MULTIPLE ROOTS

We study two classes of functions for which there is additional difficulty in calculating their roots. The first of these are functions in which the desired root has a *multiplicity* greater than 1. What does this mean?

Let α be a root of the function $f(x)$, and imagine writing it in the factored form

$$f(x) = (x - \alpha)^m \, h(x)$$

with some integer $m \geq 1$ and some continuous function $h(x)$ for which $h(\alpha) \neq 0$. Then we say that α is a root of $f(x)$ of multiplicity m. For example, the function

$$f(x) = e^{x^2} - 1$$

has $x = 0$ as a root of multiplicity $m = 2$. In particular, define

$$h(x) = \frac{e^{x^2} - 1}{x^2}$$

for $x \neq 0$.

Using Taylor polynomial approximations, we can show for $x \neq 0$ that

$$\begin{aligned} h(x) &\approx 1 + \frac{1}{2}x^2 + \frac{1}{6}x^4 \\ \lim_{x \to 0} h(x) &= 1 \end{aligned}$$

This leads us to extend the definition of $h(x)$ to

$$\begin{aligned} h(x) &= \frac{e^{x^2} - 1}{x^2}, \qquad x \neq 0 \\ h(0) &= 1 \end{aligned}$$

Thus

$$f(x) = x^2 h(x)$$

as asserted and $x = 0$ is a root of $f(x)$ of multiplicity $m = 2$.

Roots for which $m = 1$ are called *simple roots*, and the methods studied to this point were intended for such roots. We now consider the case of $m > 1$.

If the function $f(x)$ is m-times differentiable around α, then we can differentiate

$$f(x) = (x - \alpha)^m \, h(x)$$

m times to obtain an equivalent formulation of what it means for the root to have multiplicity m.
For an example, consider the case

$$f(x) = (x - \alpha)^3 \, h(x)$$

Then

$$\begin{aligned} f'(x) &= 3\,(x-\alpha)^2\,h(x) + (x-\alpha)^3\,h'(x) \\ &\equiv (x-\alpha)^2\,h_2(x) \\ h_2(x) &= 3h(x) + (x-\alpha)\,h'(x) \\ h_2(\alpha) &= 3h(\alpha) \neq 0 \end{aligned}$$

This shows α is a root of $f'(x)$ of multiplicity 2.

Differentiating a second time, we can show

$$f''(x) = (x - \alpha)\, h_3(x)$$

for a suitably defined $h_3(x)$ with $h_3(\alpha) \neq 0$, and α is a simple root of $f''(x)$.

Differentiating a third time, we have

$$f'''(\alpha) = h_3(\alpha) \neq 0$$

We can use this as part of a proof of the following: α is a root of $f(x)$ of multiplicity $m = 3$ if and only if

$$f(\alpha) = f'(\alpha) = f''(\alpha) = 0, \qquad f'''(\alpha) \neq 0$$

In general, α is a root of $f(x)$ of multiplicity m if and only if

$$f(\alpha) = \cdots = f^{(m-1)}(\alpha) = 0, \qquad f^{(m)}(\alpha) \neq 0$$

DIFFICULTIES OF MULTIPLE ROOTS

There are two main difficulties with the numerical calculation of multiple roots (by which we mean $m > 1$ in the definition).

1. Methods such as Newton's method and the secant method converge more slowly than for the case of a simple root.

2. There is a large *interval of uncertainty* in the precise location of a multiple root on a computer or calculator.

The second of these is the more difficult to deal with, but we begin with the first for the case of Newton's method.

Recall that we can regard Newton's method as a fixed point method:

$$x_{n+1} = g(x_n), \qquad g(x) = x - \frac{f(x)}{f'(x)}$$

Then we substitute

$$f(x) = (x-\alpha)^m \, h(x)$$

to obtain

$$\begin{aligned} g(x) &= x - \frac{(x-\alpha)^m \, h(x)}{m\,(x-\alpha)^{m-1}\, h(x) + (x-\alpha)^m \, h'(x)} \\ &= x - \frac{(x-\alpha)\, h(x)}{mh(x) + (x-\alpha)\, h'(x)} \end{aligned}$$

Then we can use this to show

$$g'(\alpha) = 1 - \frac{1}{m} = \frac{m-1}{m}$$

For $m > 1$, this is nonzero, and therefore Newton's method is only linearly convergent:

$$\alpha - x_{n+1} \approx \lambda \left(\alpha - x_n\right), \qquad \lambda = \frac{m-1}{m}$$

Similar results hold for the secant method.

