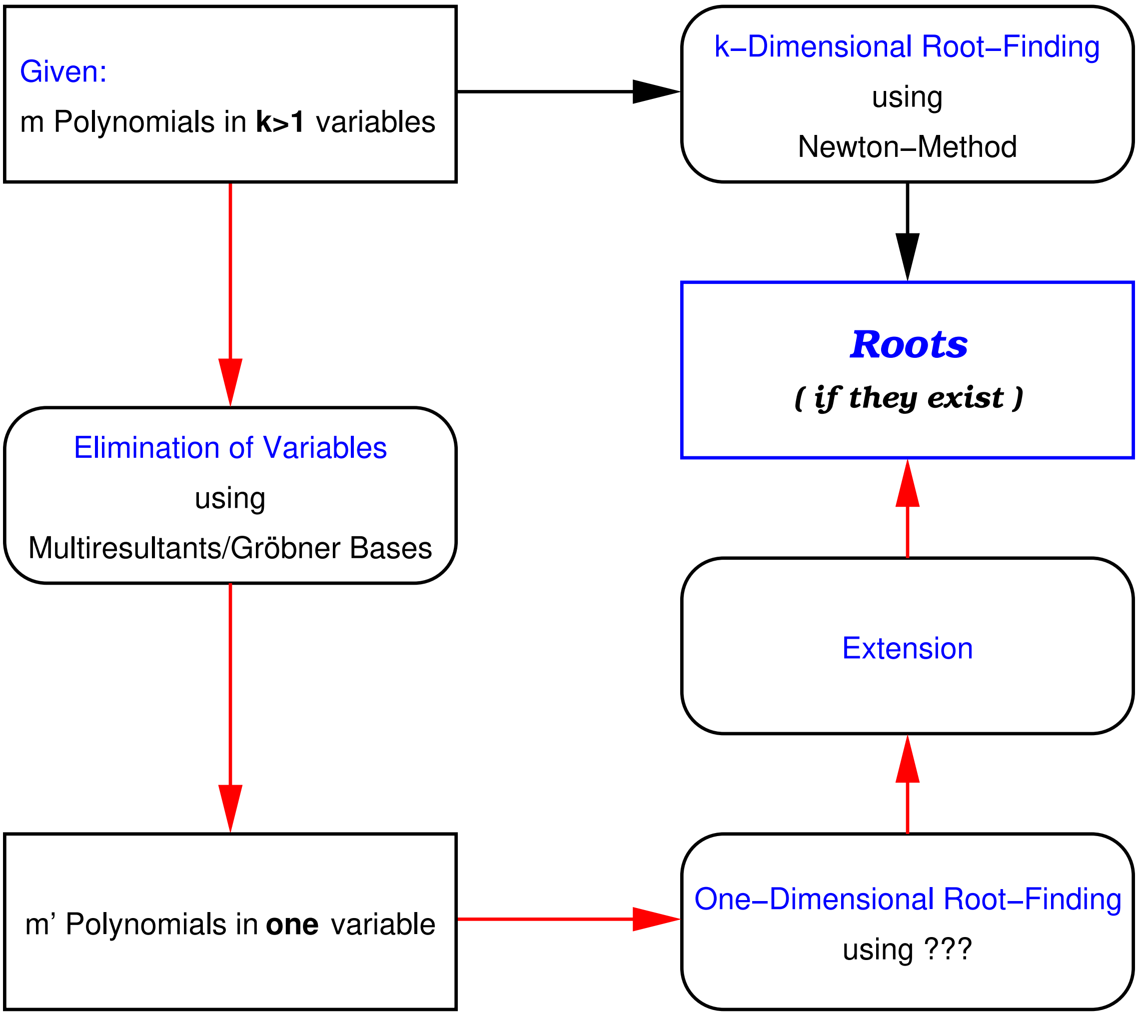
Finding Roots of Polynomials

Christian Lennerz

12th November 2002

Roots of Polynomials



Roots of Polynomials in One Variable

Definition

A polynomial P in one variable x with complex coefficients is a function, given by

$$P(x) = \sum_{i=0}^{n} a_i x^i \,,$$

where $a_0, \ldots, a_n$ are complex numbers with $a_n \neq 0$.

