Chalmers/GU
Mathematics

EXAM

TMA947/MMG621
OPTIMIZATION, BASIC COURSE

Date:	14–04–22
Time:	House V, morning, 8^{30}–13^{30}
Aids:	Text memory-less calculator, English–Swedish dictionary
Number of questions:	7; passed on one question requires 2 points out of 3 Questions are *not* numbered by difficulty. To pass requires 10 points and three passed questions
Examiner:	Michael Patriksson
Teacher on duty:	Emil Gustavsson (0703-088304)
Result announced:	14–05–14 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:

$$\begin{array}{lrl} \text{minimize} & z = 2x_1 + \ x_2, & \\ \text{subject to} & 2x_1 + \ x_2 & \geq -2, \\ & 2x_1 + 5x_2 & \leq \ \ 6, \\ & x_2 & \geq \ \ 0. \end{array}$$

(2p) a) Solve the problem using the simplex method. If the problem has an optimal solution, then present the optimal solution in both the original and in the variables used in the standard form. If the problem is unbounded, then use your calculations to find a direction of unboundedness in both the original variables and in the variables in the standard form.
Aid: Utilize the identity

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

(1p) b) Is the optimal solution obtained unique? Motivate your answer. If the optimal solution is not unique, then state all alternative optimal solutions.

(3p)
Question 2

(KKT conditions)

Let $a_1 \geq a_2 \geq \ldots \geq a_n > 0$ and consider the optimization problem to

$$\begin{array}{ll} \text{minimize} & -\log\left(\sum_{i=1}^{n} a_i x_i\right) - \log\left(\sum_{i=1}^{n} \frac{x_i}{a_i}\right) \\ \text{subject to} & \sum_{i=1}^{n} x_i = 1 \\ & x_i \geq 0, \ i = 1, \ldots, n. \end{array}$$

Show that $\mathbf{x} = (1/2, 0, \ldots, 0, 1/2)^{\mathrm{T}}$ is an optimal solution.

Question 3

(problem decomposition)

Consider the problem to minimize a convex and differentiable function f of the form

$$f(\boldsymbol{x}) := \sum_{i \in \mathcal{I}} f_i(\boldsymbol{x}_i),$$

where $\mathcal{I}$ is a finite index set and $\boldsymbol{x}_i \in \mathbb{R}^n$, subject to two types of constraints: (1) an individual feasible set for each "block" of variables $\boldsymbol{x}_i$, of the form

$$\boldsymbol{x}_i \in X_i, \qquad i \in \mathcal{I},$$

where the sets X_i are non-empty polyhedral sets in $\mathbb{R}^n$, and (2) a total resource constraint of the form

$$\sum_{i \in \mathcal{I}} \boldsymbol{x}_i \leq \boldsymbol{u},$$

for some vector $\boldsymbol{u} \in \mathbb{R}^n$.

(2p) a) Describe how a Lagrangian relaxation algorithm for this problem would appear if we Lagrangian relax the resource constraint. Describe in particular the appearance of the Lagrangian subproblem, and how you would solve it. Can you easily provide a lower bound on the optimal value of the original problem? How?

(1p) b) Suppose next that f is quadratic and that $n_i = 1$ for all $i \in \mathcal{I}$. Describe how the Lagrangian subproblem can be solved analytically.

(3p) Question 4

(Frank-Wolfe algorithm)

Consider the problem to

$$
\begin{aligned}
\underset{x_1,x_2}{\text{minimize}} \quad & f(\boldsymbol{x}) := \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 6 & 2 \\ 2 & 9 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - \begin{pmatrix} 52 & 34 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \\
\text{subject to} \quad & x_1 + 2x_2 \leq 4, \\
& x_1 + x_2 \leq 3, \\
& 2x_1 \leq 5, \\
& x_1 \geq 0, \\
& x_2 \geq 0.
\end{aligned} \tag{1}
$$

Solve problem (1) with the Frank–Wolfe algorithm. Start with initial guess $\boldsymbol{x}^{(0)} = (x_1, x_2)^{\mathrm{T}} = (2.5, 0)^{\mathrm{T}}$. Use exact minimization for line search. If necessary, you are allowed to carry out the calculations approximately with two digits of accuracy.

[*Hint:* You may find it helpful to analyze the problem and the algorithm progress in a picture, but this should be augmented with rigorous analysis.]

