Chalmers/GU
Mathematics

EXAM

TMA947/MAN280
OPTIMIZATION, BASIC COURSE

Date:	07–12–17
Time:	House V, morning
Aids:	Text memory-less calculator, English–Swedish dictionary
Number of questions:	7; passed on one question requires 2 points of 3. Questions are *not* numbered by difficulty. To pass requires 10 points and three passed questions.
Examiner:	Michael Patriksson
Teacher on duty:	Adam Wojciechowski (0762-721860)
Result announced:	08–01–08 Short answers are also given at the end of the exam on the notice board for optimization in the MV building.

Exam instructions

When you answer the questions

Use generally valid theory and methods.
State your methodology carefully.

Only write on one page of each sheet. Do not use a red pen.
Do not answer more than one question per page.

At the end of the exam

Sort your solutions by the order of the questions.
Mark on the cover the questions you have answered.
Count the number of sheets you hand in and fill in the number on the cover.

Question 1

(The Simplex method)

Consider the following linear program:

$$\begin{array}{lrcrcrcrcl} \text{minimize} & z = 2x_1 & & & & & & & & \\ \text{subject to} & x_1 & & & - & x_3 & & & = & 3, \\ & x_1 & - & x_2 & & & - & 2x_4 & = & 1, \\ & 2x_1 & & & & & + & x_4 & \leq & 7, \\ & x_1, & & x_2, & & x_3, & & x_4 & \geq & 0. \end{array}$$

(2p) a) Solve this problem by using phase I and phase II of the Simplex method.

Aid: Some matrix inverses that might come in handy are

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 0 \\ -2 & 1 & 0 \\ 1 & 2 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 0.5 & -0.5 & 0 \\ 1 & 0 & 0 \\ -2.5 & 0.5 & 1 \end{pmatrix},$$

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & -2 \\ 0 & 2 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & -0.2 & -0.4 \\ 0 & 0.2 & 0.4 \\ 0 & -0.4 & 0.2 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 2 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & 0 \\ -2 & 0 & 1 \end{pmatrix},$$

$$\begin{pmatrix} 0 & 1 & 0 \\ -1 & 1 & -2 \\ 0 & 2 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 5 & -1 & -2 \\ 1 & 0 & 0 \\ -2 & 0 & 1 \end{pmatrix}.$$

(1p) b) If a primal LP is infeasible, what can you say about its LP dual?

(3p)

Question 2

(the KKT conditions)

Consider the problem to find

$$\begin{aligned} f^* := \operatorname*{infimum}_{x} \; & f(\boldsymbol{x}), \\ \text{subject to } \; & g_i(\boldsymbol{x}) \leq 0, \qquad i = 1, \dots, m, \end{aligned}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ and $g_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, 2, \dots, m$, are given differentiable functions.

(1p) a) State the KKT conditions regarding locally optimal solutions to this problem.

(1p) b) Assume that there are two locally optimal solutions, $\boldsymbol{x}^1$ and $\boldsymbol{x}^2$, to the problem at hand. Suppose that the feasible set at $\boldsymbol{x}^1$ satisfies the linear independence constraint qualification (LICQ). Does the vector $\boldsymbol{x}^1$ satisfy the KKT conditions? Does the vector $\boldsymbol{x}^2$ satisfy the KKT conditions?

(1p) c) Assume instead that there are two vectors, $\boldsymbol{x}^1$ and $\boldsymbol{x}^2$, both satisfying the KKT conditions. Assume also that these are the only KKT points. Suppose that the feasible set, at $\boldsymbol{x}^1$, satisfies the linear independence constraint qualification (LICQ). Further, assume that there exists at least one locally optimal solution to the given problem. In terms of local or global optimality, what can be said about the vectors $\boldsymbol{x}^1$ and $\boldsymbol{x}^2$?

Question 3

(short questions on different topics)

(1p) a) Motivate whether the polyhedron in $\mathbb{R}^5$ described by the system

$$\begin{aligned} x_1 + 2x_2 - x_3 - 2x_4 + 4x_5 &= 0, \\ 2x_1 - x_2 + 2x_3 + 3x_4 + x_5 &= 4, \\ x_1, \; x_2, \; x_3, \; x_4, \; x_5 &\geq 0, \end{aligned}$$

has or has not an extreme point in $(1, 0, 1, 0, 0)^{\mathrm{T}}$.

