Computing the common zeros of two bivariate functions via Bézout resultants

Colorado State University, 26th September 2013

Alex Townsend
PhD student
Mathematical Institute
University of Oxford

(with Yuji Nakatsukasa & Vanni Noferini)

Work supported by supported by EPSRC grant EP/P505666/1.

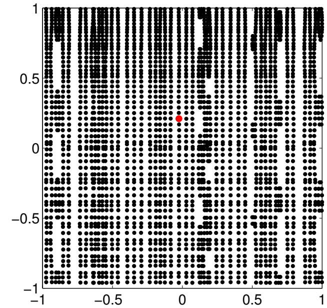
Introduction
Motivation

- Global 1D rootfinding is crucial (25% of Chebfun code needs roots)
- Chebfun2 is an extension of Chebfun for bivariate functions
- Very high degree polynomial interpolants are common

Find the global minimum of

$$f(x,y) = \left(\frac{x^2}{4} + e^{\sin(50x)} + \sin(70\sin(x))\right) + \left(\frac{y^2}{4} + \sin(60e^y) + \sin(\sin(80y))\right) - \cos(10x)\sin(10y) - \sin(10x)\cos(10y).$$

```
g = chebfun2( f );
r = roots( gradient( g ) );
```



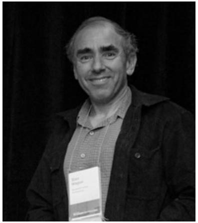
There are 2720 local extrema.

Let f and g be real-valued Lipschitz functions on $[-1,1]^2$. Solve:

$$\begin{pmatrix} f(x,y) \\ g(x,y) \end{pmatrix} = 0, \qquad (x,y) \in [-1,1]^2.$$

- **"Polynomialization":** Replace f and g with bivariate polynomials p and q
- **"Act locally":** Subdivide $[-1,1]^2$ with piecewise approximants until total degree ≤ 16, solve low degree rootfinding problems
- **"Think globally":** Do refinement and regularization to improve global stability

"**Think globally, act locally**", Stan Wagon

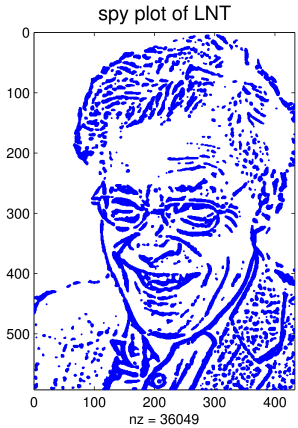


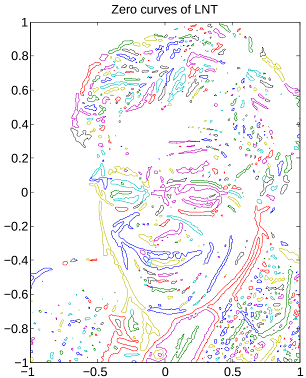
Introduction
NOT curve finding

Not to be confused with bivariate ~~rootfinding~~ curve finding:

$$f(x, y) = 0, \qquad (x, y) \in [-1, 1]^2.$$

Solutions lie along curves. Chebfun2 computes these by **Marching Squares**.





* Photo courtesy of Nick Hale.

Introduction
Talk overview

The talk follows Stan Wagon:

- "Polynomialization"
- "Act locally"
- "Think globally"
- Numerical examples

WARNING: Simple common zeros only!

