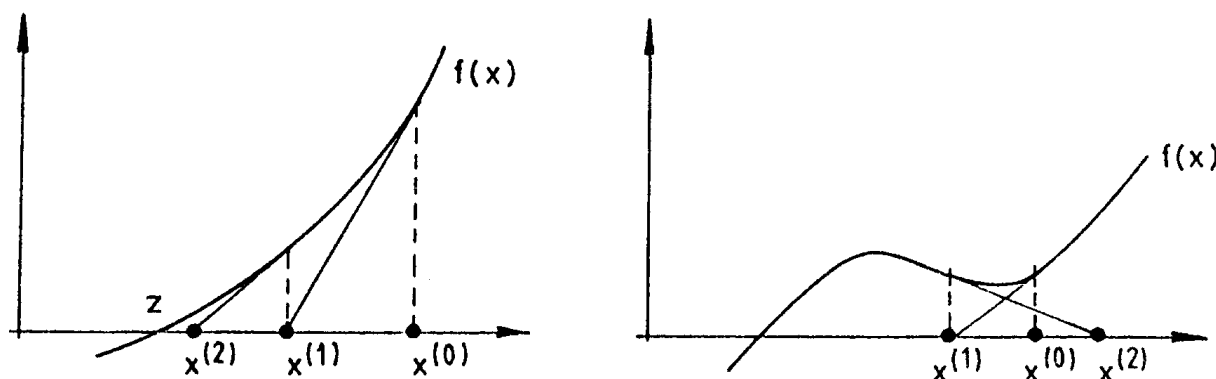


## NUMERICAL METHODS FOR NONLINEAR EQUATIONS AND SYSTEMS OF EQUATIONS

### NEWTON'S METHOD FOR EQUATIONS

Here we try to approximate a root  $z$  of the equation  $f(x)=0$  in such a way that we use the tangential straight line of at different points. Starting from  $x^{(0)}$  the left figure shows a successful trial and the right figure shows an unsuccessful one:



So,  $x^{(k+1)}$  approximation of  $z$  can be got by Taylor's series of the function  $f(x)$  used only the constant and linear term of it. Hence  $x^{(k+1)}$  is the solution of the linear equation

$$f(x^{(k)}) + f'(x^{(k)})(x - x^{(k)}) = 0.$$

If this equation has unique solution for  $k = 0, 1, 2, \dots$ , then

$$x^{(k+1)} = x^{(k)} - \frac{f(x^{(k)})}{f'(x^{(k)})},$$

and it is called the formula of Newton's method for the equation  $f(x) = 0$ .

#### THEOREM.

Let the function  $f(x) = 0$  be twice continuously differentiable on  $[a, b]$ . Assume that there is a root of the equation  $f(x) = 0$  on the interval  $[a, b]$  and  $f'(x) \neq 0$   $f''(x) \neq 0$   $\forall x \in [a, b]$ .

Then Newton's iteration converges (starting from  $x^{(0)} = a$ , if  $f'(x) \cdot f''(x) < 0$  and from  $x^{(0)} = b$ , if  $f'(x) \cdot f''(x) > 0$ ) to the unique solution on  $[a, b]$  and

$$|x^{(k+1)} - z| \leq \frac{M}{2m} |x^{(k+1)} - x^{(k)}|^2,$$

where

$$0 < m \leq |f'(x)|, M \geq |f''(x)|, \forall x \in [a, b].$$

## NUMERICAL EXAMPLE

Determine the root of the equation

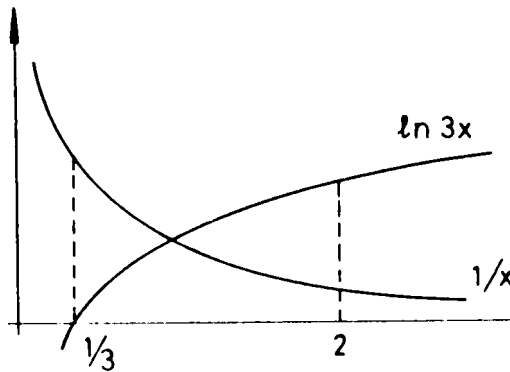
$$\ln(3x) - \frac{1}{x} = 0$$

by Newton's method and with error bound  $\varepsilon = 5 \cdot 10^{-5}$ .

(1) We can get a suitable interval (for which the conditions of our theorem are satisfied) by graphical way and (or) tabling. If we write our equation into the form

$$\ln(3x) = \frac{1}{x}$$

and represent the function of left and right sides (roughly), then we get a (rough) estimate about the root:



We can see the equation has unique root ( $z$ ) and it is on the interval  $(1/3, 2)$ . Moreover, by the table

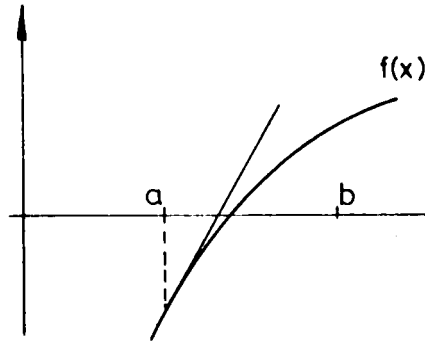
$x$	$1/3$	$2/3$	$1$
$\ln 3x$	$0$	$0.69$	$1.10$
$1/x$	$3$	$1.50$	$1$

$z$  is on the interval  $(2/3, 1)$ , too.