Question 5

(true or false)

The below three claims should be assessed. Are they true or false? Provide an answer together with a short but complete motivation.

(1p) a) Suppose that you wish to solve a linear integer program, and that you start by solving its continuous relaxation. Suppose that $\bar{\boldsymbol{x}}$ is a solution to this problem. Then, an optimal solution to the integer program can always be found by rounding, individually, each element of $\bar{\boldsymbol{x}}$ either up or down to the nearest integer value.

(1p) b) Suppose that you are able to solve a nonlinear optimization problem and that in a globally optimal solution to it there is one inequality constraint that is satisfied with *strict* inequality. Then, this inequality is redundant, and can be removed without affecting the optimal solution.

(1p) c) Suppose that you consider minimizing a convex and differentiable function f over a closed convex set S, and that you have found an optimal solution

$\boldsymbol{x}^*$. Suppose also that there is another optimal solution $\bar{\boldsymbol{x}}$. Then, all points on the line segment between $\boldsymbol{x}^*$ and $\bar{\boldsymbol{x}}$ must also be optimal.

(3p) Question 6

(the Relaxation Theorem)

Given the problem to find

$$f^* := \underset{x}{\text{infimum}} \; f(\boldsymbol{x}), \tag{1a}$$

$$\text{subject to } \boldsymbol{x} \in S, \tag{1b}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is a given function and $S \subseteq \mathbb{R}^n$, we define a *relaxation* to (1) to be a problem of the following form: find

$$f_R^* := \underset{x}{\text{infimum}} \; f_R(\boldsymbol{x}), \tag{2a}$$

$$\text{subject to } \boldsymbol{x} \in S_R, \tag{2b}$$

where $f_R : \mathbb{R}^n \to \mathbb{R}$ is a function with the property that $f_R \leq f$ on S, and where $S_R \supseteq S$. For this pair of problems, we have the following basic result. You are asked to establish it.

Theorem 1 (Relaxation Theorem)

(a) [relaxation] $f_R^* \leq f^*$.

(b) [infeasibility] *If* (2) *is infeasible, then so is* (1).

(c) [optimal relaxation] *If the problem* (2) *has an optimal solution,* $\boldsymbol{x}_R^*$*, for which it holds that*

$$\boldsymbol{x}_R^* \in S \qquad \text{and} \qquad f_R(\boldsymbol{x}_R^*) = f(\boldsymbol{x}_R^*), \tag{3}$$

then $\boldsymbol{x}_R^*$ *is an optimal solution to* (1) *as well.*

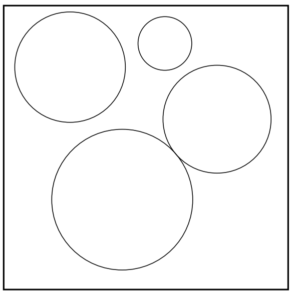
Question 7

(modelling)

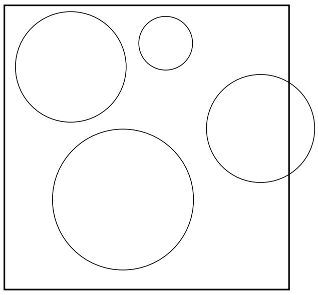
Consider a square with side length L and corners in $(0,0), (0,L), (L,0)$ and (L,L). Formulate the problem of placing n circles inside the square in such a way that no circles overlap and such that the total area covered by the circles is maximized. Note that the radius of each circle should also be a variable in the optimization model.

Tip: Let r_i be the radius of circle $i = 1, \dots, n$, and let (x_i, y_i) be the coordinate of the center point of circle $i = 1, \dots, n$. These are the only variables you will need in the optimization model.

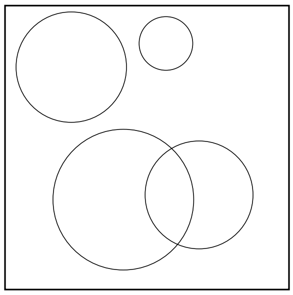
In the figures below you can see one feasible solution and two infeasible solutions for the problem when $n = 4$.



