CO496 Coursework 1

Jinsung Ha

26th October 2018

1 Differentiation

(a) Expand the completed square form by substituting the variables.

$$\textit{Let } C = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}, \vec{c} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \textit{ and } \vec{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$\begin{aligned}
f_1(\vec{x}) &= (\vec{x} - \vec{c})^T C(\vec{x} - \vec{c}) + c_0 \\
&= \begin{pmatrix} x_1 - c_1 \\ x_2 - c_2 \end{pmatrix}^T \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} \begin{pmatrix} x_1 - c_1 \\ x_2 - c_2 \end{pmatrix} + c_0 \\
&= (x_1 - c_1)(c_{11}(x_1 - c_1) + c_{21}(x_2 - c_2)) \\
&\quad + (x_2 - c_2)(c_{12}(x_1 - c_1) + c_{22}(x_2 - c_2)) + c_0 \\
&= c_{11}(x_1 - c_1)^2 + c_{22}(x_2 - c_2)^2 + (c_{21} + c_{12})(x_1 - c_1)(x_2 - c_2) + c_0 \\
&= c_{11}(x_1^2 + c_1^2 - 2c_1x_1) + c_{22}(x_2^2 + c_2^2 - 2c_2x_2) \\
&\quad + (c_{12} + c_{21})(x_1x_2 - c_2x_1 - c_1x_2 + c_1c_2) + c_0 \\
&= c_{11}x_1^2 + c_{11}c_1^2 - 2c_1c_{11}x_1 + c_{22}x_2^2 + c_2^2c_{22} - 2c_2c_{22}x_2 \\
&\quad + c_{12}x_1x_2 - c_2c_{12}x_1 - c_1c_{12}x_2 + c_1c_2c_{21} + c_0 \\
&= c_{11}x_1^2 + c_{22}x_2^2 + (c_{12} + c_{21})x_1x_2 + x_1(-2c_1c_{11} - c_2c_{12} - c_2c_{21}) \\
&\quad + x_2(-2c_2c_{22} - c_1c_{12} - c_1c_{21}) + c_1^2c_{11} + c_2^2c_{22} + c_1c_2c_{12} + c_1c_2c_{21} + c_0
\end{aligned} \tag{1}$$

Then expand the equation from the original form.

$$\begin{aligned}
f_1(\vec{x}) &= \vec{x}^T\vec{x} + \vec{x}^T B\vec{x} - \vec{a}^T\vec{x} + \vec{b}^T\vec{x} \\
&= \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}^T \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}^T \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - \begin{pmatrix} -1 \\ 0 \end{pmatrix}^T \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} 0 \\ -1 \end{pmatrix}^T \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\
&= x_1^2 + x_2^2 + x_1(3x_1 - x_2) + x_2(-x_1 + 3x_2) - x_1 - x_2 \\
&= 4x_1^2 + 4x_2^2 - 2x_1x_2 - x_1 - x_2
\end{aligned} \tag{2}$$

Now, equating the equations (1) and (2) will result in simultaneous equations. Solving them will result in $C = \begin{pmatrix} 4 & -1 \\ -1 & 4 \end{pmatrix}, \vec{c} = \frac{1}{6}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $c_0 = -\frac{1}{6}$

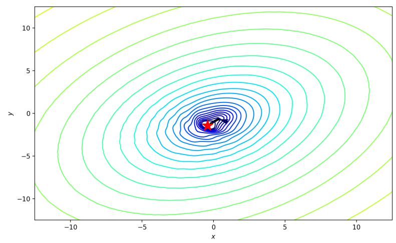

Figure 1: contour of f_2

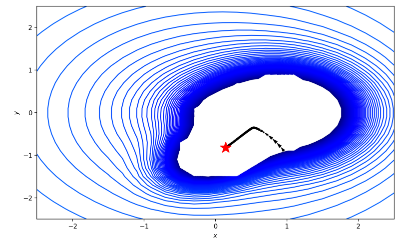

Figure 2: contour of f_3

(b) Given that the expanded form of $f_1(x)$ is $4x_1^2 + 4x_2^2 - 2x_1x_2 - x_1 - x_2$. If Hessian matrix H of f_1 is positive definite, f_1 has a local minimum.

$$Hf_1(x_1, x_2) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} \end{pmatrix} = \begin{pmatrix} 8 & -2 \\ -2 & 8 \end{pmatrix}$$

If $Hf_1(x_1, x_2)$ is positive definite, the determinant of its characteristic function should give positive eigenvalues only.

$$\begin{aligned} det(Hf_1(x_1, x_2) - \lambda I) &= \begin{vmatrix} 8-\lambda & -2 \\ -2 & 8-\lambda \end{vmatrix} \\ &= (\lambda - 8)^2 - 4 \\ &= \lambda^2 - 16\lambda + 60 \\ &= (\lambda - 6)(\lambda - 10) \ = \ 0 \end{aligned}$$

As such, eigenvalues are 6 or 10 which are both positive. Therefore, $Hf_1(x_1, x_2)$ is positive definite. Thus, the minimum value and its corresponding input can be achieved from the completed square form of $f_1(\vec{x})$.

$\therefore$ The minimum value of $f_1 = -\frac{1}{6}$ and the corresponding input $\vec{x} = \frac{1}{6}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

(f) The figure 1 and 2 show the 2d contour plot of f_2 and f_3 respectively. The step size of 0.1 is used to generate the plots.

(g)

1) f_2 has only one minima.
2) When step size is large, f_2's gradient, thus the step diverges.
3) f_3 has 2 minimas.
4) When moderate step size is used, f_3's step converges to its local minima not the global one.
5) When a large step size is used, f_3's step starts to oscillate within the function and never converges.
6) The main difference between f_2 and f_3 is that f_2 diverges far more quickly than f_3.