Math 164-1: Optimization

Instructor: Alpár R. Mészáros

Final Exam, June 9, 2016

Name (use a pen):

Student ID (use a pen):

Signature (use a pen):

Rules:

- Duration of the exam: **180 minutes**.
- By writing your name and signature on this exam paper, you attest that you are the person indicated and will adhere to the UCLA Student Conduct Code.
- **No** calculators, computers, cell phones (all the cell phones should be turned off during the exam), notes, books or other outside material are permitted on this exam. If you want to use scratch paper, you should ask for it from one of the supervisors. Do not use your own scratch paper!
- Please justify all your answers with mathematical precision and write rigorous and clear proofs. You may lose points in the lack of justification of your answers.
- Theorems from the lectures and homework assignments may be used in order to justify your solution. In this case state the theorem you are using.
- This exam has 5 problems and is worth **20 points**. Adding up the indicated points you can observe that there are **26 points**, which means that there are **6 "bonus" points**. This permits to obtain the highest score 20, even if you do not answer some of the questions. On the other hand nobody can be bored during the exam. All scores higher than 20 will be considered as 20 in the gradebook.
- I wish you success!

Problem	Score
Exercise 1	
Exercise 2	
Exercise 3	
Exercise 4	
Exercise 5	
Total	

Exercise 1 (4 points)**.**

Let us consider the following linear program

$$\min\{2x_1 + 2x_2 + 2x_3\} \quad \text{s.t.} \begin{cases} x_1 + x_2 + x_3 &= 6, \\ 2x_1 + x_3 &= 5, \quad x_1, x_2, x_3 \geq 0. \\ -x_2 + 3x_3 &= 7, \end{cases} \qquad (LP_1)$$

(1) Write a feasible solution for (LP_1)!

(2) Does (LP_1) have an optimal solution? If yes, is this solution unique? Determine all the optimal solutions and the value of the objective function as well!

(3) What would happen, if one would exchange the constraint $x_1, x_2, x_3 \geq 0$ in (LP_1) to $x_1, x_2, x_3 \leq 0$? Would this new problem have a solution?

Imagine now that one removes the last equality constraint and one considers

$$\min\{2x_1 + 2x_2 + 2x_3\} \quad \text{s.t.} \begin{cases} x_1 + x_2 + x_3 &= 6, \\ 2x_1 + x_3 &= 5, \end{cases} \quad x_1, x_2, x_3 \geq 0. \qquad (LP_2)$$

(4) What does the set of constraints in (LP_2) represent geometrically? Is it convex? Determine all the feasible solutions for (LP_2)! Justify your answer!

(5) Determine all the optimal feasible solutions for (LP_2)! Do we still have uniqueness? Determine the value of the objective function for the feasible solutions! Justify your answers!

Hint: it is easier to use some geometrical arguments and the structure of (LP_2), than to use the simplex algorithm for instance.

Solution.

(1) It is easy to check that the constraint system has precisely one solution, i.e. $(1, 2, 3)$ for which all its coordinates are nonnegative. So this is the only feasible solution for (LP_1).

(2) By (1) one knows that there exists precisely one feasible solution, in particular this is optimal and it is a unique optimizer for (LP_1). The value of the objective function at $(1, 2, 3)$ is 12.

(3) Since by (1) the constraint system has only one solution that is $(1, 2, 3)$, by changing the inequalities to the opposite ones, (LP_1) would not have any feasible solution, hence no optimal solution either.

(4) The constraint system represent the intersection of two planes in $\mathbb{R}^3$ that are not parallel, so the intersection is a line. Taking for instance x_2 as a parameter, we have $x_1 = x_2 - 1$ and $x_3 = 7 - 2x_2$, which together with the inequality constraints ($x_2 - 1 \geq 0, x_2 \geq 0$ and $7 - 2x_2 \geq 0$) provide us the set of all feasible solutions, i.e. the line segment parametrized as $(x_2 - 1, x_2, 7 - 2x_2)$ where $x_2 \in [1, 7/2]$, which is clearly a convex set.