Feasible



Infeasible
(not inside the square)



Infeasible
(overlapping)

Chalmers/GU
Mathematics

EXAM SOLUTION

TMA947/MMG621
NONLINEAR OPTIMISATION

Date: 14–04–22
Examiner: Michael Patriksson

Question 1

(the simplex method)

(2p) a) We first rewrite the problem on standard form. We introduce slack variables s_1 and s_2 and $x_1 = x_1^+ - x_1^-$. Consider the following linear program:

$$\begin{array}{llrcl}
\text{minimize} & z = & 2x_1^+ - 2x_1^- + x_2 & & \\
\text{subject to} & & -2x_1^+ + 2x_1^- - x_2 + s_1 & = & 2, \\
& & 2x_1^+ - 2x_1^- + 5x_2 + s_2 & = & 6, \\
& & x_1^+, \quad x_1^-, \quad x_2, \quad s_1, \quad s_2 & \geq & 0.
\end{array}$$

Phase II
The *Phase I* does not have to be used in this case, the starting basis is obviously (s_1, s_2).

Calculating the reduced costs, we obtain $\tilde{\mathbf{c}}_N = (2, -2, 1)^{\mathrm{T}}$, meaning that x_1^- should enter the basis. From the minimum ratio test, we get that the outgoing variable is s_1. Updating the basis we now have (x_1^-, s_2) in the basis.

Calculating the reduced costs, we obtain $\tilde{\mathbf{c}}_N = (0, 0, 1)^{\mathrm{T}} \geq 0$, meaning that the current basis is optimal. The optimal solution is thus

$$\boldsymbol{x}^* = (x_1^+, x_1^-, x_2, s_1, s_2)^{\mathrm{T}} = (0, 1, 0, 0, 8)^{\mathrm{T}},$$

which in the original variables means $\boldsymbol{x}^* = (x_1, x_2)^{\mathrm{T}} = (-1, 0)^{\mathrm{T}}$ with optimal objective value $f^\star = -2$.

(1p) b) The reduced costs of for the optimal basis of the problem are $\tilde{\mathbf{c}}_N = (0, 0, 1)^{\mathrm{T}}$ meaning that the variable x_2 can enter the basis and the optimal objective value will remain the same $f^* = -2$. The alternative optimal solution is then $\tilde{\mathbf{x}}^* = (x_1, x_2)^{\mathrm{T}} = (-2, 2)^{\mathrm{T}}$. Hence, all points lying on the line segment connecting the extreme points $\mathbf{x}^*$ and $\tilde{\mathbf{x}}^*$ are optimal, i.e., $[x_1, -2x_1 - 2], \ \forall x_1 \in [-2, -1]$ is the optimal solution.

(3p)

Question 2

(KKT conditions) The objective function is convex, as can be seen by noting that both terms are compositions of a convex function (i.e., $\sum_i a_i x_i$) and an increasing convex function $-\log(.)$. Since the constraints are linear, the problem is a convex one, and the KKT conditions are thus sufficient for global optimality.

The KKT conditions become (with λ being the multiplier associated to the equality constraint, and μ_i being the multiplier associated to the i:th non-negativity constraint)

$$\begin{aligned}
\frac{a_i}{\sum_i a_i x_i} + \frac{1/a_i}{\sum_i x_i/a_i} + \mu_i &= \lambda, \quad i = 1, \dots, n, && (1)\\
\sum_i x_i &= 1, && (2)\\
x_i &\geq 0, \quad i = 1, \dots, n, && (3)\\
\mu_i x_i &= 0, \quad i = 1, \dots, n, && (4)\\
\mu_i &\geq 0, \quad i = 1, \dots, n. && (5)
\end{aligned}$$

Inserting $\mathbf{x} = (1/2, 0, \dots, 0, 1/2)^{\mathrm{T}}$ yields a feasible solution, and show the optimality of $\mathbf{x}$ we must produce a solution (λ, μ_i) to the system

$$\begin{aligned}
\frac{a_i}{a_1 + a_n} + \frac{a_i}{\frac{1}{a_1} + \frac{1}{a_n}} + \mu_i &= \lambda, \quad i = 1, \dots, n, && (6)\\
\mu_i &\geq 0, \quad i = 1, \dots, n && (7)\\
\mu_1 = \mu_n &= 0. && (8)
\end{aligned}$$

We see that using the first equality for $i = 1$ yields that we must have

$$\begin{aligned}
\lambda &= \frac{a_1}{a_1 + a_n} + \frac{1/a_1}{\frac{1}{a_1} + \frac{1}{a_n}}\\
&= \frac{a_1(1/a_1 + 1/a_n) + 1/a_1(a_1 + a_n)}{(a_1 + a_n)(1/a_1 + 1/a_n)} \qquad (9)\\
&= \frac{2 + a_1/a_n + a_n/a_1}{(a_1 + a_n)(1/a_1 + 1/a_n)}
\end{aligned}$$

