Iteration Revisited

Examples From a General Theory

Poul Wulff Pedersen

Margerittens Kvarter 40, DK-2990 Nivaa, Denmark
p-w-p@get2net.dk

Abstract. For 'any' equation $F(x) = 0$ with an isolated solution x^* which we want to compute, it will be shown that quantities p_n can be defined and computed so that

$$xp_{n+1} := xp_n + p_n \cdot F(xp_n)$$

will 'converge' with 'any' speed needed, i.e. starting with an xp_0 in a neighbourhood of x^* , xp_N will for some small value of N give x^* within the precision you need. And it will do so even if the 'standard' iteration $x_{n+1} := x_n + F(x_n)$ is divergent.

1 Introduction

Let

$$F(x) = 0, \quad x, F(x) \in \Re \quad (1)$$

denote an equation with x^* as the solution whose value we want to compute. We assume that x^* is known to exist and that it is the only solution in some interval $I_0 = [L_0; R_0]$. The equation (1) can be brought on an "equivalent iterative form", i.e. an equation $x = f(x)$ with x isolated on one side and with x^* as the only solution in I_0 - for example by adding x on both sides, giving

$$x_{n+1} = x_n + F(x_n), \quad n = 0, 1, 2, \dots \quad (2)$$

In the ideal world of the textbooks this iteration will converge; in practice the situation is often different. And it is easy to see what happens:

Comment 1 x_{n+1} is the x -value where the line from $(x_n, F(x_n))$ with slope $\alpha = -1$ crosses the x -axis.

The usual criterium for an iteration to converge is in this case $-2 < F'(x^*) < 0$ instead of $-1 < f'(x^*) < 1$ for $x_{n+1} := f(x_n)$. The center-value for the interval $[-2; 0]$ is of course the slope $\alpha = -1$ from before.

In the one-dimensional case a suitable function f for an equivalent iteration form $x_{n+1} := f(x_n)$ can of course normally be found without too much trouble, at least as soon as the situation has been analyzed. But, and that is the real problem, you cannot leave this essential part to the average computer. And for N- and infinite-dimensional problems the situation is of course more complicated.

Example 1. The equation

$$x^3 - x - 1 = 0 \quad \text{with} \quad F(x) := x^3 - x - 1$$

has a solution x^* in $I_0 := [1.3 ; 1.4]$ and there is only this one in I_0 . Obvious equivalent iterative forms are

$$x = x^3 - 1, \quad x = \frac{1}{x^2 - 1}, \quad x = \sqrt[3]{x + 1}$$

and it is seen that only the third one gives a convergent iteration. We shall discuss an iteration based on the first one later on.

Instead of trying to change F we shall use 'any' F but define quantities p_n so that

$$xp_{n+1} = xp_n + p_n \cdot F(xp_n), \quad n = 0, 1, 2, \dots \quad (3)$$

will give - for some small N - a value xp_N as close to the solution x^* as we need.

For $\sqrt{A}$, $1/\sqrt{A}$ and division $1/A$ (defined by simple algebraic equations) we can compute an ideal p-value $p^* = p^*(x)$ as a series and an iteration of order $m + 1$ is obtained by taking m terms of this series as p . We shall also discuss how to compute interval approximations.

For a function like Arcsin , p^* will be a series in the inverse function $\sin(x)$. Or we could compute $\sin(x)$ in terms of its inverse $\text{Arcsin}(x)$, etc.

In the general case, p_n is found by [inverse] interpolation.

This paper is about DIMension 1, but much can readily be extended to DIM = N (incl. complex numbers) and to dimension ∞ .

Example 2. To compute $\sqrt{A}$ we use $F(x) := A - x^2$ and the ideal p-value $p^* = p^*(x)$ - which will take any x into $\sqrt{A}$ - is

$$p^*(x) = \frac{1}{\sqrt{A} + x} = \frac{1}{2x} - \frac{1}{8x}\epsilon + \frac{1}{16x}\epsilon^2 - + \dots, \quad \epsilon = A - x^2$$

With one term $p(x) := 1/2x$ we get Newton's formula of order 2; with two terms we get a Halley formula of order 3, etc.