(5) Since the objective function is $2x_1 + 2x_2 + 2x_3 = 2(x_1 + x_2 + x_3) = 12$ for all feasible solutions in (LP_2) (since $x_1 + x_2 + x_3 = 6$), every feasible solution determined in (4) is optimal with the same objective function value of 12. Hence we do not have uniqueness.

Exercise 2 (6 points).

Find the triangles in the plane with a fixed given area $A > 0$ that have minimal perimeter. To this aim determine the lengths of their sides $a, b, c > 0$ in terms of A using the theory of Lagrange multipliers.

(1) Write down the first order optimality conditions and select the candidates for the optimizers!

(2) Eventually using second order optimality conditions, determine which from the selected candidates in (1) are indeed optimizers.

Hint: you may use Heron's formula for the area, i.e $A = \sqrt{p(p-a)(p-b)(p-c)}$, where $p := \dfrac{a+b+c}{2}$ is the semi-perimeter. Solutions, using other formulas will be also accepted.

Solution.

The problem can we formulated as

$$\min f(a,b,c,p) := 2p$$

subsect to

$$h_1(a,b,c,p) := a+b+c-2p = 0$$

and

$$h_2(a,b,c,p) := p(p-a)(p-b)(p-c) - A^2 = 0.$$

We do not impose directly the inequality constraints $a, b, c > 0$, this will be hidden in the problem, meaning that one should be sure when we find a solution that all its coordinates are positive.

(1) By FONC in the Lagrangian theory one has that if (a,b,c,p) is an optimizer, then there exists $\lambda_1, \lambda_2 \in \mathbb{R}$ such that $\nabla f(a,b,c,p) + \lambda_1 \nabla h_1(a,b,c,p) + \lambda_2 \nabla h_2(a,b,c,p) = 0$, which give us the following system

$$\begin{cases} \lambda_1 - \lambda_2 p(p-b)(p-c) &= 0 \\ \lambda_1 - \lambda_2 p(p-a)(p-c) &= 0 \\ \lambda_1 - \lambda_2 p(p-a)(p-b) &= 0 \\ 2 - 2\lambda_1 + \lambda_2[(p-a)(p-b)(p-c) + p(p-b)(p-c) + p(p-a)(p-c) + p(p-a)(p-b)] &= 0. \end{cases}$$

Notice first that if $\lambda_2 = 0$, the first equation implies that $\lambda_1 = 0$ as well, which by the last equation $(2 = 0)$ would lead to a contradiction. So one has always that $\lambda_2 \neq 0$.

Secondly, if $\lambda_1 = 0$ and $\lambda_2 \neq 0$, since $p > 0$ the first 3 equations imply that at least two of the equations $p - a = 0, p - b = 0, p - c = 0$ should hold true, which by the fourth equation once again (2=0) lead to a contradiction.

So one always has that $\lambda_1 \neq 0$ and $\lambda_2 \neq 0$. This will immediately imply by the first 3 equations that $p \neq a$, $p \neq b$ and $p \neq c$. Continuing the argument, the first three equations easily imply now that $a = b = c$. Using this fact, one can simplify the above system using $p = \dfrac{3}{2}a$ and $p-a = p-b = p-c = \frac{1}{2}a$ as

$$\begin{cases} \lambda_1 - \lambda_2 \dfrac{3}{8}a^3 &= 0 \\ 2 - 2\lambda_1 + \lambda_2 \dfrac{5}{4}a^3 &= 0. \end{cases}$$

In addition, Heron's formula fives us $\dfrac{3}{16}a^4 = A^2$, from where knowing that $a > 0$ one has that $a = \dfrac{2\sqrt{A}}{\sqrt[4]{3}}$.

Now solving the above system for λ_1 and λ_2 one has that $\lambda_1 = -\dfrac{3}{2}$ and $\lambda_2 = -\dfrac{4}{a^3}$.

