90. Linearization. Jacobi matrix. Newton's method.

The fixed point iteration (and hence also Newton's method) works equally well for systems of equations. For example,

$$\begin{aligned} x_2\big(1-x_1^2\big)&=0,\\ 2-x_1x_2&=0, \end{aligned}$$

is a system of two equations in two unknowns. See Problem 90.5 below. If we define two functions

$$\begin{aligned} f_1(x_1,x_2)&=x_2\big(1-x_1^2\big),\\ f_2(x_1,x_2)&=2-x_1x_2, \end{aligned}$$

the equations may be written

$$\begin{aligned} f_1(x_1,x_2)&=0,\\ f_2(x_1,x_2)&=0. \end{aligned}$$

With $f=(f_1,f_2)$, $x=(x_1,x_2)$, and $0=(0,0)$, we note that $f:\mathbf{R}^2\to\mathbf{R}^2$ and we can write the equations in the compact form

$$f(x)=0.$$

In this lecture we will see how Newton's method can be applied to such systems of equations.

Note that the bisection algorithm can only be used for a single equation, but not for a system of several equations. This is because it relies on the fact the the graph of a Lipschitz continuous function $f:\mathbf{R}\to\mathbf{R}$ must pass the value zero if it is positive in one point and negative in another point. This has no counterpart for functions $f:\mathbf{R}^2\to\mathbf{R}^2$.

0.1 Function of one variable, $f:\mathbf{R}\to\mathbf{R}$

(AMBS 23) A function $f:\mathbf{R}\to\mathbf{R}$ of one variable is differentiable at $\bar{x}$ if there are constants $m(\bar{x})$, $K_f(\bar{x})$ such that

$$f(x)=f(\bar{x})+m(\bar{x})(x-\bar{x})+E_f(x,\bar{x}), \tag{1}$$

where the remainder E_f satisfies $|E_f(x,\bar{x})|\le K_f(\bar{x})(x-\bar{x})^2$ when x is close to $\bar{x}$. The constant $m(\bar{x})$ is called the derivative of f at $\bar{x}$ and we write

$$m(\bar{x})=f'(\bar{x})=Df(\bar{x})=\frac{df}{dx}(\bar{x}).$$

It is convenient to use the abbreviation $h=x-\bar{x}$, so that $x=\bar{x}+h$ and (1) becomes

$$f(x)=f(\bar{x}+h)=f(\bar{x})+f'(\bar{x})h+E_f(x,\bar{x}), \tag{2}$$

where $|E_f(x,\bar{x})|\le K_f(\bar{x})h^2$ when x is close to $\bar{x}$. Note that the first term on the right side, $f(\bar{x})$, is constant with respect to x. The second term,

$$f'(\bar{x})h=f'(\bar{x})(x-\bar{x}), \tag{3}$$

is a linear function of the increment $h=x-\bar{x}$. These terms are called the *linearization of f at $\bar{x}$*,

$$\tilde{f}_{\bar{x}}(x)=f(\bar{x})+f'(\bar{x})(x-\bar{x}). \tag{4}$$

The straight line $y=\tilde{f}_{\bar{x}}(x)$ is the tangent to the curve $y=f(x)$ at $\bar{x}$.

Example 1. Let $f(x) = x^2$. Then $f'(x) = 2x$ and the linearization at $\bar{x} = 3$ is

$$\tilde{f}_3(x) = 9 + 6(x - 3).$$

Numerical computation of the derivative. (AMBS 23.13) If we divide (2) by h, then we get

$$\frac{f(\bar{x} + h) - f(\bar{x})}{h} = f'(\bar{x}) + E_f(x, \bar{x})/h. \tag{5}$$

