The Chinese University of Hong Kong, Shenzhen
SDS · School of Data Science

Prof. Andre Milzarek · Prof. Zizhuo Wang · Prof. Haoxiang Yang
Spring Semester 2022

MAT 3007 – Optimization

Solutions – Final Exam – Sample

Please make sure to present your solutions and answers in a comprehensible way and give explanations of your steps and results!

Good Luck! Viel Glück!

Exercise 1 (KKT Conditions and Constrained Problems): (24 points)
Consider the nonlinear optimization problem

$$\min_{x\in\mathbb{R}^2} \; f(x) := \ln(1+x_2) + x_1x_2 - x_1^2x_2^2 \qquad \text{s.t.} \qquad g(x) \le 0, \tag{1}$$

where the constraint function $g : \mathbb{R}^2 \to \mathbb{R}^3$ is given by

$$g_1(x) := x_1^2 + (x_2-1)^2 - 1, \quad g_2(x) := 1 - (x_1+1)^2 - x_2^2, \quad g_3(x) := x_1 - x_2.$$

Let us further set $\bar{x} := (0,0)^\top$. (*Here,* $\ln$ *denotes the natural logarithm to the base* e).

a) Sketch the feasible set $\Omega := \{x \in \mathbb{R}^2 : g(x) \le 0\}$.

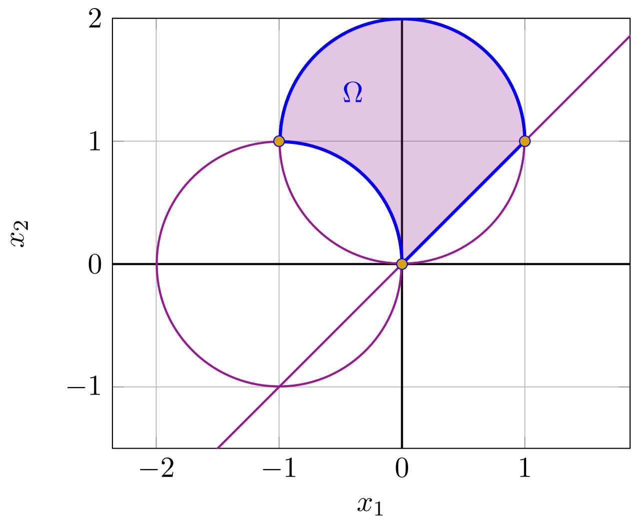
b) Is problem (1) convex? Explain your answer!

c) Determine the active set $\mathcal{A}(\bar{x})$.

d) Is $\bar{x}$ a KKT point of problem (1)? If yes, find all corresponding Lagrange multipliers $\bar{\lambda} \in \mathbb{R}^3$ such that the pair $(\bar{x}, \bar{\lambda})$ is a KKT point of (1). Is the multiplier $\bar{\lambda}$ unique?

Solution :

a) The following sketch shows the feasible region Ω:



b) Based on our sketch in part a), we immediately see that Ω is not convex. Hence, problem (1) is not a convex program.

c) We have $g_1(\bar{x}) = g_2(\bar{x}) = g_3(\bar{x}) = 0$ and hence, it follows $\mathcal{A}(\bar{x}) = \{1,2,3\}$.

d) We first calculate several derivatives:

$$\nabla f(x) = \begin{pmatrix} x_2 - 2x_1x_2^2 \\ \frac{1}{1+x_2} + x_1 - 2x_1^2x_2 \end{pmatrix}, \quad \nabla g_1(x) = \begin{pmatrix} 2x_1 \\ 2(x_2-1) \end{pmatrix},$$

$$\nabla g_2(x) = \begin{pmatrix} -2(x_1+1) \\ -2x_2 \end{pmatrix}, \quad \nabla g_3(x) = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Inserting $x = \bar{x}$, we obtain

$$\nabla f(\bar{x}) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \nabla g_1(\bar{x}) = \begin{pmatrix} 0 \\ -2 \end{pmatrix}, \quad \nabla g_2(\bar{x}) = \begin{pmatrix} -2 \\ 0 \end{pmatrix}, \quad \nabla g_3(\bar{x}) = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

We need to find $\lambda \in \mathbb{R}^3$ such that

$$\nabla f(\bar{x}) + \nabla g_1(\bar{x})\lambda_1 + \nabla g_2(\bar{x})\lambda_2 + \nabla g_3(\bar{x}) = \begin{pmatrix} -2\lambda_2 + \lambda_3 \\ 1 - 2\lambda_1 - \lambda_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Hence, we have $\lambda_3 = 1 - 2\lambda_1$ and $\lambda_2 = \frac{1}{2} - \lambda_1$. Since the multiplier need to be nonnegative, we need to have $\lambda_1 \leq \frac{1}{2}$. Furthermore, due to $\mathcal{A}(\bar{x}) = \{1, 2, 3\}$, the complementarity conditions are automatically satisfied. Thus, $\bar{x}$ is KKT point and all associated multiplier are given by $(\lambda, \frac{1}{2} - \lambda, 1 - 2\lambda)$ with $\lambda \in [0, \frac{1}{2}]$. This also shows that the multiplier is not unique in this case.