(2) Clearly the tangent space at any point is $T(a,a,a,3a/2) = \{(y_1, y_2, y_3, 0) \in \mathbb{R}^4 : y_1 + y_2 + y_3 = 0\}$. On the other hand since f and h_1 are linear functions, their Hessian matrices are 0. So, one only need to compute $D^2 h_2(a,a,a,3a/2)$ and one needs to multiply this matrix by any element from $T(a,a,a,3a/2)$. The computation of $D^2 h_2$ is rather complicated, but one can notice immediately that since we have to

multiply it by an element from the tangent space, all these vectors have 0 fourth coordinate, so actually the fourth column of D^2h_2 does not play any role, that is why we do not compute it. For the other entries one has

$$D^2h_2(a,b,c,p) = \begin{pmatrix} 0 & p(p-c) & p(p-b) & \dots \\ p(p-c) & 0 & p(p-a) & \dots \\ p(p-b) & p(p-a) & 0 & \dots \\ -[(p-b)(p-c)+p(p-b)+p(p-c)] & -[(p-a)(p-c)+p(p-a)+p(p-c)] & -[(p-a)(p-b)+p(p-a)+p(p-b)] & \dots \end{pmatrix}$$

which in particular gives us

$$D^2h_2(a,a,a,3a/2) = \begin{pmatrix} 0 & 3a^2/4 & 3a^2/4 & \dots \\ 3a^2/4 & 0 & 3a^2/4 & \dots \\ 3a^2/4 & 3a^2/4 & 0 & \dots \\ -7a^2/4 & -7a^2/4 & -7a^2/4 & \dots \end{pmatrix}.$$

This implies that for all $y \in T(a,a,a,3a/2)$ one has $y^\top D^2(a,a,a,3a/2)y = 3a^2/4(2(y_1y_2+y_2y_3+y_3y_1))$. Since $y_1+y_2+y_3=0$, taking squares of both sides one obtains that $2(y_1y_2+y_2y_3+y_3y_1) = -(y_1^2+y_2^2+y_3^2)$, thus

$$\lambda_2 y^\top D^2(a,a,a,3a/2)y = -\frac{4}{a^3}\frac{3a^2}{4}(-(y_1^2+y_2^2+y_3^2)) = \frac{3}{a}(y_1^2+y_2^2+y_3^2) > 0$$

for all $y \in T(a,a,a,3a/2)$, $y \neq 0$. In particular the SOSC is satisfied, hence the equilateral triangle with side length $a = \dfrac{2\sqrt{A}}{\sqrt[4]{3}}$ has the smallest perimeter among all triangles having area A.

Exercise 3 (6 points)**.**

Let us consider the function $f : \mathbb{R}^n \to \mathbb{R}$ defined as

$$f(x) = \frac{1}{2}x^\top Ax - b^\top x + \gamma\|x\|^2,$$

where $n \geq 1$ is an integer, $A \in \mathbb{R}^{n\times n}$ is a symmetric matrix, $b \in \mathbb{R}^n$, $\gamma \in \mathbb{R}$ and $\|x\| := \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$ denotes the Euclidean norm of the vector $x \in \mathbb{R}^n$. In the followings, we denote the eigenvalues (that are all real) of A with possible multiplicities ordered by $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$. We aim to study the minimization problem

$$\min_{x\in\mathbb{R}^n} f(x). \tag{P}$$

(1) Show that the condition $\lambda_1 + 2\gamma > 0$ is a sufficient condition that ensures that (P) has a unique solution. Determine this unique solution! *Hint:* you may use first and second order optimality conditions! Moreover you may use Rayleigh's inequality, saying that if $Q \in \mathbb{R}^{n\times n}$ is symmetric, then the minimum of Rayleigh's quotient $\dfrac{x^\top Qx}{\|x\|^2}$ is the smallest, while the maximum of this quotient is the largest eigenvalue of Q.