(2) Now let us examine the conditions concerning the convergence of the method

a)  $f'(x) = \frac{x+1}{x^2} \neq 0$ ,  $f''(x) = -\frac{x+2}{x^3} \neq 0$ ,  $\forall x \in \left[\frac{2}{3}, 1\right]$ .

b) Since  $f'(x) > 0$  and  $f''(x) < 0 \forall x \in \left[\frac{2}{3}, 1\right]$ , so on a small neighbourhood of  $z$  the graph of  $f(x)$  is the following:



Hence  $x^{(0)} = a = \frac{2}{3}$  the right choice.

c) Since  $f'(x) > 0$  and  $f'(x)$  is strictly monotone decreasing on  $\left[\frac{2}{3}, 1\right]$ ,

$$|f'(x)| \geq f'(1) = 2 = m.$$

Furthermore,  $f''(x) < 0$  and it is strictly monotone increasing on  $\left[\frac{2}{3}, 1\right]$ . Therefore,  $|f''(x)|$  is strictly monotone decreasing and

$$|f''(x)| \leq \left|f''\left(\frac{2}{3}\right)\right| = 9 = M.$$

(3) The computation of the iterations and the estimation of the errors:

$$x^{(0)} = \frac{2}{3}$$

$$x^{(1)} = 0.881827, \quad |x^{(1)} - z| \leq 0.104161$$

$$x^{(2)} = 0.948421, \quad |x^{(2)} - z| \leq 0.009978$$

$$x^{(3)} = 0.952451, \quad |x^{(3)} - z| \leq 0.000037$$

As  $|x^{(3)} - z| \leq \varepsilon$  is fulfilled,  $x^{(3)}$  can be considered a suitable approximation of the root. (If we used more rough bounds for  $m$  and  $M$  then the value  $M/2m$  would be larger and we might need more iterations to guarantee the same accuracy.)

## NEWTON'S METHOD FOR SYSTEMS OF EQUATIONS

Let the nonlinear system of equations

$$\begin{aligned}f_1(x_1, x_2, \dots, x_n) &= 0 \\f_2(x_1, x_2, \dots, x_n) &= 0 \\&\vdots \\f_n(x_1, x_2, \dots, x_n) &= 0\end{aligned}$$

be given. In shortly form:

$$f(x) = 0, \quad f: \mathbb{R}^n \rightarrow \mathbb{R}^n.$$

Starting from the initial vector  $x^{(0)} = \{x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)}\}$  make  $x^{(1)}$  on such a way that the functions  $f_1, f_2, \dots, f_n$  are made „linear" by the first  $(n + 1)$  terms of their Taylor series at  $x^{(0)}$ . It means the surfaces are replaced by their tangential planes concerning  $x^{(0)}$ . Hence, we can get  $x^{(1)}$  from the system of linear equations

$$\begin{aligned}f_1(x^{(0)}) + \partial_1 f_1(x^{(0)})(x_1 - x_1^{(0)}) + \dots + \partial_n f_1(x^{(0)})(x_n - x_n^{(0)}) &= 0 \\&\vdots \\f_n(x^{(0)}) + \partial_1 f_n(x^{(0)})(x_1 - x_1^{(0)}) + \dots + \partial_n f_n(x^{(0)})(x_n - x_n^{(0)}) &= 0\end{aligned}$$

if there exists unique solution of it. Using matrices we can write

$$\begin{bmatrix} f_1(x^{(0)}) \\ \vdots \\ f_n(x^{(0)}) \end{bmatrix} + \begin{bmatrix} \partial_1 f_1(x^{(0)}) & \dots & \partial_n f_1(x^{(0)}) \\ \vdots & & \vdots \\ \partial_1 f_n(x^{(0)}) & \dots & \partial_n f_n(x^{(0)}) \end{bmatrix} \begin{bmatrix} x_1 - x_1^{(0)} \\ \vdots \\ x_n - x_n^{(0)} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Shortly

$$f(x^{(0)}) + J(x^{(0)})(x - x^{(0)}) = 0,$$

where  $J$  is called Jacobian matrix of the function system  $f_1, f_2, \dots, f_n$ . If there exists unique solution ( $\det J(x^{(0)}) \neq 0$ ) then we can express  $x = x^{(1)}$  from this equation with using the inverse matrix of  $J(x^{(0)})$

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

So, the formula of Newton's method (written  $x^{(k)}$  and  $x^{(k+1)}$  into the places of  $x^{(0)}$  and  $x^{(1)}$  respectively)

$$x^{(k+1)} = x^{(k)} - J^{-1}(x^{(k)})f(x^{(k)}), \quad k = 0, 1, 2, \dots$$

The following theorem gives a fundamental characterization of the method.