Exercise 2 (Convexity): (16 points)
Consider the following tasks:

a) Let $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^n$ be given. Show that $\Omega := \{x \in \mathbb{R}^n : \|Ax\|^2 - b^\top x \leq 1\}$ is a convex set.

b) Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) := (1 + x^2 + y^2)\ln(1 + x^2 + y^2)$.

   Is the mapping f convex on $\mathbb{R}^2$? Explain your answer!

Solution :

a) Let us set $g(x) = \|Ax\|^2 - b^\top x - 1$, then $\Omega := \{x : g(x) \leq 0\}$. Convexity of Ω now follows from convexity of g. In particular, we have $\nabla g(x) = 2A^\top Ax - b$ an $\nabla^2 g(x) = 2A^\top A$. The positive semidefiniteness of $\nabla^2 g$ (for all x) then implies convexity of g and Ω.

b) We define $h : \mathbb{R}^2 \to [1, \infty)$, $h(x, y) := 1 + x^2 + y^2$ and $g : \mathbb{R}_+ \to \mathbb{R}$, $g(x) = x\ln(x)$. The mapping h is obviously convex and satisfies $h(x, y) \geq 1$ for all $(x, y) \in \mathbb{R}^2$. Let us now consider g; we have:

$$g'(x) = \ln(x) + 1, \quad g''(x) = \frac{1}{x} > 0 \quad \forall\, x > 0.$$

This shows that g is convex on $\mathbb{R}_{++}$ and monotonically increasing on $[1, \infty)$. Together, this implies that the composition $f = g \circ h$ is a convex function.

Exercise 3 (Gradient Descent): (10 points)
We consider the unconstrained minimization problem

$$\min_{x\in\mathbb{R}^2} \; f(x) := \frac{1}{2}x_1^2 x_2 + x_1^2 x_2^2 - x_2(x_1 - 2) + \frac{1}{4}x_2^4.$$

The gradient of f is given by (*you don't need to verify this*):

$$\nabla f(x) = \begin{pmatrix} x_1 x_2 + 2x_1 x_2^2 - x_2 \\ \frac{1}{2}x_1^2 + 2x_1^2 x_2 - (x_1 - 2) + x_2^3 \end{pmatrix}.$$

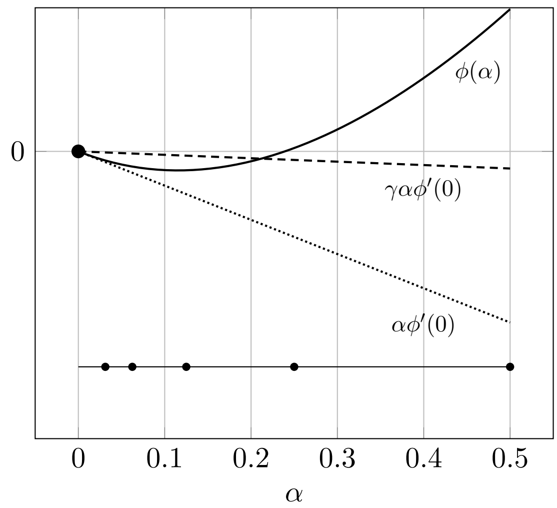


Figure 1: Plot of the functions $\alpha \mapsto \phi(\alpha)$, $\alpha \mapsto \alpha\phi'(0)$, and $\alpha \mapsto \gamma\alpha\phi'(0)$ for $\alpha \in [0, \frac{1}{2}]$.

a) We want to apply the gradient descent method with backtracking to solve $\min_x f(x)$. We choose the initial point x^0 and the Armijo parameter as follows:

$$x^0 = (2, -0.5)^\top, \quad \gamma = 0.1, \quad \sigma = 0.5.$$

We now select $d_{\mathrm{g}}^0 = -\nabla f(x^0)$ and set $\phi(\alpha) := f(x^0 + \alpha d_{\mathrm{g}}^0) - f(x^0)$. Compute the gradient iterate x_{g}^1 and the stepsize α_0 using backtracking and the plot shown in Figure 1.

b) Let $A \in \mathbb{R}^{m\times 2}$, $m \in \mathbb{N}$, and $x \in \mathbb{R}^2$ be a given with $\nabla f(x) \neq 0$. Verify whether the direction $d = -\nabla f(x) - A^\top A \nabla f(x)$ is a descent direction of f at x.

