Andrea Damura Casanova
309354
Group 1: Tuesday Tutorial

EXERCISE SHEET 6

1. Exact Newton Method

$$f(x,y) = \ln(1+x+2y)$$

a) $\nabla f(x,y) = \left\langle \frac{1}{1+x+2y}, \frac{2}{1+x+2y} \right\rangle$

$$\frac{\delta f(x,y)}{\delta x} = \frac{1}{1+x+2y}$$

$$\frac{\delta f(x,y)}{\delta y} = \frac{2}{1+x+2y}$$

$$\nabla^2 f(x,y) = \begin{pmatrix} f''_{xx} & f''_{xy} \\ f''_{yx} & f''_{yy} \end{pmatrix}$$

$$\frac{\delta}{\delta x}\left[\frac{f(x)}{g(x)}\right] = \frac{g(x)f'(x) - f(x)g'(x)}{[g(x)]^2}$$

$$\frac{\delta^2 f(x,y)}{\delta x^2} = \cancel{\frac{(1+x+2y)-1\cdot x}{(1+x+2y)^2}} = \cancel{\frac{1+2y}{(1+x+2y)^2}} = \frac{(1+x+2y)\cdot 0 - 1}{(1+x+2y)^2} = -\frac{1}{(1+x+2y)^2}$$

$$\frac{\delta^2 f(x,y)}{\delta y^2} = \frac{(1+x+2y)\cdot 0 - 2\cdot 2}{(1+x+2y)^2} = \frac{-4}{(1+x+2y)^2}$$

$$\frac{\delta^2 f(x,y)}{\delta x \delta y} = \frac{(1+x+2y)\cdot 0 - 1\cdot 2}{(1+x+2y)^2} = -\frac{2}{(1+x+2y)^2}$$

$$\frac{\delta^2 f(x,y)}{\delta y \delta x} = \frac{(1+x+2y)\cdot 0 - 2\cdot 1}{(1+x+2y)^2} = \frac{-2}{(1+x+2y)^2}$$

$$\nabla^2 f(x,y) = \begin{bmatrix} \frac{-1}{(1+x+2y)^2} & \frac{-2}{(1+x+2y)^2} \\ \frac{-2}{(1+x+2y)^2} & \frac{-4}{(1+x+2y)^2} \end{bmatrix}$$

b) $\nabla^2 f(x,y)^{-1} = \frac{1}{ad - cb} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \frac{1}{\frac{-1}{(1+x+2y)^2} \cdot \frac{-4}{(1+x+2y)^2} - \frac{-2}{(1+x+2y)^2} \cdot \frac{-2}{(1+x+2y)^2}} \cdot \begin{pmatrix} \frac{-4}{(1+x+2y)^2} & \frac{2}{(1+x+2y)^2} \\ \frac{2}{(1+x+2y)^2} & \frac{-1}{(1+x+2y)^2} \end{pmatrix} =$

$= \cancel{\frac{(1+x+2y)^4}{0}} \cdot \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$ The hessian matrix is singular

(c) The update step of the Newton Algorithm can't be computed because the Hessian matrix is singular (non-invertible).

2. Quasi-Newton Method: BFGS

$f(x_1, x_2) = x_1^2 + 0.5x_2^2 + 3$. $\nabla f(x_1, x_2) = \begin{pmatrix} 2x_1 \\ x_2 \end{pmatrix}$

$x_0 = [1, 2]^T$

First iteration:

$A^{(0)} = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$

$\Delta x^{(0)} = -\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 2 \cdot 1 \\ 2 \end{pmatrix} = -\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 2 \end{pmatrix} = \begin{pmatrix} -4 \\ -2 \end{pmatrix}$

$\mu^{(0)} = \frac{0.001}{\sqrt{\sum \nabla f(x_1, x_2)}} = \frac{0.001}{\sqrt{2^2 + 2^2}} = \frac{0.001}{2.828} = 3.53 \cdot 10^{-4}$