Here the remainder satisfies $|E_f(x, \bar{x})/h| \le K_f(\bar{x})|h|$ when h is small. This suggests that we can approximate the derivative by the difference quotient

$$f'(\bar{x}) \approx \frac{f(\bar{x} + h) - f(\bar{x})}{h}. \tag{6}$$

A better approximation is obtained if we use Taylor's formula (AMBS 24.8, 28.15):

$$\begin{aligned} f(\bar{x} + h) - f(\bar{x} - h) &= f(\bar{x}) + f'(\bar{x})h + f''(\bar{x})h^2/2 + R_2(\bar{x} + h, \bar{x}) \\ &\quad - \big(f(\bar{x}) - f'(\bar{x})h + f''(\bar{x})h^2/2 + R_2(\bar{x} - h, \bar{x})\big) \\ &= 2f'(\bar{x})h + R_2(\bar{x} + h, \bar{x}) - R_2(\bar{x} - h, \bar{x}). \end{aligned}$$

The remainders satisfy $|R_2(\bar{x} \pm h, \bar{x})| \le K(\bar{x})|h|^3$ when h is small. This suggests the symmetric difference quotient:

$$f'(\bar{x}) \approx \frac{f(\bar{x} + h) - f(\bar{x} - h)}{2h}. \tag{7}$$

The difference quotients in (6) and (7) are of the form "small number divided by small number". If this is computed with round-off error on a computer, then the total error will be large if the step h is very small. Therefore we must choose the step "moderately small" here, see (AMBS 23.13). It can be shown that in Matlab a good choice for (6) is $h = 10^{-8}$ and for (7) $h = 10^{-5}$.

0.2 Function of two variables, $f : \mathbf{R}^2 \to \mathbf{R}$

(AMBS 24.11) Let $f(x_1, x_2)$ be a function of two variables, i.e., $f : \mathbf{R}^2 \to \mathbf{R}$. We write $x = (x_1, x_2)$ and $f(x) = f(x_1, x_2)$. The function f is differentiable at $\bar{x} = (\bar{x}_1, \bar{x}_2)$, if there are constants $m_1(\bar{x})$, $m_2(\bar{x})$, $K_f(\bar{x})$ such that

$$f(x) = f(\bar{x} + h) = f(\bar{x}) + m_1(\bar{x})h_1 + m_2(\bar{x})h_2 + E_f(x, \bar{x}), \quad h = x - \bar{x}, \tag{8}$$

where the remainder E_f satisfies $|E_f(x, \bar{x})| \le K_f(\bar{x})|h|^2$ when x is close to $\bar{x}$. Here $|h| = \sqrt{h_1^2 + h_2^2}$ denotes the norm of the increment $h = (h_1, h_2) = (x_1 - \bar{x}_1, x_2 - \bar{x}_2)$.

If we take $h = (h_1, 0)$, then we get

$$f(x_1, \bar{x}_2) = f(\bar{x}_1 + h_1, \bar{x}_2) = f(\bar{x}) + m_1(\bar{x})h_1 + E_f(x, \bar{x}),$$

with $|E_f(x, \bar{x})| \le K_f(\bar{x})h_1^2$. This means that $m_1(\bar{x})$ is the derivative of the one-variable function $\hat{f}(x_1) = f(x_1, \bar{x}_2)$, obtained from f by keeping $x_2 = \bar{x}_2$ fixed. By taking $h = (0, h_2)$ we see in a similar way that $m_2(\bar{x})$ is the derivative of the one-variable function, which is obtained from f by keeping $x_1 = \bar{x}_1$ fixed. The constants $m_1(\bar{x})$, $m_2(\bar{x})$ are called the *partial derivatives* of f at $\bar{x}$ and we denote them by

$$m_1(\bar{x}) = f'_{x_1}(\bar{x}) = \frac{\partial f}{\partial x_1}(\bar{x}), \quad m_2(\bar{x}) = f'_{x_2}(\bar{x}) = \frac{\partial f}{\partial x_2}(\bar{x}). \tag{9}$$

Now (8) may be written

$$f(x) = f(\bar{x} + h) = f(\bar{x}) + f'_{x_1}(\bar{x})h_1 + f'_{x_2}(\bar{x})h_2 + E_f(x, \bar{x}), \quad h = x - \bar{x}. \tag{10}$$

It is convenient to write this formula by means of matrix notation. Let

$$a = \begin{bmatrix} a_1, a_2 \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}.$$

We say that a is a row matrix of type 1×2 (one by two) and that b is a column matrix of type 2×1 (two by one). Their product is defined by

$$ab = \begin{bmatrix} a_1, a_2 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = a_1 b_1 + a_2 b_2.$$

The result is a matrix of type 1×1 (a real number), according to the rule: 1×2 times 2×1 makes 1×1.