And (due to the symmetry between a_1 and a_n in the above we see that the first equality is also satisfied for $i = n$ with this λ. It only remains to show that

$$\mu_i = \frac{2 + a_1/a_n + a_n/a_1}{(a_1 + a_n)(1/a_1 + 1/a_n)} - \frac{a_i}{a_1 + a_n} + \frac{a_i}{\frac{1}{a_1} + \frac{1}{a_n}} \geq 0 \qquad (10)$$

For all $i = 2, \dots, n - 1$. But writing the above with a common denominator we get

$$\frac{2 + a_1/a_n + a_n/a_1}{(a_1 + a_n)(1/a_1 + 1/a_n)} - \frac{a_i}{a_1 + a_n} + \frac{a_i}{\frac{1}{a_1} + \frac{1}{a_n}} = \frac{a_i/a_1 + a_i/a_n + a_1/a_i + a_n/a_i - 2 - a_1/a_n - a_n/a_1}{(a_1 + a_n)(1/a_1 + 1/a_n)} \geq 0 \tag{11}$$

Where the final follows since

$$a_i/a_1 \geq 1, \tag{12}$$
$$a_n/a_i \geq 1, \tag{13}$$
$$a_1/a_i \geq a_1/a_n, \tag{14}$$
$$a_i/a_n \geq a_1/a_n \tag{15}$$

Thus $(1/2, 0, \ldots, 0, 1/2)^{\mathrm{T}}$ is a KKT point, and hence optimal since the problem is convex.

Question 3

(problem decomposition)

(2p) a) The Lagrangian subproblem separates into $|\mathcal{I}|$ independent subproblems of the form

$$\underset{x_i \in X_i}{\text{minimize}}\ f_i(\boldsymbol{x}_i) + \boldsymbol{\mu}^{\mathrm{T}} \boldsymbol{x}_i;$$

the value of the Lagrangian dual function $q(\boldsymbol{\mu})$ is the sum of these $|\mathcal{I}|$ optimal values minus $\boldsymbol{\mu}^{\mathrm{T}} \boldsymbol{u}$. Any such value is a lower bound on the optimal valuem by the Weak Duality Theorem 6.5.

(1p) b) In this case $f_i(x_i) = c_i x_i + \frac{q_i}{2} x_i^2$, where $q_i \geq 0$ for all i, hence the Lagrangian term for index i has the form $c_i x_i + \frac{q_i}{2} x_i^2 + \mu_i x_i$. Its minimum over the closed interval X_i is easily found by comparing objective values at the two boundary points and potentially feasible stationary points.

(3p) ## Question 4

(Frank-Wolfe algorithm)

Figure 1 shows the feasible set of the problem (i.e., the polyhedron with thick black boundary lines) and some contours of the objective function. The optimal solution is denoted by $x^\star$ (i.e., the red dot in the figure). $x^{(k)}$ for $k = 0, 1, 2$ denotes iterates visited by the Frank-Wolfe algorithm.

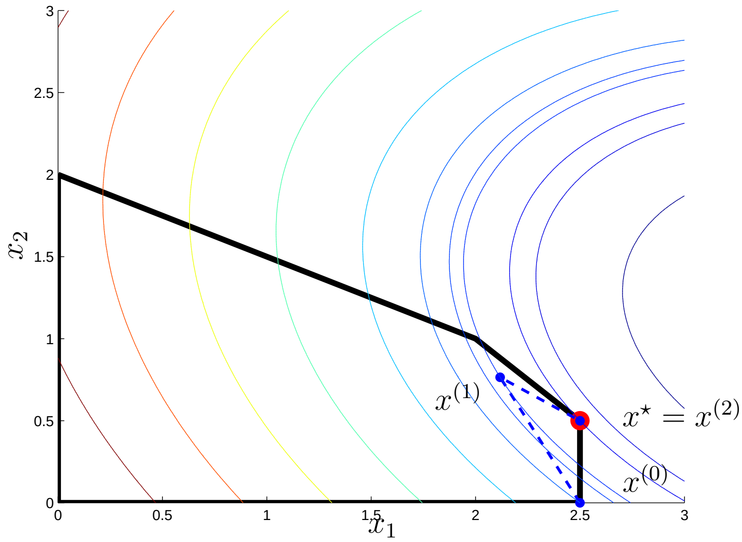


Figure 1: Illustration of the Frank-Wolfe algorithm. The feasible set is a polyhedron with boundary denoted by the thick black lines. Some contours of the objective function are shown. The optimal solution $\boldsymbol{x}^\star = (2.5, 0.5)$. The dotted lines show the Frank-Wolfe iterations, with $\boldsymbol{x}^k, k = 0, 1, 2$ denoting the iterates.

