

Optimization and Computational Linear Algebra for Data Science

Homework 11: Optimality conditions

November 26, 2019

-
- Unless otherwise stated, all answers must be mathematically justified.
 - Partial answers will be graded.
 - You can work in groups but each student must write his/her own solution based on his/her 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 (lm4271@nyu.edu) or to stop at the office hours.
-

Problem 11.1 (2 points). Let $f, g : \mathbb{R}^3 \rightarrow \mathbb{R}$ be the functions defined by

$$f(x, y, z) = 2x^2 + y^2 + \frac{1}{2}z^2 + 4x - 6y - z + 1 \quad \text{and} \quad g(x, y, z) = xyz + x + y + z.$$

Compute the critical points of f and g and determine if they are global/local maximizer, local minimizer or saddle point.

Problem 11.2 (3 points). We consider the following constrained optimization problem:

$$\text{minimize } x - y + z \quad \text{subject to } x^2 + y^2 + z^2 = 1 \quad \text{and} \quad x + y + z = 1. \quad (1)$$

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).

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

Problem 11.3 (2 points). Let $u \in \mathbb{R}^n$ be a vector whose coordinates are all distinct. We consider the constrained optimization problem

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

- (a) Show that this problem has a unique solution x^* and give the expression of x^* in terms of u (Lagrange multipliers are not needed here).
- (b) Give a graphical interpretation.

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

$$\text{maximize } x^T A x \quad \text{subject to } \|x\| = 1. \quad (2)$$

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

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

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

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

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 here a simple portfolio optimization problem. Assume that we can invest in n financial assets. Each asset i has a return of X_i (X_i is a random variable) meaning that investing w \$ in the asset i will generate a return of $w \times X_i$ \$. We introduce

$$r_i = \mathbb{E}[X_i], \\ \Sigma = \text{Cov}(X, X) = \mathbb{E}[(X - r)(X - r)^T]$$

that are respectively the average return and the covariance matrix of the returns. We assume that the covariance matrix Σ is invertible. We represent a portfolio (an investment strategy) by a vector $w = (w_1, \dots, w_n) \in \mathbb{R}^n$, where each coordinate w_i represents the amount invested on the asset i . We allow w_i to be negative, which means that it is possible to “short” a security. The expected return of the portfolio is then

$$R(w) = \mathbb{E} \left[\sum_{i=1}^n w_i X_i \right] = \sum_{i=1}^n w_i r_i$$

and its variance is

$$V(w) = \text{Var} \left(\sum_{i=1}^n w_i X_i \right) = \mathbb{E} \left[\left(\sum_{i=1}^n w_i (X_i - r_i) \right)^2 \right] = \mathbb{E} \left[\sum_{i=1}^n \sum_{j=1}^n w_i w_j (X_i - r_i)(X_j - r_j) \right] \\ = w^T \Sigma w.$$

We want here to invest a total amount of m on the assets. Given an targeted expected return μ , we would like to find a portfolio w such that $R(w) = \mu$ and $\sum_{i=1}^n w_i = m$ for which the variance $V(w)$ is minimal. That is, we aim at solving the following problem:

$$\text{minimize } V(w) \quad \text{subject to } R(w) = \mu \quad \text{and} \quad \sum_{i=1}^n w_i = m. \quad (5)$$

Let $\mathbf{1}$ denotes the all-ones vector of dimension n . We assume that $r \notin \text{Span}(\mathbf{1})$. We also assume that $m \min_i r_i < \mu < m \max_i r_i$, so that there exists a feasible w to the problem (5). Show that the matrix

$$M = \begin{pmatrix} \mathbf{1}^\top \Sigma^{-1} \mathbf{1} & \mathbf{1}^\top \Sigma^{-1} r \\ \mathbf{1}^\top \Sigma^{-1} r & r^\top \Sigma^{-1} r \end{pmatrix}$$

is invertible and that the unique solution to (5) is

$$w^* = \Sigma^{-1} \begin{pmatrix} | & | \\ \mathbf{1} & r \\ | & | \end{pmatrix} M^{-1} \begin{pmatrix} m \\ \mu \end{pmatrix}.$$