There are ways of improving the speed of convergence of Newton's method, creating a modified method that is again quadratically convergent. In particular, consider the fixed point iteration formula

$$x_{n+1} = g(x_n), \qquad g(x) = x - m\frac{f(x)}{f'(x)}$$

in which we assume to know the multiplicity m of the root α being sought. Then modifying the above argument on the convergence of Newton's method, we obtain

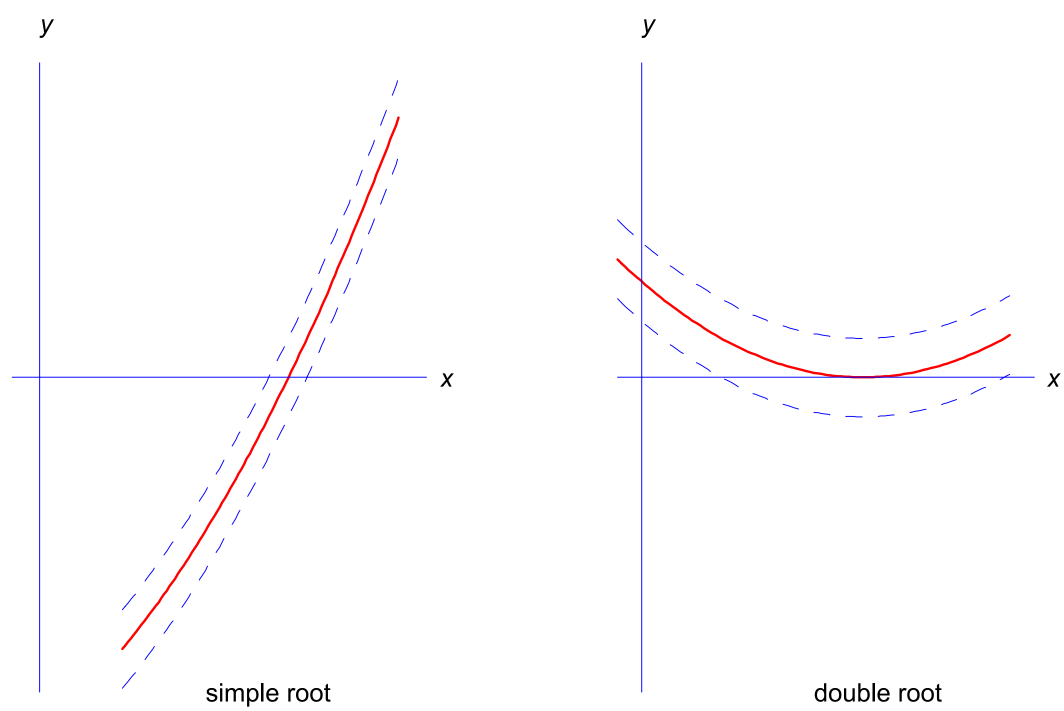
$$g'(\alpha) = 1 - m \cdot \frac{1}{m} = 0$$

and the iteration method will be quadratically convergent.

But this is not the fundamental problem posed by multiple roots.

NOISE IN FUNCTION EVALUATION

Recall the discussion of *noise* in evaluating a function $f(x)$, and in our case consider the evaluation for values of x near to α. In the following figures, the noise as measured by vertical distance is the same in both graphs.



Noise was discussed earlier in §2.2, and Figures 2.1 and 2.2 were of

$$f(x) = x^3 - 3x^2 + 3x - 1 \equiv (x-1)^3$$

Because of the noise in evaluating $f(x)$, it appears from the graph that $f(x)$ has many zeros around $x = 1$, whereas the exact function outside of the computer has only the root $\alpha = 1$, of multiplicity 3.

Any rootfinding method to find a multiple root α that uses evaluation of $f(x)$ is doomed to having a **large** interval of uncertainty as to the location of the root. If high accuracy is desired, then the only satisfactory solution is to reformulate the problem as a new problem $F(x) = 0$ in which α is a simple root of F. Then use a standard rootfinding method to calculate α. It is important that the evaluation of $F(x)$ not involve $f(x)$ directly, as that is the source of the noise and the uncertainly.

EXAMPLE

From the text, consider finding the roots of

$$f(x) = 2.7951 - 8.954x + 10.56x^2 - 5.4x^3 + x^4$$

This has a root to the right of 1. From an examination of the rate of linear convergence of Newton's method applied to this function, one can guess with high probability that the multiplicity is $m = 3$. Then form exactly the second derivative

$$f''(x) = 21.12 - 32.4x + 12x^2$$

Applying Newton's method to this with a guess of $x = 1$ will lead to rapid convergence to $\alpha = 1.1$.

In general, if we know the root α has multiplicity $m > 1$, then replace the problem by that of solving

$$f^{(m-1)}(x) = 0$$

since α is a simple root of this equation.

STABILITY

Generally we expect the world to be *stable*. By this, we mean that if we make a small change in something, then we expect to have this lead to other correspondingly small changes. In fact, if we think about this carefully, then we know this need not be true. We now illustrate this for the case of rootfinding.