(2) Let us suppose that for some of the eigenvalues λ_i one has $\lambda_i + 2\gamma < 0$. Show that in this case (P) does not have a solution! *Hint:* select a sequence of vectors $(x^k)_{k\geq 0}$ (using for instance the eigenvectors associated to the eigenvalue λ_i) such that $f(x_k) \to -\infty$, as $k \to +\infty$. You may also use the Cauchy-Schwartz inequality (i.e. $a \cdot b \leq \|a\|\|b\|, \forall a, b \in \mathbb{R}^n$).

From now on, one supposes that the assumption in (1), i.e. $\lambda_1 + 2\gamma > 0$ is satisfied, hence (P) has a unique minimizer x^*. We aim to find this minimizer numerically.

(3) Discuss why could we use *Newton's algorithm* to find x^*. How many steps are necessary using Newton's algorithm to find x^* starting from any initial guess $x^0 \in \mathbb{R}^n$? Justify your answer!

(4) Explain why can we use the *conjugate direction* algorithm developed during the lectures to find x^*! With respect to which matrix need we choose the conjugate directions? Justify your answer!

(5) Suppose that $A = I_n$ is the identity matrix. Construct a set of n vectors that are conjugate w.r.t. the matrix determined in (4)!

(6) Suppose once again that $A = I_n$. Write down the updates in the conjugate direction algorithm to find x^* starting from and initial point x^0 using the conjugate directions from (5)!

(7) Explain what is happening geometrically while proceeding the algorithm in (6)! Supposing that one knows x^*, construct an initial guess x^0 for which the algorithm in (6) terminates in precisely 2 steps.

Solution.

Notice first that we can write the objective function as

$$f(x) = \frac{1}{2}x^\top (A + 2\gamma I_n)x - b^\top x.$$

Let us use the notation $Q := A + 2\gamma I_n$.

(1) At a minimizer one has to satisfy the FONC, and if SOSC is satisfied at a point for which FONC holds, then the point is a minimizer. Since Q is also symmetric, one has $\nabla f(x) = Qx - b$ and $D^2 f(x) = Q$. By Rayleigh's inequality one has

$$x^\top Qx = x^\top Ax + 2\gamma\|x\|^2 \geq (\lambda_1 + 2\gamma)\|x\|^2 > 0, \ \forall x \in \mathbb{R}^n,$$

by the assumption that $\lambda_1 + 2\gamma > 0$. So Q is a positive definite matrix, which implies that there exists a unique minimizer, which is $x^* = Q^{-1}b$.

(2) Following the hint, set $x^k = kv^i$, $k \in \mathbb{N}$, where $\|v^i\| = 1$ is an eigenvector of A corresponding to λ_i. Clearly

$$f(x^k) = \frac{1}{2}(x^k)^\top Ax^k - b^\top x^k + \gamma k^2 = \frac{1}{2}\lambda_i k^2 + \gamma k^2 - b^\top x^k \leq \left(\frac{1}{2}\lambda_i + \gamma\right) k^2 + k\|b\|,$$

where we use the Cauchy-Schwarz inequality in the last inequality. Since $\lambda_i + 2\gamma < 0$, one has that $\frac{1}{2}\lambda_i + \gamma < 0$ and

$$\lim_{k\to+\infty} f(x^k) \leq \lim_{k\to+\infty} \left(\frac{1}{2}\lambda_i + \gamma\right) k^2 + k\|b\| = -\infty,$$

hence there exists no minimizer for f on $\mathbb{R}^n$.

(3) By the assumption in (1), f is a quadratic function with $Q > 0$, so Newton's algorithm can be used to find x^* in one step starting at arbitrary $x^0 \in \mathbb{R}^n$. Building the first update starting at x^0 one has

$$x^1 = x^0 - Q^{-1}(Qx^0 - b) = Q^{-1}b = x^*.$$

(4) f is a quadratic function with $Q > 0$, hence the conjugate direction algorithm developed during the lectures applies here. We need to choose conjugate directions w.r.t. $Q = A + 2\gamma I_n$.

