# Numerical Optimization with Python

**Dry HW 01**

**Part I: Modeling optimization problems:**

For each of the following descriptions, formulate a mathematical optimization problem (constrained or unconstrained) that models the given definitions (not solving the problem at all).

1. An ice cream manufacturer produces three flavors: vanilla, chocolate and strawberry. The profit per Kg from manufacturing vanilla is 7NIS, for chocolate is 6NIS and for strawberry is 5NIS. There's one machine, which is capable of producing no more than 100Kg of ice cream per day. Each Kg of Vanilla requires 0.5Kg of milk, 0.4Kg of cream and 0.1Kg of vanilla extract. Each Kg of chocolate requires 0.2Kg of milk, 0.7Kg of cream and 0.1Kg of chocolate. Each Kg of strawberry requires 0.4Kg of milk, 0.4Kg of cream and 0.2Kg of strawberries. The manufacturer has the following limits on daily supplies: milk 45Kg, cream 60Kg, vanilla 10Kg, chocolate 10Kg, strawberries 15Kg. Formulate an optimization problem that finds the quantities to manufacture that maximize the manufacturer profit.
2. Given a vector $x$ of $n$ real values $x_1, \dots, x_n$, a vector $p \in \mathbb{R}^n$ is called a ***probability distribution*** on $x$, if $p_i \geq 0$ and $p_1 + \cdots + p_n = 1$. In other words, the $p_i$'s are non-negative and sum to 1. Given a probability vector $p$ for $x$, the ***expected value*** (or expectation) of $x$ is the weighted average $\mathrm{E}x = \sum_{i=1}^{n} p_i x_i$. Given a probability distribution $p$ on $n$ values, the ***entropy*** of $p$ is defined to be: $\mathrm{H} = -\sum_{i=1}^{n} p_i \log p_i$ .
   For the values $x_1 = -10.2, x_2 = 0.4, x_3 = 16.6, x_4 = 10.3$, formulate an optimization problem that finds a probability distribution for $x$ with expected value zero and has maximal entropy.
3. For some dimension $n > 1$, consider upper ("northern") half of the closed unit ball (boundary as well as interior). Formulate an optimization problem that seeks the point in the above described domain, that is closest to a given point $p \in \mathbb{R}^n$ in the Euclidean norm ($L_2$).
   a. If the problem you formulated is not smooth, adjust it so it is equivalent (same solution) but smooth.
   b. Formulate the same problem but now seeking a point in the domain that is closest to the given point $p \in \mathbb{R}^n$ in $L_1$ norm. Is your problem smooth?

c. Transform the problem in (b) to a smooth problem that seeks the same solution (hint: introduce helper variables that will increase the dimension and add constraints, but provide a smooth problem with the same interpretation, regarding the original variables)

**Part II: General preview material – multivariate differentiation:**

1. Let $f: \mathbb{R}^n \to \mathbb{R}$ be defined by: $f(x) = x_1 + x_2 + \cdots + x_n$.
   a. What is the partial derivative: $\frac{\partial f(x)}{\partial x_i}$ ?
   b. What is $\nabla f(x)$? (remember – the gradient is a column vector)
   c. What is the directional derivative $\frac{\partial f}{\partial u}$ w.r.t the unit vector: $u = \left[\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right]^T$? Does it depend on $x$? Why?
   d. For a constant vector $a \in \mathbb{R}^n$, write the gradient of the linear function: $f(x) = a^T x$
2. Let $f: \mathbb{R}^2 \to \mathbb{R}$ be defined by: $f(x) = x_1^2 + 2x_2^2 + x_1 x_2$.
   a. Write $f$ as a quadratic function in matrix form, namely find a symmetric matrix $Q$ such that $f$ is given by: $f(x) = x^T Q x$.
   b. Find $\frac{\partial f}{\partial x_1}$ and $\frac{\partial f}{\partial x_2}$, put them together and write $\nabla f(x)$.
   c. Express $\nabla f(x)$ found in (b) using the matrix $Q$ from (a). Hint: it should be a generalization of the derivative of a quadratic function in one variable.
   d. Find $\frac{\partial^2 f}{\partial x_2^2}, \frac{\partial^2 f}{\partial x_1 \partial x_2}$ and $\frac{\partial^2 f}{\partial x_1^2}$. Arrange them together and write the Hessian $\nabla^2 f(x)$.
   e. Express $\nabla^2 f(x)$ found in (d) in terms of the matrix $Q$ from (a). Hint: it should again be a generalization of the second derivative of a quadratic function in one variable.
   f. The general case: for a constant matrix $Q \in \mathbb{R}^{n \times n}$, write the gradient $\nabla f(x)$ and the Hessian $\nabla^2 f(x)$ of the quadratic function $f(x) = x^T Q x$.
3. Let $f: \mathbb{R}^n \to \mathbb{R}$ and $u \in \mathbb{R}^n$, $\|u\| = 1$. Recall from class: the directional derivative $\frac{\partial f(x)}{\partial u}$ is the slope of the graph of $f$ when marching in the direction $u$ at the point $x$, we have shown this is given by $u^T \nabla f(x)$. In this exercise we develop the matrix form expression for the directional <u>second</u> derivative.
   a. Write a parametrized expression for a line $l(t)$ in the direction $u$ that through $x$.
   b. Write a function of a single variable $g(t)$, defined to be the restriction of $f$ to the line $l(t)$
   c. Write an expression for the first derivative $g'(t)$. Hint: $g$ is a composition $f \circ l: \mathbb{R} \to \mathbb{R}$, use the chain rule (repeat as done in class).

d. Using $g'(t)$ from (c), write an expression for $g''(t)$ (Hint: use the chain rule again, and recall we noted in class what the differential matrix of the vector function: $x \mapsto \nabla f$ is.)

e. Finally, use $g''$ from (d) to obtain $\frac{\partial^2 f(x)}{\partial u^2}$. Hint: which $t$ value gives $x$ along the line $l$?

f. For the quadratic function in question 2, what is $\frac{\partial^2 f(x)}{\partial u^2}$ in the unit direction $u = \left[\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right]^T$? Does it depend on $x$? Why?

g. (5 point bonus) we know that $\frac{\partial f(x)}{\partial u}$ has a geometric interpretation, as the slope at $x$ in the direction $u$. What is the geometric interpretation of $\frac{\partial^2 f(x)}{\partial u^2}$? (what does it measure? No need to be precise regarding units/scale)