Theoretical Computer Science ••• (••••) •••–•••

Contents lists available at ScienceDirect

Theoretical Computer Science

www.elsevier.com/locate/tcs

Real polynomial root-finding by means of matrix and polynomial iterations ☆

Victor Y. Pan [a,b,*], Liang Zhao [b]

[a] Department of Mathematics and Computer Science, Lehman College of the City University of New York, Bronx, NY 10468, USA
[b] Ph.D. Programs in Mathematics and Computer Science, The Graduate Center of the City University of New York, New York, NY 10036, USA

ARTICLE INFO

Article history:
Available online xxxx

Keywords:
Polynomials
Real roots
Matrices
Matrix sign iterations
Companion matrix
Frobenius algebra
Square root iterations
Root squaring

ABSTRACT

Univariate polynomial root-finding is a classical subject, still important for modern computing. Frequently one seeks just the real roots of a polynomial with real coefficients. They can be approximated at a low computational cost if the polynomial has no nonreal roots, but for high degree polynomials, nonreal roots are typically much more numerous than the real ones. The challenge is known for a long time, and the subject has been intensively studied. Nevertheless, we propose some novel ideas and techniques and obtain dramatic acceleration of the known algorithms. In order to achieve our progress we exploit the correlation between the computations with matrices and polynomials, randomized matrix computations, and complex plane geometry, extend the techniques of the matrix sign iterations, and use the structure of the companion matrix of the input polynomial. The results of our extensive tests with benchmark polynomials and random matrices are quite encouraging. In particular in our tests the number of iterations required for convergence of our algorithms grew very slowly (if it grew at all) as we increased the degree of the input polynomials and the dimension of the input matrices from 64 to 1024.

1. Introduction

Let us be given a univariate polynomial of degree n having real coefficients,

$$p(x) = \sum_{i=0}^{n} p_i x^i = p_n \prod_{j=1}^{n} (x - x_j), \quad p_n \neq 0, \tag{1.1}$$

which has r real roots $x_1, \dots, x_r$ and $s = (n-r)/2$ pairs of nonreal complex conjugate roots. In some applications, e.g., to algebraic and geometric optimization, one seeks only the r real roots, which make up just a small fraction of all roots. This is a well studied subject (see [14, Section 10.3.5], [37,44], and the bibliography therein), but the most popular packages of subroutines for root-finding such as MPSolve 2000 [5], Eigensolve 2001 [15], and MPSolve 2012 [10] approximate the r real

☆ This work has been supported by NSF Grant CCF-1116736 and CCF-1563942 and PSC CUNY Award 67699-00 45. Some of its results have been presented at CASC 2014.

* Corresponding author.
E-mail addresses: victor.pan@lehman.cuny.edu (V.Y. Pan), lzhao1@gc.cuny.edu (L. Zhao).
URL: http://comet.lehman.cuny.edu/vpan/ (V.Y. Pan).

http://dx.doi.org/10.1016/j.tcs.2017.03.032
0304-3975/

Please cite this article in press as: V.Y. Pan, L. Zhao, Real polynomial root-finding by means of matrix and polynomial iterations, Theoret. Comput. Sci. (2017), http://dx.doi.org/10.1016/j.tcs.2017.03.032

roots about as fast and as slow as all the n complex roots. It can be surprising, but we present some novel methods that accelerate the solution by a factor of n/r, which means dramatic speed up in various important applications.

The springboard for our real root-finders is the matrix sign iterations, which we apply to the companion matrix of an input polynomial. It is a well known technique [19], but never used in this direction. We show that it is particularly efficient for our purpose of real root-finders, and we combine it with various old and new techniques that support fast convergence of the iterations and their numerical validity.

Our numerical tests for benchmark polynomials are in good accordance with our formal results that show efficiency of this approach. The number of iterations required for convergence was typically quite small and grew very slowly (if it grew at all) as the polynomial degree increased from 64 to 1024.

Some of our techniques should be of independent interest, e.g., our numerical stabilization in Section 3.3, our exploitation of matrix functions and randomized matrix computations in Algorithm 3.1, and the combination of our maps of the complex plane with rational transformations of the variable and the roots. Some of our algorithms (e.g., the ones of Section 3.4) combine operations with matrices and polynomials, thus demonstrating once again the value of synergistic combinations of this kind, which we have been advocating since [25] and [6].

We hope that our progress will motivate further work on the implementation of our algorithms at the competitive level. For example, Stage 3 of our Algorithm 3.1 is reduced to the inversion or orthogonalization of Toeplitz-like matrices, and here one can benefit greatly from incorporation of the recent highly efficient numerical algorithms of the papers [51] and [52].