Polynomialization

1D Chebyshev interpolants

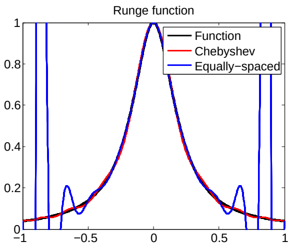
For $n \geq 1$, the **Chebyshev points** (of the 2nd kind) are given by

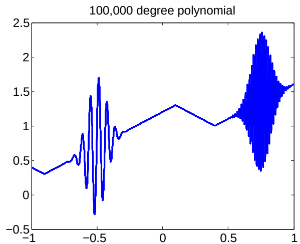
$$x_j^n = \cos\left(\frac{j\pi}{n}\right), \qquad 0 \leq j \leq n.$$

The **Chebyshev interpolant** of f is the polynomial p of degree at most n s.t.

$$p(x) = \sum_{j=0}^{n} c_j T_j(x), \qquad p(x_j^n) = f(x_j^n), \qquad 0 \leq j \leq n,$$

where $T_j(x) = \cos(j\cos^{-1}(x))$ is the Chebyshev polynomial of degree j.





Polynomialization
Tensor-product approximation

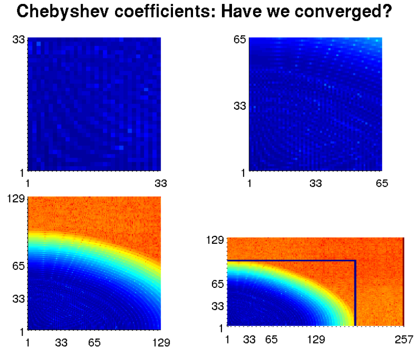
Replace f and g by their polynomial interpolants

$$p(x,y) = \sum_{i=0}^{n_p}\sum_{j=0}^{m_p} \alpha_{ij} T_i(x) T_j(y), \qquad q(x,y) = \sum_{i=0}^{n_q}\sum_{j=0}^{m_q} \beta_{ij} T_i(x) T_j(y)$$

such that $p(x_s^{n_p}, x_t^{m_p}) = f(x_s^{n_p}, x_t^{m_p})$ and $q(x_s^{n_q}, x_t^{m_q}) = g(x_s^{n_q}, x_t^{m_q})$. Select n_p, m_p and n_q, m_q large enough.

Take $n_p = 9, 17, 33, 65,$ and so on, until tail of coefficients falls below **relative** machine precision.

Chebyshev coefficients computed by fast DCT-I transform [Gentleman 72].

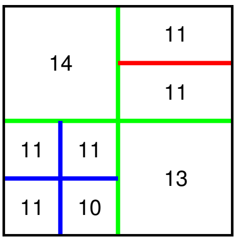


Act locally
Subdivision

Key fact: Subdivide to deal with high degree

Subdivide into subrectangles until polynomial degrees are small.

$$\sin((x-1/10)y)\cos(1/(x + (y-9/10) + 5))$$
$$= (y-1/10)\cos((x+(y+9/10)^2/4)) = 0$$



Real solutions only.

Do not bisect! Instead subdivide off-center (to avoid awkward coincidences).

Subdivide until degree 16.

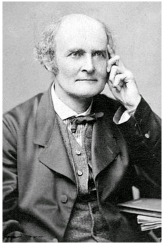
Like 1D subdivision:

Theorem (Bézout resultant theorem)

Let p_y and q_y be two univariate polynomials of degree at most n_p and n_q. The Chebyshev Bézout resultant matrix

$$B(p_y, q_y) = \left(b_{ij}\right)_{1\leq i,j\leq \max(n_p,n_q)}, \quad \frac{p_y(s)q_y(t) - p_y(t)q_y(s)}{s-t} = \sum_{i,j=1}^{\max(n_p,n_q)} b_{ij} T_{i-1}(s) T_{j-1}(t).$$

is nonsingular if and only if p_y and q_y have no common roots.