Going back to (10) we define

$$f'(\bar{x}) = Df(\bar{x}) = \begin{bmatrix} f'_{x_1}(\bar{x}) & f'_{x_2}(\bar{x}) \end{bmatrix}, \quad h = \begin{bmatrix} h_1 \\ h_2 \end{bmatrix}.$$

The matrix $f'(\bar{x}) = Df(\bar{x})$ is called the derivative (or Jacobi matrix) of f at $\bar{x}$. Then (10) may be written

$$\begin{aligned} f(x) = f(\bar{x} + h) &= f(\bar{x}) + \begin{bmatrix} f'_{x_1}(\bar{x}) & f'_{x_2}(\bar{x}) \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} + E_f(x, \bar{x}) \\ &= f(\bar{x}) + f'(\bar{x})h + E_f(x, \bar{x}), \quad h = x - \bar{x}. \end{aligned} \tag{11}$$

Note that the first term on the right side, $f(\bar{x})$, is constant with respect to x. The second term,

$$f'(\bar{x})h = f'(\bar{x})(x - \bar{x}), \tag{12}$$

is a linear function of the increment $h = x - \bar{x}$. These terms are called the *linearization of f at $\bar{x}$*,

$$\tilde{f}_{\bar{x}}(x) = f(\bar{x}) + f'(\bar{x})(x - \bar{x}). \tag{13}$$

The plane $x_3 = \tilde{f}_{\bar{x}}(x_1, x_2)$ is the tangent to the surface $x_3 = f(x_1, x_2)$ at $\bar{x}$.

Example 2. Let $f(x) = x_1^2 x_2^5$. Then

$$\frac{\partial f}{\partial x_1}(x) = \frac{\partial f}{\partial x_1}\left(x_1^2 x_2^5\right) = 2x_1 x_2^5, \quad \frac{\partial f}{\partial x_2}(x) = \frac{\partial f}{\partial x_2}\left(x_1^2 x_2^5\right) = 5x_1^2 x_2^4,$$

so that $f'(x) = \begin{bmatrix} 2x_1 x_2^5 & 5x_1^2 x_2^4 \end{bmatrix}$ and the linearization at $\bar{x} = (3, 1)$ is

$$\tilde{f}_{\bar{x}}(x) = 9 + \begin{bmatrix} 6 & 45 \end{bmatrix} \begin{bmatrix} x_1 - 3 \\ x_2 - 1 \end{bmatrix}.$$

0.3 Two functions of two variables, $f : \mathbf{R}^2 \to \mathbf{R}^2$

Let $f_1(x_1, x_2)$, $f_2(x_1, x_2)$ be two functions of two variables. We write $x = (x_1, x_2)$ and $f(x) = (f_1(x_1, x_2), f_2(x_1, x_2))$, i.e., $f : \mathbf{R}^2 \to \mathbf{R}^2$. The function f is differentiable at $\bar{x} = (\bar{x}_1, \bar{x}_2)$, if there are constants $m_{11}(\bar{x})$, $m_{12}(\bar{x})$, $m_{21}(\bar{x})$, $m_{22}(\bar{x})$, and $K_f(\bar{x})$ such that

$$\begin{aligned} f_1(x) = f_1(\bar{x} + h) = f_1(\bar{x}) + m_{11}(\bar{x})h_1 + m_{12}(\bar{x})h_2 + E_{f_1}(x, \bar{x}), \\ f_2(x) = f_2(\bar{x} + h) = f_2(\bar{x}) + m_{21}(\bar{x})h_1 + m_{22}(\bar{x})h_2 + E_{f_2}(x, \bar{x}), \end{aligned} \tag{14}$$

where $h = x - \bar{x}$ and the remainders E_{f_j} satisfy $|E_{f_j}(x, \bar{x})| \leq K_f(\bar{x})|h|^2$ when x is close to $\bar{x}$. Here $|h| = \sqrt{h_1^2 + h_2^2}$ denotes the norm of the increment $h = (h_1, h_2) = (x_1 - \bar{x}_1, x_2 - \bar{x}_2)$. From the previous subsection we recognize that the constants $m_{ij}(\bar{x})$ are the partial derivatives of the functions f_i at $\bar{x}$ and we denote them by

