Numerical Analysis

Chapter Three: Solution of Nonlinear Equations

主讲人 朱亚菲
幻灯片制作 朱亚菲

中国海洋大学 信息科学与工程学院

2013 年 9 月 27 日

Contents

Contents

Contents

3.5.3 Laguerre Iteration

Laguerre's method is also used for finding the roots of a polynomial p.

- Advantages:
  - Favorable convergence properties
  - Rather robust

The algorithm is iterative and proceeds from one approximate root z to a new one by calculating

$$A = -p'(z)/p(z)$$

$$B = A^2 - p''(z)/p(z)$$

$$C = n^{-1}[A \pm \sqrt{(n-1)(nB - A^2)}]$$

$$z_{new} = z + 1/C$$

3.5.3 Laguerre Iteration

THEOREM 10 (Theorem on Radius of Convergences in Laguerre's Method)

If p is a polynomial of degree n, if z is any complex number, and if C is computed as in Laguerre's algorithm, then p has a root in the complex plane within distance $\sqrt{n}/C$ of z.

3.5.3 Laguerre Iteration

```
function z1=Laguerre(a,M,epsilon,z0)
n=length(a);
for k=1:M
   alpha=a(n);beta=0;gamma=0;
   for j=n-1:-1:1
      gamma =z0*gamma+beta;
      beta =z0*beta+alpha;
      alpha =z0*alpha+a(j);
   end
   A=-beta/alpha;B=A^2-2*beta/alpha;
   C=[A+((n-1)^0.5)*((n*B-A^2)^0.5)]/n;
   z1(k)=z0+1/C
   if abs(z1-z0)<epsilon  break
   end
   z0=z1(k);
end
```

3.5.3 Laguerre Iteration

LEMMA 1 (First Lemma on Interval Endpoints)

Let v_1, v_2, ..., v_n be any real numbers. Put $\alpha = \sum_{i=1}^{n} v_i$ and $\beta = \sum_{i=1}^{n} v_i^2$. Then the numbers v_i lie in the closed interval whose endpoints are

$$n^{-1}[\alpha \pm \sqrt{(n-1)(n\beta - \alpha^2)}]$$

3.5.3 Laguerre Iteration

Proof.

It suffices to proof that v_1 lies in the interval described. Recall the Cauchy-Schwarz Inequality:

$$\left(\sum_{i=1}^{m} x_i y_i\right)^2 \leqslant \left(\sum_{i=1}^{m} x_i^2\right)\left(\sum_{j=1}^{m} y_j^2\right)$$

Applying this, we have

$$\alpha^2 - 2\alpha v_1 + v_1^2 = (\alpha - v_1)^2 = (v_2 + v_3 + \ldots + v_n)^2$$

$$\leqslant (1^2 + 1^2 + \ldots + 1^2)(v_2^2 + v_3^2 + \ldots + v_n^2) = (n-1)(v_2^2 + v_3^2 + \ldots + v_n^2)$$

$$= (n-1)(\beta - v_1^2) = (n-1)\beta - nv_1^2 + v_1^2$$

□

3.5.3 Laguerre Iteration

Proof.

Rearranging this inequality gives us

$$nv_1^2 - 2\alpha v_1 + \alpha^2 - (n-1)\beta \leqslant 0$$

This shows that the quadratic function $q(x) = nx^2 - 2\alpha x + \alpha^2 - (n-1)\beta$ has the property $q(v_1) \leqslant 0$. For large |x|, obviously $q(x) > 0$. hence, v_1 lies between the two roots of q, and they are the endpoints in Formula $n^{-1}[\alpha \pm \sqrt{(n-1)(n\beta - \alpha^2)}]$ □

3.5.3 Laguerre Iteration

LEMMA 2 (Second Lemma on Interval Endpoints)

Let p be a real polynomial of degree n, whose roots, r_1, r_2, ..., r_n are real. For any real x different from all the r_j, the numbers $(x - r_j)^{-1}$ lie in the interval whose endpoints are

$$[np(x)]^{-1}\{p'x \pm \sqrt{[(n-1)p'(x)]^2 - n(n-1)p(x)p''(x)}\}$$

3.5.3 Laguerre Iteration

THEOREM 11 (Theorem on Monotonic Convergence of Laguerre Methed)

Let p be a real polynomial whose roots are all real. The sequence produced by the Laguerre algorithm with an arbitrary starting point converges monotonically to a root of p.

Contents

3.5.4 Complex Newton's Method

For a polynomial having complex coefficients, Newton's method should be programmed in complex arithmetic.
After one root has been found, the deflation process (also programmed in complex arithmetic) should be used. Newton's method can then be applied to the reduced polynomial. This process can be repeated until all roots have been determined.

