Numerical Analysis 23 **Solutions for sheet 4**

Aitken's Δ^2 Method, Newton-Raphson in higher dimensions, Steepest Decent

---

1. The fixed point iteration is $x_{n+1} = \frac{1}{2}\sqrt{10 - x_n^3}$. Aitken's Δ^2 method uses the sequence of the x_n to define a new sequence in the form

$$\hat{x}_n = x_n - \frac{(\Delta x_n)^2}{\Delta^2 x_n} = x_n - \frac{(x_{n+1} - x_n)^2}{(x_{n+2} - 2x_{n+1} + x_n)} .$$

The following table shows the numerical results for the two sequences. The starting point is $x_0 = 1.5$. One can see that the sequence obtained from Aitken's method converges much

n	x_n	$\hat{x}_n$
0	1.5000000	1.361886
1	1.2869538	1.364329
2	1.4025408	1.364999
3	1.3454584	1.365169
4	1.3751703	1.365214
5	1.3600942	
6	1.3678470	

quicker to the exact solution $x^* = 1.3652300$.

2. Newton's method for a system $\mathbf{f}(\mathbf{x}) = 0$ consists of the iteration

$$\mathbf{x}^{(\mathbf{m+1})} = \mathbf{x}^{(\mathbf{m})} - J^{-1}(\mathbf{x}^{(\mathbf{m})})\,\mathbf{f}(\mathbf{x}^{(\mathbf{m})}),$$

where $J(\mathbf{x})$ is the Jacobian matrix. Its matrix elements are $J_{ij} = \partial f_i/\partial x_j$.

(a) Let A be a non singular $n \times n$ matrix. The linear system $A\mathbf{x} = \mathbf{b}$ can be written in the form $\mathbf{f}(\mathbf{x}) = A\mathbf{x} - \mathbf{b} = 0$. Alternatively, we can write it in component form as

$$f_i(x_1, \ldots, x_n) = \sum_{j=1}^{n} A_{ij} x_j - b_i = 0 , \qquad i = 1, \ldots, n ,$$

where A_{ij} are the matrix elements of A, and x_i and b_i are the components of $\mathbf{x}$ and $\mathbf{b}$, respectively. The Jacobian matrix J for this system has matrix elements

$$J_{ij}(\mathbf{x}) = \frac{\partial f_i(\mathbf{x})}{\partial x_j} = A_{ij} .$$

One sees that the Jacobian matrix J is identical to the matrix A, and it does not dependent on $\mathbf{x}$. Let us denote the initial point for Newton's method by $\mathbf{x}^{(0)}$. One step of Newton's method results in

$$\mathbf{x}^{(1)} = \mathbf{x}^{(0)} - J^{-1}(\mathbf{x}^{(0)})\,\mathbf{f}(\mathbf{x}^{(0)}) = \mathbf{x}^{(0)} - A^{-1}(A\mathbf{x}^{(0)} - \mathbf{b}) = A^{-1}\mathbf{b} .$$

This is the exact solution of the linear system $A\mathbf{x} = \mathbf{b}$.

(b) The system of equations is

$$f(x,y) = ax^2 + by + c = 0\,, \qquad\qquad g(x,y) = dx + e = 0\,,$$

from which we obtain the solution as $x^* = -e/d$ and $y^* = -ae^2/(bd^2) - c/b$. The Jacobian matrix J for this system and its inverse J^{-1} are given by

$$J = \begin{bmatrix} f_x & f_y \\ g_x & g_y \end{bmatrix} = \begin{bmatrix} 2ax & b \\ d & 0 \end{bmatrix}, \qquad J^{-1} = \frac{1}{(-bd)} \begin{bmatrix} 0 & -b \\ -d & 2ax \end{bmatrix}.$$

One step of Newton's method with initial point (x_0, y_0) results in

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} - \frac{1}{(-bd)} \begin{bmatrix} 0 & -b \\ -d & 2ax_0 \end{bmatrix} \begin{bmatrix} ax_0^2 + by_0 + c \\ dx_0 + e \end{bmatrix} = \frac{1}{bd} \begin{bmatrix} -be \\ adx_0^2 - cd + 2aex_0 \end{bmatrix}.$$

