# Optimization Algorithms Assignment 002

Nguyen Quan Ba Hong*

Students at Faculty of Math and Computer Science,

Ho Chi Minh University of Science, Vietnam

email. `nguyenquanbahong@gmail.com`

blog. `www.nguyenquanbahong.com` †

June 21, 2018

**Abstract**

This assignment aims at solving some selected problems for the final exam of the course *Optimization Algorithms*.

*Student ID: 1411103

# Contents

# 1 Problems

**Problem 1.1.** *Let $f : \mathbb{R}^2 \to \mathbb{R}$ defined by*

$$f(x, y) = \frac{1}{2}\left(x^2 + 5y^2\right) + x + y. \tag{1.1}$$

1. *Prove that $f$ is convex.*
2. *Find minimizer $(x^\star, y^\star)$ of $f$ in $\mathbb{R}^2$.*
3. *By the steepest descent method with exact linesearches, start at the point $(x_0, y_0) = (0, 0)$ and present the first iteration.*
4. *By the steepest descent method with exact linesearches, starting at the point $(x_0, y_0) = (0, 0)$, we obtain a sequence $\{(x_n, y_n)\}_{n\geq 0}$. Find the smallest $n$ such that*

$$f(x_n, y_n) - f(x^\star, y^\star) \leq 10^{-2}. \tag{1.2}$$

Solution.

1. The gradient and the Hessian matrix of $f$ are given by

$$\nabla f(x, y) = \begin{bmatrix} x+1 \\ 5y+1 \end{bmatrix}, \quad \nabla^2 f(x, y) = \begin{bmatrix} 1 & 0 \\ 0 & 5 \end{bmatrix}, \tag{1.3}$$

for all $(x, y) \in \mathbb{R}^2$. The eigenvalues of $\nabla^2 f(x, y)$ are $\lambda_1 = 1$ and $\lambda_2 = 5$. Hence $\nabla^2 f(x, y)$ is positive definite for all $(x, y) \in \mathbb{R}^2$ and thus $f$ is strictly convex.

2. Since $f$ is convex, $(x^\star, y^\star)$ is (global) minimizer of $f$ if and only if $\nabla f(x^\star, y^\star) = 0$. Solving the equation $\nabla f(x, y) = 0$ yields that $(x^\star, y^\star) = \left(-1, -\frac{1}{5}\right)$ is the unique minimizer of $f$ in $\mathbb{R}^2$.

3. We choose the starting descent direction as

$$d_0 = -\nabla f(x_0, y_0) = -\nabla f(0, 0) = \begin{bmatrix} -1 \\ -1 \end{bmatrix}. \tag{1.4}$$

We will find a step size $t_0 > 0$ such that $f((x_0, y_0) + t_0 d_0)$ attains its minimizer, i.e., $t_0 = \arg\min_{t>0} f((x_0, y_0) + t d_0)$. This is equivalent to $t_0 = \arg\min_{t>0}\left(3t^2 - 2t\right)$, which gives us $t_0 = \frac{1}{3}$. Thus, we obtain, in the first iteration of the steepest descent method with exact linesearches,

$$(x_1, y_1) = (x_0, y_0) + t_0 d_0 = \left(-\frac{1}{3}, -\frac{1}{3}\right). \tag{1.5}$$

4. Similarly, for any $n \in \mathbb{Z}_+$, the $n$th descent direction is given by

$$d_n = -\nabla f(x_n, y_n) = -\begin{bmatrix} x_n + 1 \\ 5y_n + 1 \end{bmatrix}. \tag{1.6}$$

It will be proved, after choosing a sequence $t_n$'s, that $(x_n, y_n) \neq \left(-1, -\frac{1}{5}\right)$ for all $n \in \mathbb{N}$. We also find a $n$th step size $t_n > 0$ as

$$t_n = \arg\min_{t>0} f\left((x_n, y_n) + td_n\right) \tag{1.7}$$

$$= \arg\min_{t>0} f\left(x_n - t\left(x_n + 1\right), y_n - t\left(5y_n + 1\right)\right) \tag{1.8}$$

$$= \arg\min_{t>0} g_n\left(t\right), \tag{1.9}$$

where

$$g_n\left(t\right) = \frac{1}{2}\left(\left(x_n + 1\right)^2 + 5\left(5y_n + 1\right)^2\right)t^2 \tag{1.10}$$

$$-\left(\left(x_n + 1\right)^2 + \left(5y_n + 1\right)^2\right)t + \frac{1}{2}\left(x_n^2 + 5y_n^2\right) + x_n + y_n. \tag{1.11}$$

Consider the behavior of this quadratic function with respect to the variable $t$, it is easy to verify that

$$t_n = \frac{\left(x_n + 1\right)^2 + \left(5y_n + 1\right)^2}{\left(x_n + 1\right)^2 + 5\left(5y_n + 1\right)^2}. \tag{1.12}$$

Hence, the iterations in the steepest descent method with exact line-searches have the following form

$$\left(x_{n+1}, y_{n+1}\right) = \left(x_n, y_n\right) - \frac{\left(x_n + 1\right)^2 + \left(5y_n + 1\right)^2}{\left(x_n + 1\right)^2 + 5\left(5y_n + 1\right)^2}\left(x_n + 1, 5y_n + 1\right), \tag{1.13}$$

for all $n \in \mathbb{N}$, or equivalently,

$$x_{n+1} = x_n - \frac{\left(x_n + 1\right)^2 + \left(5y_n + 1\right)^2}{\left(x_n + 1\right)^2 + 5\left(5y_n + 1\right)^2}\left(x_n + 1\right), \tag{1.14}$$

$$y_{n+1} = y_n - \frac{\left(x_n + 1\right)^2 + \left(5y_n + 1\right)^2}{\left(x_n + 1\right)^2 + 5\left(5y_n + 1\right)^2}\left(5y_n + 1\right). \tag{1.15}$$

Define $a_n := x_n + 1$, $b_n := 5y_n + 1$ for all $n \in \mathbb{N}$, then $f$ can be rewritten as

$$f\left(x_n, y_n\right) = \frac{5a_n^2 + b_n^2}{10} - \frac{3}{5}, \text{ for all } n \in \mathbb{N}, \tag{1.16}$$

and (1.14)-(1.15) becomes

$$a_{n+1} = \frac{4a_n b_n^2}{a_n^2 + 5b_n^2}, \tag{1.17}$$

$$b_{n+1} = -\frac{4a_n^2 b_n}{a_n^2 + 5b_n^2}. \tag{1.18}$$

Since $(a_0, b_0) = (1, 1)$, (1.17)-(1.18) implies that $(a_n, b_n) \neq (0, 0)$ for all $n \in \mathbb{N}$, i.e., $(x_n, y_n) \neq \left(-1, -\frac{1}{5}\right)$ for all $n \in \mathbb{N}$ as stated above. Hence, (1.12) make a sense and $t_n > 0$ for all $n \in \mathbb{N}$.

Run the following Matlab script

```
f = @(x,y) (x.^2 + 5*y.^2)/2 + x + y;
d = @(x,y) -[x + 1; 5*y + 1];
t = @(x,y) ((x+1).^2 + (5*y+1).^2)/((x+1).^2 + 5*(5*y+1).^2);
X = [0;0]; % X_n := (x_n,y_n)
n = 0;
while (abs(f(X(1),X(2)) - f(-1,-1/5)) > 1e-2)
    Xtemp = X;
    X = X + t(X(1),X(2))*d(X(1),X(2));
    n = n + 1;
end
n
```

yields that $n = 6$ is the smallest positive integer such that (1.2) holds. $\square$

**Problem 1.2.** *Let $f : \mathbb{R}^3 \to \mathbb{R}$ be a mapping defined by $f(x) = \frac{1}{2}x^T Ax - c^T x$, where $A = diag(1, 5, 25)$ and $c = [-1, -1, -1]^T$.*

1. *Prove that $f$ is convex.*
2. *Find the minimizer $x^\star$ of $f$ in $\mathbb{R}^3$.*
3. *By the steepest descent method with exact linesearches, starting at the point $x_0 = (0, 0, 0)$, present the first iteration.*

Solution.

1. The eigenvalues of $\nabla^2 f(x)$ are $\lambda_1 = 1$, $\lambda_2 = 5$, and $\lambda_3 = 25$. Hence $\nabla^2 f(x)$ is positive definite for all $x \in \mathbb{R}^3$ and thus $f$ is strictly convex.
2. Since $f$ is strictly convex, $x^\star$ is the unique minimizer of $f$ if and only if $\nabla f(x^\star) = 0$. Solving the equation $\nabla f(x) = 0$ yields that $x^\star = \left(-1, -\frac{1}{5}, -\frac{1}{25}\right)$ is the unique minimizer of $f$ in $\mathbb{R}^3$.
3. We choose the starting descent direction as $d_0 = -\nabla f(x_0) = [-1, -1, -1]^T$. The starting step size $t_0$ is chosen as
$$t_0 := \arg\min_{t>0} f(x_0 + td_0) = \arg\min_{t>0} \left(\frac{31}{2}t^2 - 3t\right) = \frac{3}{31}. \tag{1.19}$$
Thus, we obtain, in the first iteration of the steepest descent method with exact linesearches, $x_1 = x_0 + t_0 d_0 = \left[-\frac{3}{31}, -\frac{3}{31}, -\frac{3}{31}\right]^T$. $\square$

**Problem 1.3.** *Let $f, g : \mathbb{R}^2 \to \mathbb{R}$ be two mappings defined by*
$$f(x, y) = (x - y + 1)^2 + (2x - y)^2, \tag{1.20}$$
$$g(x, y) = (x + y)^2 + (y - 2x + 1)^2. \tag{1.21}$$

1. *Prove that $f$, $g$ are convex.*
2. *Find the minima of $f$ and $g$ in $\mathbb{R}^2$.*
3. *By the steepest descent method with exact linesearches, starting at the point $(x_0, y_0) = (0, 0)$, which the values of $f$ or $g$ will converge to the optimal values faster?*

SOLUTION.