We organize our paper as follows. In the next section we cover the background for our algorithms based on the modified matrix sign iterations and their extensions. We present a variety of these algorithms in Section 3. In Section 4, which is the contribution of the second author, we present the results of our numerical tests. In the Appendix we cover some auxiliary results.

2. Basic definitions and results

Hereafter "flop" stands for "floating point arithmetic operation", assumed to be performed numerically, with bounded precision, e.g., the standard IEEE double precision.

2.1. Some basic definitions for matrix computations

- $\mathbb{C}^{m\times n}$ denotes the linear space of complex $m\times n$ matrices. $\mathbb{R}^{m\times n}$ is its subspace of $m\times n$ real matrices.
- $M^T=(m_{ji})_{i,j=1}^{m,n}$ is the transpose of a matrix $M=(m_{ij})_{i,j=1}^{m,n}\in\mathbb{C}^{m\times n}$. M^H is its Hermitian transpose. $M^H=M^T$ for a real matrix M.
- $||M||=||M||_2$ denotes its spectral norm.
- $I=I_n$ is the $n\times n$ identity matrix.
- $\mathrm{diag}(b_j)_{j=1}^{s}=\mathrm{diag}(b_1,\dots,b_s)$ is the $s\times s$ diagonal matrix with the diagonal entries $b_1,\dots,b_s$.
- $\mathcal{R}(M)$ is the range of a matrix M, that is, the linear space generated by its columns.
- A matrix of full column rank is a *matrix basis* of its range.
- A matrix Q is *orthogonal* or *unitary* if $Q^HQ=I$ or $QQ^H=I$.
- Suppose that an $m\times n$ matrix M has rank n (and so $m\ge n$). Write $(Q,R)=(Q(M),R(M))$ to denote a unique pair of unitary $m\times n$ matrix Q and upper triangular $n\times n$ matrix R such that $M=QR$ and all diagonal entries of the matrix R are positive [16, Theorem 5.2.3].
- M^+ is the unique Moore–Penrose pseudo inverse of M [16, Section 5.5.2], equal to M^H if and only if the matrix M is unitary.
- An $m\times n$ matrix M has an $n\times m$ *left inverse* matrix $X=M^{(I)}$ such that $XM=I_n$ if and only if it has full column rank n. In this case M^+ is a left inverse. The left inverse is unique if and only if M is a nonsingular matrix, in which case $m=n$ and $M^{(I)}=M^{-1}$.
- The ϵ-rank of a matrix M is the minimal rank of the matrices in its ϵ-neighborhood. *Numerical rank* $\mathrm{nrank}(M)$ is the ϵ-rank where ϵ is small in context.

Definition 2.1. Eigenvalues, eigenvectors and eigenspaces.

- A scalar x is an *eigenvalue* of a matrix M associated with an *eigenvector* $\mathbf{v}$ if $M\mathbf{v}=x\mathbf{v}$.
- The eigenvectors associated with an eigenvalue x or, more generally, with any set of the eigenvalues $\mathcal{X}\in\mathcal{X}(M)$ form the *eigenspaces* $\mathcal{S}(M,x)$ and $\mathcal{S}(M,\mathcal{X})$, respectively, associated with the eigenvalue x and the set $\mathcal{X}$ of eigenvalues, respectively.
- An eigenvalue x of a matrix M is a root of the characteristic polynomial $\det(xI-M)$. The multiplicity of this root is the *algebraic multiplicity* of the eigenvalue x, denoted $am(x)$. The dimension $gm(x)=\dim(\mathcal{S}(M,x))$ is the *geometric multiplicity* of x, never exceeding $am(x)$. An eigenvalue x is *simple* if $gm(x)=1$.
- An eigenspace $\mathcal{S}=\mathcal{S}(M,\mathcal{X})$ associated with some set of its eigenvalues of a matrix M is an *eigenspace of* M. A linear subspace $\mathcal{S}$ of $\mathbb{C}^{n\times n}$ is an eigenspace of a matrix M if and only if $M\mathcal{S}=\{M\mathbf{v}:\mathbf{v}\in\mathcal{S}\}\subseteq\mathcal{S}$ (see [43, Definition 4.1.1]).

Please cite this article in press as: V.Y. Pan, L. Zhao, Real polynomial root-finding by means of matrix and polynomial iterations, Theoret. Comput. Sci. (2017), http://dx.doi.org/10.1016/j.tcs.2017.03.032