Efficient Algorithms for Computing the Nearest Polynomial With A Real Root and Related Problems

Erich Kaltofen
North Carolina State University
www.math.ncsu.edu/~kaltofen

Joint work with: Markus Hitz (North Georgia College)
Lakshman Y. N. (Bell Labs)

ISSAC 1996 – Karmarkar and Lakshman

Nearest approximate GCD in the l^2-norm:

Let $f, g \in \mathbb{C}[z]$, both monic, $\deg(f) = m$ and $\deg(g) = n$.
Assuming that $\mathrm{GCD}(f, g) = 1$, find $\tilde{f}, \tilde{g} \in \mathbb{C}[z]$ monic of the same degrees, such that

$$\begin{aligned}&\mathrm{GCD}(\tilde{f}, \tilde{g}) \text{ is non-trvial and}\\ &\mathcal{N} = \|f - \tilde{f}\|^2 + \|g - \tilde{g}\|^2 \text{ is minimized.}\end{aligned}$$

$\|f\|$ denotes a norm of the coefficient vector of f.

The *symbolic* minimum of $\mathcal{N}$ with respect to a common root $\alpha \in \mathbb{C}$ can be obtained in closed-form:

$$\mathcal{N}_{min} = \frac{\overline{f(\alpha)}f(\alpha)}{\sum_{k=0}^{m-1}(\overline{\alpha}\alpha)^k} + \frac{\overline{g(\alpha)}g(\alpha)}{\sum_{k=0}^{n-1}(\overline{\alpha}\alpha)^k}$$

The individual perturbations of the coefficients of f and g are

$$f_i - \tilde{f}_i = \frac{(\overline{\alpha})^i f(\alpha)}{\sum_{k=0}^{m-1}(\overline{\alpha}\alpha)^k} \quad \text{and} \quad g_j - \tilde{g}_j = \frac{(\overline{\alpha})^j g(\alpha)}{\sum_{k=0}^{n-1}(\overline{\alpha}\alpha)^k}$$

($\overline{\alpha}$ is the complex conjugate).

Reduced Problem: Given $f \in \mathbb{C}[z]$ and $\alpha \in \mathbb{C}$. Find $\tilde{f} \in \mathbb{C}[z]$, s.t.

$$\tilde{f}(\alpha) = 0, \quad \text{and} \quad \|f - \tilde{f}\| = \min .$$

Let

$$\begin{aligned} f(z) &= a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0 \\ \tilde{f}(z) &= (z - \alpha) \sum_{k=0}^{n-1} u_k z^k \\ &= u_{n-1} z^n + (u_{n-2} - \alpha) z^{n-1} + (u_{n-3} - \alpha u_{n-2}) z^{n-2} + \\ &\qquad \cdots \quad + (u_0 - \alpha u_1) z - \alpha u_0 \end{aligned}$$

In terms of linear algebra:

$$\|f - \tilde{f}\| = \min_{\mathbf{u} \in \mathbb{C}^n} \|\mathbf{Pu} - \mathbf{b}\| \tag{1}$$

$$\begin{aligned} \mathbf{b} &= [a_0, \ldots, a_{n-1}, a_n]^{tr} \in \mathbb{C}^{n+1} \\ \mathbf{u} &= [u_0, \ldots, u_{n-1}]^{tr} \in \mathbb{C}^n \end{aligned}$$

$$\mathbf{P} = \begin{bmatrix} -\alpha & & & 0 \\ 1 & -\alpha & & \\ & \ddots & \ddots & \\ 0 & & 1 & -\alpha \\ & & & 1 \end{bmatrix} \in \mathbb{C}^{(n+1)\times n} \tag{2}$$

(1) is an over-determined linear system of equations.