1. The gradients and the Hessian matrices of $f$ and $g$ are given by

$$\nabla f(x,y) = \begin{bmatrix} 10x - 6y + 2 \\ 4y - 6x - 2 \end{bmatrix}, \quad \nabla^2 f(x,y) = \begin{bmatrix} 10 & -6 \\ -6 & 4 \end{bmatrix}, \tag{1.22}$$

$$\nabla g(x,y) = \begin{bmatrix} 10x - 2y - 4 \\ 4y - 2x + 2 \end{bmatrix}, \quad \nabla^2 g(x,y) = \begin{bmatrix} 10 & -2 \\ -2 & 4 \end{bmatrix}, \tag{1.23}$$

for all $(x,y) \in \mathbb{R}^2$, respectively. The eigenvalues of $f$ and $g$ are $\lambda_{f,1} = 7 - 3\sqrt{5}, \lambda_{f,2} = 7 + 3\sqrt{5}$ and $\lambda_{g,1} = 7 - \sqrt{13}, \lambda_{g,2} = 7 + \sqrt{13}$, respectively. Hence, both $\nabla^2 f(x,y)$ and $\nabla^2 g(x,y)$ are positive definite for $(x,y) \in \mathbb{R}^2$ and thus $f$ and $g$ are strictly convex.

2. Since $f$ and $g$ are strictly convex, $x_{f,\star}, x_{g,\star}$ are their unique minima if and only if $\nabla f(x_{f,*}) = 0$ and $\nabla g(x_{g,*}) = 0$, respectively. Solving the equations $\nabla f(x,y) = 0$, $\nabla g(x,y) = 0$ yields that $x_{f,\star} = (1,2)$ and $x_{g,\star} = \left(\frac{1}{3}, -\frac{1}{3}\right)$ are the minima of $f$ and $g$ in $\mathbb{R}^2$, respectively.

3. Since $f$, $g$ are of class $C^2\left(\mathbb{R}^2\right)$ and $\nabla^2 f(x_{f,*})$, $\nabla^2 g(x_{g,*})$ are positive definite, we suppose that the sequences $\{x_{f,n}\}_{n\ge 0}$, $\{x_{g,n}\}_{n\ge 0}$ generated by the steepest descent method with exact linesearches converge to $x_{f,\star}$ and $x_{g,\star}$, respectively. Applying Theorem 3.2.1, [2], p. 31 to $f$ and $g$ yields

$$|f(x_{f,n+1}) - f(x_{f,*})| \le \left(\frac{\lambda_{f,2} - \lambda_{f,1}}{\lambda_{f,2} + \lambda_{f,1}}\right)^2 |f(x_{f,n}) - f(x_{f,*})| \tag{1.24}$$

$$= \frac{45}{49} |f(x_{f,n}) - f(x_{f,*})|, \tag{1.25}$$

$$|g(x_{g,n+1}) - g(x_{g,*})| \le \left(\frac{\lambda_{g,2} - \lambda_{g,1}}{\lambda_{g,2} + \lambda_{g,1}}\right)^2 |g(x_{g,n}) - g(x_{g,*})| \tag{1.26}$$

$$= \frac{13}{49} |g(x_{g,n}) - f(x_{g,*})|, \tag{1.27}$$

i.e., the rates of convergence of the gradient method with exact linesearches for $f$ and $g$ are $\frac{45}{49}$ and $\frac{13}{49}$, respectively. Theoretically, we predict that the values of $g$ will converge to its optimal value faster than those of $f$. □

**Problem 1.4.** *Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a mapping defined by*

$$f(x,y) = \frac{1}{2}x^2 + \frac{a}{2}y^2, \tag{1.28}$$

*where $a \ge 1$. By the steepest descent method with exact linesearches, starting at the point $(x_0, y_0) = (a, 1)$, prove by induction that the $n$th iteration is*

$$(x_n, y_n) = \left(\frac{a-1}{a+1}\right)^n (a, (-1)^n). \tag{1.29}$$

PROOF. The gradient and the Hessian matrix of $f$ are given by

$$\nabla f(x,y) = \begin{bmatrix} x \\ ay \end{bmatrix}, \quad \nabla^2 f(x,y) = \begin{bmatrix} 1 & 0 \\ 0 & a \end{bmatrix}, \tag{1.30}$$

for all $(x,y) \in \mathbb{R}^2$.

For arbitrary $n \in \mathbb{N}$, the $n$th descent direction is chosen as

$$d_n = -\nabla f(x_n, y_n) = \begin{bmatrix} -x_n \\ -ay_n \end{bmatrix}, \tag{1.31}$$

and the step size $t_n$ is chosen as

$$\begin{aligned} t_n &= \arg\min_{t>0} f((x_n, y_n) + td_n) && (1.32) \\ &= \arg\min_{t>0} \left( \frac{1}{2}\left(x_n^2 + a^3 y_n^2\right) t^2 - \left(x_n^2 + a^2 y_n^2\right) t + \frac{1}{2}\left(x_n^2 + a y_n^2\right) \right) && (1.33) \\ &= \frac{x_n^2 + a^2 y_n^2}{x_n^2 + a^3 y_n^2}, && (1.34) \end{aligned}$$

where we will prove that $(x_n, y_n) \neq (0,0)$ for all $n \in \mathbb{N}$. With the chosen $t_n$'s, the iterations of the steepest descent method with exact linesearches are

$$(x_{n+1}, y_{n+1}) = (x_n, y_n) + \frac{x_n^2 + a^2 y_n^2}{x_n^2 + a^3 y_n^2}(-x_n, -ay_n), \text{ for all } n \in \mathbb{N}, \tag{1.35}$$

which is equivalent to

$$x_{n+1} = \frac{a^2(a-1)x_n y_n^2}{x_n^2 + a^3 y_n^2}, \tag{1.36}$$

$$y_{n+1} = \frac{(1-a)x_n^2 y_n}{x_n^2 + a^3 y_n^2}, \tag{1.37}$$

for all $n \in \mathbb{N}$. Combining the fact that $(x_0, y_0) = (a, 1) \neq (0,0)$ with (1.36)-(1.37) yields that $(x_n, y_n) \neq (0,0)$ for all $n \in \mathbb{N}$. Thus $t_n$ defined by (1.34) makes sense and $t_n > 0$ for all $n \in \mathbb{N}$.

We now prove (1.29) by induction. The case $n = 0$ is the given starting point. Assume that (1.29) holds for some $n \geq 0$, we have

$$\begin{aligned} (x_{n+1}, y_{n+1}) &= (x_n, y_n) + \frac{x_n^2 + a^2 y_n^2}{x_n^2 + a^3 y_n^2}(-x_n, -ay_n) && (1.38) \\ &= \left(\frac{a-1}{a+1}\right)^n \left[(a, (-1)^n) + \frac{2a}{1+a}\left(-a, a(-1)^{n+1}\right)\right] && (1.39) \\ &= \left(\frac{a-1}{a+1}\right)^n \left(a - \frac{2a}{1+a}, \frac{2a}{1+a}(-1)^{n+1} - (-1)^{n+1}\right) && (1.40) \\ &= \left(\frac{a-1}{a+1}\right)^{n+1} \left(a, (-1)^{n+1}\right). && (1.41) \end{aligned}$$

By the principle of mathematical induction, we deduce that (1.29) holds for all $n \in \mathbb{N}$. □

**Problem 1.5.** *Let $f : \mathbb{R}^2 \to \mathbb{R}$ defined by*

$$f(x,y) = \frac{1}{2}(y - 2x)^2 + y^4. \tag{1.42}$$

*Determine the Newton direction of $f$ at the point $x_0 = (1, 2)$.*

SOLUTION. The gradient and the Hessian matrix of $f$ are given by

$$\nabla f(x, y) = \begin{bmatrix} 4x - 2y \\ 4y^3 + y - 2x \end{bmatrix}, \quad \nabla^2 f(x, y) = \begin{bmatrix} 4 & -2 \\ -2 & 12y^2 + 1 \end{bmatrix}, \tag{1.43}$$

for all $(x, y) \in \mathbb{R}^2$.

The Newton direction of $f$ at the point $(x, y) \in \mathbb{R}^2$ can be obtained by solving the equation

$$\nabla^2 f(x, y)\, d(x, y) = -\nabla f(x, y). \tag{1.44}$$

Solving (1.44) yields that $d(x, y) = \left(-x + \frac{y}{3}, -\frac{y}{3}\right)$ is the Newton direction of $f$ at the point $(x, y)$. In particular, $d(1, 2) = \left(-\frac{1}{3}, \frac{2}{3}\right)$. □

**Problem 1.6.** *Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a mapping defined by*

$$f(x, y) = (x - y)^2 + (2x + y - 3)^2. \tag{1.45}$$

1. *Prove that $f$ is convex.*
2. *Find the minimizer $(x^\star, y^\star)$ of $f$ in $\mathbb{R}^2$.*
3. *By the pure Newton method, starting at the point $(x_0, y_0) = (0, 0)$, present the first iteration. Comment about the point $(x_1, y_1)$.*

SOLUTION.