(5) If $A = I_n$, $Q = (1+2\gamma)I_n$. Hence the conjugate direction have to by precisely pairwise orthogonal. A such of set could be the canonical basis of $\mathbb{R}^n$, that we denote by $\{e^0, \ldots, e^{n-1}\}$.

(6) Pick an initial guess $x^0 \in \mathbb{R}^n$, then construct

$$x^{k+1} = x^k + \alpha_k e^k,$$

where

$$\alpha_k = \text{argmin}_\alpha f(x^k + \alpha e^k) = -\frac{\nabla f(x^k) \cdot e^k}{(e^k)^\top Q e^k} = -\frac{(1+2\gamma)x^k \cdot e^k - b \cdot e^k}{1 + 2\gamma}.$$

(7) Geometrically, since the conjugate direction algorithm uses the canonical basis of $\mathbb{R}^n$, at each step we "hit" the corresponding coordinate of the optimizer x^*. More precisely after k iterations, x^k has at least k coordinates matching with the coordinates of x^*.

For an initial guess that has its last $n-2$ coordinates exactly the same as x^*, the algorithm terminates in precisely two steps.

Exercise 4 (5 points)**.**

A small company in a country far-far away named Appel wants to attribute three tasks (let us say: changing the locks on some of the office doors; ordering the files in some of the offices and assembling some electronic devices) to its workers (John the locksmith and Jane the accountant) in a way that it maximizes the *productivity*. We know that both John and Jane can work on fractions of each of the tasks and based on previous experience, the company knows exactly how productive John and Jane is for the different tasks. This can be seen from the following *productivity matrix*

$$P = \begin{pmatrix} 5 & 1 & 3 \\ 0 & 4 & 2 \end{pmatrix}.$$

The two rows of this matrix represent the productivity of John and Jane, respectively in the different tasks, e.g. John has a productivity 5 when changing the locks, while 1 for ordering the files, while Jane has productivity 0 when changing locks and productivity 2 for assembling the devices, and so on.

We know moreover the total number of each tasks: there are 3 locks to be changed, 2 offices with files to be ordered and 10 electronic devices to be assembled.

The company is aiming to find the optimal values for the variables $\gamma_{ij} \geq 0$, $i \in \{1, 2\}$ and $j \in \{1, 2, 3\}$ which represent how much John and Jane works on the different tasks to obtain the highest possible global productivity. For instance if we get $\gamma_{13} = 7.4$ that means that John should assemble 7.4 electronic devices, if we get $\gamma_{22} = 1.7$ that means that Jane should order the files in 1.7 offices, etc.

So the company is looking for γ_{ij}'s that solve the problem

$$\max \sum_{i=1}^{2} \sum_{j=1}^{3} P_{ij}\gamma_{ij},$$

under the constraints that all the tasks are completed. Our job is to find these optimal quantities and let the company know how to distribute the different tasks among John and Jane.

(1) Write the above problem as a linear program, i.e. write the objective function that has to be maximized together with the constraints that should be satisfied. Explain how did you obtain them!

(2) Is the LP from (1) in standard form (in the sense of our lectures)? If not, transform it into a standard form. Justify your answer!

(3) Transform the LP from (1) into canonical form (in the sense of our lectures). Give two different set of basic variables and write a basic feasible solution for each cases.

(4) Use the simplex algorithm to solve the LP from (1), at each step write the reduced cost coefficients and a basic feasible solution. Write the optimal solution and the value of the objective function at the optimizer.

(5) Is the optimizer that you have found in (4) unique? Justify your answer! Interpret the solution that you have found in (4)! What do you observe from the "economical point of view"? For this, answer the following questions: what would be different if we would change the coefficients in the matrix P? Do you observe the same phenomena? Would be something different if one would have more than one locksmith or accountant at the company with similar skills?

