Global optimization: a model problem

Siegfried M. Rump *

April 28, 2009

Let $P, Q \in \mathbb{R}[x]$ be real polynomials. For $P(x) = \sum_{\nu=0}^{n-1} p_\nu x^\nu$ define

$$||P|| := ||p||_2 = (\sum_{\nu=0}^{n-1} p_\nu^2)^{1/2}.$$

For given $2 \leq n \in \mathbb{N}$ solve

$$||PQ||^2 = \min! \quad \text{subject to} \quad ||P||^2 = ||Q||^2 = 1, \tag{1}$$

where PQ denotes the polynomial multiplication (convolution). For given n this is an optimization problem in *2n* unknowns with polynomials P, Q of degree $n-1$.

Denote the global minimum by μ_n. We have a very efficient algorithm to compute an upper bound for μ_n, and we strongly believe that our approximations are very close to the true minimum. The challenge is to compute rigorous *lower bounds* for μ_n. The best known lower and upper bounds for μ_n are displayed in Table 1.

As has been mentioned, we have reasons to believe that the displayed upper bounds are very close to the global minimum.

By mathematical means we can show (see below) that global minimizers P, Q must have all roots on the unit circle. More precisely, for arbitrary given normed P a (normed) polynomial Q minimizing $||PQ||$ has all its roots on the unit circle.

*Institute for Reliable Computing, Hamburg University of Technology, Schwarzenbergstraße 95, Hamburg 21071, Germany, and Visiting Professor at Waseda University, Faculty of Science and Engineering, 3–4–1 Okubo, Shinjuku-ku, Tokyo 169–8555, Japan (`rump@tu-harburg.de`).

Table 1:

n	lower bound	upper bound
2	0.5	0.5
3	1/9	1/9
4	0.01742917332	0.01742917332143265289
5	0.002339595548	0.00233959554815559113
6	0.0002897318752	0.00028973187527968193
7	0.0000341850698	0.00003418506980008285
8	0.0000039054356	0.00000390543564975573
9	4.360016239e-07	0.43600165391810484613e-06
10	4.783939568e-08	0.47839395687709759327e-07
11	5.17870e-09	0.51787490974469905331e-08
12	5.54539e-10	0.55458818311631347612e-09
13	5.881019273e-11	0.58866880811866093130e-10
14	6.1e-12	0.62024449920539050220e-11
15	6.0e-13	0.64943654185809512880e-12
16	6.0e-14	0.67636042558221379058e-13
17	6.0e-15	0.70112631970143741585e-14
18	?	0.72383944796943875862e-15
19	?	0.74460039557776135597e-16
20	?	0.76350536937729609919e-17
21	?	0.78064642460112629708e-18
22	?	0.79611166679747050564e-19
23	?	0.80998543482245086492e-20
24	?	0.82234846891461875204e-21
25	?	0.83327806667912091279e-22

It follows that for polynomials P, Q realizing μ_n that the coefficient vectors must be symmetric or skewsymmetric, i.e. $p_{n-1-\nu} = \pm p_\nu$ and similarly for Q. Thus the problem can be rewritten into three optimization problems with about half the number of unknowns, where the smallest of the three minima is equal to μ_n.

The above lower bounds are taken from a paper by Kaltofen et al. [5] improving the results in [4]. They also mention a general lower bound

$$\mu_n \geq \begin{pmatrix} 2n-2 \\ n-1 \end{pmatrix}^{-2}$$

which is obtained by using a factor coefficient bound by Mignotte. Other optimization packages usually find a verified lower bound up to dimension $n \leq 4$. An exception is Martin Berz's COSY package (thanks to Kyoko Makino), which could compute bounds up to $n \leq 8$ for the modified problem.

Quite a number of interesting results on this optimization problem can be found in [3].

For convenience, we give AMPL-like problem formulations up to $n = 11$ for the original problem and for the modified problem, as well as local minimizers which we believe are close to the global ones. Those are computed by a Matlab routine. For larger values of n the upper bounds in Table 1 are computed with higher precision in a few seconds.

Mathematical background

Let a linear system $Ax{=}b$ with Toeplitz matrix A be given. A specialized solver needs only the first row and column of A as input. Thus general perturbations are not possible, only Toeplitz perturbations. Accordingly, the sensitivity of the system should be judged by the Toeplitz condition number, not by the general condition number. The question arises how small the ratio between the structured and unstructured condition number can be.