1. The gradient and the Hessian matrix of $f$ are given by

$$\nabla f(x, y) = \begin{bmatrix} 10x + 2y - 12 \\ 2x + 4y - 6 \end{bmatrix}, \quad \nabla^2 f(x, y) = \begin{bmatrix} 10 & 2 \\ 2 & 4 \end{bmatrix}, \tag{1.46}$$

   for all $(x, y) \in \mathbb{R}^2$.

   The eigenvalues of $\nabla^2 f(x, y)$ are $\lambda_1 = 7 - \sqrt{13}$ ,$\lambda_2 = 7 + \sqrt{13}$. Hence $\nabla^2 f(x, y)$ is positive definite for all $(x, y) \in \mathbb{R}^2$ and thus $f$ is strictly convex.

2. Since $f$ is strictly convex, $(x^\star, y^\star)$ is the unique minimizer of $f$ in $\mathbb{R}^2$ if and only if $\nabla f(x^\star, y^\star) = 0$. Solving the equation $\nabla f(x, y) = 0$ yields that $(x^\star, y^\star) = (1, 1)$ is the unique (global) minimizer of $f$ in $\mathbb{R}^2$.

3. The Newton direction of $f$ at the point $(x, y) \in \mathbb{R}^2$ can be obtained by solving the equation

$$\nabla^2 f(x, y)\, d(x, y) = -\nabla f(x, y). \tag{1.47}$$

   Solving (1.47) yields that $d(x, y) = (1 - x, 1 - y)$ is the Newton direction of $f$ at the point $(x, y)$. In particular, $d(0, 0) = (1, 1)$. Then the first iteration of the pure Newton method is

$$(x_1, y_1) = (x_0, y_0) + d(x_0, y_0) = (1, 1) = (x^\star, y^\star). \tag{1.48}$$

   Thus, we need only one iteration of the pure Newton method to obtain the minimizer. □

**Problem 1.7.** *By the pure Newton method, select the starting point and build an iterated sequence* $\{x_n\}_{n\in\mathbb{N}}$ *to approximate the minimizer of the following minimization problem*

$$\text{Min } f(x) \text{ s.t. } x \in \mathbb{R} \text{ with } f(x) = \frac{1}{4}x^4. \tag{1.49}$$

*Does the sequence $x_n$'s converge quadratically to the minimizer?*

SOLUTION. The first and second derivatives of $f$ are $f'(x) = x^3$, $f''(x) = 3x^2 \geq 0$ for all $x \in \mathbb{R}$. Hence, $f$ is convex, and its unique minimizer is $x^\star = 0$ obviously. The Newton direction at an arbitrary point $x \in \mathbb{R}$ can be obtained by solving the equation $f''(x)\, d(x) = -f'(x)$. Solving the last equation gives us $d(x) = -\frac{x}{3}$. Starting at a point $x_0 \in \mathbb{R}$, the iterations of the pure Newton method are

$$x_{n+1} = x_n + d(x_n) = x_n - \frac{x_n}{3} = \frac{2}{3}x_n, \text{ for all } n \in \mathbb{N}. \tag{1.50}$$

We consider the following cases depending on the starting value $x_0$.

- *Case* $x_0 = 0$. In this case, (1.50) gives us $x_n = 0$ for all $n \in \mathbb{N}$. Hence, the sequence $x_n$'s converge quadratically to $x^\star$ in this case.
- *Case* $x_0 \neq 0$. In this case, (1.50) gives us the general formula of $x_n$ as $x_n = \left(\frac{2}{3}\right)^n x_0$ for all $n \in \mathbb{N}$. Then

$$\frac{|x_{n+1} - x^\star|}{|x_n - x^\star|^2} = \frac{\left(\frac{2}{3}\right)^{n+1} |x_0|}{\left(\frac{2}{3}\right)^{2n} x_0^2} = \left(\frac{2}{3}\right)^{1-n} |x_0|^{-1} \to +\infty \tag{1.51}$$

  as $n \to +\infty$. As a consequence, $x_n$'s does not converge quadratically to the minimizer $x^\star$ of $f$ in this case. However, $x_n$'s converges linearly to $x^\star$ since $|x_{n+1} - x^\star| = \frac{2}{3}|x_n - x^\star|$ for all $n \in \mathbb{N}$. □

**Problem 1.8.** *Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a mapping defined by*

$$f(x,y) = 2x^2 + y^2 - 2xy + 2x^3 + x^4. \tag{1.52}$$

1. *Find all the stationary points*[1] *of $f$ in $\mathbb{R}^2$.*
2. *By the pure Newton method, starting at the point $(x_0, y_0) = (-1, 0)$, present the first iteration to obtain $(x_1, y_1)$.*

SOLUTION.

1. The gradient and the Hessian matrix of $f$ are given by

$$\nabla f(x,y) = \begin{bmatrix} 4x^3 + 6x^2 + 4x - 2y \\ 2y - 2x \end{bmatrix}, \tag{1.53}$$

$$\nabla^2 f(x,y) = \begin{bmatrix} 12x^2 + 12x + 4 & -2 \\ -2 & 2 \end{bmatrix}, \tag{1.54}$$

   for all $(x,y) \in \mathbb{R}^2$. Solving the equation $\nabla f(x,y) = 0$ yields that $(-1,-1)$, $\left(-\frac{1}{2}, -\frac{1}{2}\right)$, and $(0,0)$ are the only stationary points of $f$ in $\mathbb{R}^2$.

[1] "Stationary points", see [2], or "critical points", see [1].

2. Given $(x, y) \in \mathbb{R}^2$ arbitrarily, the Newton direction of $f$ at the point $(x, y)$ can be obtained by solving the equation $\nabla^2 f(x, y)\, d(x, y) = -\nabla f(x, y)$. Solving the last equation yields that

$$d(x, y) = -\frac{1}{6x^2 + 6x + 1} \begin{bmatrix} x\left(2x^2 + 3x + 1\right) \\ y + 6xy + 6x^2 y - 3x^2 - 4x^3 \end{bmatrix}. \tag{1.55}$$

Starting at the point $(x_0, y_0) = (-1, 0)$, the first iteration of the pure Newton method is

$$(x_1, y_1) = (x_0, y_0) + d(x_0, y_0) = (-1, -1), \tag{1.56}$$

which is one of the stationary points of $f$. $\square$

**Problem 1.9.** *Consider the following problem*

$$(P) \text{ Min } x_1^2 + x_2^2 \text{ s.t. } 2x_1 - x_2 - 1 \le 0. \tag{1.57}$$

1. *Prove that $(P)$ is a convex problem and the Slater condition is satisfied.*
2. *Use the KKT conditions (Karush-Kuhn-Tucker conditions), find optimal solution $x^\star$ of $(P)$.*
3. *Establish a barrier approximation problem for $(P)$, find the optimal value $x^\star(t)$ of that barrier approximation problem. Prove that $x^\star(t) \to x^\star$ as $t \to 0^+$.*

SOLUTION.

1. Set $f(x_1, x_2) = x_1^2 + x_2^2$ and $g(x_1, x_2) = 2x_1 - x_2 - 1$ for all $(x_1, x_2) \in \mathbb{R}^2$, the gradients and the Hessian matrices of $f$ and $g$ are given by

$$\nabla f(x_1, x_2) = \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix}, \quad \nabla^2 f(x_1, x_2) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, \tag{1.58}$$

$$\nabla g(x_1, x_2) = \begin{bmatrix} 2 \\ -1 \end{bmatrix}, \quad \nabla^2 g(x_1, x_2) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \tag{1.59}$$

for all $(x_1, x_2) \in \mathbb{R}^2$, respectively. It is clear that $\nabla^2 f(x_1, x_2)$ and $\nabla^2 g(x_1, x_2)$ are positive definite and semi-positive definite matrices, respectively. Thus $f$ is strictly convex and $g$ is convex. Consequently, $(P)$ is a convex problem.

Since $g(0, 0) = -1 < 0$, the Slater condition is satisfied.