$$\text{LP problem, if } \|\cdot\| \text{ is the } \begin{cases} l^\infty & \text{norm, or} \\ l^1 & \text{norm} \end{cases}$$

$$\text{LS problem, if } \|\cdot\| \text{ is the } \quad l^2 \text{ norm.}$$

Solutions for the l^2-norm in closed form:

$$\mathcal{N}_{min}(\alpha) = \|f - \tilde{f}\|^2 = \frac{\overline{f(\alpha)}f(\alpha)}{\sum_{k=0}^{n}(\overline{\alpha}\alpha)^k}, \quad f_j - \tilde{f}_j = \frac{(\overline{\alpha})^j f(\alpha)}{\sum_{k=0}^{n}(\overline{\alpha}\alpha)^k}$$

(also derived in Corless et al. [ISSAC'95] via SVD)

Constraining a Root Locus to a Curve

Let Γ be a piecewise smooth curve with finitely many segments, each having a parametrization $\gamma_k(t)$ in a single real parameter t.

For a given polynomial $f \in \mathbb{C}[z]$, we want to find a minimally perturbed polynomial $\tilde{f} \in \mathbb{C}[z]$ that has (at least) one root on Γ.

Parametric Minimization

We substitute the parametrization $\gamma_k(t)$ for the indeterminate α in $\mathcal{N}_{min}(\alpha)$. The resulting expression is a function in $t \in \mathbb{R}$.

It attains its minima at its *stationary* points. We have to compute the *real* roots of the derivative.

The derivative of the norm-expression is determined *symbolically*, the roots can be computed numerically.

Algorithm C

Input: $f \in \mathbb{C}[z]$, and a curve Γ.

Output: $\tilde{f} \in \mathbb{C}[z]$, and $\tau \in \mathbb{R}$, s.t. $\tilde{f}(\gamma_k(\tau)) = 0$ for some segment of Γ, and $\|f - \tilde{f}\|_2 = \min$.

($\mathbf{C}_1$) For each segement of Γ:

- ($\mathbf{C}_{1.1}$) Substitute $\gamma_k(t)$ for α in the symbolic minimum $\mathcal{N}_{min}(\alpha) \mapsto N(t)$.
- ($\mathbf{C}_{1.2}$) Symbolically determine the derivative $N'(t)$.
- ($\mathbf{C}_{1.3}$) Compute the *real* roots (of the numerator) of $N'(t)$. Select the one that minimizes $N(t)$.

($\mathbf{C}_2$) From all $N(\tau_k)$ of step ($\mathbf{C}_{1.3}$) determine the minimum $N(\tau)$.

($\mathbf{C}_3$) Compute the perturbations δ_j. Return $\tilde{f}$, k, and τ.

Computing the Radius of Stability in the l^2-Norm

Definition: Let $\mathcal{D} \subset \mathbb{C}$ be an open, and convex domain of the complex plane. The polynomial $f \in \mathbb{C}[z]$ is called $\mathcal{D}$-stable, if all its roots are located within $\mathcal{D}$.

Special cases: – the left half-plane: *Hurwitz* stability
– the open unit-disc: *Schur* stability

Given a $\mathcal{D}$-stable polynomial f, how much can we perturb its coefficients such that the perturbed polynomial is still $\mathcal{D}$-stable?

If we have a (piecewise) real parametrization of the boundary $\partial\mathcal{D}$ then we can apply our algorithm to find a *nearest unstable* polynomial.

Theorem: Let $f \in \mathbb{C}[z]$ be $\mathcal{D}$-stable, and

let $\hat{f} \in \mathbb{C}[z]$ be an unstable polynomial,
such that $\|f - \hat{f}\| = \varepsilon$, where $\varepsilon \in \mathbb{R}, \varepsilon > 0$.

Then, there exists $\tilde{f} \in \mathbb{C}[z]$ and $\zeta \in \partial\mathcal{D}$ such that

$\|f - \tilde{f}\| \leq \varepsilon$ and $\tilde{f}(\zeta) = 0$.

Example (of a monic polynomial)

$$\begin{aligned} f(z) &= z^3 + (2.41 - 3.50\mathbf{i})z^2 + (2.76 - 5.84\mathbf{i})z \\ &\quad -1.02 - 9.25\mathbf{i} \end{aligned}$$

is Hurwitz.