$x^{(1)} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} + 3.53 \cdot 10^{-4} \cdot \begin{pmatrix} -4 \\ -2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} - \begin{pmatrix} 1.41 \cdot 10^{-3} \\ 7.06 \cdot 10^{-3} \end{pmatrix} = \begin{pmatrix} 0.998 \\ 1.99 \end{pmatrix}$

$s^{(1)} = \begin{pmatrix} 0.998 \\ 1.99 \end{pmatrix} - \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} -0.002 \\ -0.01 \end{pmatrix}$

2.1

$g^{(1)} = \begin{pmatrix} 2 \cdot 0.998 \\ 1.99 \end{pmatrix} - \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} -0.004 \\ 0.99 \end{pmatrix}$

$A^{(1)} = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} + \frac{\left[\begin{pmatrix} -0.002 \\ -0.01 \end{pmatrix} - \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} -0.004 \\ 0.99 \end{pmatrix}\right]\left[\begin{pmatrix} -0.002 \\ -0.01 \end{pmatrix} - \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} -0.004 \\ 0.99 \end{pmatrix}\right]^T}{\left[\begin{pmatrix} -0.002 \\ -0.01 \end{pmatrix} - \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} -0.004 \\ 0.99 \end{pmatrix}\right]^T \cdot \begin{pmatrix} -0.004 \\ 0.99 \end{pmatrix}} =$

$= \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} + \frac{\left[\begin{pmatrix} -0.002 \\ -0.01 \end{pmatrix} - \begin{pmatrix} -8 \cdot 10^{-3} \\ 0.99 \end{pmatrix}\right] \cdot \left[\begin{pmatrix} -0.002 \\ -0.01 \end{pmatrix} - \begin{pmatrix} -8 \cdot 10^{-3} \\ 0.99 \end{pmatrix}\right]^T}{\left[\begin{pmatrix} -0.002 \\ -0.01 \end{pmatrix} - \begin{pmatrix} -8 \cdot 10^{-3} \\ 0.99 \end{pmatrix}\right]^T \cdot \begin{pmatrix} -0.004 \\ 0.99 \end{pmatrix}} = A^{(0)} + \frac{\begin{pmatrix} 6 \cdot 10^{-3} \\ -1 \end{pmatrix} (6 \cdot 10^{-3} \; -1)}{(6 \cdot 10^{-3} \; -1) \begin{pmatrix} -0.004 \\ 0.99 \end{pmatrix}} = A^{(0)} + \frac{\begin{pmatrix} 3.6 \cdot 10^{-5} & -6 \cdot 10^{-3} \\ -6 \cdot 10^{-3} & 1 \end{pmatrix}}{-0.99} =$

$= \begin{pmatrix} -3.69 \cdot 10^{-5} & 6.06 \cdot 10^{-3} \\ 6.06 \cdot 10^{-3} & -1.01 \end{pmatrix} + \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1.99 & 6.06 \cdot 10^{-3} \\ 6.06 \cdot 10^{-3} & -0.01 \end{pmatrix}$ $-2-$

2nd iteration

$$A^{(1)} = \begin{pmatrix} 1.99 & 6.06\cdot10^{-3} \\ 6.06\cdot10^{-3} & -0.01 \end{pmatrix}$$

$$\Delta x^{(1)} = -\begin{pmatrix} 1.99 & 6.06\cdot10^{-3} \\ 6.06\cdot10^{-3} & -0.01 \end{pmatrix}\begin{pmatrix} 2\cdot 0.998 \\ 1.99 \end{pmatrix} = \begin{pmatrix} 3.98 \\ -7.8\cdot10^{-3} \end{pmatrix}$$

$$\mu^{(1)} = \frac{3.53\cdot10^{-4}}{\sqrt{1.996^2+1.99^2}} = \frac{3.53\cdot10^{-4}}{2.82} = 1.25\cdot10^{-4}$$

