Exercise 1

✎ **1** Compute the gradient $\nabla f(x)$ and Hessian $\nabla^2 f(x)$ of the Rosenbrock function

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

Show that $x^* = (1, 1)^T$ is the only local minimizer of this function, and that the Hessian matrix at that point is positive definite.

$$\text{Sol}: \quad \nabla f_{(x)} = \begin{bmatrix} \frac{\partial f}{\partial x_1}(x) \\ \frac{\partial f}{\partial x_2}(x) \end{bmatrix} = \begin{bmatrix} -400\, x_1 (x_2 - x_1^2) + 2x_1 - 2 \\ 200\, (x_2 - x_1^2) \end{bmatrix}$$

$$\nabla^2 f_{(x)} = \begin{bmatrix} -400(x_2 - x_1^2) + (-400x_1)(-2x_1) + 2 & -400x_1 \\ -400x_1 & 200 \end{bmatrix} = \begin{bmatrix} 1200x_1^2 - 400x_2 + 2 & -400x_1 \\ -400x_1 & 200 \end{bmatrix}$$

let $\nabla f_{(x)} = 0 \Rightarrow$
$$\begin{cases} 400\, x_1^3 - 400 x_1 x_2 + 2(x_1 - 1) = 0 \\ 200\, (x_2 - x_1^2) = 0 \end{cases} \Rightarrow \begin{cases} x_1 = 1 \\ x_2 = 1 \end{cases} \quad \text{only stationary point.}$$

$$\nabla^2 f\,{(1,1)}^T = \begin{bmatrix} 802 & -400 \\ -400 & 200 \end{bmatrix} \qquad (\lambda - 802)(\lambda - 200) - 160000 = 0 \quad \Rightarrow \quad \lambda^2 - 1002\lambda + 400 = 0$$

$$\Rightarrow \quad \lambda = 501 \pm \sqrt{250601} \quad . \text{ both} > 0 \quad (501^2 = 251001 > 250601).$$

$\nabla^2 f$ at $(1,1)^T$ is p.d. $f \in C^2(U_{x^*})$. $\nabla f(x^*) = 0$, thus $(1,1)^T$ is and is the only local minimum.

✎ **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.

$$\text{Sol}: \quad \nabla f_{(x)} = \begin{bmatrix} 8 + 2x_1 \\ 12 - 4x_2 \end{bmatrix}$$
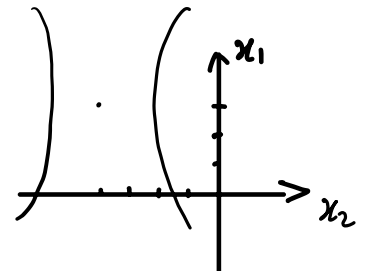
let $\nabla f_{(x)} = 0 \Rightarrow$ the only solution $\begin{cases} x_1 = -4 \\ x_2 = 3 \end{cases}$ i.e. the only stationary point $(-4, 3)$

$$\nabla^2 f = \begin{bmatrix} 2 & 0 \\ 0 & -4 \end{bmatrix} \quad \text{not definite at } x^* = (-4, 3)^T,$$

thus don't satisfy the 2nd order necessary condition. i.e. x^* is not a local extremum

$$f(x) = (x_1 + 4)^2 - [2(x_2 - 3)^2] - 16 + 18 = (x_1 + 4)^2 - 2(x_2 - 3)^2 + 2.$$

Contour line: $f(x) = C \Rightarrow (x_1 + 4)^2 - 2(x_2 - 3)^2 = C - 2$

hyperbolics centred at $(-4, 3)$

✎ **3** Let a be a given n-vector, and A be a given $n \times n$ symmetric matrix. Compute the gradient and Hessian of $f_1(x) = a^T x$ and $f_2(x) = x^T A x$.