Root locations: $-1.04 + 3.10\mathbf{i}$, $-.99 - 1.30\mathbf{i}$, $-.37 + 1.70\mathbf{i}$

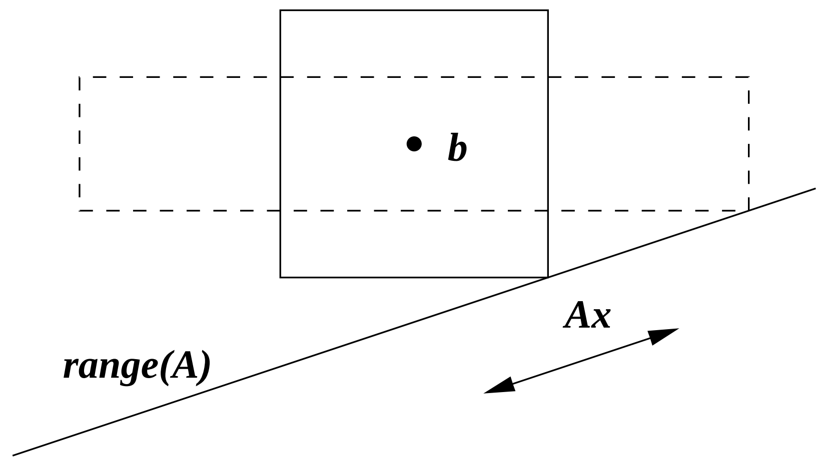
Nearest unstable polynomial:

$$\begin{aligned} \tilde{f}(z) &= z^3 + (2.7037 - 3.1492\mathbf{i})z^2 + (2.5740 - 5.6842\mathbf{i})z \\ &\quad -1.1026 - 9.3486\mathbf{i}. \end{aligned}$$

Radius of stability in the l^2-norm: 0.533567.

Tchebycheff's nearest consistency: $\mathbf{A} \in \mathbb{R}^{m \times n}, \mathbf{b} \in \mathbb{R}^m$:

$$\min_{\hat{\mathbf{x}}} \|\mathbf{b} - \mathbf{A}\hat{\mathbf{x}}\|_\infty = \min_{\hat{\mathbf{x}}} \left(\max_{1 \le i \le m} \left| b_i - \sum_{j=1}^{n} a_{i,j} \hat{x}_j \right| \right)$$



Solution by linear programmming:

minimize: δ

linear constraints: $\delta \geq \quad b_i - \sum_{j=1}^{n} a_{i,j} \hat{x}_j \;\; (1 \leq i \leq m)$

$\delta \geq -b_i + \sum_{j=1}^{n} a_{i,j} \hat{x}_j \;\; (1 \leq i \leq m)$

Special case: Stiefel's 1959 theorem

Let $\mathbf{A}$ be a matrix

$$\mathbf{A} = \begin{bmatrix} a_{0,0} & \cdots & a_{0,n-1} \\ \vdots & & \vdots \\ a_{n,0} & \cdots & a_{n,n-1} \end{bmatrix} \in \mathbb{R}^{(n+1)\times n}$$

of rank n such that no row of $\mathbf{A}$ is the zero vector, and let $\mathbf{b} = [b_0, \ldots, b_n] \in \mathbb{R}^{n+1}$ such that $\mathbf{Ax} \neq \mathbf{b}$ for all $\mathbf{x} \in \mathbb{R}^n$. Then

$$\delta = \min_{\mathbf{x}\in\mathbb{R}^n} ||\mathbf{Ax} - \mathbf{b}||_\infty = \left| \frac{\sum_{i=0}^{n} \lambda_i b_i}{\sum_{i=0}^{n} |\lambda_i|} \right|,$$