The details of the algorithm steps are as follows. Let X denote the feasible set. Let $f(x_1, x_2)$ denote the objective function. For any given iterate $\boldsymbol{x}^k = (x_1^k, x_2^k)$. The objective function gradient vector is

$$\nabla f(x_1^k, x_2^k) = \begin{bmatrix} 12 & 4 \\ 4 & 18 \end{bmatrix} \begin{bmatrix} x_1^k \\ x_2^k \end{bmatrix} - \begin{bmatrix} 52 \\ 34 \end{bmatrix}.$$

The search direction problem is

$$\underset{x \in X}{\text{minimize}} \quad \nabla f(x_1^k, x_2^k)^{\mathrm{T}} \boldsymbol{x}. \tag{1}$$

If $\min_{x \in X} \nabla f(x_1^k, x_2^k)^{\mathrm{T}} \boldsymbol{x} \geq \nabla f(x_1^k, x_2^k)^{\mathrm{T}} x^k$, then by the optimality conditions (for minimizing a convex function over a convex feasible set) $\boldsymbol{x}^k$ is optimal. Otherwise, let $\boldsymbol{y}^k$ denote an optimal solution to the search direction problem. Then the exact minimization line search problem can be expressed into

$$\underset{\alpha \in [0,1]}{\text{minimize}} \quad f(\alpha \boldsymbol{x}^k + (1-\alpha)\boldsymbol{y}^k) \iff \underset{\alpha \in [0,1]}{\text{minimize}} \quad g\alpha^2 + h\alpha,$$

where

$$\begin{aligned} g &= \left(\boldsymbol{x}^k - \boldsymbol{y}^k\right)^{\mathrm{T}} \begin{bmatrix} 6 & 2 \\ 2 & 9 \end{bmatrix} \left(\boldsymbol{x}^k - \boldsymbol{y}^k\right) \\ h &= \left(\boldsymbol{x}^k - \boldsymbol{y}^k\right)^{\mathrm{T}} \left(\begin{bmatrix} 12 & 4 \\ 4 & 18 \end{bmatrix} \boldsymbol{y}^k - \begin{bmatrix} 52 \\ 34 \end{bmatrix} \right). \end{aligned} \tag{2}$$

The minimizing value of α, denoted by α^k, can be found using the optimality condition to be

$$\alpha^k = \begin{cases} 0 & \text{if } \quad -\frac{h}{2g} < 0 \\ -\frac{h}{2g} & \text{if } \quad 0 \leq -\frac{h}{2g} \leq 1 \\ 1 & \text{if } \quad -\frac{h}{2g} > 1 \end{cases}. \tag{3}$$

The iterate update formula is

$$\boldsymbol{x}^{k+1} = \alpha^k \boldsymbol{x}^k + (1 - \alpha^k) \boldsymbol{y}^k. \tag{4}$$

Now we begin applying the Frank-Wolfe algorithm. At the first iteration with $\boldsymbol{x}^0 = (2.5, 0)^{\mathrm{T}}$, the objective function gradient is

$$\nabla f(x_1^0, x_2^0) = \begin{bmatrix} 12 & 4 \\ 4 & 18 \end{bmatrix} \begin{bmatrix} x_1^0 \\ x_2^0 \end{bmatrix} - \begin{bmatrix} 52 \\ 34 \end{bmatrix} = \begin{bmatrix} 12 & 4 \\ 4 & 18 \end{bmatrix} \begin{bmatrix} 2.5 \\ 0 \end{bmatrix} - \begin{bmatrix} 52 \\ 34 \end{bmatrix} = \begin{bmatrix} -22 \\ -24 \end{bmatrix}.$$

