Problem 1

Comment on the convexity of the following functions over $\mathbb{R}^n$. Always justify your answers.

(a) $f(x) = \frac{1}{4} (x^T A^T A x)^3 + \|x\|^2$, where $A \in \mathbb{R}^{m \times n}$ is some matrix and $\|\cdot\|_2$ is the Euclidean norm

(b) $g(x) = (\|x\|^2 - 4)^2$

(c) $h(x) = -\log(\prod_{i=1}^n x_i)$, where $\log(\cdot)$ is the natural logarithm

Problem 2

Consider the following two functions

$$f(x) = (4x_1^2 - x_2)^2 \quad \text{and} \quad g(x) = x_1^2 + 4x_1x_2 + x_2^2 + x_1 - x_2,$$

where $x = (x_1, x_2) \in \mathbb{R}^2$.

(a) Identify all stationary points of $f$. Classify them as saddle points, strict/nonstrict local minimizers, or strict/nonstrict global minimizers.

(b) Identify all stationary points of $g$. Classify them as saddle points, strict/nonstrict local minimizers, or strict/nonstrict global minimizers.

(c) Are $f$ and $g$ coercive? Justify your answer.

Hint: We say that $f$ is coercive, if $f(x) \to +\infty$ when $\|x\|_2 \to +\infty$.

(d) Using contourf and plot commands in Python's Matplotlib library, plot the functions and their critical points in the interval $-4 \leq x_1 \leq 4$ and $-4 \leq x_2 \leq 4$. Include the plot in your submission.

Problem 3

The Rayleigh quotient of a symmetric matrix $A \in \mathbb{R}^{n \times n}$ is defined as

$$R_A(x) = \frac{x^T A x}{\|x\|^2}, \quad \forall x \neq 0.$$

Using the spectral decomposition of $A$ show that

$$\lambda_{\min}(A) \leq R_A(x) \leq \lambda_{\max}(A), \quad \forall x \neq 0,$$

$\lambda_{\min}(A)$ and $\lambda_{\max}(A)$ are the smallest and largest eigenvalues of $A$, respectively.

Problem 4

Consider the following function defined over $\mathbb{R}^n$

$$f(x) = g(x) + h(x) \quad \text{where} \quad g(x) = \|y - Ax\|_2^2 \quad \text{and} \quad h(x) = \|x\|_\infty.$$

Here, $\|x\|_2 = \sqrt{x^T x}$ denotes the standard Euclidean norm and $\|x\|_\infty = \max_{i \in \{1, \ldots, n\}} |x_i|$. Assume that $A$ is an orthogonal matrix. Justify all your answers.

(a) Comment on coercivity of $g$, $h$, and $f$. 

Page 1 of 2
(b) Consider convex functions $f_1, ..., f_k$ defined over $\mathbb{R}^n$. Show that if any of $k$ functions is strongly convex, then the function

$$\sum_{i=1}^{k} f_i(x)$$

is also strongly convex.

(c) Comment on convexity of $g$, $h$, and $f$.

**Problem 5**

Consider the following objective function

$$f(x) = 4x_1^2 + 2x_1x_2 + 2x_2^2 \quad \text{with} \quad x \in \mathbb{R}^2.$$  

(a) Is $f$ coercive? Justify your answer.

(b) Find all the minimizers of $f$ analytically. Are they strict/nonstrict local/global minimizers?

(c) Download and modify the Python notebook script GradientMethod.ipynb for minimizing $f$ with the gradient method (GM). Set the parameters of GM to $x^0 = \text{xitInit} = (1, 1)$, $\gamma = \text{stepSize} = 0.01$, $\text{maxIter} = 100$, and $\text{tol} = 10^{-6}$. Submit the generated plot and the PDF printout of your code with you responses.

(d) For each of the following step parameter $\gamma > 0$ comment on the convergence of GM: (i) $\gamma = 0.0025$, (ii) $\gamma = 0.025$, and (iii) $\gamma = 0.25$. Submit the generated plots with your responses.

(e) *(Bonus)* Empirically find the largest step $\gamma^* > 0$ for which GM converges. Round $\gamma^*$ to four significant digits. Is the convergence better for $\gamma^*$ or $\gamma = 0.025$? Any ideas why?