Using that for any approximation x , x and A/x will be on each side of $\sqrt{A}$ we get an inclusion by replacing $\sqrt{A}$ in $p^*(x)$ by these.

1.1 p-iteration, deviation functions.

Before defining p it may eventually be convinient to modify F into a 'deviation' function.

A deviation function for a solution x^* to $F(x) = 0$ is a function $D = D(x)$ so that

D1 : $D(x)$ is defined and can be computed in an interval $I_0 = [L_0; R_0]$ with x^* as an interior point.

D2 : $D(x) = 0$ if and only $x = x^*$,

D3 : A value $D(x_1)$ or a set of a few values $D(x_1), D(x_2), \dots$ will give an indication as to where we should look for better approximation to x^* .

D4 : By making $|D(x)|$ sufficiently small, we can make the error $|E(x)| := |x^* - x|$ as small as we need

$$|D(x_n)| < \delta \quad \Rightarrow \quad |E(x_n)| < \epsilon,$$

and a value $\delta = \delta(\epsilon)$ can be computed from ϵ .

D3 is a little vague - examples later on will clarify the situation. Often F from $F(x) = 0$ determining x^* can be used as D . It may eventually be convinient to have two deviation functions, one indicating (mostly) how close an approximation x_n is to x^* , and the other one (mostly) indicating the direction of a continued search.

The error-function E would give an ideal iteration in the sense that

$$x_{n+1} := x_n + E[x_n]$$

gives $x_{n+1} = x^*$ for any start x_n - except that we cannot compute $E(x)$ without x^* .

Change-of-sign : If the solution x^* is not isolated to an interval with a change-of-sign some extra care is needed. A single value $F(x_1)$ will - at most - indicate how close x_1 is to x^* but we cannot tell in which direction we should search for better approximations. And $|F(x_2)| < |F(x_1)|$ does not guarantee that we should continue in the direction from x_1 towards x_2 . Think of function values of x^2 type coming from the computer, without the benefit of a drawing to help - and the "computer" has to ensure that you have points x_L, x_R on each side of x^* . Normally the p-method to be described will provide x-values so that $|D(x)|$ and thus also $|E(x)|$ are small, but for the computer it may be a difficult problem to ensure if - or when - we have an inclusion.

1.2 p-iterations.

A p-iteration for solving $D(x) = 0$ is an iteration based on a step

$$xp_n := x_n + p_n \cdot D(x_n) \tag{4}$$

where p_n is a number so that the deviation $|D(xp_n)|$ is [much] smaller than $|D(x_n)|$. The ideal p-value p_n^* is a number so that $xp = xp(x_n, p_n^*)$ is the solution x^* , i.e. p_n^* is determined by

$$x^* = x_n + p_n^* \cdot D(x_n) \tag{5}$$

Since x_n is $\neq x^*$, $D(x_n) \neq 0$, and we have

Theorem 1. *For any $x_n \neq x^*$, there exists an ideal p -value $p^* = p^*(x_n)$ determined by*

$$p^*(x_n) := \frac{x^* - x_n}{D(x_n)} \quad (6)$$

Alternatively, we can take x_n to the left side of (4) giving

$$E[x_n] = p^*(x_n) \cdot D(x_n) \quad (7)$$

showing that

Corollary 1. *$p^*(x)$ is the factor that takes deviation $D(x)$ into error $E[x]$.*

1.3 $p^* = p^*(x_n)$ as a mean-value :

From (6) [with $D(x_n) \neq 0$ since $x_n \neq x^*$]

$$p^*(x_n) = \frac{x^* - x_n}{D(x^*) + (x_0 - x^*) \cdot D'(\xi_n)}$$

and, using that $D(x^*) = 0$,

$$p^*(x_n) = \frac{-1}{D'(\xi_n^*)} \quad \xi_n^* \in I_n = ITV]x_n; x^*[\quad (8)$$

for some ξ_n^* in the open interval $ITV]x_n; x^*[$ between x_n and the solution x^* .

If we have an inclusion for $-1/D'(x)$ over I_n we can compute an inclusion for $p^*(\xi_n)$ and thus for x^* .