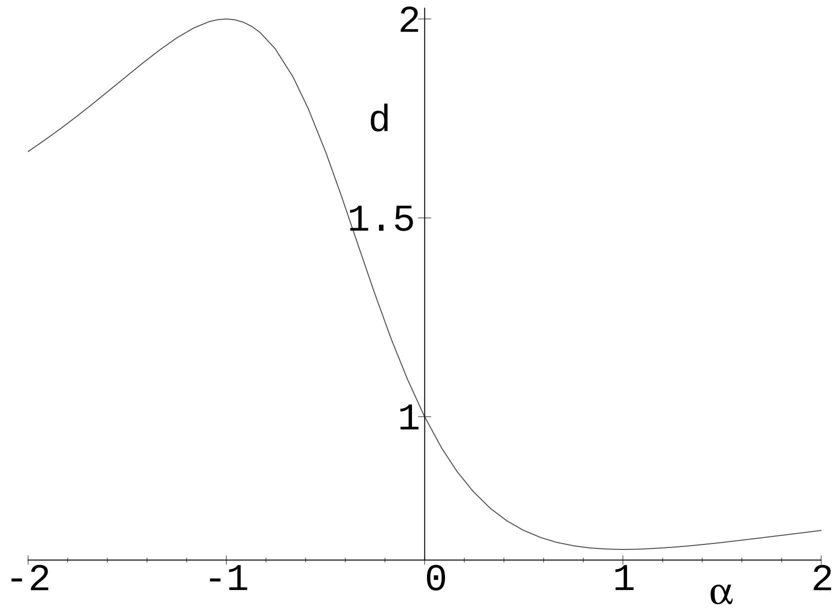
where $\Lambda = [\lambda_0, \ldots, \lambda_n]^{tr} \neq 0$ is a linear dependency among the rows of $\mathbf{A}$, i.e., $\Lambda^{tr}\mathbf{A} = 0$.

Special case: nearest polynomial with root α:

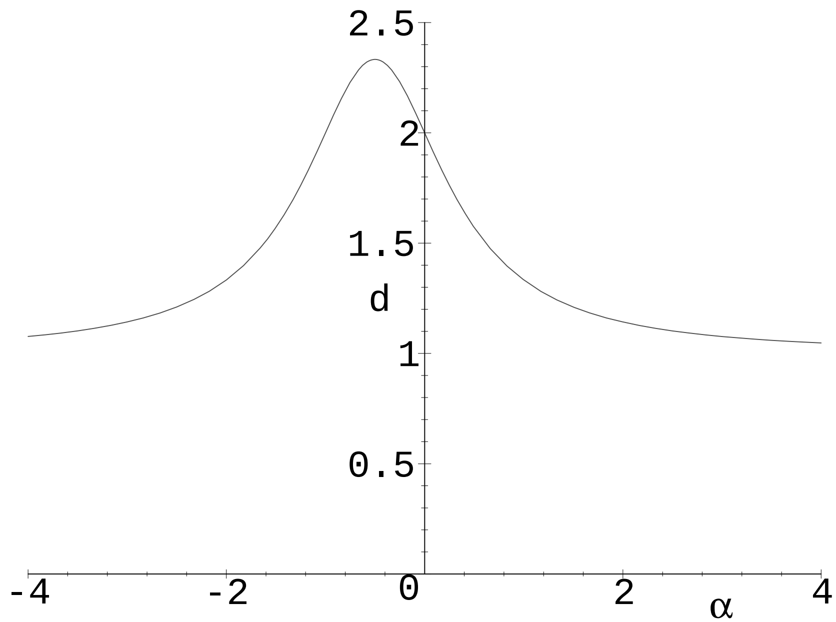
$$\delta(\alpha) = \min_{\mathbf{u}\in\mathbb{R}^n} \|\mathbf{Pu}-\mathbf{b}\|_\infty = \left|\frac{\sum_{i=0}^{n}\lambda_i a_i}{\sum_{i=0}^{n}|\lambda_i|}\right| = \left|\frac{f(\alpha)}{\sum_{i=0}^{n}|\alpha^i|}\right|. \qquad (3)$$

Stiefel's theorem also gives algorithm for finding $\mathbf{u}$.