Solution.
(1) The objective function to be maximized is $f : \mathbb{R}^6 \to \mathbb{R}$ defined as

$$f(\gamma_{11}, \gamma_{12}, \gamma_{13}, \gamma_{21}, \gamma_{22}, \gamma_{23}) = 5\gamma_{11} + \gamma_{12} + 3\gamma_{13} + 4\gamma_{22} + 2\gamma_{23}.$$

To simplify the notation one uses the notation $x \in \mathbb{R}^6$ for $(x_1, x_2, x_3, x_4, x_5, x_6) := (\gamma_{11}, \gamma_{12}, \gamma_{13}, \gamma_{21}, \gamma_{22}, \gamma_{23})$. We have 3 equality constraints coming from the fact that John and Jane have to complete all the tasks; these can be written in term of x as

$$\begin{cases} x_1 + x_4 &= 3 \\ x_2 + x_5 &= 2 \\ x_3 + x_6 &= 10, \end{cases}$$

where the first equation means for instance that the total amount that John and Jane works together on changing locks should be 3, since there are 3 locks to be changed, and so on. We know also that all the variables should be nonnegative, so $x_i \geq 0$ for all $i \in \{1, \dots, 6\}$, or $x \geq 0$ with the vectorial notation.

Introducing the matrix

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

and the vector $b := (3, 2, 10)^\top$ the linear program that we get reads as

$$\max f(x) \quad \text{s.t.} \quad Ax = b \quad \text{and} \quad x \geq 0. \tag{LP_{Appel}}$$

(2) (LP_{Appel}) is not in standard form, because it is a maximization. Introducing

$$c = (-5, -1, -3, 0, -4, -2)^\top$$

one has that $c^\top x = -f(x)$. Moreover $x \geq 0$ and $\operatorname{rank}(A) = 3 < 6$, so ($LP_{\text{Appel}}$) in standard form reads as

$$\min c^\top x \quad \text{s.t.} \quad Ax = b \quad \text{and} \quad x \geq 0. \tag{LPS_{Appel}}$$

(3) (LPS_{Appel}) is already in canonical form, since for instance either the first 3 column or last last 3 ones form I_3 and $b \geq 0$.

The two sets of basic variables are for instance $\{x_1, x_2, x_3\}$ with the corresponding basic feasible solution $(3, 2, 10, 0, 0, 0)^\top$ and $\{x_2, x_4, x_6\}$ with the corresponding basic feasible solution $(0, 2, 0, 3, 0, 10)^\top$.

(4) Since (LPS_{Appel}) contains many identity blocks, we can start the simplex algorithm for instance with $\{x_1, x_2, x_3\}$ as basic variables and $(3, 2, 10, 0, 0, 0)^\top$ as the corresponding basic solution. The canonical augmented matrix (CAM) reads at the beginning as

a_1	a_2	a_3	a_4	a_5	a_6	b
1	0	0	1	0	0	3
0	1	0	0	1	0	2
0	0	1	0	0	1	10.

Now we compute the reduced costs coefficients r_4, r_5, r_6 as

$$\begin{aligned} r_4 &= c_4 - c^\top \tilde{a}_4 = 0 + 5 = 5 \\ r_5 &= c_5 - c^\top \tilde{a}_5 = -4 + 1 = -3 \\ r_6 &= c_6 - c^\top \tilde{a}_6 = -2 + 3 = 1, \end{aligned}$$

where by $\tilde{a}_i \in \mathbb{R}^6$ we denote the column a_i "augmented w.r.t. the basic variables" to be a vector in $\mathbb{R}^6$. Since $r_5 = -3 < 0$ and all the others are positive, a_5 will enter the basis. It is clear by trying to build the fractions b_i / a_{i5} that the only possibility is $2/1$, hence a_2 has to exit the basis. Since a_5 has the same form as a_2 one does not have to perform any row operations, the CAM will remain the same.

