MAST30013 – Techniques in Operations Research

Semester 1, 2021

Tutorial 5 Solutions

1. Newton's Method

Step 0.1 $k = 0$, $\boldsymbol{x}^0 = (1,1,1)^T$.

Step 0.2 $\| \boldsymbol{\nabla} f(1,1,1) \| = \| (-2,3,3)^T \| = \sqrt{22} > 0.01$.

$$\boldsymbol{\nabla}^2 f(1,1,1) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}, \text{ which is positive definite.}$$

Therefore, $\boldsymbol{d}^0 = -\boldsymbol{\nabla}^2 f(1,1,1)^{-1} \boldsymbol{\nabla} f(1,1,1) = (1,-1,-1)^T$.

Step 0.3 $q(t) = f(1+t, 1-t, 1-t) = 4t^2 - 8t$, which is minimized when $8t - 8 = 0$ $\Longrightarrow t_0 = 1$.

Step 0.4 $\boldsymbol{x}^1 = (1,1,1)^T + 1(1,-1,-1)^T = (2,0,0)^T$, $k = 1$.

Step 1.2 $\| \boldsymbol{\nabla} f(2,0,0) \| = \| (0,0,0)^T \| = 0 < 0.01$.

We have that $\boldsymbol{x}_{\min} = (2,0,0)$.
Newton's method has found the minimum in one step since f is a quadratic function with $\boldsymbol{B}$ positive definite.

BFGS quasi-Newton Method

Step 0.1 $k = 0$, $\boldsymbol{x}^0 = (1,1,1)^T$, $\boldsymbol{H}_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{3} \end{pmatrix}$.

Step 0.2 $\| \boldsymbol{\nabla} f(1,1,1) \| = \| (-2,3,3)^T \| = \sqrt{22} > 0.01$.
$\boldsymbol{d}^0 = -\boldsymbol{H}_0 \boldsymbol{\nabla} f(1,1,1) = (2,-1,-1)^T$.

Step 0.3 $q(t) = f(1+2t, 1-t, 1-t) = 7t^2 - 10t$, which is minimized when $14t - 10 = 0$ $\Longrightarrow t_0 = \frac{5}{7}$.

Step 0.4 $\boldsymbol{x}^1 = (1,1,1)^T + \frac{5}{7}(2,-1,-1)^T = \left(\frac{17}{7}, \frac{2}{7}, \frac{2}{7}\right)^T$,
$\boldsymbol{s}^0 = \left(\frac{10}{7}, -\frac{5}{7}, -\frac{5}{7}\right)^T$, $\boldsymbol{g}^0 = \left(\frac{20}{7}, -\frac{15}{7}, -\frac{15}{7}\right)^T$, $\boldsymbol{r}^0 = \left(\frac{2}{5}, -\frac{1}{10}, -\frac{1}{10}\right)^T$,

$$\boldsymbol{H}_1 = \begin{pmatrix} \frac{29}{49} & \frac{3}{49} & \frac{3}{49} \\ \frac{3}{49} & \frac{55}{147} & \frac{2}{49} \\ \frac{3}{49} & \frac{2}{49} & \frac{55}{147} \end{pmatrix}, \ k = 1.$$

Step 1.2 $\| \boldsymbol{\nabla} f\left(\frac{17}{7}, \frac{2}{7}, \frac{2}{7}\right) \| = \| \left(\frac{6}{7}, \frac{6}{7}, \frac{6}{7}\right)^T \| = \frac{6}{7}\sqrt{3} > 0.01$.
$\boldsymbol{d}^1 = -\boldsymbol{H}_1 \boldsymbol{\nabla} f\left(\frac{17}{7}, \frac{2}{7}, \frac{2}{7}\right)^T = \left(-\frac{30}{49}, -\frac{20}{49}, -\frac{20}{49}\right)^T$.

Step 1.3 $q(t) = f\left(\frac{17}{7} - \frac{30}{49}t, \frac{2}{7} - \frac{20}{49}t, \frac{2}{7} - \frac{20}{49}t\right) = \frac{300}{343}t^2 - \frac{60}{49}t - \frac{25}{7}$, which is minimized when $\frac{600}{343}t - \frac{60}{49} = 0 \Longrightarrow t_1 = \frac{7}{10}$.

