Assignment 8

Problem 1

a) Theorem 2.3

If x^* is a local minimizer of f and $\nabla^2 f$ exists and is continuous in an open neighborhood of x^*, then $\nabla f(x^*) = 0$ and $\nabla^2 f(x^*)$ is positive semidefinite.

b) From 2.2 we have that $\nabla f(x^*) = 0$, because x^* is the lowest point (meaning it's a minimizer). This is proven by contradiction, assuming that $\nabla^2 f(x^*)$ is not positive semidefinite. In the proof we see that this can't be so, because we end up with a decreasing function after a taylor expansion around x^*. Thus, $\nabla^2 f(x^*)$ must be positive semidefinite.

c) The difference lies in 2.3 only asking $\nabla^2 f(x^*)$ to be positive semidefinite, meaning there can be other points in the open neighbourhood with the same value as x^*. This is therefore not a strict local minimizer of f. As for 2.4, this sufficient condition is fullfilled and x^* is a strict local minimizer of f.

Problem 2

a)

To derive the Newton direction we simply set the derivative of $m_k(p)$ to zero:

$$\frac{d}{dp} m_k(p) = \nabla f_k + \nabla^2 f_k p = 0$$

$$\nabla^2 f_k p = -\nabla f_k$$

$$p = -(\nabla^2 f_k)^{-1} \nabla f_k$$

■

b) In the case that $\nabla^2 f_k = 0$, p will not be defined. In the case that $\nabla^2 f_k < 0$, the Newton-direction will be increasing meaning, it's not a descent direction,

c) We have that the gradient of f is given by:

$$\nabla f = \frac{1}{2} Gx + \frac{1}{2} G^T x + c$$
$$= Gx + c$$

$$\nabla^2 f = G$$

We now find the Newton direction

$$p = -(\nabla^2 f)\nabla f$$
$$= -(G)^{-1}(Gx_k + c)$$

where $G^{-1}G = I$

$$p = -x_k - G^{-1}c$$

We then use this in the eq:

$$x_{k+1} = x_k + p = x_k - x_k - G^{-1}c$$
$$= -G^{-1}c$$

Since f is convex we also know that $x^* = -G^{-1}c = x_{k+1}$.

thus f converges in one step.

d) For $f(x)$ to be convex it must satisfy

$$f(\alpha x+(1-\alpha)y) \leq \alpha f(x)+(1-\alpha)f(y),$$

$\forall \alpha \in [0,1]$. We know also that the set $X = \{x \in \mathbb{R}^2 \mid x_1^2+x_2^2 \leq 1\}$

is convex.

We start with the left-hand side of the inequality:

$$f(\alpha x+(1-\alpha)y) = \frac{1}{2}(\alpha x+(1-\alpha)y)^T G(\alpha x+(1-\alpha)y) + (\alpha x+(1-\alpha)y)^T c$$

$$= \frac{1}{2}\alpha x^T G(\alpha x+(1-\alpha)y) + \frac{1}{2}y^T G(\alpha x+(1-\alpha)y) + \alpha x^T c + y^T c - \alpha y^T c - \frac{1}{2}\alpha y^T G(\alpha x+(1-\alpha)y)$$

$$= \frac{1}{2}\alpha^2 x^T G x + \frac{1}{2}\alpha x^T G(1-\alpha)y + \frac{1}{2}y^T G\alpha x + \frac{1}{2}y^T G(1-\alpha)y - \frac{1}{2}\alpha^2 y^T G x - \frac{1}{2}\alpha y^T G(1-\alpha)y + \alpha x^T c + (1-\alpha)y^T c$$

The right hand side is given by.

$$\alpha f(x) + (1-\alpha) f(y) = \frac{1}{2}\alpha x^T G x + \alpha x^T c$$
$$+ (1-\alpha)\frac{1}{2} y^T G y + (1-\alpha) y^T c$$

we have the same linear terms on the right and left hand side.

the inequality then becomes:

$$\frac{1}{2}\alpha^2 x^T G x + \frac{1}{2}\alpha x^T G (1-\alpha) y + \frac{1}{2} y^T G \alpha x$$
$$+ \frac{1}{2} y^T G (1-\alpha) y - \frac{1}{2}\alpha^2 y^T G x - \frac{1}{2}\alpha\, y^T G (1-\alpha) y$$
$$\leq \frac{1}{2} x^T G x + (1-\alpha)\frac{1}{2} y^T G y$$

Problem 3

2.1 $f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$

$\nabla_{x_1} f = 2 \cdot 100(x_2 - x_1)(-2x_1) + 2(1 - x_1)(-1)$

$\nabla_{x_2} f = 2 \cdot 100(x_2 - x_1) \cdot 1$

$$\nabla f = \begin{bmatrix} -400(x_2 - x_1)x_1 - 2(1 - x_1) \\ 200(x_2 - x_1) \end{bmatrix}$$

$\nabla_{x_1 x_1} f = -400(x_2 - x_1) + 1$

$\nabla_{x_1 x_2} f = -400$

$\nabla_{x_2 x_1} f = -200$

$\nabla_{x_2 x_2} f = 200$

$$\nabla^2 f = \begin{bmatrix} -400(x_2 - x_1) + 1 & -400 \\ -200 & 200 \end{bmatrix}$$