## THEOREM

Suppose that each of the partial derivatives  $\partial_i f_j$ ,  $i, j = 1, 2, \dots, n$  is continuous on a neighbourhood of a solution  $z$  and the matrix  $J(z)$  is invertable. Then Newton's method is convergent starting from sufficiently short distance. And if the partial derivatives  $\partial_i \partial_j f_k$ ,  $i, j, k = 1, 2, \dots, n$  are also continuous on the above mentioned neighbourhood then the error decreases at least quadratically, that is

$$|x^{(k+1)} - z| \leq c \cdot |x^{(k)} - z|^2,$$

where  $c$  is a positive constant.

## REMARK

It is not practical to compute  $x^{(k+1)}$  by the formula

$$x^{(k+1)} = x^{(k)} - J^{-1}(x^{(k)})f(x^{(k)})$$

(here we have to compute the inverse of a matrix of order  $n$  in every step), but it is practical to solve the original system of linear equations for the unknowns  $x^{(k+1)} - x^{(k)}$ , and then we can get the values  $x^{(k+1)}$  immediately. Namely, solving system of linear equations requires smaller work than the computing of the inverse matrix.

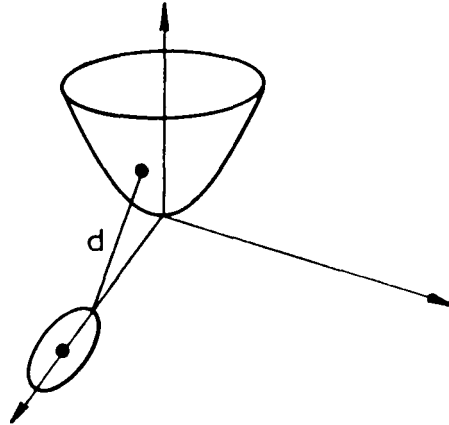
## NUMERICAL EXAMPLE.

Let a circle and a paraboloid be given by the formulas

$$\begin{array}{ll} x = 5 + 2 \cos t & x = u \\ y = 2 \sin t \quad t \in [-\pi, \pi], \text{ and } & y = v \\ z = 0 & z = u^2 + v^2. \end{array}$$

respectively. Determine the distance of the given curve and surface.

Geometrically we have to find the shortest way starting from a circle point ending at a paraboloid point:



The square of the distance

$$D = d^2 = (5 + 2 \cos t - u)^2 + (2 \sin t - v)^2 + (u^2 + v^2)^2.$$

For the extremum points it is a necessary condition that

$$\frac{\partial D}{\partial t} = 4 \sin t (u - 5) - 4v \cos t = 0$$

$$\frac{\partial D}{\partial u} = -10 - 4 \cos t + 2u + 4u^3 + 4uv^2 = 0$$

$$\frac{\partial D}{\partial v} = -4 \sin t + 2v + 4vu^2 + 4v^3 = 0.$$

This way we have a nonlinear system of equations. Now introduce the notations

$$x = \{x_1, x_2, x_3\},$$

where  $x_1 = t$ ,  $x_2 = u$ ,  $x_3 = v$  and try to solve the nonlinear system by Newton's method starting from

$$x^{(0)} = \{2, 2, 2\}.$$

We determine the vector  $e = x^{(k)} - x^{(k+1)}$  from the linear system of equations

$$J(x^{(k)})(x^{(k)} - x^{(k+1)}) = f(x^{(k)}) \quad k = 0, 1, 2, \dots,$$

again and again. For  $k = 0$  the solution of the linear system is

$$e = \{-0.895051, 0.613877, 0.700131\}$$

and

$$x^{(1)} = x^{(0)} - e = \{2.895051, 1.386123, 1.299869\}.$$

The next approximations are

$$x^{(2)} = \{2.954787, 1.112591, 0.742890\}$$

$$x^{(3)} = \{3.066840, 1.047144, 0.195795\}$$

$$x^{(4)} = \{3.131045, 1.019094, 0.042924\}$$

$$x^{(5)} = \{3.140318, 1.000788, 0.001324\}$$

$$x^{(6)} = \{3.141592, 1.000001, 0.000002\}$$

In the practice the following stop-criteria are usually used:

- a) the iteration is stopped if  $|f_i(x_1, x_2, \dots, x_n)| < \varepsilon \quad \forall i$ , where  $\varepsilon > 0$  is a given positive number;
- b) the iteration is stopped if the distance of the last two approximating vectors (in some norm) is small enough.

If we use the criteria a) and  $\varepsilon = 5 \cdot 10^{-4}$ , then  $x^{(6)}$  is the last iteration, since

$$|f_1(x^{(6)})| = 2.2 \cdot 10^{-7} < \varepsilon, \quad |f_2(x^{(6)})| = 7.1 \cdot 10^{-6} < \varepsilon, \quad |f_3(x^{(6)})| = 4.1 \cdot 10^{-6} < \varepsilon.$$