$$x^{(2)} = \begin{pmatrix} 0.998 \\ 1.99 \end{pmatrix} + 1.25\cdot10^{-4}\cdot\begin{pmatrix} 3.98 \\ -7.8\cdot10^{-3} \end{pmatrix} = \begin{pmatrix} 0.998 \\ 1.99 \end{pmatrix} - \begin{pmatrix} 4.975\cdot10^{-4} \\ 9.75\cdot10^{-4} \end{pmatrix} = \begin{pmatrix} 0.997 \\ 1.989 \end{pmatrix}$$

$$s^{(2)} = \begin{pmatrix} 0.997 \\ 1.989 \end{pmatrix} - \begin{pmatrix} 0.998 \\ 1.99 \end{pmatrix} = \begin{pmatrix} -1\cdot10^{-3} \\ -1\cdot10^{-3} \end{pmatrix}$$

$$g^{(2)} = \begin{pmatrix} 2\cdot(-1\cdot10^{-3}) \\ -1\cdot10^{-3} \end{pmatrix} - \begin{pmatrix} 1.996 \\ 1.99 \end{pmatrix} = \begin{pmatrix} -1.998 \\ -1.991 \end{pmatrix}$$ ①

the same as ①

$$A^{(2)} = \begin{pmatrix} 1.99 & 6.06\cdot10^{-3} \\ 6.06\cdot10^{-3} & -0.01 \end{pmatrix} + \frac{\left[\begin{pmatrix} -1\cdot10^{-3} \\ -1\cdot10^{-3} \end{pmatrix} - \begin{pmatrix} 1.99 & 6.06\cdot10^{-3} \\ 6.06\cdot10^{-3} & -0.01 \end{pmatrix}\cdot\begin{pmatrix} -1.998 \\ -1.991 \end{pmatrix}\right]\left[\ \circledast\ \right]^T}{\left[\ \circledast\ \right]^T\begin{pmatrix} -1.998 \\ -1.991 \end{pmatrix}} =$$

the same as ①

$$= A^{(1)} + \frac{\left[\begin{pmatrix} -1\cdot10^{-3} \\ -1\cdot10^{-3} \end{pmatrix} - \begin{pmatrix} -3.988 \\ 7.80\cdot10^{-3} \end{pmatrix}\right]\left[\begin{pmatrix} -1\cdot10^{-3} \\ -1\cdot10^{-3} \end{pmatrix} - \begin{pmatrix} -3.988 \\ 7.80\cdot10^{-3} \end{pmatrix}\right]^T}{\left[\begin{pmatrix} -1\cdot10^{-3} \\ -1\cdot10^{-3} \end{pmatrix} - \begin{pmatrix} -3.988 \\ 7.80\cdot10^{-3} \end{pmatrix}\right]^T\cdot\begin{pmatrix} -1.998 \\ -1.991 \end{pmatrix}} = A^{(1)} + \frac{\begin{pmatrix} 3.987 \\ -8.8\cdot10^{-3} \end{pmatrix}\cdot(3.987 \quad -8.8\cdot10^{-3})}{(3.987 \quad -8.8\cdot10^{-3})\cdot\begin{pmatrix} -1.998 \\ -1.991 \end{pmatrix}} =$$

$$= A^{(1)} + \frac{\begin{pmatrix} 15.89 & -0.035 \\ -0.035 & 7.74\cdot10^{-5} \end{pmatrix}}{-7.94} = \begin{pmatrix} 1.99 & 6.06\cdot10^{-3} \\ 6.06\cdot10^{-3} & -0.01 \end{pmatrix} + \begin{pmatrix} -2 & 4.41\cdot10^{-3} \\ 4.41\cdot10^{-3} & -9.76\cdot10^{-6} \end{pmatrix} = \begin{pmatrix} -0.01 & 0.01 \\ 0.01 & -0.01 \end{pmatrix}$$

3.1

Index der Kommentare

2.1 this should be -0.002

3.1 the steps are correct but the final values differs