HOMEWORK
Jun Cheng
February 17, 2016

Problem 1:

$$\begin{aligned}
f(x_1,x_2) &= x_1^3 + 2x_1x_2 - 3x_1^2x_2^2 \\
\frac{df}{dx_1} &= 2x_1^2 + 2x_2 - 6x_1x_2^2 \\
\frac{df}{dx_2} &= 2x_1 - 6x_1^2x_2 \\
f_{x_1x_1} &= 4x_1 - 6x_2^2 \\
f_{x_1x_2} &= 2 - 12x_1x_2 \\
f_{x_2x_1} &= 2 - 12x_1x_2 \\
f_{x_2x_2} &= -6x_1^2
\end{aligned}$$

Tylor expansion:

$$f(\mathbf{x}) = f(\mathbf{x}^0) + (\mathbf{x}-\mathbf{x}^0)^T \bigtriangledown f(\mathbf{x}^0) + (\mathbf{x}-\mathbf{x}^0)^T H(\mathbf{x})(\mathbf{x}-\mathbf{x}^0)$$

where $\mathbf{x}^0 = \begin{bmatrix}1\\1\end{bmatrix}$
For linear approximation:

$$\begin{aligned}
l(x_1,x_2) &= f(\mathbf{x}^0) + (\mathbf{x}-\mathbf{x}^0)^T \bigtriangledown f(\mathbf{x}^0) \\
&= f\left(\begin{bmatrix}1\\1\end{bmatrix}\right) + \begin{bmatrix}x_1-1 & x_2-1\end{bmatrix}\begin{bmatrix}-1\\-4\end{bmatrix} \\
&= -x_1 - 4x_2 + 5
\end{aligned}$$

Hessian:

$$H = \begin{bmatrix} f_{x_1x_1} & f_{x_1x_2} \\ f_{x_2x_1} & f_{x_2x_2} \end{bmatrix}$$

For quadratic approximation:

$$\begin{aligned}
f(\mathbf{x}) &= f(\mathbf{x}^0) + (\mathbf{x}-\mathbf{x}^0)^T \bigtriangledown f(\mathbf{x}^0) + (\mathbf{x}-\mathbf{x}^0)^T H(\mathbf{x})(\mathbf{x}-\mathbf{x}^0) \\
&= -x_1 - 4x_2 + 5 + \begin{bmatrix}x_1-1 & x_2-1\end{bmatrix}\begin{bmatrix}4x_1-6x_2^2 & 2-12x_1x_2 \\ 2-12x_1x_2 & -6x_1^2\end{bmatrix}\begin{bmatrix}x_1-1\\x_2-1\end{bmatrix} \\
&= -x_1 - 4x_2 + 5 + (2x_1 - 6x_1^2x_2)(x_1-1)^2 - 6x_1^2(x_2-1)^2 + (4-24x_1x_2)(x_1-1)(x_2-1)
\end{aligned}$$

Problem 2:

1.

$$
\begin{aligned}
f(\mathbf{x}) &= \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} x_1 // x_2 \end{bmatrix} - \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} -2 \\ 1 \end{bmatrix} + \pi \\
&= x_1^2 + 3x_2^2 + 6x_1x_2 + 2x_1 + x_2 + \pi \\
\frac{df}{dx_1} &= 2x_1 + 6x_2 + 2 \\
\frac{df}{dx_2} &= 6x_2 + 6x_1 + 1 \\
Df(\mathbf{x}) &= \begin{bmatrix} \frac{df}{dx_1} & \frac{df}{dx_2} \end{bmatrix} = \begin{bmatrix} 2x_1 + 6x_2 + 2 & 6x_2 + 6x_1 + 1 \end{bmatrix}
\end{aligned}
$$

2.

$$
\begin{aligned}
f(\mathbf{x}) &= \frac{1}{2} \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 9 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 2 \\ -3 \end{bmatrix} + \log 3 \\
&= x_1^2 + \frac{1}{2}x_2^2 + 6x_1x_2 + 2x_1 - 3x_2 + \log 3 \\
\frac{df}{x_1} &= 2x_1 + 6x_2 + 2 \\
\frac{df}{x_2} &= x_2 + 6x_1 - 3 \\
f_{x_1x_1} &= 2 \\
f_{x_1x_2} &= 6 \\
f_{x_2x_1} &= 6 \\
f_{x_2x_2} &= 1
\end{aligned}
$$