Step 1.4 $\boldsymbol{x}^2 = \left(\frac{17}{7}, \frac{2}{7}, \frac{2}{7}\right)^T + \frac{7}{10}\left(-\frac{30}{49}, -\frac{20}{49}, -\frac{20}{49}\right)^T = (2,0,0)^T$, $k = 2$.

Step 2.2 $\| \boldsymbol{\nabla} f(2,0,0) \| = \| (0,0,0)^T \| = 0 < 0.01$.

We have that $\boldsymbol{x}_{\min} = (2, 0, 0)$.
The BFGS quasi-Newton method has found the minimum in two steps.

2. Suppose that $\boldsymbol{H}_0$ is symmetric. Assume that for some $k \geq 1$ that $\boldsymbol{H}_k$ is symmetric. Then

$$\begin{aligned}
\boldsymbol{H}_{k+1}^T &= \left(\boldsymbol{H}_k + \frac{1 + \left(\boldsymbol{r}^k\right)^T \boldsymbol{g}^k}{\left(\boldsymbol{s}^k\right)^T \boldsymbol{g}^k} \boldsymbol{s}^k \left(\boldsymbol{s}^k\right)^T - \left(\boldsymbol{s}^k \left(\boldsymbol{r}^k\right)^T + \boldsymbol{r}^k \left(\boldsymbol{s}^k\right)^T\right)\right)^T \\
&= \boldsymbol{H}_k^T + \left(\frac{1 + \left(\boldsymbol{r}^k\right)^T \boldsymbol{g}^k}{\left(\boldsymbol{s}^k\right)^T \boldsymbol{g}^k} \boldsymbol{s}^k \left(\boldsymbol{s}^k\right)^T\right)^T - \left(\boldsymbol{s}^k \left(\boldsymbol{r}^k\right)^T + \boldsymbol{r}^k \left(\boldsymbol{s}^k\right)^T\right)^T \\
&= \boldsymbol{H}_k + \frac{1 + \left(\boldsymbol{r}^k\right)^T \boldsymbol{g}^k}{\left(\boldsymbol{s}^k\right)^T \boldsymbol{g}^k} \left(\boldsymbol{s}^k \left(\boldsymbol{s}^k\right)^T\right)^T - \left(\left(\boldsymbol{s}^k \left(\boldsymbol{r}^k\right)^T\right)^T + \left(\boldsymbol{r}^k \left(\boldsymbol{s}^k\right)^T\right)^T\right) \\
&= \boldsymbol{H}_k + \frac{1 + \left(\boldsymbol{r}^k\right)^T \boldsymbol{g}^k}{\left(\boldsymbol{s}^k\right)^T \boldsymbol{g}^k} \left(\left(\boldsymbol{s}^k\right)^T\right)^T \left(\boldsymbol{s}^k\right)^T - \left(\left(\left(\boldsymbol{r}^k\right)^T\right)^T \left(\boldsymbol{s}^k\right)^T + \left(\left(\boldsymbol{s}^k\right)^T\right)^T \left(\boldsymbol{r}^k\right)^T\right) \\
&= \boldsymbol{H}_k + \frac{1 + \left(\boldsymbol{r}^k\right)^T \boldsymbol{g}^k}{\left(\boldsymbol{s}^k\right)^T \boldsymbol{g}^k} \boldsymbol{s}^k \left(\boldsymbol{s}^k\right)^T - \left(\boldsymbol{r}^k \left(\boldsymbol{s}^k\right)^T + \boldsymbol{s}^k \left(\boldsymbol{r}^k\right)^T\right) \\
&= \boldsymbol{H}_k + \frac{1 + \left(\boldsymbol{r}^k\right)^T \boldsymbol{g}^k}{\left(\boldsymbol{s}^k\right)^T \boldsymbol{g}^k} \boldsymbol{s}^k \left(\boldsymbol{s}^k\right)^T - \left(\boldsymbol{s}^k \left(\boldsymbol{r}^k\right)^T + \boldsymbol{r}^k \left(\boldsymbol{s}^k\right)^T\right) \\
&= \boldsymbol{H}_{k+1}.
\end{aligned}$$

Therefore, for $n \geq 0$, $\boldsymbol{H}_n$ is symmetric, by induction.