Sol: $f_1 = (a_1 \cdots a_n)^T \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \sum_{i=1}^{n} a_i x_i \qquad \nabla f_1(x) = a \qquad \nabla^2 f_1(x) = 0_{n \times n}$

$f_2 = (x_1 \cdots x_n)^T \begin{bmatrix} a_{11} & & a_{1n} \\ & \ddots & \\ & & a_{nn} \end{bmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \sum_{i=1}^{n} a_{ii} x_i^2 + 2 \sum_{1 \le i < j \le n} a_{ij} x_i x_j$

$\nabla f_2(x) = \begin{bmatrix} 2\sum a_{1i} x_i \\ 2\sum a_{2i} x_i \\ \cdots \\ 2\sum a_{ni} x_i \end{bmatrix} = 2Ax \qquad \nabla^2 f_2(x) = 2A.$

✎ **4** Write the second-order Taylor expansion

$$f(x+p) = f(x) + \nabla f(x)^T p + \tfrac{1}{2} p^T \nabla^2 f(x+tp) p,$$

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

Sol: (1) for $f(x) = \cos(1/x)$.

$\nabla f(x) = \frac{1}{x^2} \sin(\frac{1}{x})$. $\qquad \nabla^2 f(x) = \frac{-\frac{1}{x^2}\cos(\frac{1}{x})\cdot x^2 - 2x \cdot \sin\frac{1}{x}}{x^4} = -\frac{\cos(\frac{1}{x}) + 2x\sin\frac{1}{x}}{x^4}$

$\cos(\frac{1}{x+p}) = \cos(\frac{1}{x}) + \frac{\sin(1/x)}{x^2}p + \frac{1}{2}\left[-\frac{2\sin(\frac{1}{x+tp})}{(x+tp)^3} - \frac{\cos(\frac{1}{x+tp})}{(x+tp)^4} \right] p^2 \qquad t \in (0,1).$

(2) for $f(x) = \cos x$.

$\cos(x+p) = \cos x - \sin x \cdot p - \frac{1}{2}\cos x \cdot p^2 + \frac{1}{6}\sin(x+tp)\cdot p^3 \qquad t \in (0,1).$

when $x = 1$. $\cos(1+p) = \cos 1 - p\sin 1 - \frac{1}{2}\cos 1 \cdot p^2 + \frac{1}{6}\sin(1+tp)\cdot p^3.$

✎ **5** Consider the function $f: R^2 \to R$ defined by $f(x) = \|x\|^2$. Show that the sequence of iterates $\{x_k\}$ defined by

$$x_k = \left(1 + \frac{1}{2^k}\right) \begin{bmatrix} \cos k \\ \sin k \end{bmatrix}$$

satisfies $f(x_{k+1}) < f(x_k)$ for $k = 0, 1, 2, \ldots$. Show that every point on the unit circle $\{x \mid \|x\|^2 = 1\}$ is a limit point for $\{x_k\}$.

Hint: Every value $\theta \in [0, 2\pi]$ is a limit point of the subsequence $\{\xi_k\}$ defined by

$$\xi_k = k \pmod{2\pi} = k - 2\pi \left\lfloor \frac{k}{2\pi} \right\rfloor,$$

where the operator $\lfloor \cdot \rfloor$ denotes rounding down to the next integer.

Sol:
$f(x_k) = \|x_k\| = \left(1 + \frac{1}{2^k}\right)^2$

Since $1 + \frac{1}{2^k} > 1 + \frac{1}{2^{k+1}} > 1$. and $f(x_{k+1}) < f(x_k)$. then $f(x_k) \to 1$. (monotonic + bounded $\Rightarrow$ conv.).

let $\xi_k = k \bmod 2\pi$ $\qquad \theta_k = (\cos\xi_k, \sin\xi_k)^T$. $\quad r_k = 1 + \frac{1}{2^k}$ $\qquad x_k = r_k \cdot \theta_k$.

$\forall \theta \in [0, 2\pi)$. $\forall \varepsilon > 0$. $\exists k \in N$. s.t. $|\xi_k - \theta| < \varepsilon$. i.e. $\exists$ subsequence $\{\xi_j\}$. $\xi_{k_j} \to \theta$. $r_{k_j} \to 1$.

thus $\exists x_{k_j} \to \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}$ on the unit circle.

6 Prove that all isolated local minimizers are strict. (*Hint: Take an isolated local minimizer x^* and a neighborhood $\mathcal{N}$. Show that for any $x \in \mathcal{N}, x \neq x^*$ we must have $f(x) > f(x^*)$.*)

Pf: x^* is isolated local min. i.e. $\exists \mathcal{N}_{x^*}$. $\forall x \in \mathcal{N}$. $f(x) > f(x^*)$. and x are not local min.