Hessian:

$$H = \begin{bmatrix} 2 & 6 \\ 6 & 1 \end{bmatrix}$$

Problem 3:

$$f(x_1, x_2) = e^{3x_1x_2^2}$$

1.

$$
\begin{aligned}
\frac{df}{x_1} &= 3x_2^2 e^{3x_1x_2^2} \\
\frac{df}{x_2} &= 6x_1x_2 e^{3x_1x_2^2}
\end{aligned}
$$

Then the gradient at $\mathbf{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is :

$$\begin{bmatrix} 3e^3 \\ 6e^3 \end{bmatrix}$$

2. Rate of increase of f at the point $\mathbf{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ in the direction $\mathbf{d}$

$$\frac{Df(\mathbf{x}) \cdot \mathbf{d}}{\|\mathbf{d}\|} = \frac{15e^3}{\sqrt{5}}$$

3. In the direction $\mathbf{d} = \begin{bmatrix} 3 \\ 6 \end{bmatrix}$, we can have the maximum rate of increase $3\sqrt{5}e^3$

Problem 4:

$$\begin{aligned} f(x_1, x_2, x_3) &= -(x_1^2 + 4\epsilon x_2^2 + 5x_3^2 - 2x_1x_3 + 2\epsilon x_1x_2 + 4x_2x_3) \\ &= -\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} 1 & \epsilon & -1 \\ \epsilon & 4\epsilon & 2 \\ -1 & 2 & 5 \end{bmatrix} \end{aligned}$$

Using principal minor method:

$$\begin{aligned} &k = 1: \\ &\quad D_1 = 20\epsilon - 4 > 0 \\ &\quad D_2 = 7 \\ &\quad D_3 = 4\epsilon - \epsilon^2 > 0 \\ &k = 2: \\ &\quad D_1 = 1 \\ &\quad D_2 = 4\epsilon > 0 \\ &\quad D_3 = 5 \\ &k = 3: \\ &\quad D = 12\epsilon - 4 - 5\epsilon^2 > 0 \end{aligned}$$

To make the function be negative semi-definite, the range of ϵ is $0.4 \leq \epsilon \leq 2$

Problem 5:

1.

$$\begin{aligned} f(x_1, x_2) &= \frac{1}{3}x_2^3 + \frac{1}{2}x_2^2 + 2x_1x_2 + \frac{1}{2}x_1^2 - x_1 + 10 \\ \frac{df}{dx_1} &= 2x_2 + x_1 - 1 = 0 \\ \frac{df}{dx_2} &= x_2^2 + x_2 + 2x_1 = 0 \end{aligned}$$

We can get two points $(-1, 1)$ and $(-3, 2)$ which satisfy the first-order necessary conditions for the extremum.

2.

$$\begin{aligned} \frac{df}{dx_1dx_1} &= 1 \\ \frac{df}{dx_1dx_2} &= 2 \\ \frac{df}{dx_2dx_1} &= 2 \\ \frac{df}{dx_2dx_2} &= 2x_2 + 1 \end{aligned}$$

Table 1: Problem 7 Golden section search.

k	a_k	b_k	f(a_k)	f(b_k)	New interval
1	-0.3	-0.14	0.27	0.0588	0.576
2	-0.14	-0.0457	0.0588	0.006	0.356
3	-0.0407	0.0137	0.006	000056	0.22
4	0.0137	0.05	0.00056	0.0075	0.135
5	0.009	0.0137	0.0002	0.0006	0.083

Then Hessian is :

$$\begin{bmatrix} 1 & 2 \\ 2 & 2x_2 + 1 \end{bmatrix}$$

For point $(-1, 1)$,
$H = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$ is indefinite For point $(-3, 2)$,
$H = \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix}$ is positive definite.
Then the point $(-3, 2)$ is a strict local minimizer.