(1p) b) Consider the unconstrained minimization of a C^2 function $f : \mathbb{R}^n \to \mathbb{R}$. Suppose that, at $\boldsymbol{x}_k$, $\nabla f(\boldsymbol{x}_k) \neq \mathbf{0}^n$. In the Levenberg–Marquardt modification of Newton's method, the Newton equation for determining the search direction $\boldsymbol{p}_k$,

$$\nabla^2 f(\boldsymbol{x}_k)\boldsymbol{p}_k = -\nabla f(\boldsymbol{x}_k),$$

is modified, whenever necessary, such that a multiple $\gamma_k > 0$ of the unit matrix is added to the Hessian in order to make the (modified) Newton equation uniqely solvable. Show that this modification of the search direction always yields a descent direction.

(1p) c) Consider the problem to

$$\begin{aligned} \text{minimize} \quad & f(\boldsymbol{x}), \\ \text{subject to} \quad & g_i(\boldsymbol{x}) = 0, \quad i = 1, \dots, m, \\ & x \in X, \end{aligned}$$

where f and g_i, $i = 1, \dots, m$ are convex functions and where $X \subseteq \mathbb{R}^n$ is a convex set. Is it true that each local minimum also is a global minimum? If so, motivate carefully. If not, present a counterexample.

(3p) ## Question 4

(the separation theorem)

Given a closed and convex set $C \subset \mathbb{R}^n$ and a vector $\boldsymbol{y} \in \mathbb{R}^n$ that does not belong to C, the separation theorem states a result on the existence of a separating hyperplane. State the separation theorem precisely, and establish its correctness with a proof.

Question 5

(LP duality and derivatives)

Consider the LP problem to find

$$\begin{aligned} v(\boldsymbol{b}) := \underset{x \in \mathbb{R}^n}{\text{minimum}} \quad & \boldsymbol{c}^{\mathrm{T}} \boldsymbol{x}, \\ \text{subject to} \quad & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \\ & \boldsymbol{x} \geq \boldsymbol{0}^n, \end{aligned} \tag{1}$$

where $\boldsymbol{c} \in \mathbb{R}^n$, $\boldsymbol{A} \in \mathbb{R}^{m \times n}$, and $\boldsymbol{b} \in \mathbb{R}^m$.

(1p) a) Establish that the function v is convex.

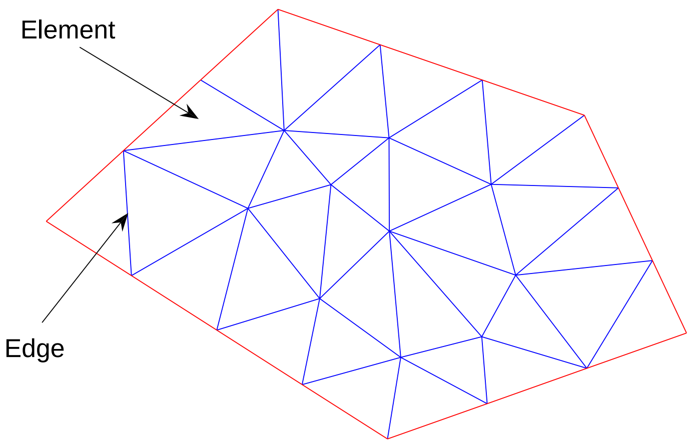
(2p) b) Suppose that locally around the vector $\boldsymbol{b}$, v is finite; that is, suppose that the LP problem (1) has finite optimal solutions for all right-hand side vectors close to $\boldsymbol{b}$. Suppose, further, that for the given value of $\boldsymbol{b}$, $\boldsymbol{y}^* \in \mathbb{R}^m$ is an optimal solution to the corresponding LP dual problem. Prove that $\boldsymbol{y}^*$ is a subgradient of v at $\boldsymbol{b}$. In particular, supposing that $\boldsymbol{y}^*$ is the *unique* optimal solution, establish that then v is differentiable at $\boldsymbol{b}$, and $\nabla v(\boldsymbol{b}) = \boldsymbol{y}^*$.