Assume the converse. $\exists x' \in \mathcal{N} \setminus \{x^*\}$. s.t. $f(x') \leq f(x^*)$, by the def of local min of x^*. the only possible case is $f(x') = f(x^*)$.

let $\varepsilon > 0$. s.t. $\mathcal{N}_\varepsilon(x') \subseteq \mathcal{N}(x^*)$. $\forall x'' \in \mathcal{N}_\varepsilon(x')$. $f(x'') \geq f(x')$ which means x' is also a local minimum which against the "isolation" of x^*. thus. x^* must be strict local minimum.

7 Suppose that $f(x) = x^T Q x$, where Q is an $n \times n$ symmetric positive semidefinite matrix. Show using the definition that $f(x)$ is convex on the domain $\mathbb{R}^n$.

Hint: It may be convenient to prove the following equivalent inequality:

$$f(y + \alpha(x - y)) - \alpha f(x) - (1 - \alpha)f(y) \leq 0,$$

for all $\alpha \in [0, 1]$ and all $x, y \in \mathbb{R}^n$.

Pf: $\forall \alpha \in [0,1]$. $x, y \in \mathbb{R}^n$. (since Q is symmetric. $y^T Q x = x^T Q y$).

$f(y + \alpha(x-y)) = (y + \alpha(x-y))^T Q (y + \alpha(x-y)) = y^T Q (y + \alpha(x-y)) + \alpha(x-y)^T Q (y + \alpha(x-y))$

$\qquad = y^T Q y + \alpha y^T Q (x-y) + \alpha(x-y)^T Q y + \alpha(x-y)^T Q (x-y)$

$\qquad = (1-2\alpha) y^T Q y + \alpha y^T Q x + \alpha x^T Q y + \alpha^2 x^T Q x - \alpha^2 y^T Q x + \alpha^2 y^T Q y - \alpha^2 x^T Q y$.

$\qquad = (1-\alpha)^2 y^T Q y + 2\alpha(1-\alpha) x^T Q y + \alpha^2 x^T Q x$.

$f(y - \alpha(x-y)) - \alpha f(x) - (1-\alpha) f(y) = \alpha(1-\alpha) x^T Q x - \alpha(1-\alpha) y^T Q y + 2\alpha(1-\alpha) x^T Q y$

$\qquad\qquad = \alpha(1-\alpha) [-x^T Q x - y^T Q y + 2 x^T Q y]$

$\qquad\qquad = -\alpha(1-\alpha)(x-y)^T Q (x-y) = -\alpha(1-\alpha) f(x-y)$.

since Q semi positive definite. $f(x-y) \geq 0$. thus. $f(y - \alpha(x-y)) - \alpha f(x) - (1-\alpha) f(y) \leq 0$

8 Suppose that f is a convex function. Show that the set of global minimizers of f is a convex set.

Pf: f is convex $\forall x, y \in \mathbb{R}^n$. $\alpha \in [0,1]$. $\alpha f(x) + (1-\alpha) f(y) \geq f(\alpha x + (1-\alpha) y)$

denote $M = \{x^* \in \mathbb{R}^n \mid \forall x \in \mathbb{R}^n. f(x) \geq f(x^*)\}$ thus we have $\forall x_1^*. x_2^* \in M$ $f(x_1^*) = f(x_2^*)$

$\forall x_1^*. x_2^* \in M$ $f(\alpha x_1^* + (1-\alpha) x_2^*) \leq \alpha f(x_1^*) + (1-\alpha) f(x_2^*) = f(x_1^*) = f(x_2^*)$.

by def of global minimum. $f(\alpha x_1^* + (1-\alpha)x_2^*) = f(x_1^*) \Rightarrow \alpha x_1^* + (1-\alpha) x_2^* \in M$.

thus M is convex.