- Usually, this theorem is stated using the Sylvester resultant
- Usually, stated in terms of the monomial basis
- There are stable ways to form $B(p_y, q_y)$. We use [T., Noferini, Nakatsukasa, 13a]

Act locally

Hidden-variable resultant method

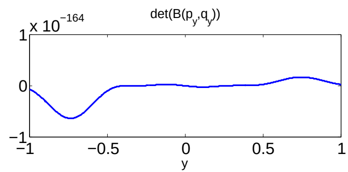
The hidden-variable resultant method "hides" one of the variables:

$$p_y(x) = p(x,y) = \sum_{i=0}^{n_p} \alpha_i(y) T_i(x), \qquad q_y(x) = q(x,y) = \sum_{i=0}^{n_q} \beta_i(y) T_i(x).$$

- $B(p_y, q_y)$ is a **symmetric** matrix of size $\max(n_p, n_q)$
- Each entry of $B(p_y, q_y)$ is a polynomial in y, of degree $m_p + m_q$
- For the y-values of $p(x,y) = q(x,y) = 0$ we want to solve

$$\det\big(B(p_y, q_y)\big) = 0, \qquad y \in [-1, 1].$$

Problem! Determinant is numerically zero:



Act locally

Matrix polynomial linearization

Key fact: Inherit robustness from eigenvalue solver

$B(p_y, q_y)$ is a matrix-valued polynomial in y: $B(p_y, q_y) = \sum_{i=0}^{M} A_i T_i(y) \in \mathbb{R}^{N \times N}$.

The colleague matrix [Specht 1960,Good 1961]:

$$yX + Y = y \begin{bmatrix} A_M & & & \\ & I_N & & \\ & & \ddots & \\ & & & I_N \end{bmatrix} - \frac{1}{2} \begin{bmatrix} -A_{M-1} & I_N - A_{M-2} & -A_{M-3} & \cdots & -A_0 \\ I_N & 0 & I_N & & \\ & \ddots & \ddots & \ddots & \\ & & I_N & 0 & I_N \\ & & & 2I_N & 0 \end{bmatrix}.$$

- Similar to companion, but for Chebyshev.
- **Inherited robustness** from eigenvalue solver.
- Strong linearization.

Act locally
Univariate rootfinding

Key point: Use univariate rootfinder for *x*-values

We use Chebfun's 1D rootfinder for the x-values, once we have the y-values.

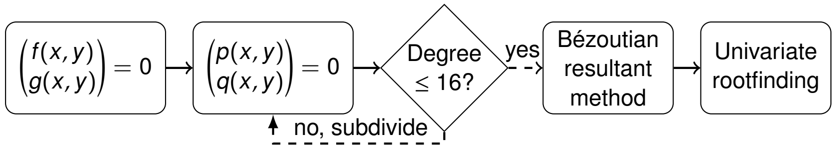
We independently solve for each y_*

$$p(x, y_*) = 0, \quad x \in [-1, 1] \qquad \text{and} \qquad q(x, y_*) = 0, \quad x \in [-1, 1].$$

- Based on the colleague matrix ($\approx$ companion matrix)
- Gets its robustness from eigenvalue solver
- Originally Boyd's algorithm from [Boyd 02]
- 1D subdivision is not needed for us

Act locally
Reviewing the algorithm

Flowchart of the algorithm:



Collect together the solutions from the subdomains.

Keep solutions in $[-1,1]^2$, throw away the rest. Perturb some if necessary.

Further questions:

1. Should we hide the x- or y-variable in the hidden-variable resultant method?
2. What is the operational cost of the algorithm?
3. Is the algorithm stable?

Think globally

Stability of the Bézout resultant method

Let $p(x_*, y_*) = q(x_*, y_*) = 0$ with $\|p\|_\infty = \|q\|_\infty = 1$. The Jacobian matrix is

$$J = J(x_*, y_*) = \begin{bmatrix} \frac{\partial p}{\partial x}(x_*, y_*) & \frac{\partial p}{\partial y}(x_*, y_*) \\ \frac{\partial q}{\partial x}(x_*, y_*) & \frac{\partial q}{\partial y}(x_*, y_*) \end{bmatrix}.$$

Absolute condition number of problem at (x_*, y_*): $\kappa_* = \|J^{-1}\|_2$

Absolute condition number of y_* for Bézout: $\kappa(y_*, B) \geq \frac{1}{2}\frac{\kappa_*^2}{\kappa_2(J)} \geq \frac{\kappa_*}{\|\mathrm{adj}(J)\|_2}$ [1]

The Bézout resultant method is unstable: If entries of J are small then,

$$\kappa(y_*, B) \gg \kappa_*$$

This is BAD news!