(3p) **Question 6**

(modelling)

Load balancing is a technique used to spread work between computers in order to get optimal resource utilization and decrease computing time.

You are about to numerically solve a partial differential equation, which has been discretized on a computational mesh, consisting of n elements. The amount of work should be distributed among a set of computers. In more detail, the elements of the computational mesh need to be assigned to the different computers. The amount of work per element is η flops (floating point operations). Obviously, an element can only be assigned to one computer.



To construct the final solution, the computers need to communicate with each other. The amount of work for communication depends on the boundary between the elements of the different computers. For each edge between two elements assigned to two different computers, the amount of work is ρ flops for each of the two computers. The communication between the computers can only be done after all have completed the work on the elements. This means that if one computer finishes early with the elements, it has to wait for the others.

For the sake of simplicity, assume that you only have two computers. Both can do ν flops per second. You have access to a list of all elements in the mesh, as well as a list of all m edges between elements. For example, the list can be represented by a m-by-2 matrix E, where each row of E contains two indices to elements sharing an edge.

Your job is to formulate an optimization problem which assigns the elements to the two computers, so that you minimize the computing time (this includes both the work on the elements and the communication work). Your optimization problem can contain continuous, integer or binary variables, but the constraints and the objective function must be linear.

(3p) Question 7

(Lagrangian Duality) By studying the non-linear program to

$$\begin{aligned} \text{minimize} \quad & z = \sum_{i=1}^{n} x_i^2, \\ \text{subject to} \quad & \sum_{i=1}^{n} x_i = b, \end{aligned}$$

where $b > 0$, use Lagrangian duality theory to derive the (special case of the Cauchy-Schwarz) inequality

$$n \sum_{i=1}^{n} x_i^2 \geq \left(\sum_{i=1}^{n} x_i \right)^2 .$$

Show also that equality holds if and only if $x_1 = x_2 = \ldots = x_n$.

Good luck!

Chalmers/Gothenburg University
Mathematical Sciences

EXAM SOLUTION

TMA947/MAN280
APPLIED OPTIMIZATION

Date: 07–12–17
Examiner: Michael Patriksson

Question 1

(The Simplex method)

(2p) a) By introducing a slack variable x_5 and two artificial variables a_1 and a_2, we get the Phase I problem to

$$
\begin{array}{rrrrrrrrcl}
\text{minimize} \quad w = & & & & & & a_1 & +a_2 & & \\
\text{subject to} & x_1 & & -x_3 & & & +a_1 & & = & 3\,, \\
 & x_1 & -x_2 & & -2x_4 & & & +a_2 & = & 1\,, \\
 & 2x_1 & & & +x_4 & +x_5 & & & = & 7\,, \\
 & x_1\,, & x_2\,, & x_3\,, & x_4\,, & x_5\,, & a_1\,, & a_2 & \geq & 0\,.
\end{array}
$$

Let $\boldsymbol{x}_B^{\mathrm{T}} = (a_1, a_2, x_5)$ and $\boldsymbol{x}_N^{\mathrm{T}} = (x_1, x_2, x_3, x_4)$ be the initial basic and nonbasic vector. The reduced costs of the nonbasic variables are

$$\boldsymbol{c}_N^{\mathrm{T}} - \boldsymbol{c}_B^{\mathrm{T}} \boldsymbol{B}^{-1} \boldsymbol{N} = (-2, 1, 1, 2),$$

which means that x_1 is the entering variable. Further, we have

$$\boldsymbol{B}^{-1}\boldsymbol{b} = (3, 1, 7)^{\mathrm{T}},$$

$$\boldsymbol{B}^{-1}\boldsymbol{N}_1 = (1, 1, 2)^{\mathrm{T}},$$

which gives

$$\operatorname{argmin}_{j:(B^{-1}N_1)_j>0} \frac{(\boldsymbol{B}^{-1}\boldsymbol{b})_j}{(\boldsymbol{B}^{-1}\boldsymbol{N}_1)_j} = 2,$$