✎ **9** Consider the function $f(x_1, x_2) = (x_1 + x_2^2)^2$. At the point $x^T = (1, 0)$ we consider the search direction $p^T = (-1, 1)$. Show that p is a descent direction and find all minimizers of the problem:
$$\min_{\alpha > 0} f(x_k + \alpha p_k).$$

Sol: $\nabla f = \begin{bmatrix} 2(x_1 + x_2^2) \\ 4x_2(x_1 + x_2^2) \end{bmatrix}$ $\nabla^2 f = \begin{bmatrix} 2 & 4x_2 \\ 4x_2 & 4x_1 + 12x_2^2 \end{bmatrix}$

$\nabla f \big|_{(1,0)} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$ $\nabla f(x)^T p = (2,0)\begin{pmatrix} -1 \\ 1 \end{pmatrix} = -2 < 0$ → descent direction

let $m_k(\alpha) = f\left[\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \alpha \begin{pmatrix} -1 \\ 1 \end{pmatrix}\right] = f\begin{pmatrix} 1-\alpha \\ \alpha \end{pmatrix} = (1 - \alpha + \alpha^2)^2$

$m_k'(\alpha) = 2(1-\alpha+\alpha^2)(2\alpha-1) = 4(\alpha - \frac{1}{2})((\alpha - \frac{1}{2})^2 + \frac{3}{4})$.

$m_k'(\alpha)$ has single real root $\alpha = \frac{1}{2}$. $m_k''(\alpha) = 2\left[2(\alpha^2 - \alpha + 1) + (2\alpha-1)(2\alpha-1)\right] > 0$.

$m_k(\alpha)_{min} = m_k(\frac{1}{2})$ when $\alpha = \frac{1}{2}$. $x_k + \alpha p_k = (\frac{1}{2}, \frac{1}{2})^T$

✎ **10** Consider the sequence $\{x_k\}$ defined by
$$x_k = \begin{cases} (\frac{1}{4})^{2^k}, & k \text{ even,} \\ (x_{k-1})/k, & k \text{ odd.} \end{cases}$$

2^{-2^k}

$2^{-2^{k+1}} \cdot 2^{-2^{k+1}}$

$\dfrac{(\frac{1}{4})^{2^{2k}}}{\left((\frac{1}{4})^{2^{k-2}}\right)^4}$

Is this sequence Q-superlinearly convergent? Q-quadratically convergent? R-quadratically convergent?

Sol: $x_k \to 0$. denote the diff $\|x_k - 0\| = e_k$.

① consider $\lim\limits_{k\to\infty} \dfrac{e_{k+1}}{e_k}$

$k = $ even. $\lim \dfrac{e_{k+1}}{e_k} = \lim \dfrac{1}{k+1} = 0$

$k = $ odd $\lim\limits_{k\to\infty} \dfrac{(\frac{1}{4})^{2^{k+1}}}{(x_{k-1})/k} = \lim \dfrac{(\frac{1}{4})^{2^{k+1}} \cdot k}{(\frac{1}{4})^{2^{k-1}}} = \lim (\frac{1}{4})^{3 \cdot 2^{k-1}} \cdot k = 0$.

thus $\{x_k\}$. Q-superlinearly conv.

② $k = $ even. $e_{k+1} = \dfrac{x_k}{k+1}$ we need C s.t. $e_{k+1} \le C e_k^2 \Rightarrow \frac{1}{k+1} \le C e_k$.

but $e_k \to 0$. impossible to find const C. Thus $\{x_k\}$ not Q-quadratically conv

③ consider $\varepsilon_{2m} = (e_{2m})^{\frac{1}{3}}$ $\varepsilon_{2m+1} = (e_{2m})^{\frac{2}{3}}$. $\varepsilon_k \to 0$.

$e_{2m} < 1$ $e_{2m} < \varepsilon_{2m}$ $e_{2m+1} = \dfrac{e_{2m}}{2m+1} < (e_{2m})^{\frac{2}{3}}$ thus $e_k \le \varepsilon_k$.

$\dfrac{\varepsilon_{2m+1}}{\varepsilon_{2m}^2} = 1$ $\dfrac{\varepsilon_{2m}}{\varepsilon_{2m-1}^2} = \dfrac{(e_{2m})^{\frac{1}{3}}}{((e_{2m-2})^{2/3})^2} = \left(\dfrac{e_{2m}}{e_{2m-2}^4}\right)^{\frac{1}{3}} = 1$.

thus ∃ $C = 1$. $\varepsilon_{k+1} \le 1 \varepsilon_k^2$ i.e. $\{x_k\}$ R-quadratically conv.