[1] Nakatsukasa, Noferini, & T., 2013b.

Think globally
Local refinement

Key fact: Local refinement can improve stability

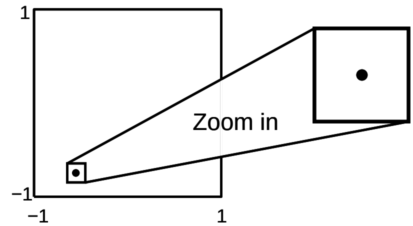
Redo Bézout resultant in Ω near (x_*, y_*).

Let $\Omega = [x_{\min}, x_{\max}] \times [y_{\min}, y_{\max}]$

$$|\Omega| = x_{\max} - x_{\min} \approx y_{\max} - y_{\min}$$

$$\kappa_\Omega(y_*, B) \approx |\Omega|^2 \kappa(y_*, B)$$

- Shrinking $|\Omega|$ improves stability (in a think globally sense).
- Get $\mathcal{O}(\kappa_* u)$ error from polynomialization.
- Also do local refinement in detected ill-conditioned regions.



Key fact: Regularize the problem by projecting

The Bézout resultant is symmetric. Partition such that

$$B(p_y, q_y) = \begin{bmatrix} B_1(y) & E(y)^T \\ E(y) & B_0(y) \end{bmatrix}, \qquad B_1(y) \in \mathbb{R}^{k\times k}, \quad B_0(y) \in \mathbb{R}^{(N-k)\times(N-k)},$$

with

$$\|B_0(y)\|_2 = \mathcal{O}(u), \qquad \|E(y)\|_2 = \mathcal{O}(u^{1/2}).$$

The eigenvalues of $B_1(y)$ and $B(p_y, q_y)$ in $[-1, 1]$ are usually within $\mathcal{O}(u)$.

Effectively this step removes large eigenvalues.

More details

Many other approaches

Homotopy continuation method

Solve a problem, make it harder.

$$H(\lambda, z) + Q(z)(1-\lambda) + P(z)\lambda,$$

$$\lambda \in (0,1).$$

Two-parameter eigenvalue problem

Use EIG to solve x and y together.

$$A_1 v = xB_1 v + yC_1 v,$$

$$A_2 w = xB_2 w + yC_2 w.$$

Contour algorithms

Solve two curve finding problems:

$$f(x,y) = 0, \qquad g(x,y) = 0.$$

Find intersection of curves.

Other resultant methods

- Sylvester resultants
- u-resultants
- Inverse iteration, Newton-like

More details

Which variable should the resultant method hide?

Let p and q be of degree (n_p, m_p, n_q, m_q).

If we solve for the y-variable first,

$$B(p_y, q_y) = \sum_{i=0}^{M} A_i T_i(y) \in \mathbb{R}^{N\times N}, \qquad \underbrace{NM = \max(n_p, n_q)(m_p + m_q)}_{\text{Size of eigenvalue problem}}.$$

If we solve for the x-variable first,

$$B(p_x, q_x) = \sum_{i=0}^{M} B_i T_i(x) \in \mathbb{R}^{N\times N}, \qquad \underbrace{NM = \max(m_p, m_q)(n_p + n_q)}_{\text{Size of eigenvalue problem}}.$$

- Solve for y-variable first if $\max(n_p, n_q)(m_p + m_q) \leq \max(m_p, m_q)(n_p + n_q)$.
- **Important:** It does not change stability issues.

More details

What is the observed computational cost?