Observation

In general a the roots can be real or complex, single or multiple...

Theorem 1

If all coefficients of P are real numbers, then the complex roots occur in pairs that are conjugate and both roots have the same multiplicity.

Finding Roots by Polynomial Deflation

Theorem 2

Every polynomial P of degree n with complex coefficients and $a_n \neq 0$ has the following product representation:

$$P(x) = a_n(x - \xi_1)(x - \xi_2) \cdots (x - \xi_n),$$

where $\xi_1, \ldots, \xi_n$ are the roots of P.

Idea (Successive Deflation of P)

Given a root ξ_i of P, $1 \leq i \leq n$, the polynomial can be factored into the following product:

$$P(x) = (x - \xi_i)Q(x).$$

Then the following properties hold:

1. The reduced polynomial Q has degree one less than P.
2. The roots of Q are exactly the remaining roots of P.

Remarks

- Deflation is simply polynomial division.
- The effort of finding a root hopefully decreases in each step.
- The method cannot converge twice to the same non multiple root.
- Roots become more and more inaccurate, when not polished up.
- Successive Deflation is numerical stable, if the root of smallest absolute value is divided out in each step.
- In our context, we don't need complex arithmetics:

  $$[x-(a+bi)]\cdot[x-(a-bi)] = x^2-2ax+a^2+b^2 \in \mathbb{R}.$$

Bracketing

Definition

A root is bracketed in the interval (a, b) if $f(a)$ and $f(b)$ have different signs.

Motivation

According to the Intermediate Value Theorem there must be at least one root in (a, b), unless a singularity is present.

Remark

With standard arithmetic there is no sure way of bracketing all roots of an arbitrary function:

$$f(x) = 3x^2 + \frac{1}{\pi^4} \ln[(\pi - x)^2] + 1$$

dips below zero only in the interval $\pi \pm 10^{-667}$.

Bisection

Idea

- Precondition: A bracketed range is given as starting point.
- Evaluate the function at the midpoint of the interval and examine its sign.
- Use the midpoint to replace whichever limit has the same sign.

Remarks

- The method converges linearly because the width of the bracketed range decreases by a factor of two after each iteration.
- If the interval contains two or more roots, the method will find only one of them.
- The method does not distinguish singularities from roots.

Regula Falsi and Secant Method

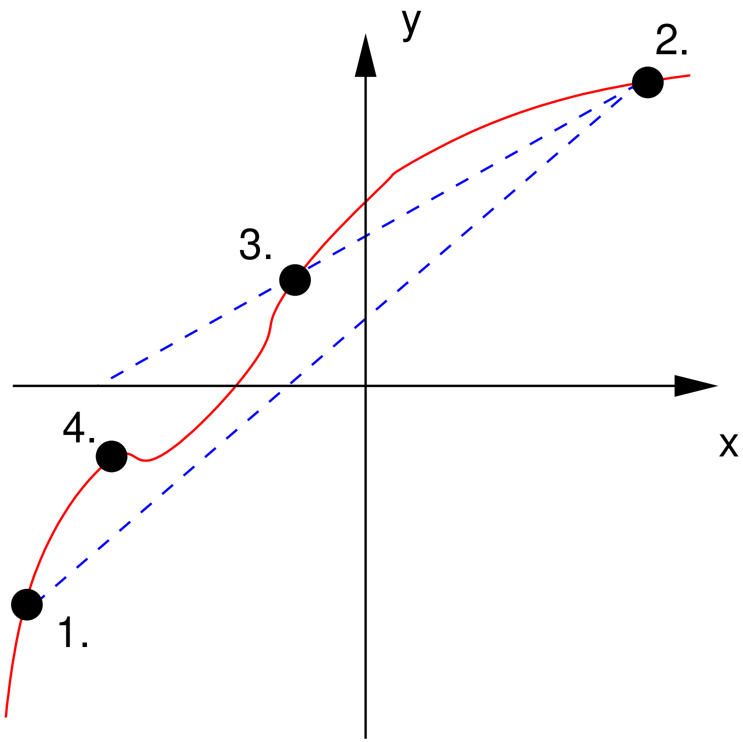
Assumption

The function is approximately linear in the local region of interest.