The new basic variables are $\{x_1, x_5, x_3\}$ and the corresponding basic feasible solution is $(3, 0, 10, 0, 2, 0)^\top$.

The new reduced cost coefficients are

$$\begin{aligned} r_2 &= c_2 - c^\top \tilde{a}_2 = -1 + 4 = 3 \\ r_4 &= c_4 - c^\top \tilde{a}_4 = 0 + 5 = 5 \\ r_6 &= c_6 - c^\top \tilde{a}_6 = -2 + 3 = 1. \end{aligned}$$

Since all the reduced cost coefficients are nonnegative, we reached an optimal solution, which is

$$x^* = (3, 0, 10, 0, 2, 0)^\top.$$

The value of the objective function at this point is $f(3, 0, 10, 0, 2, 0) = 53$.

(5) The interpretation of the solution found in (4) is the following: $\gamma_{11} = 3$, $\gamma_{13} = 10$ and $\gamma_{22} = 2$ which means that the maximal productivity for the company will be achieved, provided John changes all the 3 locks and assembles all the 10 devices, while Jane orders the files in both offices. This means from the economical point of view that both of them will do those tasks fully in which they are more productive than their colleague. With other words it would not be productive for the company a situation when John assembles only 9 devices and Jane assembles the remaining one, for instance.

By the mathematical structure of the problem, since the matrix A has only columns which are columns of the identity matrix, even if a basic variable exits and a nonbasic enters, the table will be always unchanged. Which means in particular that the last column (for b) is always unchanged. This means that always one person will complete all the parts of a task fully, one cannot observe ever that they both work on the same kind of task, if we want to achieve the highest productivity.

This last phenomenon will be the same, even if we change the coefficients in the productivity matrix P, what will change will be that they may work on different tasks, on the ones in which they are more productive, than the other.

By the fact that the entries in all the columns of P are different, and by the fact that we know that always a person completes a task entirely, one has uniqueness of solution in the current scenario. Giving the task to the other person will result a lower objective function value. On the other hand, it worth to mention that if the entries of a column in P would be the same (meaning both of them have the same productivity while working on a task), then the uniqueness would not hold true, we would have the same objective function value either when John completes that particular task, or when Jane does.

If one would have more than one locksmith or more than one accountant, the situation would be the same, if one represents the productivity of everyone in a different row of P (every new person would introduce 3 new variables into the problem and the identity block structure would remain the same just in higher dimensions). If one would redefine the problem, writing the productivity of a group of people for each task, instead of for everyone separately, the situation could be different.

Exercise 5 (5 points).

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a C^1 function. We define the function $f^* : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ as

$$f^*(y) = \max_{x \in \mathbb{R}^n} \{y \cdot x - f(x)\},$$

whenever the maximizer exists and $+\infty$ otherwise. We aim to study some properties of f^*.

(1) Show that if $f(0) = 0$ then $f^*(y) \geq 0$ for all $y \in \mathbb{R}^n$.

(2) Show that $x \cdot y \leq f(x) + f^*(y)$ for all $x, y \in \mathbb{R}^n$.

(3) Let $p > 1$ and $f(x) = \dfrac{1}{p}\|x\|^p$. Show that f^* is well-defined (i.e. its value is always finite) and $f^*(y) = \dfrac{1}{q}\|y\|^q$ where $\dfrac{1}{p} + \dfrac{1}{q} = 1$. *Hint:* use for instance the Cauchy-Schwartz inequality for the first part (i.e. $a \cdot b \leq \|a\|\|b\|, \forall a, b \in \mathbb{R}^n$) and the first order necessary optimality condition for the second part. For the second part you should first understand what is happening when $n = 1$.

(4) Using the previous points, show Young's inequality, i.e. $x \cdot y \leq \dfrac{1}{p}\|x\|^p + \dfrac{1}{q}\|y\|^q$ for all $x, y \in \mathbb{R}^n$ and $\dfrac{1}{p} + \dfrac{1}{q} = 1$.