Problem 6:

$$\begin{aligned} f(\mathbf{x}) &= \mathbf{x}^T \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \mathbf{x} \\ &= x_1^2 + x_1 x_2 + x_2^2 \end{aligned}$$

The direction of travel is:

$$\mathbf{d} = -\mathbf{g} = -\frac{df}{d\mathbf{x}} = (\boldsymbol{A} + \boldsymbol{A}^T)\mathbf{x} = -\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \mathbf{x} = \begin{bmatrix} -2x_1 - x_2 \\ -x_1 - 2x_2 \end{bmatrix}$$

when $\epsilon_1 = 0.1; f(x^{(0)}) = 1.065$

$$\begin{aligned} \mathbf{x}^{(1)} &= \mathbf{x}^{(0)} + \epsilon_1 \cdot \mathbf{d} \\ &= \begin{bmatrix} 0.55 \\ 0.55 \end{bmatrix} \end{aligned}$$

when $\epsilon_2 = 0.2; f(x^{(1)}) = 1.65$

$$\begin{aligned} \mathbf{x}^{(2)} &= \mathbf{x}^{(1)} + \epsilon_2 \cdot \mathbf{d} \\ &= \begin{bmatrix} -0.11 \\ -0.11 \end{bmatrix} \end{aligned}$$

Problem 7:

See Table 1 for Golden Section search.

Table 2: Problem 8: Fibonacci search

k	a_k	b_k	f(a_k)	f(b_k)	New interval
1	-0.42	-0.19	0.53	0.108	0.58
2	-0.2	-0.05	0.12	0.0075	0.36
3	-0.056	0.016	0.009	0.0007	0.072

Problem 8:

See Table 2 for Fibonacci search.

For Newton Method :

$$
\begin{aligned}
f(x) &= 3x^2 \\
f'(x) &= 6x \\
f''(x) &= 6 \\
\frac{f'(x)}{f''(x)} &= x
\end{aligned}
$$

Therefore, we can pick up a random staring point within the initial range, then one step would be enough for a quadratic equation.

$$
\begin{aligned}
x^{(0)} &= 0.46 \\
x^{(1)} &= x^{(0)} - \frac{f'(x^{(0)})}{f''(x^{(0)})} \\
&= 0.46 - 0.46 \\
&= 0 \\
x^{(2)} &= x^{(1)} - \frac{f'(x^{(1)})}{f''(x^{(1)})} \\
&= 0
\end{aligned}
$$

Problem 9:

```
[X, Y] = meshgrid(-1:0.1:1, -1:0.1:1);
Z = (X-Y).^4 + 12*X.*Y - Y + X + 5;
mesh(X, Y, Z) ;
contour(X, Y, Z, 20 ) ;
%xlabel(X1) ;
%ylabel(X2);
box on;
```

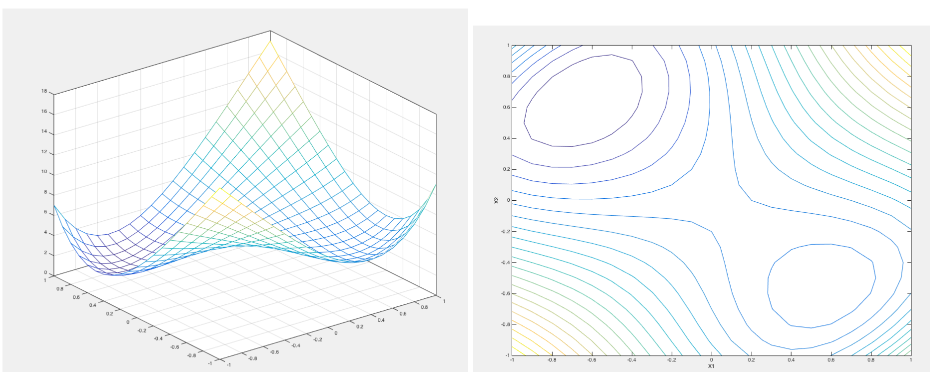

Figure 1: Left: Problem 9(a) Mesh plot . Right: Problem 9(b) Contour plot.