2. The feasible set of $(P)$ is given by

$$C := \left\{(x_1, x_2) \in \mathbb{R}^2; g(x_1, x_2) = 2x_1 - x_2 - 1 \le 0\right\}. \tag{1.60}$$

The constraint qualification hypothesis $T_C(x^\star) = L_C(x^\star)$ is guaranteed by (CQ1) or (CQ2) (since $g$ is convex and affine, see, e.g., [2], pp. 89-90). Combining this with the convexity of $(P)$, applying Theorem 8.2.1, [2], p. 89, and Theorem 8.2.2, [2], p. 91 yields that $x^\star$ is an optimal solution of

$(P)$ if and only if the vectors $(x^\star, \lambda^\star)$ satisfy the KKT conditions. The KKT conditions for $(P)$ are given by

$$(KKT) \begin{cases} \nabla f(x^\star) + \lambda^\star \nabla g(x^\star) = 0, \\ \lambda^\star g(x^\star) = 0, \\ \lambda^\star \geq 0, \\ g(x^\star) \leq 0, \end{cases} \tag{1.61}$$

which is equivalent to

$$2x_1^\star + 2\lambda^\star = 0, \tag{1.62}$$
$$2x_2^\star - \lambda^\star = 0, \tag{1.63}$$
$$\lambda^\star (2x_1^\star - x_2^\star - 1) = 0, \tag{1.64}$$
$$\lambda^\star \geq 0, \tag{1.65}$$
$$2x_1^\star - x_2^\star - 1 \leq 0. \tag{1.66}$$

Solving the first two equations in this system yields $(x_1^\star, x_2^\star) = \left(-\lambda^\star, \frac{\lambda^\star}{2}\right)$. Substituting these into the others gives us

$$\lambda^\star \left(-\frac{5\lambda^\star}{2} - 1\right) = 0, \tag{1.67}$$
$$\lambda^\star \geq 0, \tag{1.68}$$
$$\frac{5\lambda^\star}{2} + 1 \geq 0, \tag{1.69}$$

which has the unique root $\lambda = 0$. Then $x^\star = (x_1^\star, x_2^\star) = (0, 0)$ is the unique optimal solution of $(P)$.

3. The logarithmic barrier function of $(P)$ is

$$B(x_1, x_2, t) := f(x_1, x_2) - t \log(-g(x_1, x_2)) \tag{1.70}$$
$$= x_1^2 + x_2^2 - t \log(1 + x_2 - 2x_1), \tag{1.71}$$

for all $(x_1, x_2) \in C_B$ and $t > 0$, where $C_B$ is the feasible set of $B$ and is given by

$$C_B := \left\{(x_1, x_2) \in \mathbb{R}^2; g(x_1, x_2) = 2x_1 - x_2 - 1 < 0\right\}. \tag{1.72}$$

We now prove that $B$ is convex.

$\star$ *First proof of the convexity of $B$.* We use the following two well-known properties of convex function calculus[2]:

- *The sum of convex functions is a convex function.*
- *If $F$ is concave and $G$ is convex and non-increasing over a univariate domain, then $G \circ F$ is convex.*

Applying the former for $f$ and $h := -t\log(-g)$, and the later for $F := -g$ and $G = -t \log x$ yields that $B$ is convex for all $t > 0$. $\triangle$

[2]See, e.g., `https://en.wikipedia.org/wiki/Convex_function`.

⋆ *Second proof of the convexity of $B$.* The spatial gradient and the spatial Hessian matrix of $B$ are given by

$$\nabla_x B(x_1, x_2, t) = \begin{bmatrix} 2x_1 + \frac{2t}{x_2 - 2x_1 + 1} \\ 2x_2 - \frac{t}{x_2 - 2x_1 + 1} \end{bmatrix}, \tag{1.73}$$

$$\nabla_x^2 B(x_1, x_2, t) = \begin{bmatrix} 2 + \frac{4t}{(x_2 - 2x_1 + 1)^2} & -\frac{2t}{(x_2 - 2x_1 + 1)^2} \\ -\frac{2t}{(x_2 - 2x_1 + 1)^2} & 2 + \frac{t}{(x_2 - 2x_1 + 1)^2} \end{bmatrix}, \tag{1.74}$$

for all $(x_1, x_2) \in C_B$ and $t > 0$. The eigenvalues of $\nabla_x^2 B(x_1, x_2, t)$ are $\lambda_1 = 2$ and $\lambda_2 = 2 + \frac{5t}{(y - 2x + 1)^2}$. Hence, $\nabla_x^2 B(x_1, x_2, t)$ is positive definite for all $(x_1, x_2) \in C_B$ and thus $B$ is strictly convex for all $t > 0$. △

Since $B$ is strictly convex, $x^\star$ is the unique minimizer of $B$ in $C_B$ if and only if $g(x^\star) < 0$ and $\nabla_x B(x^\star, t) = 0$. The roots of the equation $\nabla_x B(x_1, x_2, t) = 0$ are $\left(\frac{1-\sqrt{10t+1}}{5}, \frac{-1+\sqrt{10t+1}}{10}\right)$ and $\left(\frac{1+\sqrt{10t+1}}{5}, -\frac{1+\sqrt{10t+1}}{10}\right)$. The former is taken and the later is omitted since

$$g\left(\frac{1-\sqrt{10t+1}}{5}, \frac{-1+\sqrt{10t+1}}{10}\right) = \frac{1-\sqrt{10t+1}}{2} - 1 < 0, \tag{1.75}$$

$$g\left(\frac{1+\sqrt{10t+1}}{5}, -\frac{1+\sqrt{10t+1}}{10}\right) = \frac{1+\sqrt{10t+1}}{2} - 1 > 0, \tag{1.76}$$

for all $t > 0$. Thus, $x^\star(t) = \left(\frac{1-\sqrt{10t+1}}{5}, \frac{-1+\sqrt{10t+1}}{10}\right)$ is the optimal solution of the barrier approximation problem for all $t > 0$. It is evident that $x^\star(t) \to x^\star$ as $t \to 0^+$. □

**Problem 1.10.** *Consider the following problem*

$$(P) \text{ Min } x_1 - x_2 \text{ s.t. } x_1^2 + x_2^2 \leq 1. \tag{1.77}$$

1. *Prove that $(P)$ is a convex problem, and the Slater condition is satisfied.*
2. *Use the KKT conditions, find the optimal solution $x^\star$ of $(P)$.*
3. *Establish the barrier approximation problem for $(P)$, find the optimal solution $x^\star(t)$ of that barrier approximation problem. Prove that $x^\star(t) \to x^\star$ as $t \to 0^+$.*

Solution.

1. Set $f(x_1, x_2) = x_1 - x_2, g(x_1, x_2) = x_1^2 + x_2^2 - 1$ for all $(x_1, x_2) \in \mathbb{R}^2$, the gradients and the Hessian matrices of $f$ and $g$ are given by

$$\nabla f(x_1, x_2) = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad \nabla^2 f(x_1, x_2) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \tag{1.78}$$

$$\nabla g(x_1, x_2) = \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix}, \quad \nabla^2 g(x_1, x_2) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, \tag{1.79}$$

for all $(x_1, x_2) \in \mathbb{R}^2$, respectively. It is clear that $\nabla^2 f(x_1, x_2)$ and $\nabla^2 g(x_1, x_2)$ are semi-positive definite and positive definite matrices, respectively. Thus, $f$ is convex and $g$ is strictly convex. Consequently, $(P)$ is a convex problem.

Since $g(0, 0) = -1 < 0$, the Slater condition is satisfied.

2. The feasible set of $(P)$ is given by

$$C := \left\{ (x_1, x_2) \in \mathbb{R}^2 ; g(x_1, x_2) = x_1^2 + x_2^2 - 1 \leq 0 \right\}. \tag{1.80}$$

The constraint qualification hypothesis $T_C(x^\star) = L_C(x^\star)$ is guaranteed by the Slater constraint qualification (CQ2). Combining this with the convexity of $(P)$, applying Theorem 8.2.1, and Theorem 8.2.2, [2], yields that $x^\star$ is an optimal solution of $(P)$ if and only if the vectors $(x^\star, \lambda^\star)$ satisfy the KKT conditions. The KKT conditions for $(P)$ are given by (1.61), i.e.,

$$2\lambda^\star x_1^\star + 1 = 0, \tag{1.81}$$
$$2\lambda^\star x_2^\star - 1 = 0, \tag{1.82}$$
$$\lambda^\star \left( (x_1^\star)^2 + (x_2^\star)^2 - 1 \right) = 0, \tag{1.83}$$
$$\lambda^\star \geq 0, \tag{1.84}$$
$$(x_1^\star)^2 + (x_2^\star)^2 - 1 \leq 0. \tag{1.85}$$

The first equation guarantees that $\lambda^\star \neq 0$. Solving the first two equations in this system yields $(x_1^\star, x_2^\star) = \left( -\frac{1}{2\lambda^\star}, \frac{1}{2\lambda^\star} \right)$. Plugging these into the others gives us

$$\lambda^\star \left( \frac{1}{2(\lambda^\star)^2} - 1 \right) = 0, \tag{1.86}$$
$$\lambda^\star \geq 0, \tag{1.87}$$
$$\frac{1}{2(\lambda^\star)^2} \leq 1, \tag{1.88}$$

which has the unique root $\lambda^\star = \frac{1}{\sqrt{2}}$. Then $(x_1^\star, x_2^\star) = \left( -\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right)$ is the unique optimal solution of $(P)$.

3. The logarithmic barrier function of $(P)$ is

$$B(x_1, x_2, t) := f(x_1, x_2) - t \log(-g(x_1, x_2)) \tag{1.89}$$
$$= x_1 - x_2 - t \log\left(1 - x_1^2 - x_2^2\right), \tag{1.90}$$

for all $(x_1, x_2) \in C_B$ and $t > 0$, where $C_B$ is the feasible set of $B$ and is given by

$$C_B := \left\{ (x_1, x_2) \in \mathbb{R}^2 ; g(x_1, x_2) = x_1^2 + x_2^2 - 1 < 0 \right\}. \tag{1.91}$$

We now prove that $B$ is convex in $C_B$.

