**DSA3102 Convex optimisation**

Assignment 1

**Instruction: please type or scan your results into a single pdf file and submit the softcopy Canvas. Please do not submit codes. Penalty will apply if you do not follow these rules.**

1. (Linear Regression) Suppose that $\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}, \cdots, \begin{pmatrix} x_n \\ y_n \end{pmatrix}$ are $n$ points in $xy$-plane and suppose that we want to 'fit' a straight line $y = a+bx$ through these points in such a way that the sum of squares of the vertical deviations of the given points from the line are as small as possible. In other words, we want to choose $a$ and $b$ so that

$$f(a,b) = \sum_{i=1}^{n} (a + bx_i - y_i)^2$$

is as small as possible. The resulting line is called the *linear regression line* for the points $\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}, \cdots, \begin{pmatrix} x_n \\ y_n \end{pmatrix}$.

Show that the coefficients $a$ and $b$ of the linear regression line are given by

$$a = \bar{y} - b\bar{x}, \quad b = \frac{n\bar{x}\bar{y} - \sum_{i=1}^{n} x_i y_i}{n(\bar{x})^2 - \sum_{i=1}^{n} (x_i)^2},$$

where

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \quad \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$$

are the averages of the $x$- and $y$- coordinates of the given points. Explain why these this solution is optimal.

2. Consider $f : \mathbf{R}^3 \to \mathbf{R}$ where

$$f(\mathbf{x}) = \ln(x_1^2 + 1) + x_1^2 - 2x_1x_2 + 4x_2^2 + x_3^4 - 8x_3^3 + 16x_3^2.$$

   (i) Determine all stationary points of $f$.

   (ii) Assume that global minimizers of $f$ exist, find all of them.

3. Program the gradient descent and Newton algorithms using the backtracking line search (Procedure below, set $\beta = 0.9$, $\sigma = 0.6$). Use them to minimize the Rosenbrock function

$$f(\mathbf{x}) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2.$$

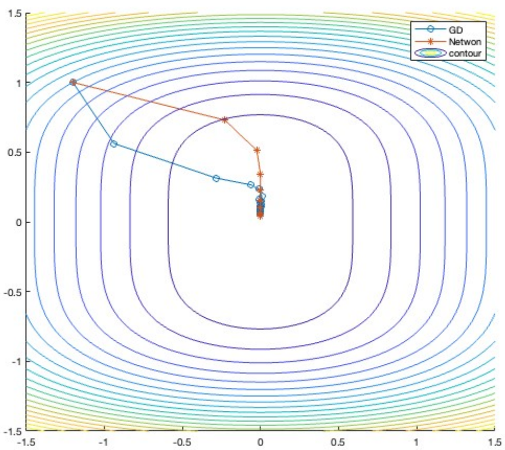


Figure 1: Q3: Plot with $f(x, y) = x^2 + y^4$

First try the initial point $\mathbf{x}_0 = (1.2; 1.2)$ and then the more difficult point $\mathbf{x}_0 = (-1.2; 1)$. Plot the contour of the function and then plot the path of iterates obtained by each method on the contour. (You can use 'help contour' and 'help plot' to find instructions of using 'contour' and 'plot'. A possible way to plot with one initial point is given below.)

**Procedure** (Backtracking Line Search).

Choose $\beta, \sigma \in (0, 1)$; Set $\alpha \leftarrow 1$;

**repeat** until $f(\mathbf{x}_k + t\mathbf{p}_k) \leq f(\mathbf{x}_k) + \sigma t \nabla f(\mathbf{x}_k)^T \mathbf{p}_k$

$t \leftarrow t\beta$;

**end (repeat)**

Terminate with $t_k = t$.