so a_2 is the leaving variable. The new basic and nonbasic vectors are $\boldsymbol{x}_B^{\mathrm{T}} = (a_1, x_1, x_5)$ and $\boldsymbol{x}_N^{\mathrm{T}} = (a_2, x_2, x_3, x_4)$, and the reduced costs are

$$\boldsymbol{c}_N^{\mathrm{T}} - \boldsymbol{c}_B^{\mathrm{T}} \boldsymbol{B}^{-1} \boldsymbol{N} = (2, -1, 1, -2),$$

so x_4 is the entering variable, and

$$\boldsymbol{B}^{-1}\boldsymbol{b} = (2, 1, 5)^{\mathrm{T}},$$

$$\boldsymbol{B}^{-1}\boldsymbol{N}_4 = (2, -2, 5)^{\mathrm{T}},$$

which gives

$$\operatorname{argmin}_{j:(B^{-1}N_4)_j>0} \frac{(\boldsymbol{B}^{-1}\boldsymbol{b})_j}{(\boldsymbol{B}^{-1}\boldsymbol{N}_4)_j} = 1,$$

and thus a_1 is the leaving variable. The new basic and nonbasic vectors are $\boldsymbol{x}_B^{\mathrm{T}} = (x_4, x_1, x_5)$ and $\boldsymbol{x}_N^{\mathrm{T}} = (a_2, x_2, x_3, a_1)$, and the reduced costs are

$$\boldsymbol{c}_N^{\mathrm{T}} - \boldsymbol{c}_B^{\mathrm{T}} \boldsymbol{B}^{-1} \boldsymbol{N} = (1, 0, 0, 1),$$

so $\boldsymbol{x}_B^{\mathrm{T}} = (x_4, x_1, x_5)$ is an optimal basic feasible solution of the Phase I problem. Since $w^* = 0$, $\boldsymbol{x}_B$ is a basic feasible solution of the Phase II problem to

$$
\begin{array}{rrrrrrcl}
\text{minimize} & z = 2x_1 & & & & & & \\
\text{subject to} & x_1 & & -x_3 & & & = & 3\,, \\
& x_1 & -x_2 & & -2x_4 & & = & 1\,, \\
& 2x_1 & & & +x_4 & +x_5 & = & 7\,, \\
& x_1\,, & x_2\,, & x_3\,, & x_4\,, & x_5 & \geq & 0\,.
\end{array}
$$

If $\boldsymbol{x}_B^{\mathrm{T}} = (x_4, x_1, x_5)$ and $\boldsymbol{x}_N^{\mathrm{T}} = (x_2, x_3)$, we get the reduced costs

$$\boldsymbol{c}_N^{\mathrm{T}} - \boldsymbol{c}_B^{\mathrm{T}} \boldsymbol{B}^{-1} \boldsymbol{N} = (0, 2).$$

This means that $\boldsymbol{x}_B$ is an optimal basic feasible solution for the Phase II problem, and we are done! $\boldsymbol{x}^* = (3, 0, 0, 1)^{\mathrm{T}}$ and $z^* = 6$.

(1p) b) If the primal is infeasible, the dual cannot have an optimal solution. Thus it is either infeasible or unbounded.

Question 2

(the KKT conditions)

(1p) a) See the Book, system (5.9).

(1p) b) The vector $\boldsymbol{x}^1$ satisfies the KKT conditions (5.9).

(1p) c) Nothing. (Under the conditions given, there may be optimal solutions that do not satisfy the KKT conditions.)

Question 3

(short questions on different topics)

(1p) a) Yes it is. $(1, 0, 1, 0, 0)^{\mathrm{T}}$ is feasible and the columns of A corresponding to the positive entries are linearly independent.