$\star$ *First proof of the convexity of $B$.* As in the proof of Problem 1.9, the convexity of $f$, $g$ and $-\log x$ yields the that of $B$. $\triangle$

$\star$ *Second proof of the convexity of $B$.* The spatial gradient and the spatial Hessian matrix of $B$ are given by

$$\nabla_x B(x_1, x_2, t) = \begin{bmatrix} 1 - \dfrac{2tx_1}{x_1^2 + x_2^2 - 1} \\ -1 - \dfrac{2tx_2}{x_1^2 + x_2^2 - 1} \end{bmatrix}, \tag{1.92}$$

and

$$\nabla_x^2 B(x_1,x_2,t) = \frac{2t}{\left(x_1^2+x_2^2-1\right)^2}\begin{bmatrix} x_1^2-x_2^2+1 & 2x_1x_2 \\ 2x_1x_2 & x_2^2-x_1^2+1 \end{bmatrix}, \quad (1.93)$$

for all $(x_1,x_2)\in C_B$ and $t>0$. The eigenvalues of $\nabla_x^2 B(x_1,x_2,t)$ are

$$\lambda_1 = \frac{2t\left(x_1^2+x_2^2+1\right)}{\left(x_1^2+x_2^2-1\right)^2}, \quad (1.94)$$

$$\lambda_2 = -\frac{2t}{x_1^2+x_2^2-1}, \quad (1.95)$$

which are positive for all $(x_1,x_2)\in C_B$ and $t>0$. Hence, $\nabla_x^2 B(x_1,x_2,t)$ is positive definite for all $(x_1,x_2)\in C_B$ and $t>0$, and thus $B$ is strictly convex for all $t>0$. △

Since $B$ is strictly convex, $x^\star$ is the unique minimizer of $B$ in $C_B$ if and only if $g(x^\star)<0$ and $\nabla_x B(x^\star,t)=0$. The roots of the equation $\nabla_x B(x,t)=0$ are $\left(\frac{t-\sqrt{t^2+2}}{2}, -\frac{t-\sqrt{t^2+2}}{2}\right)$, $\left(\frac{t+\sqrt{t^2+2}}{2}, -\frac{t+\sqrt{t^2+2}}{2}\right)$. The former is taken and the later is omitted since

$$g\left(\frac{t-\sqrt{t^2+2}}{2}, -\frac{t-\sqrt{t^2+2}}{2}\right) = \frac{2t\left(t-\sqrt{t^2+2}\right)}{2} < 0, \quad (1.96)$$

$$g\left(\frac{t+\sqrt{t^2+2}}{2}, -\frac{t+\sqrt{t^2+2}}{2}\right) = \frac{2t\left(t+\sqrt{t^2+2}\right)}{2} > 0, \quad (1.97)$$

for all $t>0$. Thus, $x^\star(t) = \left(\frac{t-\sqrt{t^2+2}}{2}, -\frac{t-\sqrt{t^2+2}}{2}\right)$ is the optimal solution of the barrier approximation problem for all $t>0$. It is evident that $x^\star(t)\to x^\star$ as $t\to 0^+$. □

**Problem 1.11.** *Consider the following problem*

$$(P) \ \text{Min } x \text{ s.t. } 0\le x\le 1. \quad (1.98)$$

1. *Prove that $(P)$ is a convex problem and the Slater condition is satisfied.*
2. *Use the KKT conditions, find the optimal solution $x^\star$ of $(P)$.*
3. *Establish the barrier approximation problem for $(P)$, find the optimal solution $x^\star(t)$ of the barrier approximation problem. Prove that $x^\star(t)\to x^\star$ as $t\to 0^+$.*

Solution.

1. Set $f(x)=x$, $g_1(x)=-x$, and $g_2(x)=x-1$ for all $(x_1,x_2)\in\mathbb{R}^2$, the first and second derivatives of $f$, $g_1$, and $g_2$ are given by $f'(x)={g_2}'(x)=1$, ${g_1}'(x)=-1$, $f''(x)={g_1}''(x)={g_2}''(x)=0$. Hence, these functions are convex and thus $(P)$ is a convex problem.

   Since $g_1\left(\frac{1}{2}\right)=g_2\left(\frac{1}{2}\right)=-\frac{1}{2}<0$, the Slater condition is satisfied.

2. The feasible set of $(P)$ is given by

$$C := \{x \in \mathbb{R}; g_i(x) \leq 0,\ i = 1, 2\} = [0, 1]. \tag{1.99}$$

The constraint qualification hypothesis $T_C(x^\star) = L_C(x^\star)$ is guaranteed by (CQ1) or (CQ2) since $g_1$, $g_2$ are convex and affine. Combining this with the convexity of $(P)$, applying Theorem 8.2.1 and Theorem 8.2.2, [2] yields that $x^\star$ is an optimal solution of $(P)$ if and only if the vector $(x^\star, \lambda^\star)$ satisfy the KKT conditions. The KKT conditions for $(P)$ are given by

$$(KKT) \begin{cases} f'(x^\star) + \lambda_1^\star {g_1}'(x^\star) + \lambda_2^\star {g_2}'(x^\star) = 0, \\ \lambda_1^\star g_1(x^\star) = \lambda_2^\star g_2(x^\star) = 0, \\ \lambda_1^\star \geq 0,\ \lambda_2^\star \geq 0, \\ g_1(x^\star) \leq 0,\ g_2(x^\star) \leq 0, \end{cases} \tag{1.100}$$

which is equivalent to

$$1 - \lambda_1^\star + \lambda_2^\star = 0, \tag{1.101}$$
$$\lambda_1^\star x^\star = \lambda_2^\star (x^\star - 1) = 0, \tag{1.102}$$
$$\lambda_1^\star \geq 0, \lambda_2^\star \geq 0, \tag{1.103}$$
$$0 \leq x^\star \leq 1, \tag{1.104}$$

We have $\lambda_1 = 1 + \lambda_2 \geq 1 > 0$, thus the second equation gives us $x^\star = 0$ and then $\lambda_2 = 0$, $\lambda_1 = 1$. Thus, $x^\star = 0$ is the optimal solution of $(P)$.

3. The logarithmic barrier function of $(P)$ is

$$B(x, t) := f(x) - t\log(-g_1(x)) - t\log(-g_2(x)) \tag{1.105}$$
$$= x - t\log(x(1 - x)) \tag{1.106}$$

for all $x \in C_B$ and $t > 0$, where $C_B$ is the feasible set of $B$ and is given by

$$C_B := \{x \in \mathbb{R}; g_i(x) < 0,\ i = 1, 2\} = (0, 1). \tag{1.107}$$

We now prove that $B$ is convex in $C_B$.

$\star$ *First proof of the convexity of $B$.* As in the proof of Problem 1.9, the convexity of $f$, $g_1$, $g_2$ and $-\log x$ yields that of $B$. $\triangle$

$\star$ *Second proof of the convexity of $B$.* The first and second $x$-derivatives of $B$ are given by

$$\frac{\partial B}{\partial x}(x, t) = 1 + \frac{t(2x - 1)}{x(1 - x)},\ \frac{\partial^2 B}{\partial x^2}(x, t) = t\left(\frac{1}{x^2} + \frac{1}{(1 - x)^2}\right), \tag{1.108}$$

for all $x \in (0, 1)$ and $t > 0$. Thus, $B$ is strictly convex for all $t > 0$. $\triangle$

Since $B$ is strictly convex, $x^\star$ is the unique minimizer of $B$ in $C_B$ if and only if $\frac{\partial B}{\partial x}(x^\star, t) = 0$ and $x^\star \in (0, 1)$. The roots of the equation $\frac{\partial B}{\partial x}(x, t) = 0$ are $x_1 = \frac{1}{2}\left(2t + 1 - \sqrt{4t^2 + 1}\right)$, $x_2 = \frac{1}{2}\left(2t + 1 + \sqrt{4t^2 + 1}\right)$. The former is taken and the later is omitted since $x_2 > 1$ and $x_1 \in (0, 1)$:

$$\frac{2t + 1 - \sqrt{4t^2 + 4t + 1}}{2} \leq \frac{2t + 1 - \sqrt{4t^2 + 1}}{2} \leq \frac{2t + 1 - 2t}{2}, \tag{1.109}$$

for all $t > 0$. Thus $x^\star(t) = \frac{1}{2}\left(2t + 1 - \sqrt{4t^2 + 1}\right)$ is the optimal solution of the barrier approximation problem for all $t > 0$. It is evident that $x^\star(t) \to x^\star$ as $t \to 0^+$. $\square$

**Problem 1.12.** *Consider the following problem*

$$(P)\ \text{Min}\ \left(x_1+\frac{3}{2}\right)^2+\left(x_2-\frac{3}{2}\right)^2\ \text{s.t. } x_1\geq -1 \text{ and } x_2\geq 1. \tag{1.110}$$

1. *Use the methods presented in [2], find the optimal solution $x^\star$ of $(P)$.*
2. *Establish the barrier approximation problem for $(P)$, find the optimal solution $x^\star(t)$ of that barrier approximation problem. Prove that $x^\star(t)\to x^\star$ as $t\to 0^+$.*

SOLUTION.