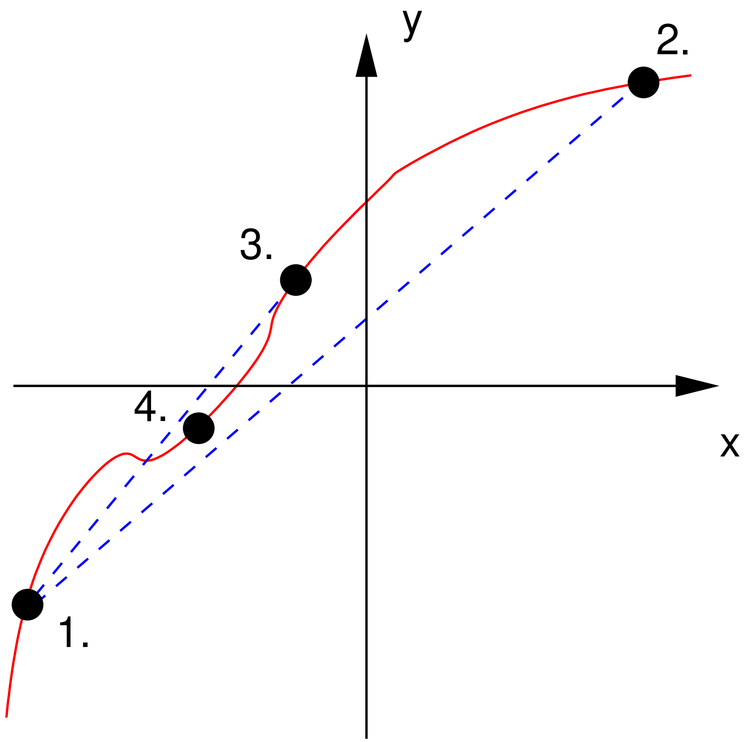
Idea

- Evaluate the function at the point where the line through both interval limits crosses the axis.
- Secant Method: Retain the most recent of the prior estimates and replace the other by the new estimate.
- Regula Falsi Method: Retain the prior estimate for which the function value has opposite sign from the current estimate of the root.

Example



Secant Method



Regula Falsi Method

Conclusions

- Secant Method:
  - The root does not always remain bracketed. There is no convergence guaranty.
  - Near the root of a sufficient continous function the convergence order is the "golden ratio" $1.618\ldots$.
- Regula Falsi Method:
  - Convergence can be guaranteed since the root remains bracketed.
  - Convergence order is lower as in the case of the Secant Method.

Improvements

- RIDDER's Method: Variant of the Regula Falsi Method that uses exponential instead of linear interpolation.
- BRENT's Method: Combines Secant Method, Bisection and quadratic interpolation.
- MULLER's Method: Generalization of the Secant Method using quadratic interpolation.

LAGUERRE's Method

Main Idea

- The root ξ_1 that we seek is assumed to be some distance a from a our current guess $\hat{\xi}_1$.
- All other roots are assumed to be located at a distance b.
- Use the polynomial P, P', P'' to solve for a, then take $\hat{\xi}_1 - a$ as the next guess.
- Continue this process until a becomes small enough.