(1p) b) By multiplying with $\boldsymbol{p}_k$ from the left we get

$$\boldsymbol{p}_k^{\mathrm{T}}(\nabla^2 f(\boldsymbol{x}_k) + \gamma_k \boldsymbol{I}^n)\boldsymbol{p}_k = -\boldsymbol{p}_k^{\mathrm{T}}\nabla f(\boldsymbol{x}_k).$$

Since γ_k is chosen such that $\nabla^2 f(\boldsymbol{x}_k) + \gamma_k \boldsymbol{I}^n$ is positive definite [that is, $\boldsymbol{u}^{\mathrm{T}}(\nabla^2 f(\boldsymbol{x}_k) + \gamma_k \boldsymbol{I}^n)\boldsymbol{u} > 0$ holds for all $\boldsymbol{u} \in \mathbb{R}^n \setminus \{\boldsymbol{0}^n\}$], it follows that $\boldsymbol{p}_k^{\mathrm{T}}\nabla f(\boldsymbol{x}_k) < 0$ and $\boldsymbol{p}_k$ is therefore a direction of descent.

(1p) c) It is not true. Consider for example the problem to

$$\begin{aligned} &\text{minimize} && x_1, \\ &\text{subject to} && x_1^2 + x_2^2 - 1 = 0, \\ & && x \in X = \{\, \boldsymbol{x} \in \mathbb{R}^2 \mid x_1 + x_2 \geq 0 \,\}, \end{aligned}$$

which has the two local minima $\left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$ and $\left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$, of which only the latter is a global minimum.

(3p) Question 4

(the separation theorem) See the Book, Theorem 4.28.

Question 5

(LP duality and derivatives)

(1p) a) If $v(\boldsymbol{b})$ is finite, then by LP duality, we have that

$$\begin{aligned} v(\boldsymbol{b}) := \underset{y \in \mathbb{R}^m}{\text{maximum}} \quad & \boldsymbol{b}^{\mathrm{T}}\boldsymbol{y}, \\ \text{subject to} \quad & \boldsymbol{A}^{\mathrm{T}}\boldsymbol{y} \leq \boldsymbol{c}, \\ & \boldsymbol{y} \text{ free.} \end{aligned} \tag{1}$$

At least one maximum in (1) is attained at an extreme point of the dual polyhedron. Therefore, we can write $v(\boldsymbol{b}) = \text{maximum}_{k \in \mathcal{K}}\, \boldsymbol{b}^{\mathrm{T}}\boldsymbol{y}_k$, where $\{\boldsymbol{y}_k\}_{k \in \mathcal{K}}$ is the (finite) set of extreme points of the dual polyhedron. The convexity of v follows simply by using the definition: for $\lambda \in (0,1)$ and arbitrary vectors $\boldsymbol{b}^1$ and $\boldsymbol{b}^2$ in $\mathbb{R}^m$ it holds that

$$\max_{k \in \mathcal{K}} [\lambda \boldsymbol{b}^1 + (1-\lambda)\boldsymbol{b}^2]^{\mathrm{T}}\boldsymbol{y}_k \leq \lambda \max_{k \in \mathcal{K}} (\boldsymbol{b}^1)^{\mathrm{T}}\boldsymbol{y}_k + (1-\lambda) \max_{k \in \mathcal{K}} (\boldsymbol{b}^2)^{\mathrm{T}}\boldsymbol{y}_k,$$

the inequality being a consequence of the added freedom of choice when separating the optimization problem on the left-hand side of the inequality with the two optimization problems in the right-hand side. Hence,

$$v(\lambda \boldsymbol{b}^1 + (1-\lambda)\boldsymbol{b}^2) \leq \lambda v(\boldsymbol{b}^1) + (1-\lambda) v(\boldsymbol{b}^2),$$

and we are done.

(2p) b) Consider the following inequality:

$$v(\boldsymbol{p}) \geq v(\boldsymbol{b}) + \boldsymbol{\xi}^{\mathrm{T}}(\boldsymbol{p} - \boldsymbol{b}), \qquad \forall \boldsymbol{p} \in \mathbb{R}^m,$$

where $\boldsymbol{\xi} \in \mathbb{R}^m$. This inequality is the definition of the vector $\boldsymbol{\xi}$ being a subgradient of the convex function v at $\boldsymbol{b}$; it in fact characterizes v as being convex, whenever it is sub-differentiable. Our task is to establish that this inequality holds when we let $\boldsymbol{\xi} = \boldsymbol{y}^*$. Since $v(\boldsymbol{b}) = \boldsymbol{b}^{\mathrm{T}}\boldsymbol{y}^*$ by assumption, the inequality reduces to stating that

$$v(\boldsymbol{p}) \geq \boldsymbol{p}^{\mathrm{T}}\boldsymbol{y}^*, \qquad \forall \boldsymbol{p} \in \mathbb{R}^m.$$

But this is true: by definition, $v(\boldsymbol{p})$ equals the supremum of $\boldsymbol{p}^{\mathrm{T}}\boldsymbol{y}$ over all feasible vectors $\boldsymbol{y}$, and $\boldsymbol{y}^*$ is just one out of all the possible choices of dual feasible vectors.