The second step in Newton's iteration can be obtained from this result by replacing x_1 and y_1 by x_2 and y_2, and also x_0 and y_0 by x_1 and y_1. We find

$$\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \frac{1}{bd} \begin{bmatrix} -be \\ adx_1^2 - cd + 2aex_1 \end{bmatrix} = \begin{bmatrix} -e/d \\ -ae^2/(bd^2) - c/b \end{bmatrix},$$

where we inserted the value $x_1 = -e/d$ from the previous equation. This agrees with the exact solution given before.

(c) Here we consider the system

$$f(x,y) = x^2 - y^2 = 0\,, \qquad\qquad g(x,y) = 1 + xy = 0\,.$$

From the first equation we obtain $y = \pm x$ and from the second equation $y = -1/x$. These relations are only compatible if we chose the negative sign in the first relation: $y = -x$. Then we obtain from the second relation $x^2 = 1$. We conclude that the solutions are $(x,y) = (1,-1)$ and $(x,y) = (-1,1)$. The Jacobian matrix J for this system and its inverse J^{-1} are given by

$$J = \begin{bmatrix} f_x & f_y \\ g_x & g_y \end{bmatrix} = \begin{bmatrix} 2x & -2y \\ y & x \end{bmatrix}, \qquad J^{-1} = \frac{1}{2(x^2+y^2)} \begin{bmatrix} x & 2y \\ -y & 2x \end{bmatrix}.$$

The initial point is $(x_0, y_0) = (\alpha, \alpha)$, and one step of Newton's method results in

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} \alpha \\ \alpha \end{bmatrix} - \frac{1}{4\alpha^2} \begin{bmatrix} \alpha & 2\alpha \\ -\alpha & 2\alpha \end{bmatrix} \begin{bmatrix} 0 \\ 1 + \alpha^2 \end{bmatrix} = \frac{\alpha^2 - 1}{2\alpha} \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

This shows that if one iteration point is on the line $y = x$ then the following iteration point is also on the line $y = x$. Hence the iteration can never converge to one of the two solutions because they are not on this line.

3. (a) The function $g(x,y) = x^2 + y^2 + axy$ with $0 < |a| < 2$ can be written in the form

$$g(x,y) = x^2 + y^2 + axy = \left(x + \frac{ay}{2}\right)^2 + \left(1 - \frac{a^2}{4}\right) y^2 \,,$$

from which we see that the minimum is at $(x,y) = (0,0)$.

We start at an initial point (x_0, y_0) and apply the method of steepest descent. This involves going in the direction of the gradient of $g(x,y)$

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} - \alpha \nabla g(x_0, y_0) = \begin{bmatrix} x_0 - \alpha(2x_0 + ay_0) \\ y_0 - \alpha(2y_0 + ax_0) \end{bmatrix} . \tag{1}$$

The value of α is determined by requiring that it minimizes the function $g(x_1, y_1)$. If the initial point is of the form $(x_0, y_0) = (\beta, \beta)$ then we find

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} \beta(1 - \alpha(2+a)) \\ \beta(1 - \alpha(2+a)) \end{bmatrix} .$$

One can see that the choice $\alpha = 1/(2+a)$ results in $(x_1, y_1) = (0,0)$ which is the absolute minimum of $g(x,y)$.

Can we arrive at the point $(x_1, y_1) = (0,0)$ from a general starting point? We conclude from equation (1) that this wound require $\alpha = x_0/(2x_0 + ay_0)$ and simultaneously $\alpha = y_0/(2y_0 + ax_0)$. This is only possible if

$$0 = \frac{x_0}{2x_0 + ay_0} - \frac{y_0}{2y_0 + ax_0} = \frac{a(x_0^2 - y_0^2)}{(2x_0 + ay_0)\,(2y_0 + ax_0)} \,,$$

i.e. for $x_0 = \pm y_0$. This is not true for a general starting point.