Some Details

Remember the following relations between P and its derivatives:

$$
\begin{aligned}
P(x) &= \prod_{i=1}^{n}(x-\xi_i) \\
Q(x) &:= \frac{P'(x)}{P(x)} = \sum_{i=1}^{n}\frac{1}{x-\xi_i} \\
R(x) &:= \left[\frac{P'(x)}{P(x)}\right]^2 - \frac{P''(x)}{P(x)} = \sum_{i=1}^{n}\frac{1}{(x-\xi_i)^2}
\end{aligned}
$$

Using our ”rather drastic set of assumptions”

$$\xi_1 = \hat{\xi}_1 - a\,, \qquad \xi_i = \hat{\xi}_1 - b\,, \quad 2 \le i \le n.$$

we obtain for $Q(\hat{\xi}_1)$ and $R(\hat{\xi}_1)$:

$$Q(\hat{\xi}_1) = \frac{1}{a} + \frac{n-1}{b} \qquad\qquad R(\hat{\xi}_1) = \frac{1}{a^2} + \frac{n-1}{b^2}.$$

Solving for a leads to:

$$a = \frac{n}{Q(\hat{\xi}_1) \pm \sqrt{(n-1)(nR(\hat{\xi}_1) - Q(\hat{\xi}_1)^2}}$$

Remarks

- There are two possibilities for a. Take the sign, such that a is minimal.
- For polynomials with all real roots the method is guaranteed to converge to a root for any starting point.
- For polynomials with some complex roots convergence cannot be guaranteed.
- When the method converges to a simple complex root the convergence is third order.
- The method requires complex arithmetic, even while converging to real roots.

Eigenvalue Methods

Facts

- The eigenvalues of a matrix A are the roots of the characteristic polynomial

$$Q_A(x) = \det(A - xI).$$

- There are efficient and numerical stable non-root-finding methods to compute the eigenvalues of a matrix.

Question

Is it possible to reduce the polynomial root-finding problem to the problem of computing the eigenvalues of a matrix ?

Answer

The characteristic polynomial of the following companion matrix

$$A = \begin{pmatrix} -\frac{a_{n-1}}{a_n} & -\frac{a_{n-2}}{a_n} & \cdots & -\frac{a_1}{a_n} & -\frac{a_0}{a_n} \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix}$$

has the same roots as the polynomial

$$P(x) = \sum_{i=0}^{n} a_i x^i.$$

Proof

Expansion by the first row gives us

$$Q_A(x) = \det(A - xI) = (-1)^n \sum_{i=0}^{n} \frac{a_i}{a_n} x^i.$$

Remarks

- The eigenvalues can be computed using the QR-Algorithm, an efficient eigenvalue method when the input is an upper Hessenberg matrix.
- *Advantage:*
  More robust technique than LAGUERRE's Method.
- *Disadvantage:*
  Typically a factor 2 slower than LAGUERRE's Method.

Simultaneous Inclusion of Real Roots

Given: $P(x) = \sum_{i=0}^{n} a_i x^i, \quad a_n = 1$

Assumptions

1. P has n real roots $\xi = (\xi^{(1)}, \ldots, \xi^{(n)})$, where multiple roots are entered according to their multiplicity.
2. We collect multiple roots as $(\xi^{(m+1)}, \ldots, \xi^{(n)})$ and forget about them.
3. For all roots we know including intervals

$$X^{(0,i)} = [x_1^{(0,i)}, x_2^{(0,i)}], \quad 1 \le i \le m.$$

4. Including intervals are pairwise disjoint:

$$X^{(0,i)} \cap X^{(0,j)} = \emptyset, \quad 1 \le i < j \le m.$$

We consider the equivalent polynomial Q where every root has multiplicity 1:

$$Q(x) = \frac{P(x)}{GCD(P(x), P'(x))} = \prod_{j=1}^{m} (x - \xi^{(i)}).$$

Extracting $(x - \xi^{(i)})$ leads to

$$\xi^{(i)} = x - Q(x) / \prod_{j=1, j \neq i}^{m} (x - \xi^{(j)}).$$

If we choose $x = x^{(0,i)} \in X^{(0,i)}$, we have $\xi^{(i)} \in$

$$X^{(1,i)} := \left\{ x^{(0,i)} - \frac{Q(x^{(0,i)})}{\prod_{j=1, j \neq i}^{m} (x^{(0,i)} - X^{(0,j)})} \right\} \cap X^{(0,i)}$$