In [6] we proved that this ratio is bounded below by

$$\frac{\kappa^{\mathrm{Toep}}(A,x)}{\kappa(A,x)} = \alpha \frac{\|A^{-1} J \Psi_x\|}{\|A^{-1}\|\,\|x\|} \geq \frac{1}{\sqrt{n}} \frac{\sigma_{\min}(\Psi_x)}{\|x\|},$$

where the matrix Ψ_x is defined by

$$\Psi_x := \begin{pmatrix} x_1 & x_2 & \dots & x_n & & & \\ & x_1 & x_2 & \dots & x_n & & \\ & & & \dots & & & \\ & & x_1 & x_2 & \dots & x_n & \end{pmatrix} \in \mathbb{R}^{n\times(2n-1)},$$

$J \in \mathbb{R}^{n\times n}$ is the permutation matrix mapping $(1,\dots,n)^T$ into $(n,\dots,1)^T$ and $\frac{1}{\sqrt{n}} \le \alpha \le \sqrt{2}$.

Surprisingly, the lower bound depends only on the solution x, not on the matrix A. However, no closed formula for the minimum ratio was known, and it was also not clear for some time whether actually Toeplitz matrices exist realizing anything close to the lower bound. This was solved in [7]. We proved that

$$\sqrt{\mu_n} = \min_{\|x\|=1} \sigma_{\min}(\Psi_x) > 0$$

for all n, and that the minimum ratio of the condition numbers satisfies

$$\sqrt{2\mu_n} \ge \inf\{\frac{\kappa^{\text{Toep}}(A,x)}{\kappa(A,x)} \ : \ A \in \mathbb{R}^{n\times n} \text{ Toeplitz}, \ 0 \ne x \in \mathbb{R}^n\} \ge \sqrt{\frac{\mu_n}{n}}$$

for all n, where μ_n is the solution of the optimization problem (1). We mention that when replacing the norm in (1) by the maximum modulus of the polynomial on the unit circle the global minimum of (1) is explicitly known [1], [2].

From this interpretation the algorithm to approximate μ_n becomes clear. Denote for any $0 \ne x \in \mathbb{R}^n$ the left singular vector of Ψ_x to the smallest singular value $\sigma_{min}(\Psi_x)$ by y. Then the entries of

$$y^T\Psi_x$$

are the coefficients of the polynomial $x(t)y(t)$, where $x(t)$ and $y(t)$ are the polynomials with coefficients x and y, respectively. By construction we have $||y^T\Psi_x|| = \sigma_{min}(\Psi_x)$. Since polynomial multiplication is commutative it follows

$$y^T\Psi_x = x^T\Psi_y \quad \text{and} \quad \sigma_{min}(\Psi_y) \le ||x^T\Psi_y||.$$

Replacing x by y we calculate the left singular vector y of the new Ψ_x and so forth generating a monotonically decreasing sequence of upper bounds for μ_n^2. The following Matlab routine implements this algorithm.

Conjecture Let P, Q be minimizing polynomials for (1). Numerical evidence suggests that without loss of generality we may assume all coefficients of P to be positive, and the coefficients of Q those of P with alternating signs.

The computed local minimizers (in double precision) display this property. A proof for that would again reduce the number of unknowns significantly.

References

[1] D.W. Boyd. Two sharp inequalities for the norm of a factor of a polynomial. *Mathematika*, 39:341–349, 1992.

[2] D.W. Boyd. Sharp Inequalities for the Product of Polynomials. *Bull. London Math. Soc.*, 26:449–454, 1994.

[3] F. Bünger. The euclidean norm of the product of two polynomials, 2009.

[4] E. Kaltofen, L. Bin, Y. Zhengfeng, and Z. Lihong. Exact certification of global optimality of approximate factorizations via rationalizing sums-of-squares with floating point scalars. *ISSAC*, pages 155–163, 2008.

[5] E. Kaltofen, L. Bin, Y. Zhengfeng, and Z. Lihong. Exact certification in global polynomial optimization via sums-of-squares of rational functions with rational coefficients. submitted for publication, 2009.

[6] S.M. Rump. Structured Perturbations Part I: Normwise Distances. *SIAM J. Matrix Anal. Appl. (SIMAX)*, 25(1):1–30, 2003.

[7] S.M. Rump and H. Sekigawa. The ratio between the Toeplitz and the unstructured condition number. accepted for publication 2008.