$$\begin{aligned} m_{11}(\bar{x}) &= f'_{1,x_1}(\bar{x}) = \frac{\partial f_1}{\partial x_1}(\bar{x}), \quad m_{12}(\bar{x}) = f'_{1,x_2}(\bar{x}) = \frac{\partial f_1}{\partial x_2}(\bar{x}), \\ m_{21}(\bar{x}) &= f'_{2,x_1}(\bar{x}) = \frac{\partial f_2}{\partial x_1}(\bar{x}), \quad m_{22}(\bar{x}) = f'_{2,x_2}(\bar{x}) = \frac{\partial f_2}{\partial x_2}(\bar{x}). \end{aligned}$$

It is convenient to use matrix notation. Let

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}.$$

We say that A is a matrix of type 2×2 (two by two) and that b is a column matrix of type 2×1 (two by one). Their product is defined by

$$Ab = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} a_{11}b_1 + a_{12}b_2 \\ a_{21}b_1 + a_{22}b_2 \end{bmatrix}.$$

The result is a matrix of type 2×1 (column matrix), according to the rule: 2×2 times 2×1 makes 2×1.

Going back to (14) we define

$$f(x) = \begin{bmatrix} f_1(x) \\ f_2(x) \end{bmatrix}, \quad f'(\bar{x}) = Df(\bar{x}) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\bar{x}) & \frac{\partial f_1}{\partial x_2}(\bar{x}) \\ \frac{\partial f_2}{\partial x_1}(\bar{x}) & \frac{\partial f_2}{\partial x_2}(\bar{x}) \end{bmatrix}, \quad h = \begin{bmatrix} h_1 \\ h_2 \end{bmatrix}. \tag{15}$$

The matrix $f'(\bar{x}) = Df(\bar{x})$ is called the derivative (or Jacobi matrix) of f at $\bar{x}$. Then (14) may be written

$$\begin{bmatrix} f_1(x) \\ f_2(x) \end{bmatrix} = \begin{bmatrix} f_1(\bar{x}+h) \\ f_2(\bar{x}+h) \end{bmatrix} = \begin{bmatrix} f_1(\bar{x}) \\ f_2(\bar{x}) \end{bmatrix} + \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\bar{x}) & \frac{\partial f_1}{\partial x_2}(\bar{x}) \\ \frac{\partial f_2}{\partial x_1}(\bar{x}) & \frac{\partial f_2}{\partial x_2}(\bar{x}) \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} + \begin{bmatrix} E_{f_1}(x,\bar{x}) \\ E_{f_2}(x,\bar{x}) \end{bmatrix}, \tag{16}$$

or in more compact form

$$f(x) = f(\bar{x}+h) = f(\bar{x}) + f'(\bar{x})h + E_f(x,\bar{x}), \quad h = x - \bar{x}. \tag{17}$$

Note that the first term on the right side, $f(\bar{x})$, is constant with respect to x. The second term,

$$f'(\bar{x})h = f'(\bar{x})(x - \bar{x}), \tag{18}$$

is a linear function of the increment $h = x - \bar{x}$. These terms are called the *linearization of f at $\bar{x}$*,

$$\tilde{f}_{\bar{x}}(x) = f(\bar{x}) + f'(\bar{x})(x - \bar{x}). \tag{19}$$

Example 3. Let $f(x) = \begin{bmatrix} x_1^2 x_2^5 \\ x_2^3 \end{bmatrix}$. Then

$$f'(x) = Df(x) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(x) & \frac{\partial f_1}{\partial x_2}(x) \\ \frac{\partial f_2}{\partial x_1}(x) & \frac{\partial f_2}{\partial x_2}(x) \end{bmatrix} = \begin{bmatrix} 2x_1x_2^5 & 5x_1^2x_2^4 \\ 0 & 3x_2^2 \end{bmatrix}$$

and the linearization at $\bar{x} = (3,1)$ is

$$\tilde{f}_{\bar{x}}(x) = \begin{bmatrix} 9 \\ 1 \end{bmatrix} + \begin{bmatrix} 6 & 45 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} x_1 - 3 \\ x_2 - 1 \end{bmatrix}.$$