... and the following total step iteration scheme:

$$\begin{aligned}
X^{(k+1,i)} &:= \left\{ x^{(k,i)} - Q(x^{(k,i)}) / B^{(k,i)} \right\} \cap X^{(k,i)} \\
B^{(k,i)} &:= \prod_{j=1, j \neq i}^{m} (x^{(k,i)} - X^{(k,j)}) \\
x^{(k,i)} &\in X^{(k,i)}, \quad 1 \leq i \leq m, \quad k \geq 0.
\end{aligned}$$

Improvements

Before computing $X^{(k+1,j)}$, $j \geq i$, we already know $X^{(k+1,j)}$, $j < i$. Therefore we can replace $B^{(k,i)}$ by

$$C^{(k,i)} := \prod_{j=1}^{i-1} (x^{(k,i)} - X^{(k+1,j)}) \cdot \prod_{j=i+1}^{m} (x^{(k,i)} - X^{(k,j)}).$$

to get a tighter denominator and the following single step iteration scheme:

Init: $\quad x^{(0,i)} \in X^{(0,i)}$

Step:

$$X^{(k+1,i)} := \left\{ x^{(k,i)} - Q(x^{(k,i)})/C^{(k,i)} \right\} \cap X^{(k,i)}$$

$$C^{(k,i)} := \prod_{j=1}^{i-1} (x^{(k,i)} - X^{(k+1,j)}) \cdot \prod_{j=i+1}^{m} (x^{(k,i)} - X^{(k,j)}),$$

$$x^{(k+1,i)} \in X^{(k+1,i)}, \quad 1 \leq i \leq m, \quad k \geq 0.$$

Heuristic: Choose $x^{(k,i)} = \frac{1}{2}(x_1^{(k,i)} + x_2^{(k,i)})$

Conclusions

- **Advantages**
  - Simultaneous determination of polynomial roots.
  - Reliable information about root location.
  - Always converging under the assumptions made above.
  - Early sign prediction of roots possible.
  - Convergence order ≥ 2 for the total step method and > 2 in the case of the single step method.
- **Disadvantages**
  - Interval arithmetic is expensive.
  - All roots of the polynomial have to be real.

Simultaneous Inclusion of Complex Roots

Assumption

We are using circular regions as complex intervals.

The following relationship between P and P':

$$Q(z) = \frac{P'(z)}{P(z)} = \frac{\sum_{i=1}^{m} \prod_{j=1, j\neq i}^{m} (z - \xi^{(j)})}{\prod_{j=1}^{m} (z - \xi^{(j)})} = \sum_{i=1}^{m} \frac{1}{z - \xi^{(i)}}$$

let us write $\xi^{(i)}$ as follows:

$$\xi^{(i)} = z - (z - \xi^{(i)}) = z - \frac{1}{\frac{1}{z-\xi^{(i)}}} = z - \frac{1}{Q(z) - \sum_{j=1, j\neq i}^{m} \frac{1}{z-\xi^{(j)}}}.$$

This leads to the following total step iteration scheme:

$$\begin{aligned}
Z^{(k+1,i)} &= \langle z^{(k+1,i)}, r^{(k+1,i)} \rangle \\
&= z^{(k,i)} - \frac{1}{Q(z^{(k,i)}) - C^{(k,i)}} \\
z^{(k,i)} &= m(Z^{(k,i)}), \\
Q(z^{(k,i)}) &= \frac{P'(z^{(k,i)})}{P(z^{(k,i)})} \quad \text{for } P(z^{(k,i)}) \neq 0, \\
C^{(k,i)} &= \sum_{j=1, j\neq i}^{m} \frac{1}{z^{(k,i)} - Z^{(k,i)}}, \\
& \quad 1 \leq i \leq m, \quad k \geq 0.
\end{aligned}$$

Remark

- A single step method can be formulated as in the case of real roots.
- The convergence is at least third order.