(5) Let us define $f^{**} = (f^*)^*$. Suppose that f^* and f^{**} are well-denied (i.e. finite) everywhere in $\mathbb{R}^n$. Show that $f^{**}(x) \leq f(x)$ for all $x \in \mathbb{R}^n$.

Solution.

(1) By the definition of f^* (since it is a max) one has $f^*(y) \geq y \cdot x - f(x)$ for all $x \in \mathbb{R}^n$. In particular if $x = 0$, by the fact that $f(0) = 0$ one has the desired lower bound.

(2) Once again using the definition of f^* one has that $f^*(y) \geq y \cdot x - f(x)$ for all $x, y \in \mathbb{R}^n$. Rewriting the terms one obtains $f^*(y) + f(x) \geq y \cdot x$ for all $x, y \in \mathbb{R}^n$.

(3) Since $f(0) = 0$ one has by (1) that $f^*(y) \geq 0$ for all $y \in \mathbb{R}^n$. So one has to show only that $f^*(y)$ is bounded from above for all $y \in \mathbb{R}^n$. Let us pick $x, y \in \mathbb{R}^n$, one has by the Cauchy-Schwarz inequality that

$$x \cdot y - f(x) = x \cdot y - \frac{1}{p}\|x\|^p \leq \|x\|\|y\| - \frac{1}{p}\|x\|^p = \|y\|t - \frac{1}{p}t^p,$$

where $t := \|x\|$. Let us use the notation $c := \|y\| \geq 0$. We need to study the growth properties of the function

$$g : [0, +\infty) \to \mathbb{R}, \; g(t) = ct - \frac{1}{p}t^p.$$

Clearly $g'(t) = c - t^{p-1} = 0$ has a unique solution, when $p > 1$, namely $t = c^{\frac{1}{p-1}}$ which is a maximizer of g. This means that $g(t) \leq g(c^{\frac{1}{p-1}})$ for all $t \in [0, +\infty)$. This implies in particular that

$$x \cdot y - f(x) \leq cc^{\frac{1}{p-1}} - \frac{1}{p}c^{\frac{p}{p-1}},$$

and the right hand side is independent of x, so taking the maximum of both sides w.r.t. $x \in \mathbb{R}^n$, one obtains that

$$f^*(y) \leq c^{\frac{p}{p-1}} - \frac{1}{p}c^{\frac{p}{p-1}} = \frac{1}{q}c^q,$$

where $c = \|y\|$ and $q = \dfrac{p}{p-1}$.

Thus one obtained $f^*(y) \leq \dfrac{1}{q}\|y\|^q$, hence in particular f^* is well-defined for all $y \in \mathbb{R}^n$.

We need to show now the opposite inequality, i.e. $f^*(y) \geq \frac{1}{q}\|y\|^q$. Since $f^*(y) \geq x \cdot y - f(x)$ for all $x \in \mathbb{R}^n$, take $x = \|y\|^{\frac{1}{p-1}-1}y$ (which is coming from the FONC, since $\nabla f(x) = \|x\|^{p-2}x$) then one has

$$f^*(y) \geq \|y\|^{1+\frac{1}{p-1}} - \frac{1}{p}\|y\|^{\frac{p}{p-1}} = \frac{1}{q}\|y\|^q.$$

This shows the desired equality.

(4) Is a simple consequence of (2) and (3).

(5) One has

$$f^{**}(x) = \max_y\{y\cdot x - f^*(y)\} = \max_y\{y\cdot x - \max_z\{y\cdot z - f(z)\}\} = \max_y\{y\cdot x + \min_z\{-y\cdot z + f(z)\}\} \leq \max_y\{y\cdot x - y\cdot x + f(x)\} = f(x),$$

where we used the fact that $\min_z\{-y \cdot z + f(z)\} \leq -y \cdot x + f(x)$.