Cost of rootfinding is function-dependent

Assume $n = m_p = m_q = n_p = n_q$.

$$\mathcal{O}(n^6) \qquad \text{vs.} \qquad \mathcal{O}(16^6 n^{-\log 4/\log \tau})$$

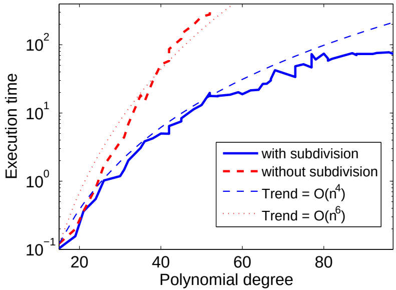
$\tau =$ average degree reduction.

$\tau \approx 0,\quad |x||y|$

$\tau = \frac{1}{2},\quad sin(Mx)sin(My),\quad M \gg 1$

$\tau = \frac{1}{\sqrt{2}},\quad \sin(M(x-y)),\quad M \gg 1$

$\tau \approx 1,\quad |\sin(M(x-y))|,\quad M \gg 1$

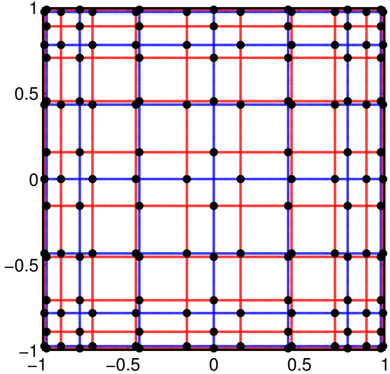


$$\begin{pmatrix} \sin(\omega(x+y)) \\ \cos(\omega(x-y)) \end{pmatrix} = 0, \quad 1 \le \omega \le 50$$

Numerical examples
Coordinate alignment

Solve

$$\begin{pmatrix} T_7(x)T_7(y)\cos(xy) \\ T_{10}(x)T_{10}(y)\cos(x^2y) \end{pmatrix} = 0.$$



Degrees are very small,

$$(m_p, n_p, m_q, n_q) = (20, 20, 24, 30),$$

but solutions aligned with grid.

Abs error $= 8 \times 10^{-16}$

$B(y)$ has **semisimple** eigenvalues with multiplicity 7 or 10. Numerically fine.

Numerical examples

Very high degree example

Find the global minimum of

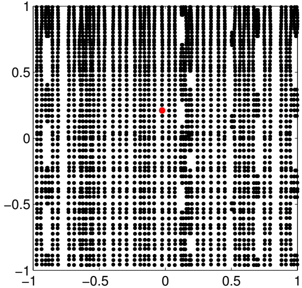
$$f(x,y) = \left(\frac{x^2}{4} + e^{\sin(50x)} + \sin(70\sin(x))\right) + \left(\frac{y^2}{4} + \sin(60e^y) + \sin(\sin(80y))\right) - \cos(10x)\sin(10y) - \sin(10x)\cos(10y).$$

This example is of high degree,

$$(m_p, n_p, m_q, n_q) = (901, 625, 901, 625).$$

There are 2720 local extrema.

$$\tau \approx 0.53 \quad \Rightarrow \quad \mathcal{O}(n^{2.2})$$



$$\text{Error} = 1.1 \times 10^{-15}$$

$$\text{Time} = 257\text{s}.$$

Numerical examples
Very high degree example

Find the global minimum of

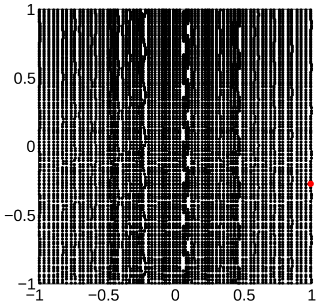
$$f(x,y) = \left(\frac{x^2}{4} + e^{\sin(100x)} + \sin(140\sin(x))\right) + \left(\frac{y^2}{4} + \sin(120e^y) + \sin(\sin(160y))\right) - \cos(20x)\sin(20y) - \sin(20x)\cos(20y).$$

This example as of high degree,

$$(1781, 1204, 1781, 1204).$$

There are 9318 local extrema.

$$\tau \approx 0.5 \quad \Rightarrow \quad \mathcal{O}(n^{2.1})$$



Time = 1300s.

