sense that $||(f-\hat{r})/\sqrt{\hat{g}}||_S$ is minimal.

If, in addition, $f \in T^{RS}_{MN}$ and $\hat{g} \in T^{RS}_L$, then $\hat{r} \in T^{RS}_{m'',n'}$, and $\hat{r}$ is optimal in $T^{RS}_{m''+\ell,n'+\ell}$.

Remarks. (i) In the degenerate case we do not claim that $\hat{r}$ is best out of T^R_{mn}. However, the latter is true if $n - n' \leq m - m''$ and $m = M - 1$ or $m'' = m'$ and $n = N - 1$.

(ii) Note that the weight function is $1/\sqrt{\hat{g}}$.

Proof. We substitute $\tilde{e} = f - \tilde{r}^*$ given by (2) and (3) into (6) and multiply through by $q(z)\overline{q}(1/z)$:

$$\begin{aligned}\hat{r}(z)q(z)\overline{q}(1/z) &= f(z)q(z)\overline{q}(1/z) \\ &\quad - \sigma z^{J-L+\gamma}\frac{b^*(z)}{b(z)}q^*(z)\overline{q}(1/z)g(z) \qquad (7)\\ &\quad - \sigma z^{-J+L-\gamma}\frac{b^*(1/z)}{\overline{b}(1/z)}\overline{q^*}(1/z)q(z)\overline{g}(1/z).\end{aligned}$$

According to the definition of $\tilde{r}^*$ as the solution of the approximation problem for f and g the sum of the first two terms on the right-hand side of (7) equals $\tilde{r}^*(z)q(z)\overline{q}(1/z) = \tilde{p}(z)\overline{q}(1/z)$, which is analytic in $1 \leq |z| < \infty$ and of order $O(z^{m'})$ as $z \to \infty$. The third term is also analytic there (all zeros of $\overline{b}(1/z)$ lie in D) and is of order $O(z^{-J+L-\gamma+n'})$ as $z \to \infty$. Now $\hat{p}(z) := \hat{r}(z)\overline{q}(z)q(1/z)$ takes conjugate values at z and $1/\overline{z}$, hence it is also analytic in $\overline{D}\setminus\{0\}$ and grows as fast for $z \to 0$ as for $z \to \infty$. In fact $\hat{p} \in T_{m"}$ under our assumptions on m, n, M, and N.

To prove the optimality we note that the complex error curve $\tilde{e}(S)$ has winding number

$$j := J+N+n'+\gamma = m+n'+\gamma+1 = m+n-\mu+1$$

(cf. (2),(3)) and that $\tilde{e}/|g| = e/\sqrt{\hat{g}}$ has constant modulus σ on S, so that

$$(f-\hat{r})/\sqrt{\hat{g}} = 2\,\mathrm{Re}\{\tilde{e}\}/\sqrt{\hat{g}} = 2\,\mathrm{Re}\{\tilde{e}/|g|\}$$

has a set of alternation points of length 2j, which is known to be characteristic for optimality in $T^R_{m"+\ell,n'+\ell}$ [7].

References

1. Achieser, N. I., Ueber ein Tschebyscheffsches Extremumproblem. Math. Ann. 104, 739-744 (1931).

2. Adamjan, V. M., Arov, D. Z., Krein, M. G., Analytic properties of Schmidt pairs for a Hankel operator and the generalized Schur-Takagi problem. Math. USSR Sbornik 15, 31-73 (1971).

3. Gutknecht, M. H., Rational Carathéodory-Féjer approximation on a disk, a circle, and an interval. J. Approx. Theory (in print).

4. Gutknecht, M. H., On complex rational approximation, Part 2. Proceedings NATO-ASI on Computational Aspects of Complex Analysis, Reidel Publ. Co., Dordrecht (in print).

5. Kung, S., Optimal Hankel-norm model reduction--scalar systems. Proceedings 1980 Joint Automatic Control Conference San Francisco, IEEE, New York, 1980.

6. Lam, B., A note on best uniform rational approximation. SIAM J. Numer. Anal. 13, 962-965 (1976).

7. Loeb, H. L., Approximation by generalized rationals. J. SIAM Numer. Anal. 3, 34-55 (1966).

8. Mirakyan, G., Sur une nouvelle fonction qui s'écarte le moins possible de zéro. Comm. Soc. Math. Kharkof, sér. 4, t. 12, 41-48 (1935).

9. Talbot, A., On a class of Tchebysheffian approximation problems solvable algebraically. Proc. Cambridge Philos. Soc. 58, 244-267 (1962).

10. Talbot, A., The Tchebysheffian approximation problem of one rational function by another. Proc. Cambridge Philos. Soc. 60, 877-890 (1964).

11. Trefethen, L. N., Rational Chebyshev approximation on the unit disk. Numer. Math. 37, 297-320 (1981).

12. Trefethen, L. N., Gutknecht, M. H., The Carathéodory-Fejér method for real rational approximation. SIAM J. Numer. Anal. 20, No. 2 (1983) (in print).

Martin H. Gutknecht
Seminar für angewandte Mathematik
Eidg. Technische Hochschule
ETH-Zentrum
CH-8092 Zürich
Switzerland