To solve the search direction problem in (1), it is sufficient to restrict the feasible set to the set of all extreme points. That is,

$$\underset{x \in V}{\text{minimize}} \, \nabla f(x_1^0, x_2^0)^{\mathrm{T}} \boldsymbol{x}, \tag{5}$$

where V is the set of all extreme points defined as

$$V = \Big\{ (0,0), (0,2), (2,1), (2.5, 0.5), (2.5, 0) \Big\}.$$

This amounts to finding the minimum among five numbers: 0, -48, -68, -67, -55. The result is that $\boldsymbol{y}^0 = (2, 1)$. Applying the formula in (2) yields

$$\begin{aligned} g &= \left(\begin{bmatrix} 2.5 \\ 0 \end{bmatrix} - \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right)^{\mathrm{T}} \begin{bmatrix} 6 & 2 \\ 2 & 9 \end{bmatrix} \left(\begin{bmatrix} 2.5 \\ 0 \end{bmatrix} - \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right) &= 8.5 \\ h &= \left(\begin{bmatrix} 2.5 \\ 0 \end{bmatrix} - \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right)^{\mathrm{T}} \left(\begin{bmatrix} 12 & 4 \\ 4 & 18 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} - \begin{bmatrix} 52 \\ 34 \end{bmatrix} \right) &= -4 \end{aligned}$$

According to (3), $\alpha^0 = \frac{4}{17}$. Hence, by (4)

$$\boldsymbol{x}^1 = \frac{4}{17}(\frac{5}{2}, 0) + (1 - \frac{4}{17})(2, 1) = (\frac{36}{17}, \frac{13}{17}) \approx (2.12, 0.76).$$

This is shown in Figure 1.

At the next iteration with $x^1 = (\frac{36}{17}, \frac{13}{17})$, we have

$$\nabla f(x_1^1, x_2^1) = \begin{bmatrix} 12 & 4 \\ 4 & 18 \end{bmatrix} \begin{bmatrix} x_1^1 \\ x_2^1 \end{bmatrix} - \begin{bmatrix} 52 \\ 34 \end{bmatrix} = \frac{1}{17} \begin{bmatrix} -400 \\ -200 \end{bmatrix} \approx \begin{bmatrix} -23.53 \\ -11.76 \end{bmatrix}.$$

Solving (5) amounts to finding the minimum of 0, -4, -10, -11, -10. This leads to $y^1 = (2.5, 0.5)$. Applying (2) leads to

$$\begin{aligned} g &= \tfrac{1275}{1156} \approx 1.10 \\ h &= \tfrac{125}{34} \approx 3.68. \end{aligned}$$

Thus, according to (3) $\alpha^1 = 0$, and from (4) $x^2 = y^1 = (2.5, 0.5)^{\mathrm{T}}$ as shown in Figure 1.

At the final iteration with $x^2 = (2.5, 0.5)^{\mathrm{T}}$, we have

$$\nabla f(x_1^2, x_2^2) = \begin{bmatrix} -20 \\ -15 \end{bmatrix}.$$

Solving (5) leads to $\boldsymbol{y}^2 = \boldsymbol{x}^2 = (2.5, 0.5)^{\mathrm{T}}$. Thus, it holds that

$$\min_{x \in X} \nabla f(x_1^2, x_2^2)^{\mathrm{T}} \boldsymbol{x} \geq \nabla f(x_1^2, x_2^2)^{\mathrm{T}} x^2.$$

By the optimality conditions, $\boldsymbol{x}^2 = (2.5, 0.5)^{\mathrm{T}}$ is the optimal solution to our problem.

Question 5

(true or false)

(1p) a) False. It is not necessarily so that *any* such rounding, up or down, of individual variables, lead to a feasible solution.

(1p) b) False. In the non-convex case there may be “better points” outside of the feasible set.

(1p) c) True. This is Proposition 4.26.

(3p) Question 6

(the Relaxation Theorem)

This is Theorem 6.1.

(3p) Question 7

(modelling)

Let (x_i, y_i) be the coordinates of the center poitn of circle $i = 1, \ldots, n$, and let r_i be the radius of cirlce $i = 1, \ldots, n$. Then the optimization problem can be formulated as the following:

$$\begin{aligned}
\text{maximize} \quad & \sum_{i=1}^{n} \pi r_i^2, \\
\text{subject to} \quad & \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \geq r_i + r_j,\ i \neq j, \\
& r_i \leq x_i \leq L - r_i, \quad i = 1, \ldots, n, \\
& r_i \leq y_i \leq L - r_i, \quad i = 1, \ldots, n, \\
& r_i \geq 0, \quad i = 1, \ldots, n.
\end{aligned}$$