(b) If $g(x,y)$ is a function of $(x^2 + y^2)$, that is $g(x,y) = G(x^2 + y^2)$, then equation (1) is replaced by

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} - \alpha \nabla g(x_0, y_0) = \begin{bmatrix} x_0 - 2\alpha x_0 G'(x_0^2 + y_0^2) \\ y_0 - 2\alpha y_0 G'(x_0^2 + y_0^2) \end{bmatrix} .$$

and the choice $\alpha = [2G'(x_0^2 + y_0^2)]^{-1}$ yields $(x_1, y_1) = (0,0)$ for an arbitrary starting point (x_0, y_0).

4. We consider the system $f(x,y) = x + y = 0$ and $g(x,y) = x + 1 = 0$. This is transformed into a minimum finding problem by defining

$$G(x,y) = f^2(x,y) + g^2(x,y) = (x+y)^2 + (x+1)^2 \,.$$

An application of the steepest descent method to this problem yields the following iteration scheme

$$\begin{bmatrix} x_{m+1} \\ y_{m+1} \end{bmatrix} = \begin{bmatrix} x_m \\ y_m \end{bmatrix} - \alpha \nabla G(x_m, y_m) = \begin{bmatrix} x_m - \alpha(4x_m + 2y_m + 2) \\ y_m - \alpha(2x_m + 2y_m) \end{bmatrix} . \tag{2}$$

where α is determined at each step by requiring that it minimizes $G(x_{m+1}, y_{m+1})$.

We will show by induction that the iteration scheme with starting point $(x_0, y_0) = (0, 0)$ yield iterates (x_m, y_m) that have different forms for even and odd m and are given by

$$(x_{2n}, y_{2n}) = (-1 + 2^{-n}, 1 - 2^{-n}) \tag{3}$$

and

$$(x_{2n+1}, y_{2n+1}) = (-1 + 2^{-n-1}, 1 - 2^{-n}) \tag{4}$$

for $n \geq 0$.

Firstly, we see that equation (3) for $n = 0$ yields the correct starting point $(x_0, y_0) = (-1 + 1, 1 - 1) = (0, 0)$.

Next we assume that equation (3) is correct for some n. Formula (2) with $m = 2n$ has the form

$$\begin{bmatrix} x_{2n+1} \\ y_{2n+1} \end{bmatrix} = \begin{bmatrix} -1 + 2^{-n} - \alpha 2^{-n+1} \\ 1 - 2^{-n} \end{bmatrix}.$$

The constant α is chosen to minimize

$$G(x_{2n+1}, y_{2n+1}) = \alpha^2 2^{-2n+2} + (2^{-n} - \alpha 2^{-n+1})^2 = 2^{-2n}(8\alpha^2 - 4\alpha + 1).$$

This function has a minimum at $\alpha = 1/4$. With this value of α we find

$$\begin{bmatrix} x_{2n+1} \\ y_{2n+1} \end{bmatrix} = \begin{bmatrix} -1 + 2^{-n-1} \\ 1 - 2^{-n} \end{bmatrix},$$

in agreement with (4).

It remains to show that if we start with equation (4) that we will then obtain in the next iteration step an expression that agrees with equation (3). Hence we assume that equation (4) is correct for some n and apply formula (2) with $m = 2n + 1$

$$\begin{bmatrix} x_{2n+2} \\ y_{2n+2} \end{bmatrix} = \begin{bmatrix} -1 + 2^{-n-1} \\ 1 - 2^{-n} + \alpha 2^{-n} \end{bmatrix}.$$

The constant α is chosen to minimize

$$G(x_{2n+2}, y_{2n+2}) = (-2^{-n-1} + \alpha 2^{-n})^2 + 2^{-2n-2} = 2^{-2n-2}(4\alpha^2 - 4\alpha + 2).$$

The minimum is obtained for $\alpha = 1/2$. With this value of α we find

$$\begin{bmatrix} x_{2n+2} \\ y_{2n+2} \end{bmatrix} = \begin{bmatrix} -1 + 2^{-n-1} \\ 1 - 2^{-n-1} \end{bmatrix},$$

in agreement with (3). This completes the proof.