Conclusion

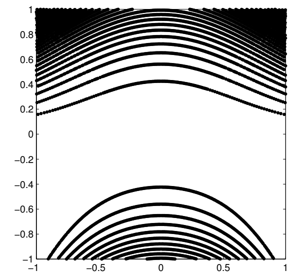
For high degree rootfinding:

- **"Polynomialization"**
- **"Act locally":** Subdivide!
- **"Think globally":** Stability.

For Bézout resultant:

- **Robustness from EIG**
- **Local refinement**
- **Regularization**

$$\begin{pmatrix} \mathrm{Ai}(-13(x^2y + y^2)))) \\ J_0(500x)y + xJ_1(500y) \end{pmatrix} = 0$$



$$(m_p, n_p, m_q, n_q) = (171, 120, 569, 568)$$

5932 solutions

time taken $= 501$s

Thank you

Special thanks to...

Nick Trefethen

Nick Higham

Françoise Tisseur

...and to you for listening.

- Y. Nakatsukasa, V. Noferini, and A. Townsend, *Computing the common zeros of two bivariate functions via Bézout resultants*, submitted, 2013.
- A. Townsend, V. Noferini, and Y. Nakatsukasa, *Vector spaces of linearizations for matrix polynomials: A bivariate polynomial approach*, submitted, 2013.
- A. Townsend and L. N. Trefethen, *An extension of Chebfun to two dimensions*, to appear in SISC, 2013.

Terminology: The eigenvalues of $B(p_y, q_y)$ satisfy

$$\det\left(B(p_y, q_y)\right) = 0.$$

If y_* is an eigenvalue then $p(x_*, y_*) = q(x_*, y_*) = 0$ for some x_*.

Assuming simple, isolated common zeros:

- **Finite common zeros:** $p(x, y_*) \neq 0$, $q(x, y_*) \neq 0$ with a common finite zero, then y_* is an eigenvalue of $B(y)$ with eigenvector $[T_0(x_*), \ldots, T_{N-1}(x_*)]^T$.
- **Common zero at infinity:** $p(x, y_*) \neq 0$, $q(x, y_*) \neq 0$ with leading coefficient $0T_N(x)$. y_* eigenvalue with $[0, \ldots, 0, 1]^T$.

If $p(x, y_*)$ and $q(x, y_*)$ have many common zeros $\Rightarrow B(y)$ has a **semisimple** eigenvalue of high multiplicity.

Extra slides
Travelling waves

Solve

$$\begin{pmatrix} \sin(\omega x - y/\omega) + y \\ \sin(x/\omega - \omega y) - x \end{pmatrix} = 0, \qquad \omega = 30.$$

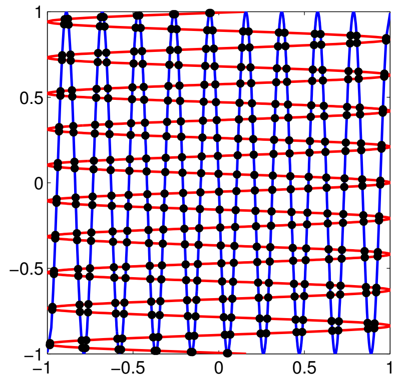
Degrees are small

$$(m_p, n_p, m_q, n_q) = (7, 63, 62, 6)$$

$$\tau \approx 0.72 \quad \Rightarrow \quad \mathcal{O}(n^{4.2})$$

Subdivision in x and y independently.

Qu: Hide x- or y-variable first?



$$\text{Abs error} = 1.3 \times 10^{-13}$$

$$\text{Time} = 10.8\text{s}$$