3.5.4 Complex Newton's Method

Further analysis and experience indicate that in general, the procedure is satisfactory provided that two precautions are taken:

1. The roots should be computed in order of increasing magnitude.
2. Any root obtained by using Newton's method on a reduced polynomial should be immediately refined by applying Newton's method to the original polynomial with the best estimate of the root as the starting value. Only after this has been done should the next step of deflation be carried out.

Contents

Contents

3.6.1 Basic Concepts

- In general, a homotopy can be any continuous connection between f and g.
- An example of a **homotopy** that connects the two functions f and g.

$$h(t, x) = tf(x) + (1 - t)g(x)$$

The parameter t runs over the inteval [0,1].

3.6.1 Basic Concepts

The basic idea of the continuation method is to embed the given problem in a one-parameter family of problems, using a parameter t that runs over the inteval [0,1]. The original problem is made to correspond to $t = 1$, and another problem with known solution is made to correspond to $t = 0$. For example, we can define

$$h(t, x) = tf(x) + (1 - t)[f(x) - f(x_0)] = f(x) + (t - 1)f(x_0)$$

Here x_0 can be any point in X, and it is clear that x_0 will be a solution of the problem when $t = 0$.

3.6.1 Basic Concepts

If the equation $h(t, x) = 0$ has a unique root for each $t \in [0, 1]$, then that root is a function of t, and we can write $x(t)$ as the unique member of X that makes the equation $h(t, x(t)) = 0$ true. The set

$$\{x(t) : 0 \leqslant t \leqslant 1\}$$

can be interpreted as an arc or curve in X, paramatrized by t. This arc leads from the known point $x(0)$ to the solution of our problem, $x(1)$. The continuation method attempts to determine this curve by computing points on it, $x(t_0), x(t_1), \ldots, x(t_m)$.

3.6.1 Basic Concepts

If the function $t \mapsto x(t)$ is differentiable and if h is differentiable, then the Implicit Function Theorem enables us to compute $x'(t)$. By pursuing this idea, we can describe the curve in $\{x(t) : 0 \leqslant t \leqslant 1\}$ by a differential equation. Assuming an arbitrary homotopy, we have

$$0 = h\big(t, x(t)\big)$$

3.6.1 Basic Concepts

On differentiating with respect to t, we obtain

$$0 = h_t\big(t, x(t)\big) + h_x\big(t, x(t)\big)x'(t)$$

in which subscripts denote partial derivatives. Thus,

$$x'(t) = -[h_x\big(t, x(t)\big)]^{-1}h_t\big(t, x(t)\big)$$

This is a differential equation for x. It has a known initial value because $x(0)$ is supposedly known. On integrating this differential equation, we shall have the value $x(1)$, which is the solution.

3.6.1 Basic Concepts

EXAMPLE 1

We let $X = Y = R^2$, and define

$$f(x) = \begin{bmatrix} \xi_1^2 - 3\xi_2^2 + 3 \\ \xi_1\xi_2 + 6 \end{bmatrix} \qquad x = (\xi_1, \xi_2) \in R^2$$

3.6.1 Basic Concepts

THEOREM 1 (Theorem on Continuously Differentiable Solution)

If $f : R^n \to R^n$ is continuously differential and if $\|[f'(x)]^{-1}\| \leqslant M$ on R^n, then for any $x_0 \in R^n$ there is a unique curve $x(t) : 0 \leqslant t \leqslant 1$ in R^n such that $f(x(t)) + (t-1)f(x_0) = 0$, with $0 \leqslant t \leqslant 1$. The function $t \mapsto x(t)$ is a continuously differentiable solution of the initial-value problem $x' = -[f'(x)]^{-1}f(x_0)$, where $x(0) = x_0$.

Contents

3.6.2 Tracing the Path

Another way of tracing the path $x(t)$:
We start with the equation $h(t, x) = 0$, supposing that $x \in R^n$ and $t \in [0, 1]$. A vector $y \in R^{n+1}$ is defined by

$$y = (t, \xi_1, \xi_2, \ldots, \xi_n)$$

where $\xi_1, \xi_2, \ldots, \xi_n$ are the components of x. Hence, our equation is simply $h(y) = 0$. Each component of y, including t, is now allowed to be a function of an independent variable s, and we write $h(y(s)) = 0$.

3.6.2 Tracing the Path

Differentiating with respect to s, we obtain the basic differential equation

$$h'(y(s))y'(s) = 0$$

The variable s starts at 0, as does t. The initial value of x is $x(0) = x_0$. Thus, suitable starting values are available for the differential equation.