Solution :

a) The Armijo condition is satisfied whenever $\phi(\alpha) \le \gamma\alpha\phi'(0)$. Since $\sigma = \frac{1}{2}$, this is the case for $\alpha_0 = \frac{1}{8}$. We then obtain

$$x_{\mathrm{g}}^1 = x^0 + \alpha_0 d_{\mathrm{g}}^0 = \begin{pmatrix} 2 \\ -\frac{1}{2} \end{pmatrix} - \frac{1}{8}\begin{pmatrix} \frac{1}{2} \\ -\frac{17}{8} \end{pmatrix} = \begin{pmatrix} \frac{31}{16} \\ -\frac{15}{64} \end{pmatrix} = \frac{1}{16}\begin{pmatrix} 31 \\ -\frac{15}{4} \end{pmatrix}.$$

b) It holds that $\nabla f(x)^\top d = -\|\nabla f(x)\|^2 - \|A\nabla f(x)\|^2 \le -\|\nabla f(x)\|^2 < 0$. Thus, d is a descent direction of f at x.

Exercise 4 (Wine Cellar): (20 points)

Lady Dimitrescu has a wine cellar with three separate temperature controlled areas.

The first area has a temperature of $8°C$ and can store up to 35 bottles (this area is mainly used for sparkling wine and champagne). The second area has a temperature of 10–$12°C$ and can store up to 55 bottles. In the third area, 110 bottles can be stored under a temperature of 16–$18°C$.

Lady Dimitrescu plans to invest in different sorts of wine and to resell the wine after three years of storage and maturation at a higher price. The following table summarizes the data and requirements of the different purchasable sorts of wine.

	Dom Perignon (Champagne)	D.R.M. Grand Cru (White Wine)	Opus One (Red Wine)
Temperature	$< 9°C$	8–12$°C$	$> 15°C$
Price per Bottle	19	14	31
Price per Crate (6 Bottles)	102	80	175
Reselling Price per Bottle (After three years)	20	14.7	33.5

Table 1: Wine Data (prices are in 100 RMB per item/bottle/crate)

	Area 1	Area 2	Area 3
Operating Costs (per Year)	6	3	0.5

Table 2: Operating Costs of the Wine Cellar (costs are in 100 RMB)

Each of the areas also has different fixed operating costs per year in case it is used. The costs are summarized in Table 2.

Formulate an integer program to determine the number of bottles of the different sorts of wine to be purchased so as to maximize the total profit (revenue minus costs) after three years when the wine is stored properly under the correct temperature. You can assume that all stored bottles can be sold after three years (using the prices in Table 1).

Solution : We introduce the following decision variables:

- Number of single bottles of champagne: x_1; number of crates of champagne: x_2.
- Number of single bottles of white wine: y_1; number of crates of white wine: y_2.
- Total number of bottles of white wine in area 1: w_1; total number of bottles of white wine in area 2: w_2.
- Number of single bottles of red wine: z_1; number of crates of red wine: z_2.
- Binary variables to indicate if the areas 1, 2, or 3 are used: d_1, d_2, d_3.

We next introduce our constraints. All variables need to be nonnegative integers. The variables x_1, y_1, and z_1 are also upper bounded by 5 (otherwise we can just buy a crate). This yields:

$$\begin{aligned} x_1,\, x_2,\, y_1,\, y_2,\, z_1,\, z_2,\, w_1,\, w_2 &\in \mathbb{Z}, \\ x_1,\, x_2,\, y_1,\, y_2,\, z_1,\, z_2,\, w_1,\, w_2 &\geq 0, \\ d_1,\, d_2,\, d_3 &\in \{0,1\}, \\ x_1 \leq 5,\, y_1 \leq 5,\, z_1 &\leq 5 \\ w_1 + w_2 &= y_1 + 6y_2. \end{aligned}$$

The last constraint is a consistency constraint between w_i and y_i. We now formulate the constraints on the total number of bottles in each area:

$$\begin{aligned} x_1 + 6x_2 + w_1 &\leq 35 \\ w_2 &\leq 55 \\ z_1 + 6z_2 &\leq 110. \end{aligned}$$

Finally, we add one more set of constraints:

$$\frac{x_1 + 6x_2 + w_1}{35} \le d_1, \quad \frac{w_2}{55} \le d_2, \quad \frac{z_1 + 6z_2}{110} \le d_3.$$

Since d_i is a binary variable, these constraints imply that d_i must be 1 whenever a single bottle is stored in area i. The objective function is then given by

$$\max_{(x,y,z,w,d)} \; x_1 + (120-102)x_2 + 0.7y_1 + (6\cdot 14.7 - 80)y_2 + 2.5z_1 + (6\cdot 33.5 - 175)z_2 - 18d_1 - 9d_2 - 1.5d_3.$$