Finally, differentiability of v at $\boldsymbol{b}$ is equivalent, given its convexity, to the existence of a unique subgradient of v at $\boldsymbol{b}$. From the above it is clear that if there is only one optimal solution to the problem (1) then that must also be the gradient of v at $\boldsymbol{b}$.

(3p) Question 6

(modelling) Introduce the variables:

x_i is 0 if element i is assigned to computer 1
and it is 1 if assigned to computer 2. $i = 1, \ldots, n$

y_k is 1 if edge k is between to elements assigned to different computers.
It is 0 otherwise. $k = 1, \ldots m$

The computing time for the elements is equal to

$$\max\left\{\frac{\eta}{\nu}\sum_{i=1}^{n} x_i,\ \frac{\eta}{\nu}\left(n - \sum_{i=1}^{n} x_i\right)\right\},$$

which can be modelled using an auxilary variable t and linear inequalities. The optimization problem reads:

$$\begin{aligned}
\text{minimize} \quad z = \frac{\eta t}{\nu} + \frac{\rho}{\nu}\sum_{k=1}^{m} \quad & y_k \\
\text{subject to} \qquad \sum_{i=1}^{n} x_i \leq \ & t \\
n - \sum_{i=1}^{n} x_i \leq \ & t \\
x_{E_{k,1}} - x_{E_{k,2}} \leq \ & y_k \ , \ k = 1, \dots, m \\
x_{E_{k,2}} - x_{E_{k,1}} \leq \ & y_k \ , \ k = 1, \dots, m \\
\boldsymbol{x} \in \ & \mathbb{B}^n \\
\boldsymbol{y} \in \ & \mathbb{B}^m \\
t \in \ & \mathbb{R}
\end{aligned}$$

(3p) Question 7

(Lagrangian Duality) Lagrangian relax the contraint to get

$$L(\boldsymbol{x}, \lambda) = -\lambda b + \sum_{i=1}^{n} x_i^2 + \lambda(\sum_{i=1}^{n} x_i - b).$$

L is differentiable and we find the Lagrangian dual function

$$q(\lambda) = \min_{x \in \mathbb{R}^n} L(\boldsymbol{x}, \lambda)$$

by setting the gradient of L with respect to $\boldsymbol{x}$ equal to zero (convex unconstrained problem, function in C^1). $\nabla_x L(\boldsymbol{x}, \lambda) = \mathbf{0} \Rightarrow x_i^* = -\frac{\lambda}{2}, \ \forall i$. We get $q(\lambda) = -\lambda b - n\frac{\lambda^2}{4}$.

In the Lagrangian dual problem we wish to maximize $q(\lambda)$ over $\mathbb{R}$ (no sign restrictions since the multiplier corresponds to an equality constraint). Also here, q is differentiable and we set the gradient equal to zero $\Rightarrow \lambda^* = -\frac{2b}{n}$ (we know that this is a maximum, since q is always concave) $\Rightarrow x_i^* = \frac{b}{n}, \ \forall i$.

Thus, for any faesible vector $\boldsymbol{x}$,

$$z^* = \sum_i \left(\frac{b}{n}\right)^2 = \frac{b^2}{n} \leq \sum_i x_i^2 \Leftrightarrow b^2 \leq n \sum_i x_i^2 \Leftrightarrow \left(\sum_{i=1}^{n} x_i\right)^2 \leq n \sum_i x_i^2.$$

The objective function is strictly convex, whence the inequality above holds with equality iff $x_i^* = \frac{b}{n}$, $\forall i$, i.e., if $x_1 = x_2 = \ldots = x_n$.