Homework no. 7

Let P be a polynomial of degree n with real coefficients:

$$P(x) = a_0x^n + a_1x^{n-1} + \cdots + a_kx^{n-k} + \cdots + a_{n-1}x + a_n \ , \quad a_0 \neq 0$$

and $\epsilon = 10^{-p}$ the computation error and $k_{\max}$ the maximum number of iterations in Halley's method.

Compute the interval $[-R, R]$ where all the real roots of the polynomial P can be found. Implement Halley's method for approximating the real roots of a polynomial. In all computations use Horner's algorithm for calculating the value of the polynomial in a point. Approximate as many as possible real roots of the polynomial P with Halley's method, starting from distinct values for x_0. Display the results on screen and also write them in a file. Write in the file only the distinct roots (2 real values v_1 and v_2 are considered distinct if $|v_1 - v_2| > \epsilon$).

Bonus 20 pt.: Implement the methods (N^4) and (N^5) from the article posted here, for approximating the root of a general function f, not necessarily polynomial. The derivatives of function f will be declared in your program as the function f is.

Halley's Method for Approximating Real Roots of Polynomials

Let P be a polynomial of degree n:

$$P(x) = a_0x^n + a_1x^{n-1} + \cdots + a_n \ , \quad (a_0 \neq 0) \tag{1}$$

A root for a polynomial P is a real or complex number $r \in \mathbb{R}$ or $r \in \mathbb{C}$ for which:

$$P(r) = a_0r^n + a_1r^{n-1} + \cdots + a_n = 0.$$

All the polynomials of degree n, with real coefficients have n real and/or complex roots. If $r = c + id$ is a complex root for polynomial P then $\bar{r} = c - id$ is also a complex root of the polynomial P. If $r \in \mathbb{R}$ is a real root for polynomial P then:

$$P(x) = (x - r)Q(x) \ , \ Q \text{ is a polynomial of degree } n - 1.$$

If the complex numbers $c \pm id$ are complex roots for polynomial P then:

$$P(x) = (x^2 - 2c\,x + c^2 + d^2)Q(x)\ ,\ Q \text{ is a polynomial of degree } n-2.$$

(the second degree polynomial $x^2 - 2c\,x + c^2 + d^2$ has the roots $c \pm id$)

All the real roots of the polynomial P are in the interval $[-R,\ R]$ where R is given by:

$$R = \frac{|a_0| + A}{|a_0|} \quad , \quad A = \max\{|a_i|\ ;\ i = \overline{1,n}\} \tag{2}$$

The real root $x^* \in [-R,\ R]$ of polynomial P defined by (1), is approximated by building a sequence of real numbers, $\{x_k\}$, that converges to a real root of P ($x_k \longrightarrow x^*$ for $k \to \infty$) .
In order to build the sequence $\{x_k\}$, one needs the first element x_0, then, the other elements are computed in the following way (x_{k+1} is computed using x_k):

$$\begin{aligned} x_{k+1} &= x_k - \frac{1}{a_k}\ , \\ a_k &= \frac{P'(x_k)}{P(x_k)} - \frac{P''(x_k)}{2P'(x_k)}. \end{aligned} \tag{3}$$

This method is called Halley's method and can be applied not only for approximating roots of polynomials but also for roots of any nonlinear twice differentiable function.

Important remark: The way the first element of the sequence, x_0, is selected can influence the convergence of the sequence x_k to x^* (or the divergence). Usually, a selection of the initial iteration x_0 in the neighborhood of a root x^*, guarantees the convergence $x_k \longrightarrow x^*$ for $k \to \infty$.

Not all the elements of the sequence $\{x_k\}$ must be memorized, in order to obtain an approximation for the root we only need the 'last' computed value x_{k_0}. A value $x_{k_0} \approx x^*$ approximates a root (thus, is the 'last' computed element of the sequence) when the difference between two successive elements of the sequence is sufficiently small, i.e.:

$$|x_{k_0} - x_{k_0-1}| < \epsilon$$

where ϵ is the precision with which we want to approximate the root x^*.

