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) By showing that if $\nabla^2 f < 0$, i.e. not positive semidefinite, there exist a direction from x^* s.t. the value of f decreases, x^* is not a local minimizer and the theorem is proven by contradiction.

c) In theorem 2.4 it's required that $\nabla^2 f$ is positive definite, i.e. $\nabla^2 f > 0$. This is so $\nabla^2 f = 0$ and we have multiple solution. If $\nabla^2 f > 0 \ \forall \ x \in \mathcal{N}(x^*)$, only x^* can be the local minimum.

Problem 2

$$m_k(p) = f_k + p^T \nabla f_k + \frac{1}{2} p^T \nabla^2 f_k \, p \approx f(x_k + p)$$

a) $\nabla_p m_k = \nabla f_k + \nabla^2 f_k \, p = 0$

$$\Rightarrow \quad p_k^N = -(\nabla^2 f_k)^{-1} \nabla f_k \quad \square.$$

b) If $\nabla^2 f_k = 0$, the Newton direction is undefined.

If $\nabla^2 f_k < 0$, the Newton direction doesn't satisfy the descent property

$$\nabla f_k^T p_k^N = \underbrace{-\nabla f_k^T \, \nabla^2 f_k^{-1} \, \nabla f_k}_{<0} \not< 0$$

and is therefore not suitable for a descent direction.

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

$\nabla f_k = G x_k + c = 0 \Rightarrow \underline{x^* = -G^{-1} c}$

$\nabla^2 f_k = G > 0 \qquad \nabla^2 f^{-1} = G^{-1} > 0$

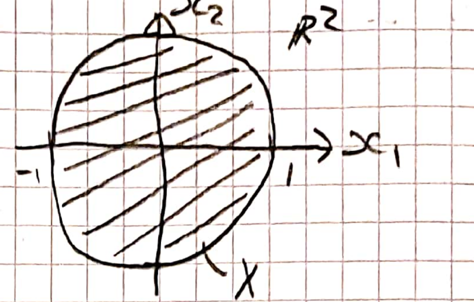
$p_k^N = -G^{-1}(G x_k + c) = -x_k - G^{-1} c$

$x_{k+1} = x_k + p_k^N = x_k - x_k - G^{-1} c = \underline{-G^{-1} c} = x^*$

Since $\nabla^2 f_k = G > 0$, the local minimizer is simply found by solving $\nabla f(x^*) = 0 \Rightarrow x^* = -G^{-1} c$. As shown, regardless of x_k, this value is obtained after a single iteration using the Newton direction.

d) $f(x) = \frac{1}{2} x^T G x + x^T c$ $\qquad x \in \{ x \in \mathbb{R}^2 / x_1^2 + x_2^2 \leq 1 \} = X$

$$G = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \qquad c = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$



Definition of convexity

$$f(\alpha x + (1-\alpha) y) \leq \alpha f(x) + (1-\alpha) f(y) \quad \forall \alpha \in [0,1]$$
$$x, y \in X \rightarrow \alpha x + (1-\alpha) y \in X.$$

Convexity of domain X:

- Let $x, y \in X$: $x_1^2 + x_2^2 \leq 1$, $y_1^2 + y_2^2 \leq 1$.
- $\alpha (x_1^2 + x_2^2) + (1-\alpha)(y_1^2 + y_2^2)$

$$= \underbrace{\alpha [\underbrace{x_1^2 + x_2^2}_{\leq 1} - \underbrace{(y_1^2 + y_2^2)}_{\leq 1}]}_{} + \underbrace{(y_1^2 + y_2^2)}_{\leq 1} \leq 1 \quad \forall \alpha \in [0,1].$$

$\downarrow$ $\qquad \square$

at most $= 1$ $(x = 1, y = 0)$ $\qquad \alpha x + (1-\alpha) y \in X$.

at least $= -1$ $(x = 0, y = 1)$

Convexity of f:

- let $x, y \in X$.
- $G > 0$ because $\begin{vmatrix} \lambda - 2 & -1 \\ -1 & \lambda - 2 \end{vmatrix} = (\lambda - 2)^2 - 1 = 0$

$$\lambda = 2 \pm 1 > 0.$$

- $f(\alpha x + (1-\alpha) y) = (LHS)$

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

- $\alpha f(x) + (1-\alpha) f(y) = (RHS)$

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

$\longrightarrow$

Since $\alpha \in [0,1] \rightarrow \alpha^2 < \alpha$. Since the difference between (RHS) and (LHS) is α^2 and $(1-\alpha)^2$ coefficients of the quadratic terms, and since $G > 0 \Rightarrow$ (LHS) $\leq$ (RHS). $\square$

It's shown that both the function and its domain in convex.

Problem 3

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

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

$$\nabla^2 f(x) = \begin{bmatrix} -400(x_2 - x_1^2) + 800x_1^2 + 2 & -400x_1 \\ -400x_1 & 200 \end{bmatrix}$$

$$= \begin{bmatrix} 1200x_1^2 - 400x_2 + 2 & -400x_1 \\ -400x_1 & 200 \end{bmatrix}$$

$x^* = [1 \quad 1]^T$: $\nabla f(x^*) = \begin{bmatrix} -400 \cdot 1(1-1^2) - 2(1-1) \\ 200(1-1^2) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ ✓

$$\nabla^2 f(x^*) = \begin{bmatrix} 1200 \cdot 1^2 - 400 \cdot 1 + 2 & -400 \cdot 1 \\ -400 \cdot 1 & 200 \end{bmatrix} = \begin{bmatrix} 802 & -400 \\ -400 & 200 \end{bmatrix}$$

Both leading principal minors $> 0 \Rightarrow \nabla^2 f(x^*) > 0$ ✓

Both first-order necessary condition and second order sufficient condition satisfied $\Rightarrow$ x^* is local minimizer.