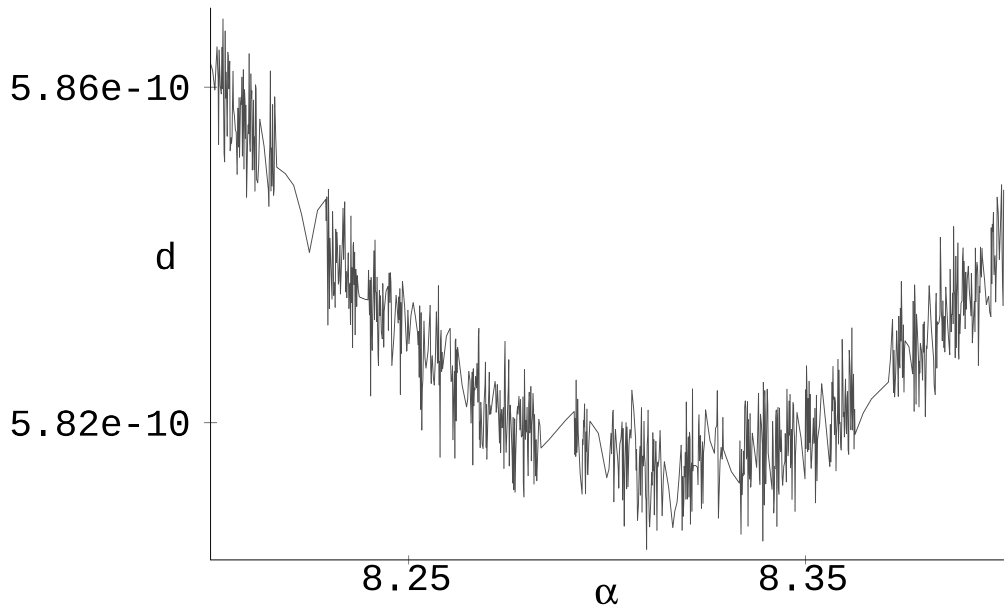
Parametric α: must minimize rational function (3).



$$f(x) = x^2 + 1,\ \tilde{f}(x) = \frac{1}{3}x^2 - \frac{2}{3}x + \frac{1}{3},\ \delta = \frac{2}{3}.$$



$$f(x) = x^2 + x + 2,\ \tilde{f}(x) = x + 2 = \ldots,\ \delta = 1.$$



$$f(x) = \prod_{k=1}^{10}(x-k-\mathbf{i})(x-k+\mathbf{i}),\ \delta \leq 5.8210^{-10}.$$

Nearest matrix with a given eigenvalue (Eckart and Young 1936):
Let $\mathbf{A} \in \mathbb{C}^{n\times n}$ and $\mu \in \mathbb{C}$:

$$\delta_{\mathbf{A}}(\mu) = \min_{\tilde{\mathbf{A}}:\ \mu \text{ is an eigenvalue of } \tilde{\mathbf{A}}} \|\mathbf{A} - \tilde{\mathbf{A}}\| = \frac{1}{\|(\mu\mathbf{I} - \mathbf{A})^{-1}\|}$$

where $\|\cdot\| = \|\cdot\|_{p,p}$ is an **induced matrix norm**.

For

$$\|\mathbf{B}\|_{\infty,\infty} = \max_i \sum_j |b_{i,j}|, \quad \|\mathbf{B}\|_{1,1} = \max_j \sum_i |b_{i,j}|,$$

we can solve optimization problem for **real** entries and parameter μ in **polynomial-time**.

Homework: Given $f = \sum f_{i,j} x^i y^j \in \mathbb{C}[x,y]$ **absolute irreducible**, find $\tilde{f} = (c_0 + c_1 x + c_2 y) u(x,y) \in \mathbb{C}[x,y]$, $\deg(\tilde{f}) \le \deg(f)$, such that

$$\|f - \tilde{f}\|_2 \text{ is minimal}$$

("nearest polynomial with a linear factor").

Hint: minimize parametric least square solution in the real and imaginary parts of the c_i.