3. Newton's Method

Step 0.1 $k = 0$, $\boldsymbol{x}^0 = (0, 0)^T$.

Step 0.2 $\| \boldsymbol{\nabla} f(0,0) \| = \| (1, -1)^T \| = 1.414 > 0.01$.
$\boldsymbol{\nabla}^2 f(0,0) = \begin{pmatrix} 2 & -1 \\ -1 & 0 \end{pmatrix}$, which is not positive definite. Therefore, $\boldsymbol{d}^0 = -\boldsymbol{\nabla} f(0,0) = (-1, 1)^T$.

Step 0.3 $q(t) = f(-t, t) = t^3 + 2t^2 - 2t$, which is minimized when $3t^2 + 4t - 2 = 0 \implies t_0 = 0.3874$.

Step 0.4 $\boldsymbol{x}^1 = (0, 0)^T + 0.3874(-1, 1)^T = (-0.3874, 0.3874)^T$, $k = 1$.

Step 1.2 $\| \boldsymbol{\nabla} f(-0.3874, 0.3874) \| = \| (-0.1623, -0.1623)^T \| = 0.2295 > 0.01$.
$\boldsymbol{\nabla}^2 f(-0.3874, 0.3874) = \begin{pmatrix} 2 & -1 \\ -1 & 2.3246 \end{pmatrix}$, which is positive definite.
$\boldsymbol{d}^1 = -\boldsymbol{\nabla}^2 f(-0.3874, 0.3874)^{-1} \boldsymbol{\nabla} f(-0.3874, 0.3874)$
$= -\begin{pmatrix} 2 & -1 \\ -1 & 2.3246 \end{pmatrix}^{-1} \begin{pmatrix} -0.1623 \\ -0.1623 \end{pmatrix} = (0.1478, 0.1334)^T$.

Step 1.3 $q(t) = f(-0.3874 + 0.1478t, 0.3874 + 0.1334t) = 0.0024t^3 + 0.0228t^2 - 0.0456t - 0.4165$, which is minimized when $0.0071t^2 + 0.0456t - 0.0456 = 0 \implies t_1 = 0.8793$.

Step 1.4 $\boldsymbol{x}^2 = (-0.3874, 0.3874)^T + 0.8793(0.1478, 0.1334)^T$
$= (-0.2574, 0.5047)^T$, $k = 2$.

Step 2.2 $\| \boldsymbol{\nabla} f(-0.2574, 0.5047) \| = \| (-0.0196, 0.0217)^T \| = 0.0292 > 0.01$.
$\boldsymbol{\nabla}^2 f(-0.2574, 0.5047) = \begin{pmatrix} 2 & -1 \\ -1 & 3.0284 \end{pmatrix}$, which is positive definite.
$\boldsymbol{d}^2 = -\boldsymbol{\nabla}^2 f(-0.2574, 0.5047)^{-1} \boldsymbol{\nabla} f(-0.2574, 0.5047)$
$= \begin{pmatrix} 2 & -1 \\ -1 & 3.6618 \end{pmatrix}^{-1} \begin{pmatrix} -0.0196 \\ 0.0217 \end{pmatrix} = (0.0074, -0.0047)^T$.

Step 2.3 $q(t) = f(-0.2574 + 0.0074t, 0.5047 - 0.0047t) = -0.0000001t^3 + 0.0001239t^2 - 0.0002479t - 0.4373760$, which is minimized when $-0.0000003t^2 + 0.000248t - 0.000248 = 0 \implies t_2 = 1.0013$.

Step 2.4 $\boldsymbol{x}^3 = (-0.2574, 0.5047)^T + 1.0013(0.0074, -0.0047)^T$
$= (-0.2500, 0.5000)^T$, $k = 3$.

Step 3.2 $\| \boldsymbol{\nabla} f(-0.2500, 0.5000) \| = \| (0.0000, 0.0000)^T \| = 0.0000 < 0.01$.

We have that $\boldsymbol{x}_{\min} = (-0.2500, 0.5000)$.