If we replace ξ_n^* by the approximation x_n we get the Newton iteration

$$NEWT[x_n] = xp_n := x_n - \frac{D(x_n)}{D'(x_n)}$$

The Newton iteration is a for-the-lack-of-better iteration in the sense that we use an endpoint x_n of the interval $ITV]x_n; x^*[$ in which the unknown point ξ_n is known to lie. On the other hand, 'any' point in $]x^*; x_n[$ would not necessarily be better than the end-point x_n .

Concerning ξ , we know [Pof] that in the limit, ξ^* will be midway between x_n and x^* . However, we don't know x^* and we don't know what happens at the beginning of an iteration.

With one more term in Taylor we have

$$p^*(x_n) = \frac{-1}{D'(x^*) + \frac{1}{2} \cdot (x_n - x^*)D''(\xi_2^*)}, \quad \xi_2^* \in ITV]x_0; x^*[\quad (9)$$

etc.

Slope α vs p Let x_α denote the x-value where the line from $P_0 = (x_0; D(x_0))$ with slope α intersect the x-axis

$$x_\alpha = x_0 + \frac{-1}{\alpha} \cdot D(x_0)$$

On the other hand, using a p-value p we get

$$xp = x_0 + p \cdot D(x_0)$$

so

$$xp = x_\alpha \quad \text{for} \quad p = \frac{-1}{\alpha}, \quad \alpha = \frac{-1}{p}.$$

Thus, using a p-value p_α is equivalent to using a slope $\alpha = -1/p$. Cf. Newton's formula.

Linear Interpolation. Using the x_α - formula, we can compute the x-value x_0 where the line through $(x_1, D(x_1))$, $(x_2, D(x_2))$ intersects the x-axis and more generally, the x-value x_c where the value is c

$$x_0 = x_1 - \frac{1}{\alpha} D(x_1) \quad x_c = x_1 - \frac{1}{\alpha} \cdot (D(x_1) - c), \quad \alpha = \frac{D(x_1) - D(x_2)}{x_1 - x_2}$$

Geometrical interpretation : $x_{n+1} := f(x_n)$ in terms of $F(x_n)$.

The 'standard' iteration

$$x_{n+1} := x_n + F(x_n) \tag{10}$$

obtained by adding x on both sides of $F(x) = 0$, is a p-iteration with $p = 1$, corresponding to $\alpha = -1$. x_{n+1} is therefore the x-value where the line with slope -1 through $(x_n, F(x_n))$ intersects the x-axis (cf. Introduction).

Higher order formulas. For solving $F(x) = 0$ we have [Fr 1]

$$x^* = x_0 - \left[\alpha + \frac{1}{2} a_2 \alpha^2 + \frac{1}{6} (3 \cdot a_2^2 - a_3) \alpha^3 + \frac{1}{24} (15a_2^3 - 10a_2 a_3 + a_4) \alpha^4 + \dots \right]$$

where

$$\alpha = \frac{F(x_0)}{F'(x_0)} \quad a_2 = \frac{F''(x_0)}{F'(x_0)}, \quad a_3 = \frac{F'''(x_0)}{F'(x_0)}, \quad a_4 := \frac{F''''(x_0)}{F'(x_0)}, \dots$$

Taking $F(x_0)$ from α we have an expression for $p^*(x_0)$. With α only from [...] we get Newton's formula

$$xp_1(x_0) := x_0 - \frac{1}{F'(x_0)} \cdot F(x_0)$$

With 2 terms we get a Halley formula of order 3

$$xp_2(x_0) := x_0 - \left[\frac{1}{F'(x_0)} + \frac{F''(x_0) \cdot F(x_0)}{2(F'(x_0))^3} \right] F(x_0)$$

etc.

An error theorem. Let $x p_0$ denote the value obtained using x_0, p_0

$$x p_0 := x_0 + p_0 \cdot D(x_0)$$

Since

$$x^* = x_0 + p^*(x_0) \cdot D(x_0)$$

we have

$$x^* - x p_0 = (p^*(x_0) - p_0) \cdot D(x_0)$$

or

$$E[x p_0] = EP[p_0] \cdot D(x_0),$$

Here, $E[x]$ is the error $x^* - x$ and $EP[p]$ is the 'error' on p : $EP[p] := p^* - p$. To reduce $E[x p]$ it would f.ex. be enough to compute a better p -value (cf. Ex. 2).