1. Set $f(x_1,x_2)=\left(x_1+\frac{3}{2}\right)^2+\left(x_2-\frac{3}{2}\right)^2$, $g_1(x_1,x_2)=-1-x_1$, and $g_2(x_1,x_2)=1-x_2$ for all $(x_1,x_2)\in\mathbb{R}^2$, the gradients the Hessian matrices of $f$, $g_1$, and $g_2$ are given by

$$\nabla f(x_1,x_2)=\begin{bmatrix}2x_1+3\\2x_2-3\end{bmatrix},\ \nabla^2 f(x_1,x_2)=\begin{bmatrix}2&0\\0&2\end{bmatrix}, \tag{1.111}$$

and

$$\nabla g_1(x_1,x_2)=\begin{bmatrix}-1\\0\end{bmatrix},\ \nabla^2 g_1(x_1,x_2)=\begin{bmatrix}0&0\\0&0\end{bmatrix}, \tag{1.112}$$

$$\nabla g_2(x_1,x_2)=\begin{bmatrix}0\\-1\end{bmatrix},\ \nabla^2 g_1(x_1,x_2)=\begin{bmatrix}0&0\\0&0\end{bmatrix}, \tag{1.113}$$

for all $(x_1,x_2)\in\mathbb{R}^2$, respectively. It is clear that $\nabla^2 f(x_1,x_2)$ is positive definite and $\nabla^2 g_i(x_1,x_2)$, for $i=1,2$, are semi-positive definite. Thus, $f$ is strictly convex, and $g_i$, $i=1,2$, are convex. Consequently, $(P)$ is a convex problem.

Since $g_1(0,2)=g_2(0,2)=-1$, the Slater condition is satisfied.

The feasible set of $(P)$ is given by

$$C:=\left\{(x_1,x_2)\in\mathbb{R}^2; g_i(x_1,x_2)\leq 0,\ i=1,2\right\}=[-1,+\infty)\times[1,+\infty). \tag{1.114}$$

The constraint qualification hypothesis $T_C(x^\star)=L_C(x^\star)$ is guaranteed by (CQ1) or (CQ2) since $g_1$, $g_2$ are convex and affine. Combining this with the convexity of $(P)$, applying Theorem 8.2.1 and Theorem 8.2.2, [2] yields that $x^\star$ is an optimal solution of $(P)$ if and only if the vectors $(x^\star,\lambda^\star)$ satisfy the KKT conditions. The KKT conditions for $(P)$ are given by

$$(KKT)\begin{cases}\nabla f(x^\star)+\lambda_1^\star\nabla g_1(x^\star)+\lambda_2^\star\nabla g_2(x^\star)=0,\\ \lambda_1^\star g_1(x^\star)=\lambda_2^\star g_2(x^\star)=0,\\ \lambda_1^\star\geq 0,\lambda_2^\star\geq 0,\\ g_1(x^\star)\leq 0, g_2(x^\star)\leq 0,\end{cases} \tag{1.115}$$

which is equivalent to

$$2x_1^\star+3-\lambda_1^\star=0, \tag{1.116}$$

$$2x_2^\star - 3 - \lambda_2^\star = 0, \tag{1.117}$$
$$\lambda_1^\star\left(1 + x_1^\star\right) = 0, \tag{1.118}$$
$$\lambda_2^\star\left(1 - x_2^\star\right) = 0, \tag{1.119}$$
$$\lambda_1^\star \geq 0,\ \lambda_2^\star \geq 0, \tag{1.120}$$
$$x_1^\star \geq -1,\ x_2^\star \geq 1. \tag{1.121}$$

Solving the first two equations gives us $x_1^\star = \frac{\lambda_1^\star - 3}{2}$, $x_2^\star = \frac{\lambda_2^\star + 3}{2}$. Substituting these into the others yields

$$\lambda_1^\star\left(\lambda_1^\star - 1\right) = 0, \tag{1.122}$$
$$\lambda_2^\star\left(\lambda_2^\star + 1\right) = 0, \tag{1.123}$$
$$\lambda_1^\star \geq 0, \lambda_2^\star \geq 0, \tag{1.124}$$
$$\lambda_1^\star \geq 1, \lambda_2^\star \geq -1. \tag{1.125}$$

which implies $\lambda_1^\star = 1$, $\lambda_2^\star = 0$. Thus $x^\star = \left(x_1^\star, x_2^\star\right) = \left(-1, \frac{3}{2}\right)$ is the unique optimal solution of $(P)$.

2. The logarithmic barrier function of $(P)$ is

$$\begin{aligned} B\left(x_1, x_2, t\right) &:= f\left(x_1, x_2\right) - t\log\left(g_1\left(x_1, x_2\right) g_2\left(x_1, x_2\right)\right) && (1.126)\\ &= \left(x_1 + \frac{3}{2}\right)^2 + \left(x_2 - \frac{3}{2}\right)^2 - t\log\left(\left(1 + x_1\right)\left(x_2 - 1\right)\right), && (1.127)\end{aligned}$$

for all $\left(x_1, x_2\right) \in C_B$ and $t > 0$, where $C_B$ is the feasible set of $B$ and is given by

$$C_B := \left\{\left(x_1, x_2\right) \in \mathbb{R}^2; g_i\left(x_1, x_2\right) < 0,\ i = 1, 2\right\} = (-1, +\infty) \times (1, +\infty). \tag{1.128}$$

We now prove that $B$ is convex in $C_B$.

$\star$ *First proof of the convexity of $B$.* As in the proof of Problem 1.9, the convexity of $f$, $g_1$, $g_2$ and $-\log x$ yields that of $B$. $\triangle$

$\star$ *Second proof of the convexity of $B$.* The spatial gradient and the spatial Hessian matrix of $B$ are given by

$$\nabla_x B\left(x_1, x_2, t\right) = \begin{bmatrix} 2x_1 + 3 - \frac{t}{x_1+1} \\ 2x_2 - 3 - \frac{t}{x_2-1} \end{bmatrix}, \tag{1.129}$$
$$\nabla_x^2 B\left(x_1, x_2, t\right) = \begin{bmatrix} 2 + \frac{t}{(x_1+1)^2} & 0 \\ 0 & 2 + \frac{t}{(x_2-1)^2} \end{bmatrix}, \tag{1.130}$$

for all $\left(x_1, x_2\right) \in C_B$ and $t > 0$. It is evident that $\nabla_x^2 B\left(x_1, x_2, t\right)$ is positive definite and thus $B$ is strictly convex for all $t > 0$. $\triangle$

Since $B$ is strictly convex, $x^\star$ is the unique minimizer of $B$ in $C_B$ if and only if $\nabla_x B\left(x^\star, t\right) = 0$ and $x^\star \in C_B$. The four roots of the equation $\nabla_x B(x, t) = 0$ are

$$\left(x_1, x_2\right) = \left\{\left(\frac{-5 \pm \sqrt{8t+1}}{4}, \frac{5 \pm \sqrt{8t+1}}{4}\right)\right\}. \tag{1.131}$$

The only solution belonging to $C_B$ is $\left(\frac{-5+\sqrt{8t+1}}{4}, \frac{5+\sqrt{8t+1}}{4}\right)$. Thus $x^\star(t) = \left(\frac{-5+\sqrt{8t+1}}{4}, \frac{5+\sqrt{8t+1}}{4}\right)$ is the optimal solution of the barrier approximation problem for all $t > 0$. It is evident that $x^\star(t) \to x^\star$ as $t \to 0^+$. □

**Problem 1.13.** *Consider the following problem*

$$(P)\ \text{Min } x_1 + x_2 \text{ s.t. } -x_1^2 + x_2 \geq 0 \text{ and } x_1 \geq 0. \tag{1.132}$$

1. *Use the methods presented in [2], find the optimal solution $x^\star$ of $(P)$.*
2. *Establish the barrier approximation problem for $(P)$, find the optimal solution $x^\star(t)$ of that barrier approximation problem. Prove that $x^\star(t) \to x^\star$ as $t \to 0^+$.*

SOLUTION.

1. Set $f(x_1, x_2) = x_1 + x_2$, $g_1(x_1, x_2) = x_1^2 - x_2$, and $g_2(x_1, x_2) = -x_1$ for all $(x_1, x_2) \in \mathbb{R}^2$, the gradients and the Hessian matrices of $f$, $g_1$, and $g_2$ are given by

$$\nabla f(x_1, x_2) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \nabla^2 f(x_1, x_2) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \tag{1.133}$$

and

$$\nabla g_1(x_1, x_2) = \begin{bmatrix} 2x_1 \\ -1 \end{bmatrix}, \quad \nabla^2 g_1(x_1, x_2) = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}, \tag{1.134}$$

$$\nabla g_2(x_1, x_2) = \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \quad \nabla^2 g_2(x_1, x_2) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \tag{1.135}$$

for all $(x_1, x_2) \in \mathbb{R}^2$, respectively. It is clear that these three Hessian matrices are semi-positive definite. Thus, $f$, $g_1$, and $g_2$ are convex. Consequently, $(P)$ is a convex problem.

Since $g_1(1,2) = g_2(1,2) = -1 < 0$, the Slater condition is satisfied.

