Rules:

- Unless otherwise stated, all answers must be mathematically justified.
- Partial answers will be graded.
- Your submission has to be uploaded to Gradescope. In Gradescope, indicate the page on which each problem is written.
- You can work in groups but each student must write his/her/their own solution based on his/her/their own understanding of the problem. Please list on your submission the students you work with for the homework (this will not affect your grade).
- Problems with a ($\star$) are extra credit, they will not (directly) contribute to your score of this homework. However, for every 4 extra credit questions successfully answered your lowest homework score get replaced by a perfect score.
- If you have any questions, feel free to contact me (`mgabrie@nyu.edu`) or to stop at the office hours.

Problem 11.1 (2 points)**.** *Compute critical points of f, g and h and determine if they are global/local maximizers/minimizers or saddle points. To determine the signs of eigenvalues it might useful to remember that for $M \in \mathbb{R}^{n\times n}$ symmetric, $\mathrm{tr}(M) = \sum_{i=1}^n M_{i,i} = \sum_{i=1}^n \lambda_i$.*

(a) $f : \mathbb{R} \to \mathbb{R}$ *with* $f(x) = (x^2 - 1)^2$.

(b) $g : \mathbb{R}^3 \to \mathbb{R}$ *with* $g(x, y, z) = (x^2 - z^2)y + 2$

(c) $h : \mathbb{R}^3 \to \mathbb{R}$ *with* $h(x, y, z) = x^2 + y^2 + z^2 - 6x + 10y - 2z + 35$

Problem 11.2 (2 points)**.** *We consider the following constrained optimization problem in $\mathbb{R}^2$:*

$$\textit{minimize} \quad x^2 + y^2 \quad \textit{subject to} \quad 2x + y = 4. \tag{1}$$

We admit that this minimization problem has (at least) one solution (this comes from the fact that a continuous function on a compact set attains its minimum).

(a) *Using Lagrange multipliers, show that (1) has a unique solution and compute its coordinates.*

(b) *Can you draw a picture in $\mathbb{R}^2$ representing the problem?*

Problem 11.3 (2 points)**.** *Let $u \in \mathbb{R}^n$ be a vector such that for all $i \neq j$, $|u_i| \neq |u_j|$. We consider the constrained optimization problem*

$$\textit{maximize} \quad \langle u, x\rangle \quad \textit{subject to} \quad \|x\|_1 \leq 1.$$

(a) *Calling i_* the index at which $|u_i|$ is maximum, give a solution for the optimization problem (no Lagrange multiplier needed).*

(b) *By contradiction, show that this solution is unique.*

(c) *Give a graphical interpretation in the case $n = 2$. You should consider the orthogonal projector onto* $\mathrm{Span}(u)$.

Problem 11.4 (3 points)**. We will prove the spectral theorem in this problem: you are therefore not allowed to use the spectral theorem and its consequences to solve this exercise.**

Let A be an $n \times n$ symmetric matrix. We consider the following optimization problem

$$\text{maximize} \quad x^\top A x \quad \text{subject to} \quad \|x\| = 1. \tag{2}$$

This optimization problem admits a solution (this comes from the fact that a continuous function on a compact set achieved its maximum) that we denote by v_1.

(a) *Using Lagrange multipliers, show that v_1 is an eigenvector of A.*

(b) *We now consider the optimization problem*

$$\text{maximize} \quad x^\top A x \quad \text{subject to} \quad \|x\| = 1 \quad \text{and} \quad \langle x, v_1 \rangle = 0. \tag{3}$$

For the same reason as above, this problem admits a solution that we denote by v_2. Show that v_2 is an eigenvector of A that is orthogonal to v_1.

(c) *We now consider the optimization problem*

$$\text{maximize} \quad x^\top A x \quad \text{subject to} \quad \|x\| = 1 \quad \text{and} \quad \langle x, v_1 \rangle = 0 \quad \text{and} \quad \langle x, v_2 \rangle = 0. \tag{4}$$

Again, this problem admits a solution that we denote by v_3. Show that v_3 is an eigenvector of A that is orthogonal to v_1 and v_2.

Conclusion: by repeating this procedure, we obtain an orthonormal family $v_1, \dots, v_n$ of eigenvectors of A. This proves the spectral theorem (without using any linear algebra result!).

Problem 11.5 ($\star$)**.** *We consider the problem with physics motivation of finding the maximal entropy distribution of a random variable (see last slides of Lecture 09) constraining values of some moments.*

To keep things simple, we consider X that can take n different values $x_1, \cdots, x_n$ in $\mathbb{R}$. We wish to infer the probabilities $p_1, \cdots, p_n$ such that the entropy is maximal and the expected value of X is equal to a previously known scalar $\mu \in \mathbb{R}$. This corresponds to solving the contrained optimization problem

$$\text{maximize} \quad -\sum_i p_i \ln p_i \quad \text{subject to} \quad p_i \leq 1 \text{ for all } i \quad \text{and} \quad \sum_{i=1}^n p_i = 1 \quad \text{and} \quad \sum_{i=1}^n p_i x_i = \mu. \tag{5}$$

(a) *Rewrite the problem as a convex minimization problem (justify).*

(b) *Using KKT theorem, give the expression of the probability vector solution $p \in \mathbb{R}^n$ as a function of Lagrange multipliers and values x_i. Give also the relations between the Lagrange multipliers, μ and values x_i.*

(c) *In the case where $n = 2$ and $x_1 = 0$ and $x_2 = 1$, solve for the values of the Lagrange multipliers and $p \in \mathbb{R}^2$. Could you have used an easier way to solve the problem in this simple case?*