A divergent iteration made convergent. The iteration for $\sqrt{A}$

$$x_{new} := x + x^2 - A$$

is divergent ($f'(x) = 1 + 2 \cdot x > 1$). But the p -iteration

$$x_{new} := x + p \cdot (x^2 - A), \quad p = \frac{-1}{2x}$$

converges for all A and all starting-values $x_0 > 0$ (Newton).

1.4 p^* for some of the standard functions.

Computing $\sqrt{A}$: Define (for $x > 0$)

$$D(x) := A - x^2, \quad E(x) := \sqrt{A} - x$$

and for

$$x p := x + p \cdot (A - x^2), \quad p^*(x) = \frac{\sqrt{A} + x}{A - x^2} = \frac{1}{\sqrt{A} + x} \quad (11)$$

For an approximation x to $\sqrt{A}$, $A/x^2 \simeq 1$ and we can define a [small] ϵ by

$$\frac{A}{x^2} = 1 + \epsilon, \quad \epsilon := \frac{A}{x^2} - 1 = \frac{A - x^2}{x^2}$$

From (11)

$$p^*(x) = \frac{1}{\sqrt{A} + x} = \frac{1}{x} \left[\frac{1}{2} - \frac{1}{8}\epsilon + \frac{1}{16}\epsilon^2 - \frac{5}{128}\epsilon^3 + \dots \right] \quad (12)$$

Alternatively (cf Corollary 1)

$$A - x^2 = (\sqrt{A} + x) \cdot (\sqrt{A} - x),$$

$$D(x) = (\sqrt{A} + x) \cdot E(x), \quad E(x) = \frac{1}{\sqrt{A} + x} \cdot D(x), \quad (13)$$

Intervals. For any approximation x to $\sqrt{A}$, A/x is also an approximation and x , A/x will be on each side of $\sqrt{A}$. Thus, replacing $\sqrt{A}$ in p^* with $ITV[x; A/x]$

$$\sqrt{A} \in ITV \left[x + \frac{1}{[x; A/x] + x} \cdot (A - x^2) \right]$$

giving

$$\sqrt{A} \in ITV \left[x + \frac{1}{2x} \cdot (A - x^2) ; x + \frac{x}{A + x^2} \cdot (A - x^2) \right] \quad (14)$$

with length

$$LGT(A, x) = \frac{1}{2x \cdot (A + x^2)} (A - x^2)^2 \simeq \frac{1}{4A\sqrt{A}} \cdot (A - x^2)^2$$

The midpoint is

$$x_{mid} := x + \frac{A + 3x^2}{4x(A + x^2)} (A - x^2) \quad \text{with error} \quad = \frac{-1}{4x(A + x^2)} \cdot (E[x])^4$$

With $A = 2$, $x_0 := 1.414$ we get an interval of length $0.323 \cdot 10^{-7}$ and the midpoint has error $-0.920 \cdot 10^{-16}$. With x_0 as a Newton-value

$$A := 2, \quad x_{00} := NEWT[NEWT[1.414]]$$

we get an interval of length $0.652 \cdot 10^{-65}$ and a midpoint with error $-0.353 \cdot 10^{-131}$.

From a mathematical point of view the two expressions for $p^*(x)$ in (11) are equivalent. But from a numerical point of view, they are completely different

Let $I_0 := [L_0; R_0]$ denote an inclusion for $\sqrt{A}$. Using I_0 in the two p^* expressions, we have

$$p^*(x_0) \in ITV \left[\frac{[L_0; R_0] - x_0}{A - x_0^2} \right], \quad p^*(x_0) \in ITV \left[\frac{1}{[L_0; R_0] + x_0} \right]$$

and we get two inclusions. The first one

$$\sqrt{A} \in x_0 + ITV \left[\frac{[L_0; R_0] - x_0}{D(x_0)} \right] \cdot D(x_0)$$

where $D(x_0)$ cancels out, giving trivially $[L_0; R_0]$. The other one

$$\sqrt{A} \in x_0 + ITV \left[\frac{1}{[L_0; R_0] + x_0} \right] \cdot (A - x_0^2)$$

with a length

$$LGT_{new} := \frac{R_0 - L_0}{(L_0 + x_0)(R_0 + x_0)} |A - x_0^2| \simeq \frac{|A - x_0^2|}{4A} \cdot LGT_{orig}$$

which is made as small as we want by reducing $A - x_0^2$ - without changing I_0 .