0.4 Several functions of several variables, $f : \mathbf{R}^n \to \mathbf{R}^m$

(AMBS 53) It is now easy to generalize to any number of functions in any number of variables. Let f_i be m functions of n variables x_j, i.e., $f : \mathbf{R}^n \to \mathbf{R}^m$. As in (15) we define

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad h = \begin{bmatrix} h_1 \\ \vdots \\ h_n \end{bmatrix} = \begin{bmatrix} x_1 - \bar{x}_1 \\ \vdots \\ x_n - \bar{x}_n \end{bmatrix},$$

$$f(x) = \begin{bmatrix} f_1(x_1, \ldots, x_n) \\ \vdots \\ f_m(x_1, \ldots, x_n) \end{bmatrix}, \quad f'(\bar{x}) = Df(\bar{x}) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\bar{x}) & \cdots & \frac{\partial f_1}{\partial x_n}(\bar{x}) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(\bar{x}) & \cdots & \frac{\partial f_m}{\partial x_n}(\bar{x}) \end{bmatrix}.$$

The $m \times n$ matrix $f'(\bar{x}) = Df(\bar{x})$ is called the derivative (or Jacobi matrix) of f at $\bar{x}$. In a similar way to (17) we get

$$f(x) = f(\bar{x} + h) = f(\bar{x}) + f'(\bar{x})h + E_f(x, \bar{x}), \quad h = x - \bar{x}. \tag{20}$$

The linearization of f at $\bar{x}$ is

$$\tilde{f}_{\bar{x}}(x) = f(\bar{x}) + f'(\bar{x})(x - \bar{x}). \tag{21}$$

Numerical computation of the derivative. In order to compute the j-th column $\frac{\partial f}{\partial x_j}(\bar{x})$ of the Jacobi matrix, we choose the increment h such that $h_j = \delta$ and $h_i = 0$ for $i \neq j$, i.e.,

$$h = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \delta \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \delta \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \delta e_j, \quad e_j = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \text{— row number } j.$$

Here the steplength δ is a small positive number and e_j is the j-th standard basis vector. If we use this increment in a symmetric difference quotient, see (7), we get

$$\frac{\partial f}{\partial x_j}(\bar{x}) \approx \frac{f(\bar{x} + \delta e_j) - f(\bar{x} - \delta e_j)}{2\delta}. \tag{22}$$

Remember that the steplength δ should be small, but not too small.

0.5 Newton's method for $f(x) = 0$

Consider a system of n equations with n unknowns:

$$\begin{aligned} f_1(x_1, \ldots, x_n) &= 0, \\ &\vdots \\ f_n(x_1, \ldots, x_n) &= 0. \end{aligned}$$

If we define

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad f = \begin{bmatrix} f_1 \\ \vdots \\ f_n \end{bmatrix}, \quad 0 = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix},$$

then $f : \mathbf{R}^n \to \mathbf{R}^n$, and we can write our system of equations in the compact form

$$f(x) = 0. \tag{23}$$

Suppose that we have found an approximate solution $\bar{x}$. We want to find a better approximation $x = \bar{x} + h$. Instead of solving (23) directly, which is usually impossible, we solve the *linearized equation at* $\bar{x}$:

$$\tilde{f}_{\bar{x}}(\bar{x} + h) = f(\bar{x}) + f'(\bar{x})h = 0. \tag{24}$$

We must solve for the increment h. Rearranging the terms we get

$$f'(\bar{x})h = -f(\bar{x}). \tag{25}$$

Remember that the Jacobi matrix is of type $n \times n$ and the increment is of type $n \times 1$. Therefore we have to solve a linear system of n equations with n unknowns to get the increment h. It is of the form $Ah = b$ with $A = f'(\bar{x})$ and $b = -f(\bar{x})$. Then we set $x = \bar{x} + h$.

