Exercise 1

N & W, Chapter 2: Either {2.1 & 2.10 } or {2.2 & 2.3}.

I choose to make the exercise {2.2 & 2.3}.

1. {2.2} Show that the function $f(x) = 8x_1+12x_2+x_1^2-2x_2^2$ has only one stationary point, and that it is neither a maximum or minimum, but a saddle point. Sketch the contour lines of f.

   Svar:

   The definition of a stationary point say *let* $f : V \to \mathbb{R}^n$ *be differentiable. Then a point* $a \in V$ *is a stationary point when* $\nabla f(a) = 0$.

   $\nabla f(x)$ is defind as:

$$\nabla f(x) = \begin{bmatrix} \frac{\partial f}{\partial x_1}(x) \\ \frac{\partial f}{\partial x_2}(x) \\ \vdots \\ \frac{\partial f}{\partial x_n}(x) \end{bmatrix}$$

   then the $\nabla f(x)$ for $f(x) = 8x_1 + 12x_2 + x_1^2 - 2x_2^2$ is;

$$\nabla f(x) = \begin{bmatrix} 8 + 2x_1 \\ 12 - 4x_2 \end{bmatrix}$$

   then the stationary point will be;

$$\nabla f(x) = \begin{bmatrix} 8 + 2x_1 \\ 12 - 4x_2 \end{bmatrix} = 0,$$

   $x_1 = -4, x_2 = 3$ is the only solution, so the stationary point is $(-4, 3)$.

   To show this is a saddle ponit we look at the definition for at saddle ponit, it says that if the functions Hessian Matrix is indefinite, which indicate that the matrix is neither positiv nor negativ definite, so the egenvalues is both positive and negative.

   The Hessian Matrix is defind as;

$$\nabla^2 f(a) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2}(a) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(a) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(a) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(a) & \frac{\partial^2 f}{\partial x_2^2}(a) & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n}(a) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(a) & \frac{\partial^2 f}{\partial x_n \partial x_2}(a) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(a) \end{bmatrix}.$$

Then the Hessian Matrix for $f(x) = 8x_1 + 12x_2 + x_1^2 - 2x_2^2$ is;

$$\nabla^2 f(x) = \begin{bmatrix} 2 & 0 \\ 0 & -4 \end{bmatrix}.$$

Which give the egenvalues $\lambda_1 = -4$ and $\lambda_2 = 2$, so the Hessian Matrix is indefinite and the point is only a saddle point.

Sketch the contour lines of f, unfortunate I can't sketch the contour lines in R, so this may be made later before the examination.

2. $\{2.3\}$ Let a be a given n-vector, and A be a given nxn symmetric matrix. Compute the gradient $\nabla f(x)$ and Hessian $\nabla^2 f(x)$ of $f_1(x) = a^T x$ and $f_2(x) = x^T Ax$

Svar:

$f_1(x)$ kan be defind as;

$$f_1(x_1, x_2, \cdots, x_n) = a^T x = a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = \sum_{i=1}^{n} a_i x_i$$

Therefor the gradient will be;

$$\nabla f_1(x) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} \\ \frac{\partial f_1}{\partial x_2} \\ \vdots \\ \frac{\partial f_1}{\partial x_n} \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = a$$

The Hessian is;

$$\nabla^2 f_1(x) = \begin{bmatrix} \frac{\partial^2 f_1}{\partial x_1^2} & \frac{\partial^2 f_1}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f_1}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f_1}{\partial x_2 \partial x_1} & \frac{\partial^2 f_1}{\partial x_2^2} & \cdots & \frac{\partial^2 f_1}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f_1}{\partial x_n \partial x_1} & \frac{\partial^2 f_1}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f_1}{\partial x_n^2} \end{bmatrix} = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}.$$

$f_2(x)$ kan be defind as;

$$f_2(x_1, x_2, \cdots, x_n) = x^T Ax = x_1 A_{11} x_1 + x_1 A_{12} x_2 + \cdots + x_n A_{nn} x_n = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i A_{ij} x_j$$

Then the gradient will be;

$$\nabla f_2(x) = \begin{bmatrix} \frac{\partial f_2}{\partial x_1} \\ \frac{\partial f_2}{\partial x_2} \\ \vdots \\ \frac{\partial f_2}{\partial x_s} \end{bmatrix}_{s=1 \cdots n} = \begin{bmatrix} \sum_j^n A_{1j} x_j + \sum_i^n A_{i1} x_i \\ \sum_j^n A_{2j} x_j + \sum_i^n A_{i2} x_i \\ \vdots \\ \sum_j^n A_{sj} x_j + \sum_i^n A_{is} x_i \end{bmatrix}_{s=1 \cdots n}.$$

Since A is Symmetric this can be defind as;

$$\begin{bmatrix} \sum_j^n A_{1j}x_j + \sum_i^n A_{i1}x_i \\ \sum_j^n A_{2j}x_j + \sum_i^n A_{i2}x_i \\ \vdots \\ \sum_j^n A_{sj}x_j + \sum_i^n A_{is}x_i \end{bmatrix}_{s=1\cdots n} = \begin{bmatrix} 2\sum_j^n A_{1j}x_j \\ 2\sum_j^n A_{2j}x_j \\ \vdots \\ 2\sum_j^n A_{sj}x_j \end{bmatrix}_{s=1\cdots n} = 2Ax$$