Division by A only. We define ϵ

$$\frac{x^2}{A} = 1 - \epsilon, \quad \epsilon := 1 - \frac{x^2}{A} = \frac{A - x^2}{A}$$

From this

$$\sqrt{A} = x \frac{1}{\sqrt{1 - \epsilon}} = x + \frac{x}{A} \left[\frac{1}{2} + \frac{3}{8}\epsilon + \frac{5}{16}\epsilon^2 + \frac{35}{128}\epsilon^3 + \dots \right] (A - x^2) \quad (15)$$

Here $D(x) = A - x^2$ so we have the same $p^*(x)$ as before but obviously written in a different form in (15).

Computing $1/\sqrt{A}$. Let x denote an approximation to $1/\sqrt{A}$ so that $1 - A \cdot x^2 \simeq 0$. and we can use

$$D(x) := 1 - A \cdot x^2 = \epsilon$$

Then

$$\frac{1}{A} = x^2 \cdot \frac{1}{1 - \epsilon}, \quad \frac{1}{\sqrt{A}} = x \cdot \frac{1}{\sqrt{1 - \epsilon}}$$

and

$$p^*(x) = \frac{1}{A} \frac{1}{\frac{1}{\sqrt{A}} + x}, \quad \frac{1}{\sqrt{A}} = x + x \left[\frac{1}{2} + \frac{3}{8}\epsilon + \frac{5}{16}\epsilon^2 + \frac{35}{128}\epsilon^3 + \dots \right] (1 - A x^2)$$

It is seen that x and $x(2 - A x^2)$ will be on each side of $1/\sqrt{A}$ and using

$$x_{n+1} = x_n + \frac{1}{2} \cdot x_n \epsilon_n \quad (16)$$

we get an inclusion

$$x_1 := x + 0.5 x (1 - A x^2), \quad x_2 := x + 0.5 x (2 - A x^2) (1 - A x^2)$$

For

$$x_w(x) := \frac{1}{4} x_1(x) + \frac{3}{4} x_2(x)$$

E^2 error terms will cancel and we get the 3.order formula.

2 Computing Arcsin(t) using its inverse function sin(x).

To compute $\arcsin(t_0)$ for t_0 in $]0 ; 1[$ we should 'solve'

$$x = \arcsin(t_0), \quad t_0 = \sin(x)$$

so we have a deviation function and a p-iteration

$$D(x) := t_0 - \sin(x) \quad xp = x + p \cdot [t_0 - \sin(x)]$$

with

$$p^*(x) = \frac{\arcsin(t_0) - x}{t_0 - \sin(x)}$$

Let

$$\epsilon := D(x) = t_0 - \sin(x), \quad t_0 = \sin(x) + \epsilon$$

Then

$$p^*(x) = \frac{\arcsin(\sin(x) + \epsilon) - x}{\epsilon} = \sum c_n \cdot \epsilon^n$$

with

$$c_0 = \frac{1}{\cos(x)}, \quad c_1 = \frac{\sin(x)}{2 \cdot \cos(x)^3}, \quad c_2 = \frac{3 - 2 \cdot \cos(x)^2}{6 \cdot \cos(x)^5}$$

$$c_3 = \frac{5 \cdot \sin(x) - 2 \cdot \sin(x) \cdot \cos(x)^2}{8 \cdot \cos(x)^7} \quad \dots$$

Here,

$$xp_0(x) := x + \frac{1}{\cos(x)}(t_0 - \sin(x)) \quad (p = c_0)$$

is the Newton-iteration of order 2. With $p = c_0 + \dots + c_n \cdot \epsilon^n$, i.e.

$$xp_n(x) := x + (c_0 + c_1 \cdot \epsilon + \dots + c_n \cdot \epsilon^n) \cdot (t_0 - \sin(x)), \quad \epsilon = t_0 - \sin(x)$$

the order is $n + 2$. Only one value of $\sin(x)$ is needed at each iteration step.