In algorithmic form Newton's method can be formulated:

```
while |h|>tol
   evaluate the residual    b=-f(x)
   evaluate the Jacobian    A=f'(x)
   solve the linear system  Ah=b
   update                   x=x+h
end
```

You will implement this algorithm in the studio exercises. You will use the MATLAB command

```
h=A\b
```

to solve the system. But later in this course we will study linear systems of equations of the form $Ah = b$ and we will answer important questions such as:

- Is there a unique solution h for each b?
- How do you compute the solution?

These questions can be answered for linear systems $Ah = b$, but not for the more general nonlinear systems $f(x) = 0$. Thus, Newton's method transforms the task of solving a difficult equation to the task of solving an easier equation many times. The study of systems of linear equations is an important part of the subject "linear algebra" which we will study in ALA-B.

90 Problems

Problem 90.1. Let

$$a = \begin{bmatrix} 1 & 2 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}.$$

Compute the products $\quad ab, \quad ba, \quad Ab, \quad Aa, \quad aA, \quad bA.$

Problem 90.2. Compute the Jacobi matrix $f'(x)$ (also denoted $Df(x)$). Compute the linearization of f at $\bar{x}$.

$$\text{(a)} \quad f(x) = \begin{bmatrix} \sin(x_1) + \cos(x_2) \\ \cos(x_1) + \sin(x_2) \end{bmatrix}, \quad \bar{x} = 0; \qquad \text{(b)} \quad f(x) = \begin{bmatrix} 1 \\ 1 + x_1 \\ 1 + x_1 e^{x_2} \end{bmatrix}, \quad \bar{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Problem 90.3. Compute the gradient vector $\nabla f(x)$ (also denoted $f'(x) = Df(x)$). Compute the linearization of f at $\bar{x}$.

$$\text{(a)} \quad f(x) = e^{-x_1} \sin(x_2), \quad \bar{x} = 0; \qquad \text{(b)} \quad f(x) = |x|^2 = x_1^2 + x_2^2 + x_3^2, \quad x \in \mathbf{R}^3, \quad \bar{x} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Problem 90.4. Here $f : \mathbf{R} \to \mathbf{R}^2$. Compute $f'(t)$. Compute the linearization of f at $\bar{t}$.

$$\text{(a)} \quad f(t) = \begin{bmatrix} \cos(t) \\ \sin(t) \end{bmatrix}, \quad \bar{t} = \pi/2; \qquad \text{(b)} \quad f(t) = \begin{bmatrix} t \\ 1+t^2 \end{bmatrix}, \quad \bar{t} = 0.$$

Problem 90.5. (a) Write the system

$$\begin{aligned} u_2\big(1 - u_1^2\big) &= 0, \\ 2 - u_1 u_2 &= 0 \end{aligned}$$

in the form $f(u) = 0$. Find the all the solutions by hand calculation.

(b) Compute the Jacobi matrix $Df(u)$.

(c) Perform the first step of Newton's method for the equation $f(u) = 0$ with initial vector $u^{(0)} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

(d) Solve the equation $f(u)$ with your MATLAB program `newton.m`.

Problem 90.6. (a) Write the system

$$\begin{aligned} u_1\big(1 - u_2\big) &= 0, \\ u_2\big(1 - u_1\big) &= 0, \end{aligned}$$

in the form $f(u) = 0$. Find the all the solutions by hand calculation.

(b) Compute the Jacobi matrix $Df(u)$.

(c) Perform the first step of Newton's method for the equation $f(u) = 0$ with initial vector $u^{(0)} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$.

(d) Solve the equation $f(u)$ with your MATLAB program `newton.m`.

Answers and solutions

90.1. Use MATLAB to check your answers.

90.2.

(a)

$$f'(x) = \begin{bmatrix} \cos(x_1) & -\sin(x_2) \\ -\sin(x_1) & \cos(x_2) \end{bmatrix}, \qquad \tilde{f}(x) = f(\bar{x}) + f'(\bar{x})(x - \bar{x}) = \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

(b)

$$f'(x) = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ e^{x_2} & x_1 e^{x_2} \end{bmatrix}, \qquad \tilde{f}(x) = f(\bar{x}) + f'(\bar{x})(x - \bar{x}) = \begin{bmatrix} 1 \\ 2 \\ 1+e \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ e & e \end{bmatrix} \begin{bmatrix} x_1 - 1 \\ x_2 - 1 \end{bmatrix}.$$

90.3.