Consider the polynomial

$$\begin{aligned} f(x) &= x^7 - 28x^6 + 322x^5 - 1960x^4 \\ &\quad +6769x^3 - 13132x^2 + 13068x - 5040 \end{aligned}$$

This has the exact roots $\{1, 2, 3, 4, 5, 6, 7\}$. Now consider the *perturbed* polynomial

$$\begin{aligned} F(x) &= x^7 - 28.002x^6 + 322x^5 - 1960x^4 \\ &\quad +6769x^3 - 13132x^2 + 13068x - 5040 \end{aligned}$$

This is a relatively small change in one coefficient, of relative error

$$\frac{-.002}{-28} = 7.14 \times 10^{-5}$$

What are the roots of $F(x)$?

Root of $f(x)$	Root of $F(x)$	Error
1	1.0000028	$-2.8E-6$
2	1.9989382	$1.1E-3$
3	3.0331253	-0.033
4	3.8195692	0.180
5	$5.4586758 + .54012578i$	$-.46 - .54i$
6	$5.4586758 - .54012578i$	$-.46 + .54i$
7	7.2330128	-0.233

Why have some of the roots departed so radically from the original values? This phenomena goes under a variety of names. We sometimes say this is an example of an *unstable* or *ill-conditioned* rootfinding problem. These words are often used in a casual manner, but they also have a very precise meaning in many areas of numerical analysis (and more generally, in all of mathematics).

A PERTURBATION ANALYSIS

We want to study what happens to the root of a function $f(x)$ when it is perturbed by a small amount. For some function $g(x)$ and for all small ε, define a perturbed function

$$F_\varepsilon(x) = f(x) + \varepsilon g(x)$$

The polynomial example would fit this if we use

$$g(x) = x^6, \qquad \varepsilon = -.002$$

Let α_0 be a simple root of $f(x)$. It can be shown (using the *implicit differentiation theorem* from calculus) that if $f(x)$ and $g(x)$ are differentiable for $x \approx \alpha_0$, and if $f'(\alpha_0) \neq 0$, then $F_\varepsilon(x)$ has a unique simple root $\alpha(\varepsilon)$ near to $\alpha_0 = \alpha(0)$ for all small values of ε. Moreover, $\alpha(\varepsilon)$ will be a differentiable function of ε. We use this to estimate $\alpha(\varepsilon)$.

The linear Taylor polynomial approximation of $\alpha(\varepsilon)$ is given by

$$\alpha(\varepsilon) \approx \alpha(0) + \varepsilon\, \alpha'(0)$$

We need to find a formula for $\alpha'(0)$. Recall that

$$F_\varepsilon(\alpha(\varepsilon)) = 0$$

for all small values of ε. Differentiate this as a function of ε and using the *chain rule*. Then we obtain

$$\begin{aligned} F_\varepsilon'(\alpha(\varepsilon)) = f'(\alpha(\varepsilon))\alpha'(\varepsilon) \\ +g(\alpha(\varepsilon)) + \varepsilon\, g'(\alpha(\varepsilon))\alpha'(\varepsilon) = 0 \end{aligned}$$

for all small ε. Substitute $\varepsilon = 0$, recall $\alpha(0) = \alpha_0$, and solve for $\alpha'(0)$ to obtain

$$f'(\alpha_0)\alpha'(0) + g(\alpha_0) = 0$$
$$\alpha'(0) = -\frac{g(\alpha_0)}{f'(\alpha_0)}$$

This then leads to

$$\begin{aligned} \alpha(\varepsilon) &\approx \alpha(0) + \varepsilon\, \alpha'(0) \\ &= \alpha_0 - \varepsilon \frac{g(\alpha_0)}{f'(\alpha_0)} \end{aligned} \qquad (*)$$

Example: In our earlier polynomial example, consider the simple root $\alpha_0 = 3$. Then

$$\alpha(\varepsilon) \approx 3 - \varepsilon\frac{3^6}{48} \doteq 3 - 15.2\varepsilon$$

With $\varepsilon = -.002$, we obtain

$$\alpha(-.002) \approx 3 - 15.2(-.002) \doteq 3.0304$$

This is close to the actual root of 3.0331253.

However, the approximation (*) is not good at estimating the change in the roots 5 and 6. By observation, the perturbation in the root is a complex number, whereas the formula (*) predicts only a perturbation that is real. The value of ε is too large to have (*) be accurate for the roots 5 and 6.

DISCUSSION

Looking again at the formula

$$\alpha(\varepsilon) \approx \alpha_0 - \varepsilon \frac{g(\alpha_0)}{f'(\alpha_0)}$$

we have that the size of

$$\varepsilon \frac{g(\alpha_0)}{f'(\alpha_0)}$$

is an indication of the *stability* of the solution α_0.

If this quantity is large, then potentially we will have difficulty. Of course, not all functions $g(x)$ are equally possible, and we need to look only at functions $g(x)$ that will possibly occur in practice.

One quantity of interest is the size of $f'(\alpha_0)$. If it is very small relative to $\varepsilon g(\alpha_0)$, then we are likely to have difficulty in finding α_0 accurately.