Example : $t_0 = 0.7$, $x_0 = 0.75$. Errors are

$n = 0$	$= 1$	$= 2$
$E_1 = 0.105(-4)$	$= -0.640(-7)$	$= 0.439(-9)$
$E_2 = 0.536(-10)$	$= -0.169(-21)$	$= 0.354(-37)$
$E_3 = 0.141(-20)$	$= -0.314(-65)$	$= 0.150(-149)$

with orders 2, 3, 4.

We can also construct algorithms of any order for π , with only one value of $\sin(x)$ needed for each step - essentially by solving $\sin(x) = 0$.

3 Inverse interpolation.

To find a solution x^* to $F(x) = 0$ we replace F by a deviation function D and define

$$DP(p) := D(x_0 + p \cdot D(x_0))$$

i.e. deviation as a function of p for a fixed x_0 . The goal is to find a p -value p_0^* which will make DP so small that the corresponding x -value $x_0 + p_0^* \cdot D(x_0)$ is so close to the exact solution x^* as we need. For any value of p we can compute $DP(p)$. The inverse problem, to compute the p -value for which DP has a certain

value [in this case 0], is solved using [inverse] interpolation : Let $p_1, p_2, \dots, p_N$ denote a set of p-values so that the points

$$x_i := x_0 + p_i \cdot D(x_0), \quad i = 1, 2, \dots, N$$

are 'close' to x^* [as shown by D being small] and let

$$DP_i := DP(p_i) = D(x_i)$$

be these deviations. Finally, let $POL_N = POL_N(DP)$ denote the interpolation polynomial through (DP_i, p_i) . The value

$$p_{00} := POL_N(0)$$

is then used as an approximation to the value p_0^* , and the validity of this choice is controlled by computing $DP(p_{00})$. If $|DP(p_{00})|$ is small enough we have the point $x_{00} := x_0 + p_{00} \cdot D(x_0)$ we wanted; otherwise we modify the set of collocation points, for example by replacing one of the p-values by the new value p_{00} , for which we already have DP . But we may also use a new set of p-values centered around P_{00} - as used below.

Example 1.1 continued.

$$D(x) = x^3 - x - 1, \quad D(1.3) = -0.103, \quad D(1.4) = 0.344$$

$$a_1 := POL_1(0; 1.3, 1.4), \quad D(a_1) \simeq -0.007134.$$

$$a_2 := POL_2(0; a_1 - \Delta, a_1, a_1 + \Delta), \quad \Delta = 10^{-3} \quad D(a_2) \simeq -0.195010^{-7}.$$

$$a_3 := POL_2(0; a_2 - \Delta, a_2, a_2 + \Delta), \quad \Delta = 10^{-9}, \quad D(a_3) \simeq -0.612210^{-24}.$$

a_1 is obtained by linear interpolation (POL_1 of degree 1). a_2 is obtained by interpolation of degree 2 with points [actually p-values corresponding to] $a_1 - \Delta, a_1, a_1 + \Delta$; a_1 being the point found with POL_1 . a_3 similarly with a polynomial of degree 2 using a_2 .

It may be convenient to replace p by $p_1 \cdot APd$ where APd is some approximation to $-1/D'(\xi)$ making $p_1 \simeq 1$. APd can also be used for estimating Error from Deviation and $-1/D'(x)$ could be used for inclusions.

References

1. W. Kahan, An example given during an IBM conference Symbolic Mathematical Computation, Oberlech, Austria 1991.
2. Oliver Abert: Precision Basic in *Precise Numerical Analysis*, 1988. More recent: *Precise Numerical Analysis Using C++*, Academic Press, 1998.
3. Ward Cheney and David Kincaid: Numerical Mathematics and Computing, 1999. Brooks/Cole Publishing Company. (Problem 3.2, 37).
4. Carl-Erik Fröberg: *Numerical Mathematics*, 1985. The 'Higher order formulas' are in Schröder's 1870 paper but I prefer Fröberg's version !.
5. Esteban I. Poffald, *The Remainder in Taylor's Formula*, American Mathematical Monthly, March 1990.