A possible approximation scheme for Halley's method for approximating the root x^*, is the following:

Halley's Method

x randomly chosen ; $k = 1$;
(for the convergence of the sequence $\{x_k\}$ is preferable to choose x_0 in the neighborhood of a root)
do
{
$\star$ compute $A = 2[P'(x)]^2 - P(x)P''(x)$;
$\star$ if ($|A| < \epsilon$) STOP;
//(try restarting the algorithm, changing x_0)
$\star$ compute $\Delta = \dfrac{2P(x)P'(x)}{A}$;
$\star$ $x = x - \Delta$;
$\star$ $k++$;
}
while ($|\Delta| \geq \epsilon$ and $k \leq k_{\max}$ and $|\Delta| \leq 10^8$)
if ($|\Delta| < \epsilon$) $x \approx x^*$;
else *divergence* ; //(try new values for x_0)

$k_{\max} \in \{1000, 2000, 5000, ...\}$

Horner's method for computing $P(v)$

Let P be a polynomial of degree p:

$$P(x) = c_0x^p + c_1x^{p-1} + \cdots + c_{p-1}x + c_p \ , \quad (c_0 \neq 0)$$

We can write polynomial P also as:

$$P(x) = ((\cdots(((c_0x + c_1)x + c_2)x + c_3)x + \cdots)x + c_{p-1})x + c_p$$

Taking into account this grouping of the coefficients, one obtains an efficient way to compute the value of polynomial P in any point $v \in \mathbf{R}$, this procedure is knowns as *Horner's method*:

$$\begin{aligned} d_0 &= c_0 \, , \\ d_i &= c_i + d_{i-1}v \ , \quad i = \overline{1, p} \end{aligned} \tag{4}$$

In the above sequence:

$$P(v) := d_p$$

the rest of the computed elements of the sequence $d_i, i = 1, ..., p-1$, are the coefficients of the quotient polynomial Q, obtained in the division:

$$\begin{aligned} P(x) &= (x - v)Q(x) + r \ , \\ Q(x) &= d_0x^{p-1} + d_1x^{p-2} \cdots + d_{p-2}x + d_{p-1} \ , \\ r &= d_p = P(v). \end{aligned}$$

Computing $P(v)$ (d_p) with formula (4) can be performed using only one real variable $d \in \mathbf{R}$ instead of using a vector $d \in \mathbf{R}^p$.

Examples

$$
\begin{aligned}
P(x) &= (x-1)(x-2)(x-3) = x^3 - 6x^2 + 11x - 6 \ , \\
& \quad a_0 = 1.0\,, \quad a_1 = -6.0\,, \quad a_2 = 11.0\,, \quad a_3 = -6. \\
P(x) &= (x-\frac{2}{3})(x-\frac{1}{7})(x+1)(x-\frac{3}{2}) \\
&= \frac{1}{42}(42x^4 - 55x^3 - 42x^2 + 49x - 6) \\
& \quad a_0 = 42.0\,, \quad a_1 = -55.0\,, \quad a_2 = -42.0\,, \quad a_3 = 49.0\,, \quad a_4 = -6.0. \\
P(x) &= (x-1)(x-\frac{1}{2})(x-3)(x-\frac{1}{4}) \\
&= \frac{1}{8}(8x^4 - 38x^3 + 49x^2 - 22x + 3) \\
& \quad a_0 = 8.0\,, \quad a_1 = -38.0\,, \quad a_2 = 49.0\,, \quad a_3 = -22.0\,, \quad a_4 = 3.0. \\
P(x) &= (x-1)^2(x-2)^2 \\
&= x^4 - 6x^3 + 13x^2 - 12x + 4 \\
& \quad a_0 = 1.0\,, \quad a_1 = -6.0\,, \quad a_2 = 13.0\,, \quad a_3 = -12.0\,, \quad a_4 = 4.0, \\
f(x) &= e^x - \sin(x) \quad , \quad f'(x) = e^x - \cos(x) \quad , \\
f''(x) &= e^x + \sin(x) \quad , \quad x^* = -3.18306301193336.
\end{aligned}
$$