Dr. Peter Arndt

Master AI and Data Science
Universität Düsseldorf

Math for AI
Winter term 2023/24

Exercise sheet 7

Exercise 25 (8 points)

Consider the multi-objective optimization problem

$$\begin{array}{ll} \text{minimize}_x & f_1(x), \ldots, f_\ell(x) \\ \text{subject to} & g_i(x) \leq 0 \ (i = 1 \ldots m) \\ & h_j(x) = 0 \ (j = 1 \ldots p) \end{array}$$

Suppose that the f_k, g_i are differentiable and convex and that the h_j are affine. For a point $c \in \mathbb{R}^\ell$ we can form the single objective function $f_c(x) := c_1 f_1(x) + \ldots + c_\ell f_\ell(x)$

Show that $x^* \in \mathbb{R}^n$, $c^* \in \mathbb{R}^\ell$, $\lambda^* \in \mathbb{R}^m$ and $\mu^* \in \mathbb{R}^p$ satisfy the following conditions:

$$\begin{aligned} \lambda_i^* &\geq 0 \\ c_k^* &\geq 0 \\ \nabla f_{c^*}(x^*) + \sum_i \lambda_i^* \nabla g_i(x^*) + \sum_j \mu_j^* \nabla h_j(x^*) &= 0 \\ \sum_i \lambda_i^* g_i(x^*) &= 0 \\ g_i(x^*) &\leq 0 \\ h_j(x^*) &= 0 \end{aligned}$$

Then show that x^* is a Pareto optimal point for the multiobjective problem.

Exercise 26 (☠, possibly tricky, 10 points)

In this exercise, modified from Exercise 5.19 of this supplement for Boyd/Vandenberghe's "Convex Optimization", you should try to reconstruct the colors of an image, given its grey scale version together with the colors of a few selected pixels. You can look at that exercise under the link for some more background information, but the explanations below and in the file should also suffice.

A colored picture, in Python, consists of three matrices R, G, B of the same size, which store the intensity of the colors red, green and blue (as numbers between 0 and 1). To turn this into a grey scale picture, one computes a matrix storing the brightness of each pixel by the formula

$$M = 0.299R + 0.587G + 0.114B.$$

The coefficients come from the perceived brightness of the respective colors.

The program in the file `ColorReconstruction`, provided in the exercise folder, takes a colored image of a flower, produces such a grey scale version, and additionally produces vectors which remember the colors of some selected pixels.

To reconstruct all colors, we are looking for the red, green and blue values for each of the 50×50 pixels, i.e. we are trying to find values in the unit interval $[0, 1]$ for variables r_{ij}, g_{ij}, b_{ij} $(0 \leq i, j \leq 49)$

One operates on the assumption that the colors do not change drastically from one pixel to the next. This means that we are trying to minimize the "color difference" function

$$\sum_{i=0}^{49}\sum_{j=0}^{49}\left((r_{ij} - r_{i+1,j})^2 + (r_{ij} - r_{i,j+1})^2 + (g_{ij} - g_{i+1,j})^2 + (g_{ij} - g_{i,j+1})^2 + (b_{ij} - b_{i+1,j})^2 + (b_{ij} - b_{i,j+1})^2\right)$$

subject to the constraints

$$0 \leq r_{ij}, g_{ij}, b_{ij} \leq 1$$

and to some equalities for those pixels where the colors are known. (hm, actually: does the sum index go up to 49 or only to 48? I leave it to you to figure it out)

Your task is this:

1. Install the Cvxopt package (installation instruction are on the linked page).
2. Write a program which minimizes this function, using cvxopt.solvers.qp

In case the installation of `Cvxopt` does not work for you for some reason, you can use scipy.optimize.lsq_linear instead.

Hint: The suggested function `cvxopt.solvers.qp` is for solving quadratic optimization problems, i.e. minimizing a function of the form $x \mapsto x^T P x$ for some matrix P. The given problem is actually of this kind: The function to be minimized gives back the length of a vector, which arises as the image of the (r_{ij}, g_{ij}, b_{ij})-vector under a certain linear map. If R is the matrix of this linear map, one can take $P = R^T R$ – convince yourself that this is indeed the same optimization problem (but you don't need to write anything about it)! You also need to think how to formulate the constraints in the form required by the `cvxopt.solvers.qp` API.

More information is to be found in the jupyter notebook file in the exercise folder.

[This exercise is a bit tricky. The problem is to get the matrix right that you give to the quadratic problem solver: It is probably clear enough what the matrix should look like within one region taking care of just, for example, the red values. But you also have to take care that the matrix doesn't introduce any dependencies between red and green for example – you might have to set a few extra coefficients to zero after setting up the matrix in a homogeneous way.]

Exercise 27 (10 points)

(a) Consider an $\mathbb{R}^3$-valued random variable (X_1, X_2, X_3) with density function $f(x_1, x_2, x_3) := (x_1^2 + x_2^2 + x_3^2)\chi_I$, where χ_I is the indicator function of the unit cube $I := [0,1]^3$, i.e. $\chi_I(x) = 1$ if $x \in I$ and $= 0$ otherwise.

(i) (2 points) Compute the probability $\mathrm{P}(X_1 \leq \frac{1}{2}, X_3 \geq \frac{1}{2})$

(ii) (2 points) Compute the density function of the $\mathbb{R}^2$-valued random variable (X_1, X_2).

(b) (2 points) Consider the $\mathbb{R}^2$-valued random variable (X, Y) with density function $f(x, y) := e^{-x-y}$ for $x, y \geq 0$ and 0 otherwise. Are the random variables X and Y independent?

(c)

(i) (2 points) Consider two independent random variables X and Y taking values 1 or -1 each with probability $\frac{1}{2}$. Let $Z := X \cdot Y$. Show that X and Z are independent, and that Y and Z are independent.

(ii) (2 points) Show that X, Y and Z are not jointly independent, in the sense that $P(X = a, Y = b, Z = c)$ is not always equal to $P(X = a) \cdot P(Y = b) \cdot P(Z = c)$.

Exercise 28 (12 points)

In the following all random variables are $\mathbb{R}^n$-valued and should be defined on the same fixed probability space (Ω, P). That is: a random variable is a map $\Omega \to \mathbb{R}^n$.

Two random variables X, Y are called *independent*, if they satisfy $P(X \in A, Y \in B) = P(X \in A) \cdot P(Y \in B)$ for all $A, B \subseteq \mathbb{R}^n$

One can add and scalar multiply random variables according to the rules $(X+Y)(\omega) := X(\omega)+Y(\omega)$ and $(\lambda X)(\omega) := \lambda \cdot X(\omega)$. With this, random variables on (Ω, P) form a vector space, so the notion of *linear independence* makes sense. In this exercise you should explore the relationship between the linear independence and the independence of random variables.

An $\mathbb{R}^n$-valued random variable Z is called *almost surely constant*, if there is a $v \in \mathbb{R}^n$ such that $P(Z = v) = 1$.

Let X, Y, Z be $\mathbb{R}$-valued random variables.

(a) Show that $\mathrm{Var}(Z) = 0$ if and only if Z is almost surely constant.

(b) Show that the almost surely constant random variables form a subvector space of all random variables.

(c) Show that if Z is not almost surely constant and Z and Y are linearly dependent, then their correlation coefficient is 1 or -1

(d) Does linear independence of two random variables X, Y imply that they are independent?

(e) Does independence of two random variables X, Y imply that they are linearly independent?

(f) Does the answer in part (e) change, if one assumes X, Y to be not almost surely constant?

[2 points per item]

Deadline: Friday 1st of December, 10:00.
Upload your solution to this link.