$\boxed{1}$

ⓐ 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 $\nabla^2 f(x^*) \geq 0$.

ⓑ The proof simply says that if $\nabla^2 f(x^*) < 0$ then there exists a direction in the neighborhood of x^* where we can move some small amount and go "downwards" on the surface, by using the Taylor expansion.

ⓒ T2.3 only assumes that $\nabla^2 f(x^*) \geq 0$, i.e. positive semi-definiteness, which can include saddle points, i.e. not strict minimizers.

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

ⓐ

$$\frac{\partial m_k}{\partial p} = \nabla f_k + \nabla^2 f_k p = 0 \implies \underline{p_k^N = -(\nabla^2 f_k)^{-1} \nabla f_k} \quad \square$$

ⓑ $p_k^{N^T} \nabla f_k = -\nabla f_k^T (\nabla^2 f_k)^{-T} \nabla f_k$

$$= -\nabla f_k^T \nabla^2 f_k^{-T} \nabla^2 f_k^T \nabla^2 f_k^{-T} \nabla f_k$$

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

$$= -y^T \nabla^2 f_k^T y > 0 \quad \text{since } \nabla^2 f_k < 0$$

$$\implies \underline{p_k^{N^T} \nabla f_k > 0}$$

If $\nabla^2 f_k < 0$, then p_k^N is not a descent direction.

ⓒ $f(x) = \frac{1}{2} x^T G x + x^T c$, $x \in \mathbb{R}^n$, $G = G^T > 0$

$\nabla f_k = Gx$, $\nabla^2 f_k = G$,

$p_k^N = -G^{-1} G x_k = -x_k \implies \underline{x_{k+1}} = x_k - x_k \underline{= 0}$

Since the optimum is always in the origin, regardless of G and c, we can just simply go to 0 - duh!

ⓓ $f(x) = \frac{1}{2}x^T G x + x^T c$, $G = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$, $c = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$,

$x \in X$, $X = \{x \in \mathbb{R}^2 \mid x_1^2 + x_2^2 \leq 1\}$

Convex if: $f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha) f(y)$,
$\alpha = [0, 1]$

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

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

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

$$\Rightarrow \alpha^2 x^T G x + \alpha(1-\alpha) x^T G y + \alpha(1-\alpha) y^T G x + (1-\alpha)^2 y^T G y \leq \alpha x^T G x + y^T G y - \alpha y^T G y$$

$$\alpha^2 x^T G x + \alpha(1-\alpha) x^T G y + \alpha(1-\alpha) y^T G x - \alpha y^T G y + \alpha^2 y^T G y \leq \alpha x^T G x$$

$$\Rightarrow (\alpha - 1)(x^T G x + y^T G y - x^T G y - y^T G x) \leq 0$$

$$\Rightarrow x^T G x + y^T G y \geq x^T G y + y^T G x$$

$$\Leftrightarrow (x-y)^T G (x-y) \geq 0$$

Since $z^T G z \geq 0$ for any z per definition, this holds. Which means $f(x)$ is convex. Finally the domain X is obviously convex, as it is just a circle.

$\boxed{3}$ $f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$

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

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

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

$\nabla f = 0 \Rightarrow x_2 = x_1^2 \Rightarrow x_1 = 1 \Rightarrow x_2 = 1$

$\Rightarrow x^* = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

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

$\det(\lambda I - \nabla^2 f(x^*)) = (\lambda - 802)(\lambda - 200) - 160000$

$= \lambda^2 - 1002\lambda + 400 = 0$

$\Rightarrow \lambda_1 \approx 0{,}3994, \quad \lambda_2 \approx 1001{,}6$

$\Rightarrow \underline{\nabla^2 f(x^*) > 0}$

Scanned by CamScanner