Exercise 5 (True and False): (15 points)

State whether each of the following statements is *True* or *False*. For each part, only your answer, which should be one of *True* or *False*, will be graded. Explanations are not required and will not be read.

a) Let $f : \mathbb{R}^n \to \mathbb{R}$ be a smooth (twice continuously differentiable) and convex function. We want to apply the globalized Newton method to solve the problem $\min_x f(x)$. Let us assume that the method generates a sequence $\{x^k\}$ with $\nabla f(x^k) \neq 0$ for all k. Then, for every iteration k, the Newton direction is well-defined and a descent direction of f at x^k.

b) Let $f : \mathbb{R}^n \to \mathbb{R}$ be bounded from below and let x^* be a stationary point of f. Then, x^* is a global minimizer of f.

c) Consider the nonlinear program $\min_{x\in\Omega} f(x)$ with linear constraints $\Omega := \{x \in \mathbb{R}^n : Ax \le b,\ Cx = d\}$. We assume that $f : \mathbb{R}^n \to \mathbb{R}$ is convex on $\mathbb{R}^n$ and that the set Ω is nonempty. Then, x^* is a global minimizer of this problem, if and only if x^* satisfies the KKT conditions.

d) We use the branch-and-bound method to solve an integer problem (maximization). Suppose we have split the problem into two branches $(S1)$ and $(S2)$ and we continue branching on the subproblem $(S1)$. The method stops branching within $(S1)$ as soon as we recover a feasible integer solution of one of the subproblems. In such a case, the branch-and-bound process continues with the branch $(S2)$.

e) The matrix

$$A = \begin{pmatrix} 1 & 0 & 1 & 0 & -1 & 1 & 0 & 0 & 1 \\ 0 & 1 & -1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & -1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \end{pmatrix}$$

is totally unimodular.

Solution :

a) False.

b) False.

c) True.

d) False.

e) False.

Exercise 6 (Projection Onto an Integer Set): (15 points)
We consider the integer optimization problem

$$\text{minimize}_x \ \frac{1}{2}\|x-z\|^2 \quad \text{subject to} \quad \mathbb{1}^\top x = 1, \quad x_i \in \{-1,1\}, \quad \forall\, i, \tag{2}$$

where $z \in \mathbb{R}^n$ is given, $\mathbb{1} \in \mathbb{R}^n$ is the vector of all ones, and n is assumed to be odd.

a) Verify that the problem (2) can be equivalently written as a special linear problem with binary constraints, i.e., as a problem of the form

$$\min_x \ c^\top x \quad \text{subject to} \quad a^\top x = b, \quad x_i \in \{0,1\}, \quad \forall\, i,$$

with suitable choices of $a \in \mathbb{Z}^n$, $b \in \mathbb{Z}$, and $c \in \mathbb{R}^n$.

Hint: The identity $\|a+b\|^2 = \|a\|^2 + 2a^\top b + \|b\|^2$, $a, b \in \mathbb{R}^n$, can be helpful.

b) Show that there is no integrality gap between the binary problem derived in part a) and its LP relaxation.

Solution :

a) Using the constraints $x_i \in \{\pm 1\}$, $i = 1, \dots, n$, we have $\frac{1}{2}\|x-z\|^2 = \frac{1}{2}\|x\|^2 - x^\top z + \frac{1}{2}\|z\|^2 = \frac{n}{2} - x^\top z + \frac{1}{2}\|z\|^2$. Thus, the objective function is linear in x. We now set $y = \frac{1}{2}(x + \mathbb{1})$ (this implies $x = 2y - \mathbb{1}$). Then, it follows $y_i \in \{0,1\}$ for all i and we have:

$$\mathbb{1}^\top x = 1 \quad \Longleftrightarrow \quad \mathbb{1}^\top(2y - \mathbb{1}) = 1 \quad \Longleftrightarrow \quad \mathbb{1}^\top y = \frac{n+1}{2}.$$

Consequently, problem (2) is equivalent to

$$\min_y \ -2z^\top y \quad \text{subject to} \quad \mathbb{1}^\top y = \frac{n+1}{2}, \quad y_i \in \{0,1\}, \quad \forall\, i. \tag{3}$$

b) Since n is odd, it follows $\frac{n+1}{2} \in \mathbb{Z}$. Since the constraint matrix $\mathbb{1}^\top$ is totally unimodular, there is no integrality gap between the optimal between the binary problem (3) and its LP relaxation. In particular, every optimal BFS of the LP relaxation will correspond to a solution of the initial problem (2).