3.6.2 Tracing the Path

Since f and g are maps of R^n into R^n, h is a map of R^{n+1} into R^n. The derivative $h'(y)$ is therefore represented by an $n \times (n+1)$ matrix, A. The vector $y(s)$ has $n+1$ components, which we denote by $\eta_1, \eta_2, \ldots, \eta_{n+1}$. By appealing to the lemma below, we can obtain another form for the Equation; namely,

$$n_j' = (-1)^{j+1} det(A_j) \qquad (1 \leqslant j \leqslant n+1)$$

where A_j is the $n \times n$ matrix that results from A by deleting the jth column.

3.6.2 Tracing the Path

EXAMPLE 2

Taking f and x_0 as in Example 1, we have

$$h(t,x) = \begin{bmatrix} \xi_1^2 - 3\xi_2^2 + 2 + t \\ \xi_1\xi_2 - 1 + 7t \end{bmatrix}$$

3.6.2 Tracing the Path

LEMMA 1 (Lemma on Solution of Homogeneous Equation)

Let A be an $n \times (n+1)$ matrix. A solution of the homogeneous equation $Ax = 0$ is given by $x_j = (-1)^j det(A_j)$, where A_j is A without column j.

Contents

3.6.3 Relation to Newton's Method

The connection between the homotopy methods and Newton's method is deeper than may be seen at first glance. Let us start with the homotopy

$$h(t, x) = f(x) - e^{-t} f(x_0)$$

In this equation, t will run from 0 to ∞. We seek a curve or path, $x = x(t)$, on which

$$0 = h\big(t, x(t)\big) = f\big(x(t)\big) - e^{-t} f(x_0)$$

3.6.3 Relation to Newton's Method

As usual, differentiation with respect to t will lead to a differential equation describing the path:

$$0 = f'\big(x(t)\big)x'(t) + e^{-t}f(x_0) = f'\big(x(t)\big)x'(t) + f\big(x(t)\big)$$

$$x'(t) = -[f'\big(x(t)\big)]^{-1}f\big(x(t)\big)$$

In this differential equation is integrated with Euler's method, with step size 1, the result is the formula

$$x_{n+1} = x_n - [f'(x_n)]^{-1}f(x_n)$$

Contents

3.6.4 Linear Programming

Consider the standard linear programming problem

$$\begin{cases} maximize\ c^T x \\ subject\ to\ Ax = b\ and\ x \geqslant 0 \end{cases}$$

Here, $c \in R^n, x \in R^n, b \in R^m$, and A is an $m \times n$ matrix.We start with a **feasible point**—that is, a point x^0 that satisfies the constraints. The **feasible set** is

$$F = \{x \in R^n : Ax = b\ and\ x \geqslant 0\}$$

3.6.4 Linear Programming

We shall try to find a curve $t \mapsto x(t)$ in the feasible set, starting at x^0 and leading to a solution of the extremal problem.
Our requirements are

1. $x(t) \geqslant 0$ for $t \geqslant 0$
2. $Ax(t) = b$ for $t \geqslant 0$
3. $c^T x(t)$ is increasing for $t \geqslant 0$.

3.6.4 Linear Programming

The curve will be defined by an initial-value problem:

$$\begin{cases} x' = f(x) \\ x(0) = x^{(0)} \end{cases}$$

The task facing us is to determine a suitable f.

3.6.4 Linear Programming

To satisfy Condition 1, we shall arrange that whenever a component x_i approaches 0, its velocity $x_i'(t)$ will also approach 0.
This can be accomplished by putting

$$D(x) = \begin{bmatrix} x_1 & & & 0 \\ & x_2 & & \\ & & \ddots & \\ 0 & & & x_n \end{bmatrix}$$

and assuming that for some bounded function G,

$$f(x) = D(x)G(x)$$

Then we have $x_i' = x_i G_i(x)$
and clearly $x_i' \to 0$ if $x_i \to 0$.

3.6.4 Linear Programming

To satisfy Requirement 2, it suffices to require $Ax' = 0$.
Since $x' = f = DG$, we must require $ADG = 0$.
This is most conveniently arranged by letting $G = PH$ where H is any funtion and P is the orthogonal projection onto the null space of AD.

3.6.4 Linear Programming

To satisfy Property 3, we should select H so that $c^T x(t)$ is increasing. Thus, we want

$$0 < (c^T x(t))' = c^T x' = c^T f(x) = c^T DG = c^T DPH$$

A convenient choice for H is Dc because then we have, with $v = Dc$,

$$c^T DPH = c^T DPDc = v^T Pv = \langle v, Pv \rangle = \langle v - Pv + Pv, Pv \rangle = \langle Pv, Pv \rangle \geqslant 0$$

Notice that $v - Pv$ is orthogonal to the range of P, and $\langle v - Pv, Pv \rangle = 0$.

3.6.4 Linear Programming

The final version of our initial-value problem is

$$x' = D(x)P(x)D(x)c \quad x(0) = x^{(0)}$$