The feasible set of $(P)$ is given by

$$C := \left\{(x_1, x_2) \in \mathbb{R}^2; g_i(x_1, x_2) \leq 0,\ i = 1, 2\right\}. \tag{1.136}$$

The constraint qualification hypothesis $T_C(x^\star) = L_C(x^\star)$ is guaranteed by the Slater constraint qualification (CQ2). Combining this with the convexity of $(P)$, applying Theorem 8.2.1 and Theorem 8.2.2, [2] yields that $x^\star$ is an optimal solution of $(P)$ if and only if the vectors $(x^\star, \lambda^\star)$ satisfy the KKT conditions. The KKT conditions for $(P)$ are given by (1.115), i.e.,

$$1 + 2\lambda_1^\star x_1^\star - \lambda_2^\star = 0, \tag{1.137}$$

$$1 - \lambda_1^\star = 0, \tag{1.138}$$

$$\lambda_1^\star \left((x_1^\star)^2 - x_2^\star\right) = 0, \tag{1.139}$$

$$\lambda_2^\star x_1^\star = 0, \tag{1.140}$$

$$\lambda_1^\star \geq 0,\ \lambda_2^\star \geq 0, \tag{1.141}$$

$$\left(x_1^\star\right)^2 \le x_2^\star, \tag{1.142}$$
$$x_1^\star \ge 0. \tag{1.143}$$

The second equation gives us $\lambda_1^\star = 1$, then the third one implies that $\left(x_1^\star\right)^2 = x_2^\star$. Hence, we obtain

$$1 + 2x_1^\star - \lambda_2^\star = 0, \tag{1.144}$$
$$\left(x_1^\star\right)^2 = x_2^\star \tag{1.145}$$
$$\lambda_2^\star x_1^\star = 0, \tag{1.146}$$
$$\lambda_2^\star \ge 0, \tag{1.147}$$
$$x_1^\star \ge 0. \tag{1.148}$$

The first equation and the last one implies that $\lambda_2^\star = 1 + 2x_1^\star \ge 1$. Combining this with the third equation yields that $x_1^\star = 0$, then $x_2^\star = 0$ and $\lambda_2^\star = 1$. Thus $x^\star = (x_1^\star, x_2^\star) = (0, 0)$ is the unique optimal solution of $(P)$.

2. The logarithmic barrier function of $(P)$ is

$$B\left(x_1, x_2, t\right) := f\left(x_1, x_2\right) - t \log\left(g_1\left(x_1, x_2\right) g_2\left(x_1, x_2\right)\right) \tag{1.149}$$
$$= x_1 + x_2 - t \log\left(x_1\left(x_2 - x_1^2\right)\right), \tag{1.150}$$

for all $(x_1, x_2) \in C_B$ and $t > 0$, where $C_B$ is the feasible set of $B$ and is given by

$$C_B := \left\{(x_1, x_2) \in \mathbb{R}^2; g_i\left(x_1, x_2\right) < 0,\ i = 1, 2\right\}. \tag{1.151}$$

We now prove that $B$ is convex in $C_B$.

$\star$ *First proof of the convexity of $B$.* As in the proof of Problem 1.9, the convexity of $f$, $g_1$, $g_2$ and $-\log x$ yields that of $B$. $\triangle$

$\star$ *Second proof the convexity of $B$.* The spatial gradient and the spatial Hessian matrix of $B$ are given by

$$\nabla_x B\left(x_1, x_2, t\right) = \begin{bmatrix} 1 - \frac{t\left(x_2 - 3x_1^2\right)}{x_1\left(x_2 - x_1^2\right)} \\ 1 - \frac{t}{x_2 - x_1^2} \end{bmatrix}, \tag{1.152}$$
$$\nabla_x^2 B\left(x_1, x_2, t\right) = \frac{1}{\left(x_2 - x_1^2\right)^2} \begin{bmatrix} \frac{t\left(3x_1^4 + x_2^2\right)}{x_1^2} & -2tx_1 \\ -2tx_1 & t \end{bmatrix}, \tag{1.153}$$

for all $(x_1, x_2) \in C_B$ and $t > 0$. The eigenvalues of $\nabla_x^2 B\left(x_1, x_2, t\right)$, denoted by $\lambda_1$ and $\lambda_2$ satisfy

$$\lambda_1 + \lambda_2 = \operatorname{trace}\left(\nabla_x^2 B\left(x_1, x_2, t\right)\right) = \frac{t\left(3x_1^4 + x_1^2 + x_2^2\right)}{x_1^2 \left(x_2 - x_1^2\right)^2} > 0, \tag{1.154}$$
$$\lambda_1 \lambda_2 = \det\left(\nabla_x^2 B\left(x_1, x_2, t\right)\right) = \frac{t^2\left(x_1^2 + x_2\right)}{x_1^2 \left(x_2 - x_1^2\right)^3} > 0. \tag{1.155}$$

This implies that $\lambda_1$ and $\lambda_2$ are positive, and thus $\nabla_x^2 B\left(x_1, x_2, t\right)$ is positive definite for all $(x_1, x_2) \in C_B$ and $t > 0$. Consequently, $B$ is strictly convex in $C_B$ for all $t > 0$. $\triangle$

Since $B$ is strictly convex, $x^\star$ is the unique minimizer of $B$ in $C_B$ if and only if $\nabla_x B(x^\star, t) = 0$ and $x^\star \in C_B$. The roots of the equation $\nabla_x B(x,t) = 0$ are $\left(\frac{-1-\sqrt{8t+1}}{4}, \frac{12t+1+\sqrt{8t+1}}{8}\right)$, $\left(\frac{-1+\sqrt{8t+1}}{4}, \frac{12t+1-\sqrt{8t+1}}{8}\right)$. The former is omitted and the later is taken since $\frac{-1-\sqrt{8t+1}}{4} < 0$ and $\frac{-1+\sqrt{8t+1}}{4} > 0$ for all $t > 0$. Thus $x^\star(t) = \left(\frac{-1+\sqrt{8t+1}}{4}, \frac{12t+1-\sqrt{8t+1}}{8}\right)$ is the optimal solution of the barrier approximation problem for all $t > 0$. It is evident that $x^\star(t) \to x^\star$ as $t \to 0^+$. □

**Problem 1.14.** *Consider the following problem*

$$(P) \text{ Min } x^2 + xy + \frac{1}{2}y^2 \text{ s.t. } 1 - x^2 - xy = 0. \tag{1.156}$$

1. *Prove that the Mangasarian-Fromovitz constraint qualification is satisfied.*
2. *Solve the KKT system, find candidate solutions.*
3. *Find the optimal solution of $(P)$.*

SOLUTION.

1. Set $f(x,y) = x^2 + xy + \frac{1}{2}y^2$, $h(x,y) = 1 - x^2 - xy$ for all $(x,y) \in \mathbb{R}^2$, the gradients and the Hessian matrices of $f$ and $h$ are given by

$$\nabla f(x,y) = \begin{bmatrix} 2x+y \\ x+y \end{bmatrix}, \quad \nabla^2 f(x,y) = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}, \tag{1.157}$$

$$\nabla h(x,y) = \begin{bmatrix} -2x-y \\ -x \end{bmatrix}, \quad \nabla^2 h(x,y) = \begin{bmatrix} -2 & -1 \\ -1 & 0 \end{bmatrix}, \tag{1.158}$$

for all $(x,y) \in \mathbb{R}^2$. The feasible set of $(P)$ is given by

$$C := \left\{(x,y) \in \mathbb{R}^2; h(x,y) = 1 - x^2 - xy = 0\right\}. \tag{1.159}$$

We check the validity of the Mangasarian-Fromovitz constraint qualification for $(P)$. Suppose that $a\nabla h(x,y) = 0$. Since $(x,y) \in C$, we have $x \neq 0$ (otherwise, $h(0,y) = 1$). Thus $\nabla h(x,y) \neq (0,0)$, and the given linear equation implies that $a = 0$, i.e., $\{\nabla h(x,y)\}$ is linearly independent for all $(x,y) \in C$. Moreover, $\nabla h(x,y)^T(0,0) = 0$ for all $(x,y) \in C$. In particular, $\{\nabla h(x^\star, y^\star)\}$ is linearly independent and $\nabla h(x^\star, y^\star)^T(0,0) = 0$ for any local minimizer of $(P)$, i.e., the Mangasarian-Fromovitz constraint qualification is satisfied.

2. Consider the KKT conditions

$$(KKT)\begin{cases} \nabla f(x^\star, y^\star) + \mu^\star \nabla h(x^\star, y^\star) = 0, \\ h(x^\star, y^\star) = 0, \end{cases} \tag{1.160}$$

i.e.,

$$(1-\mu^\star)(2x^\star + y^\star) = 0, \tag{1.161}$$

$$(1-\mu^\star)x^\star + y^\star = 0, \tag{1.162}$$

$$(x^\star)^2 + x^\star y^\star = 1, \tag{1.163}$$

If $\mu^\star = 1$, the second equation gives us $y^\star = 0$ and then the third one implies that $x^\star = \pm 1$. If $\mu^\star = -1$, the first two equations gives $2x^\star + y^\star = 0$. Plugging $y^\star = -2x^\star$ into the third ones yields $(x^\star)^2 = -1$, which is absurd for $x^\star \in \mathbb{R}$. If $\mu \neq \pm 1$, solving the first two equations yields that $(x^\star, y^\star) = (0,0)$, but this contradicts the third one.

Thus, we have two candidate solutions $x_1^\star = (1,0)$, $x_2^\star = (-1,0)$.