The Hessian is;

$$\nabla^2 f_2(x) = \begin{bmatrix} \frac{\partial^2 f_2}{\partial x_1^2} & \frac{\partial^2 f_2}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f_2}{\partial x_1 \partial x_t} \\ \frac{\partial^2 f_2}{\partial x_2 \partial x_1} & \frac{\partial^2 f_2}{\partial x_2^2} & \cdots & \frac{\partial^2 f_2}{\partial x_2 \partial x_t} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f_1}{\partial x_s \partial x_1} & \frac{\partial^2 f_1}{\partial x_s \partial x_2} & \cdots & \frac{\partial^2 f_1}{\partial x_s \partial x_t} \end{bmatrix}_{s=1\cdots n, t=1\cdots n} = \begin{bmatrix} \frac{\partial^2 \sum_i \sum_j A_{ij}x_i x_j}{\partial x_s \partial x_t} \end{bmatrix} = \begin{bmatrix} A_{st} + A_{ts} \end{bmatrix} = 2A$$

Exercise 2

N & W, Chapter 2: one of the following {2.4}, {2.5}, {2.7}, {2.9}, {2.11}.

I have choosen the exercise {2.4}

1. {2.4}

   Write the second-order Taylor expansion (2.6) for the function $cos(1/x)$ around a nonzero point x, and the third-order Taylpr expansion of $cos(x)$ around any point x. Evaluate the second expansion for the specific case of $x = 1$

   (2.6) $$f(x+p) = f(x) + \nabla f(x)^T p + \frac{1}{2} p^T \nabla^2 f(x+tp)p, \quad for\ some\ t \in (0,1).$$

   Svar:

   To use (2.6) vi know that for any funktion $f(x)$ with only one variable tha Second order Taylor expansion is;

   $$f(x+\nabla x) = f(x) + f^{(1)}(x)\nabla x + \frac{1}{2} f^{(2)}(x+t\nabla x)\nabla x^2,$$

   and the Third order Taylor expansion is;

   $$f(x+\nabla x) = f(x) + f^{(1)}(x)\nabla x + \frac{1}{2} f^{(2)}(x)\nabla x^2 + \frac{1}{6} f^{(3)}(x+t\nabla x)\nabla x^3$$

   where $t \in (0,1)$, and $\nabla x = (x - x_0)$ where $x_0 < x$.

   For the funktion $f_1(x) = cos(1/x)$ and any nonzero point x, we know that

   $$f_1^{(1)} = \frac{1}{x^2} sin\left(\frac{1}{x}\right), \quad f_1^{(2)} = -\frac{1}{x^4}\left(2xsin\left(\frac{1}{x}\right) + cos\left(\frac{1}{x}\right)\right).$$

So the Second order Taylor expansion for $f_1(x)$ is;

$$cos\left(\frac{1}{x+\nabla x}\right) = cos\left(\frac{1}{x}\right) + \frac{1}{x^2} sin\left(\frac{1}{x}\right)\nabla x - \frac{1}{2(x+\nabla x)^4}\left(cos\left(\frac{1}{x+\nabla x}\right) - 2(x+\nabla x)sin\left(\frac{1}{x+\nabla x}\right)\right)\nabla x^2,$$

where $t \in (0, 1)$.

similary, for $f_2(x) = cos(x)$, we have;

$$f^{(1)}(x) = -sin(x), \quad f^{(2)}(x) = -cos(x), \quad f^{(3)}(x) = sin(x)$$

So the Third order Taylor expansion for $f_2(x)$ is;

$$f(x+\nabla x) = cos(x) - (sin(x))\nabla x - \frac{1}{2}(cos(x))\nabla x^2 + \frac{1}{6}(sin(x+t\nabla x))\nabla x^3,$$

where $t \in (0, 1)$.

the specific case $x = 1$;

Exercise 3

N & W, Chapter 3: 3.1

1. a. Steepest Descent

   Svar:

   Is written in R

2. b. Newton

   Svar:

   Is written in R

Exercise 5

N & W, Chapter 3: One of the following: {3.3, 3.5, 3.6, 3.13, 3.15, 3.15, 3.SD}.

I have choosen exercise {3.3}

1. {3.3}

   Show that the one-demensional minimizer of a strongly convex quadratic function is given by (3.55).

$$(3.55) \qquad \alpha_k = -\frac{\nabla f_k^T p_k}{p_k^T Q p_k}$$

Svar:

From Side 56

Supose that p is a decent direction definde a one-dimensional function

$$\phi(\alpha) = f(x + \alpha p)$$

Any minimizer α^* of the function $\phi(\alpha)$ satisfies

$$\phi'(\alpha^*) = f(x + \alpha^* p)^T p = 0$$

From the book we have a strongly convex quadratic function $f(x) = \frac{1}{2}x^T Qx + b^t$, where $Q > 0$, symmetic and positive definite. The gradient is given by $\nabla f(x) = Qx + b$ (side 42). The one-dimensinal minimizer is unique, and it can be expressed as satisfying

$$[Q(x + \alpha^* p) + b]^T p = 0$$

$$(Qx + b)^T + \alpha^* p^T Qp = 0$$

By isolate α^* we get;

$$\alpha^* = -\frac{(Qx + b)^T}{p^T Qp}$$

Since $\nabla f(x) = Qx + b$ we get the result.

$$\alpha^* = -\frac{\nabla f^T p}{p^T Qp}$$