(a)

$$\nabla f(x) = \begin{bmatrix} -e^{-x_1}\sin(x_2), & e^{-x_1}\cos(x_2) \end{bmatrix},$$
$$\tilde{f}(x) = f(\bar{x}) + f'(\bar{x})(x - \bar{x}) = 0 + \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_2.$$

(b)

$$\nabla f(x) = \begin{bmatrix} 2x_1 & 2x_3 & 2x_3 \end{bmatrix},$$
$$\tilde{f}(x) = f(\bar{x}) + f'(\bar{x})(x - \bar{x}) = 3 + \begin{bmatrix} 2 & 2 & 2 \end{bmatrix} \begin{bmatrix} x_1 - 1 \\ x_2 - 1 \\ x_3 - 1 \end{bmatrix} = -3 + 2x_1 + 2x_2 + 2x_3.$$

90.4.

(a)

$$f'(t) = \begin{bmatrix} -\sin(t) \\ \cos(t) \end{bmatrix},$$
$$\tilde{f}(t) = f(\bar{t}) + f'(\bar{t})(t - \bar{t}) = \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \begin{bmatrix} -1 \\ 0 \end{bmatrix} (t - \pi/2).$$

(b)

$$f'(t) = \begin{bmatrix} 1 \\ 2t \end{bmatrix},$$
$$\tilde{f}(t) = f(\bar{t}) + f'(\bar{t})(t - \bar{t}) = \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} t = \begin{bmatrix} t \\ 1 \end{bmatrix}.$$

90.5. (a) The solutions are given by

$$f(u) = \begin{bmatrix} u_2(1 - u_1^2) \\ 2 - u_1 u_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

We find two solutions $\bar{u} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\bar{u} = \begin{bmatrix} -1 \\ -2 \end{bmatrix}$.

(b) The Jacobian is

$$Df(u) = \begin{bmatrix} -2u_1u_2 & 1-u_1^2 \\ -u_2 & -u_1 \end{bmatrix}.$$

(c) The first step of Newton's method:

evaluate $\quad A = Df(1,1) = \begin{bmatrix} -2 & 0 \\ -1 & -1 \end{bmatrix} \quad \text{and} \quad b = -f(1,1) = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$

solve $\quad Ah = b, \quad \begin{bmatrix} -2 & 0 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$

$$\begin{cases} -2h_1 = 0, \\ -h_1 - h_2 = -1, \end{cases} \quad h = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

update $\quad u^{(1)} = u^{(0)} + h = \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \bar{u}$

Bingo! We found one of the solutions.

90.6. (a) The solutions are given by

$$f(u) = \begin{bmatrix} u_1(1-u_2) \\ u_2(1-u_1) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

We find two solutions $\bar{u} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ and $\bar{u} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

(b) The Jacobian is

$$Df(u) = \begin{bmatrix} 1-u_2 & -u_1 \\ -u_2 & 1-u_1 \end{bmatrix}.$$

(c) The first step of Newton's method:

evaluate $\quad A = Df(2,2) = \begin{bmatrix} -1 & -2 \\ -2 & -1 \end{bmatrix} \quad \text{and} \quad b = -f(2,2) = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$

solve $\quad Ah = b, \quad \begin{bmatrix} -1 & -2 \\ -2 & -1 \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix},$

$$\begin{cases} -h_1 - 2h_2 = 2, \\ -2h_1 - h_2 = 2, \end{cases} \quad h = \begin{bmatrix} -2/3 \\ -2/3 \end{bmatrix}$$

update $\quad u^{(1)} = u^{(0)} + h = \begin{bmatrix} 2 \\ 2 \end{bmatrix} + \begin{bmatrix} -2/3 \\ -2/3 \end{bmatrix} = \begin{bmatrix} 4/3 \\ 4/3 \end{bmatrix}$

Getting closer to one of the solutions $\bar{u}$!

`2005-10-10 /stig`