3. Consider the Lagrangian function

$$\begin{aligned} L(x,y,\mu) &:= f(x,y) + \mu h(x,y) &(1.164)\\ &= x^2 + xy + \frac{1}{2}y^2 + \mu\left(1 - x^2 - xy\right), &(1.165)\end{aligned}$$

for all $(x,y) \in C$ and $\mu \in \mathbb{R}$. Thanks to the gradients and the Hessian matrices of $f$ and $h$ computed above, (spatial) those of $L$ are given by

$$\nabla_{\mathbf{x}} L(x,y,\mu) = \begin{bmatrix} (1-\mu)(2x+y) \\ (1-\mu)x + y \end{bmatrix}, \tag{1.166}$$

$$\nabla_{\mathbf{x}}^2 L(x,y,\mu) = \begin{bmatrix} 2(1-\mu) & 1-\mu \\ 1-\mu & 1 \end{bmatrix}, \tag{1.167}$$

for all $(x,y) \in C$ and $\mu \in \mathbb{R}$. We have that $x_1^\star$ and $x_2^\star$ are feasible (belong to $C$), and $\nabla_{\mathbf{x}} L(x_1^\star, 1) = \nabla_{\mathbf{x}} L(x_2^\star, 1) = 0$. We have

$$\nabla_{\mathbf{x}}^2 L(x,y,1) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \tag{1.168}$$

for all $(x,y) \in \mathbb{R}^2$ and

$$\begin{aligned} T(x_1^\star) &:= \left\{d \in \mathbb{R}^2; \nabla h(x_1^\star)^T d = 0\right\} &(1.169)\\ &= \left\{(d_1,d_2) \in \mathbb{R}^2; -2d_1 - d_2 = 0\right\} &(1.170)\\ &= \{(d,-2d); d \in \mathbb{R}\}, &(1.171)\\ T(x_2^\star) &:= \left\{d \in \mathbb{R}^2; \nabla h(x_2^\star)^T d = 0\right\} &(1.172)\\ &= \left\{(d_1,d_2) \in \mathbb{R}^2; 2d_1 + d_2 = 0\right\} &(1.173)\\ &= \{(d,-2d); d \in \mathbb{R}\}, &(1.174)\end{aligned}$$

Thus,

$$\begin{aligned} (d,-2d)^T \nabla_{\mathbf{x}}^2 L(x_1^\star)(d,-2d) &= (d,-2d)^T \nabla_{\mathbf{x}}^2 L(x_2^\star)(d,-2d) &(1.175)\\ &= 4d^2 > 0 \text{ for all } d \in \mathbb{R}, d \neq 0, &(1.176)\end{aligned}$$

and then we can apply Theorem 8.2.3 (second-order sufficient optimality conditions), [2], p. 92, to deduce that two points $(\pm 1, 0)$ are local minimizers for $(P)$. Since $f(-1,0) = f(1,0) = 1$, both $(\pm 1, 0)$ are the optimal solutions of $(P)$. □

**Problem 1.15.** *Consider the following problem*

$$(P) \ \text{Min } xy \text{ s.t. } x^2 + y^2 \leq 2 \text{ and } x + y \geq 0. \tag{1.177}$$

1. *Solve the systems of KKT conditions, find candidate solutions for* $(P)$.
2. *Find the optimal solution of* $(P)$.

SOLUTION.

1. Set $f(x,y) = xy$, $g_1(x,y) = x^2 + y^2 - 2$, and $g_2(x,y) = -x - y$ for all $(x,y) \in \mathbb{R}^2$, the gradients and the Hessian matrices of $f$, $g_1$, and $g_2$ are given by

$$\nabla f(x,y) = \begin{bmatrix} y \\ x \end{bmatrix}, \quad \nabla^2 f(x,y) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \tag{1.178}$$

and

$$\nabla g_1(x,y) = \begin{bmatrix} 2x \\ 2y \end{bmatrix}, \quad \nabla^2 g_1(x,y) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, \tag{1.179}$$

$$\nabla g_2(x,y) = \begin{bmatrix} -1 \\ -1 \end{bmatrix}, \quad \nabla^2 g_2(x,y) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \tag{1.180}$$

for all $(x,y) \in \mathbb{R}^2$. The feasible set of $(P)$ is given by

$$C := \left\{(x,y) \in \mathbb{R}^2; g_i(x,y) \le 0,\ i = 1,2\right\}. \tag{1.181}$$

We consider the KKT conditions

$$(KKT) \begin{cases} \nabla f(x^\star, y^\star) + \lambda_1^\star \nabla g_1(x^\star, y^\star) + \lambda_2^\star \nabla g_2(x^\star, y^\star) = 0, \\ \lambda_1^\star g_1(x^\star, y^\star) = \lambda_2^\star g_2(x^\star, y^\star) = 0 \\ \lambda_1^\star \ge 0,\ \lambda_2^\star \ge 0, \\ g_1(x^\star, y^\star) \le 0,\ g_2(x^\star, y^\star) \le 0, \end{cases} \tag{1.182}$$

i.e.,

$$y^\star + 2\lambda_1^\star x^\star - \lambda_2^\star = 0, \tag{1.183}$$
$$x^\star + 2\lambda_1^\star y^\star - \lambda_2^\star = 0, \tag{1.184}$$
$$\lambda_1^\star \left((x^\star)^2 + (y^\star)^2 - 2\right) = 0, \tag{1.185}$$
$$\lambda_2^\star (x^\star + y^\star) = 0, \tag{1.186}$$
$$\lambda_1^\star \ge 0,\ \lambda_2^\star \ge 0, \tag{1.187}$$
$$(x^\star)^2 + (y^\star)^2 \le 2, \tag{1.188}$$
$$x^\star + y^\star \ge 0. \tag{1.189}$$

The fourth equation implies that $\lambda_2^\star = 0$ or $x^\star = -y^\star$. We consider the following cases depending on the values of $\lambda_2^\star$.

- *Case* $\lambda_2^\star = 0$. The above system becomes

$$y^\star + 2\lambda_1^\star x^\star = 0, \tag{1.190}$$
$$x^\star + 2\lambda_1^\star y^\star = 0, \tag{1.191}$$
$$\lambda_1^\star \left((x^\star)^2 + (y^\star)^2 - 2\right) = 0, \tag{1.192}$$
$$\lambda_1^\star \ge 0, \tag{1.193}$$

$$(x^\star)^2 + (y^\star)^2 \leq 2. \tag{1.194}$$

If $\lambda_1^\star = \frac{1}{2}$, solving the first three equations gives us $(x^\star, y^\star) = (-1, 1)$ or $(x^\star, y^\star) = (1, -1)$. These solutions also satisfies the others. Hence, we obtain two (in their "full forms") candidates $(x^\star, y^\star, \lambda_1^\star, \lambda_2^\star) = \left(-1, 1, \frac{1}{2}, 0\right)$, $(x^\star, y^\star, \lambda_1^\star, \lambda_2^\star) = \left(1, -1, \frac{1}{2}, 0\right)$.
If $\lambda_1^\star \neq \frac{1}{2}$ and $\lambda_1^\star \geq 0$, solving the first two equations yields $x^\star = y^\star = 0$. This solution also satisfies the others. Hence, we obtain candidates $(x^\star, y^\star, \lambda_1^\star, \lambda_2^\star) = (0, 0, a, 0)$ for arbitrary $a \geq 0$ and $a \neq \frac{1}{2}$.

- *Case* $\lambda_2^\star > 0$. The fourth equation in the (KKT) system gives us $x^\star = -y^\star$. But, adding the first two equations in the (KKT) system yields that $2\lambda_2^\star = (2\lambda_1^\star + 1)(x^\star + y^\star) = 0$, which is absurd.

Hence, we have three candidate solutions: $(-1, 1)$, $(1, -1)$ (with $\lambda_1^\star = \frac{1}{2}$ and $\lambda_2^\star = 0$), and $(0, 0)$ (with $\lambda_1^\star = a$, $\lambda_2^\star = 0$, where $a \geq 0$, $a \neq \frac{1}{2}$).

2. The eigenvalues of $\nabla^2 f(x, y)$ are $\lambda_1 = -1$, $\lambda_2 = 1$, i.e., $f$, which is a quadratic function, is nonconvex. Thus, we can not apply Theorem 8.2.2, [1] to our problem. Moreover, we have $g_2(-1, 1) = g_2(1, -1) = g_2(0, 0) = 0$ but $\lambda_2^\star = 0$, i.e., the strict complementarity condition in Theorem 8.2.3, [1] fails for these candidate solutions. Thus, we also can not apply Theorem 8.2.3, [1] to our problem.

   Here is an elementary solution for finding the optimal solution of $(P)$.

   $\star$ *Elementary solution.* By Cauchy inequality, we have $|xy| \leq \frac{x^2+y^2}{2} \leq 1$. Thus, $-1 \leq f(x, y) = xy \leq 1$. The equality $f(x, y) = -1$ holds if and only if $(x, y) = (-1, 1)$ or $(x, y) = (1, -1)$. These points also satisfy the constraint $x + y \geq 0$. Therefore, $(-1, 1)$ and $(1, -1)$ are the only global minimizers of $(P)$. $\square$

THE END

# References

[1] O. Güler. *Foundations of Optimization.* Graduate Texts in Mathematics 258, Springer.

[2] Strodiot, J-J. *Numerical Methods in Optimization.* Natural Sciences University, Ho Chi Minh City, Viet Nam, April 2007.