The steepest descent method (from Tutorial 4) and Newton's method both found the minimum in 3 iterations. Although Newton's method requires more calculations it found the minimum of f to a greater degree of accuracy.

4. BFGS quasi-Newton Method

Step 0.1 $k=0$, $\boldsymbol{x}^0=(0,0)^T$, $\boldsymbol{H}_0=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$.

Step 0.2 $\parallel \boldsymbol{\nabla} f(0,0) \parallel = \parallel (1,-1)^T \parallel = 1.414 > 0.01$.
$\boldsymbol{d}^0 = -\boldsymbol{H}_0 \boldsymbol{\nabla} f(0,0) = (-1,1)^T$.

Step 0.3 $q(t) = f(-t,t) = t^3+2t^2-2t$, which is minimized when $3t^2+4t-2=0$ $\Longrightarrow t_0 = 0.3874$.

Step 0.4 $\boldsymbol{x}^1 = (0,0)^T + 0.3874(-1,1)^T = (-0.3874, 0.3874)^T$,
$\boldsymbol{s}^1 = (-0.3874, 0.3874)^T$, $\boldsymbol{g}^1 = (-1.1623, 0.8377)^T$, $\boldsymbol{r}^1 = (-1.5000, 1.0811)^T$,
$\boldsymbol{H}_1 = \begin{pmatrix}0.5446 & 0.2931\\ 0.2931 & 0.8692\end{pmatrix}$, $k=1$.

Step 1.2 $\parallel \boldsymbol{\nabla} f(-0.3874, 0.3874) \parallel = \parallel (-0.1623, -0.1623)^T \parallel = 0.2295 > 0.01$.
$\boldsymbol{d}^1 = -\boldsymbol{H}_1 \boldsymbol{\nabla} f(-0.3874, 0.3874) = (0.1359, 0.1886)^T$.

Step 1.3 $q(t) = f(-0.3874+0.1358t, 0.3874+0.1886t) = 0.0067094t^3 + 0.0341868t^2 - 0.0526675t - 0.4165020$, which is minimized when $0.0201283t^2 + 0.0683736t - 0.0526675 = 0 \Longrightarrow t_1 = 0.6471$.

Step 1.4 $\boldsymbol{x}^2 = (-0.3874, 0.3874)^T + 0.6471(0.1359, 0.1886)^T = (-0.2995, 0.5095)^T$.
$\boldsymbol{s}^2 = (0.0880, 0.1220)^T$, $\boldsymbol{g}^2 = (0.0539, 0.2404)^T$, $\boldsymbol{r}^2 = (2.9289, 6.5948)^T$,
$\boldsymbol{H}_2 = \begin{pmatrix}0.6522 & 0.2197\\ 0.2197 & 0.4584\end{pmatrix}$, $k=2$.

Step 2.2 $\parallel \boldsymbol{\nabla} f(-0.2995, 0.5095) \parallel = \parallel (-0.1084, 0.0781)^T \parallel = 0.1336 > 0.01$.
$\boldsymbol{d}^2 = -\boldsymbol{H}_2 \boldsymbol{\nabla} f(-0.2995, 0.5095) = (0.0535, -0.0120)^T$.

Step 2.3 $q(t) = f(-0.2995+0.0535t, 0.5095-0.0120t) = -0.0000017t^3 + 0.0037274t^2 - 0.0067396t - 0.4344498$, which is minimized when $-0.0000051t^2 + 0.0074549t - 0.0067396 = 0 \Longrightarrow t_2 = 0.9046$.

Step 2.4 $\boldsymbol{x}^3 = (-0.2995, 0.5095)^T + 0.9046(0.0535, -0.0120)^T = (-0.2510, 0.4986)^T$, $k = 3$.

Step 3.2 $\parallel \boldsymbol{\nabla} f(-0.2510, 0.4986) \parallel = \parallel (0.0007, 0.0031)^T \parallel = 0.0031 < 0.01$.

We have that $\boldsymbol{x}_{\min} = (-0.2510, 0.4986)$.

All three algorithms found the minimum in 3 iterations. Newton's method required the most involved calculations (for example, the calculation of the inverse of the Hessian at each iteration), and the steepest descent method the least. However, Newton's method was the most accurate, and the steepest descent method the least.