Applied Mathematics and Computation 237 (2014) 31–59

Contents lists available at ScienceDirect

Applied Mathematics and Computation

journal homepage: www.elsevier.com/locate/amc

Geometric views of the generalized Fischer–Burmeister function and its induced merit function

Huai-Yin Tsai, Jein-Shan Chen *,[1]

Department of Mathematics, National Taiwan Normal University, Taipei 11677, Taiwan

ARTICLE INFO

Keywords:
Curvature
Surface
Level curve
NCP-function
Merit function

ABSTRACT

In this paper, we study geometric properties of surfaces of the generalized Fischer–Burmeister function and its induced merit function. Then, a visualization is proposed to explain how the convergent behaviors are influenced by two descent directions in merit function approach. Based on the geometric properties and visualization, we have more intuitive ideas about how the convergent behavior is affected by changing parameter. Furthermore, geometric view indicates how to improve the algorithm to achieve our goal by setting proper value of the parameter in merit function approach.

1. Introduction

The nonlinear complementarity problem (NCP) is to find a point $x \in \mathbb{R}^n$ such that

$$x \geqslant 0, \quad F(x) \geqslant 0, \quad \langle x, F(x)\rangle = 0, \tag{1}$$

where $\langle \cdot, \cdot \rangle$ is the Euclidean inner product and $F = (F_1, \ldots, F_n)^T$ is a map from $\mathbb{R}^n$ to $\mathbb{R}^n$. We assume that F is continuously differentiable throughout this paper. The NCP has attracted much attention because of its wide applications in the fields of economics, engineering, and operations research [8,11,16], to name a few.

Many methods have been proposed to solve the NCP; see [1,14,16,20,22,25] and the references therein. One of the most powerful and popular approach is to reformulate the NCP as a system of nonlinear equations [21,23,28], or an unconstrained minimization problem [9,10,12,15,18,19,24,27]. The objective function that can constitute an equivalent unconstrained minimization problem is called a merit function, whose global minima are coincident with the solutions of the original NCP. To construct a merit function, a class of functions, called NCP-functions and defined below, plays a significant role.

A function $\phi : \mathbb{R}^2 \to \mathbb{R}$ is called an NCP-function if it satisfies

$$\phi(a, b) = 0 \Longleftrightarrow a \geqslant 0, \quad b \geqslant 0, \quad ab = 0. \tag{2}$$

Equivalently, ϕ is an NCP-function if the set of its zeros is the two nonnegative semiaxes. An important NCP-function, which plays a central role in the development of efficient algorithms for the solution of the NCP, is the well-known Fischer–Burmeister (FB) NCP-function [12,13] defined as

* Corresponding author.
E-mail addresses: tasiwhyin@gmail.com (H.-Y. Tsai), jschen@math.ntnu.edu.tw (J.-S. Chen).

[1] Member of Mathematics Division, National Center for Theoretical Sciences, Taipei Office. The author's work is supported by Ministry of Science and Technology, Taiwan.

http://dx.doi.org/10.1016/j.amc.2014.03.089
0096-3003/

$$\phi(a,b)=\sqrt{a^2+b^2}-(a+b). \tag{3}$$

With the NCP function, we can obtain an equivalent formulation of the NCP by a system of equations:

$$\Phi(x)=\begin{pmatrix}\phi(x_1,F_1(x))\\ \cdot\\ \cdot\\ \cdot\\ \phi(x_n,F_n(x))\end{pmatrix}=0. \tag{4}$$

In other words, we have

$$x \text{ solves the NCP} \Longleftrightarrow \Phi(x)=0.$$

In view of this, we define a real-valued function $\Psi:\mathbb{R}^n\to\mathbb{R}_+$

$$\Psi(x):=\frac{1}{2}\|\Phi(x)\|^2=\frac{1}{2}\sum_{i=1}^{n}\phi^2(x_i,F_i(x)). \tag{5}$$

It is known that Ψ a merit function of the NCP, i.e., the NCP is equivalent to an unconstrained minimization problem:

$$\min_{x\in\mathbb{R}^n}\Psi(x). \tag{6}$$

Merit functions is frequently used in designing numerical algorithms for solving the NCP. In particular, we can apply an iterative algorithm to minimize the merit function with hope of obtaining its global minimum.

Recently, the so-called generalized Fischer–Burmeister function was proposed in [3,4]. More specifically, they considered $\phi_p:\mathbb{R}^2\to\mathbb{R}$ and

$$\phi_p(a,b):=\|(a,b)\|_p-(a+b), \tag{7}$$

where $p>1$ is an arbitrary fixed real number and $\|(a,b)\|_p$ denotes the p-norm of (a,b), i.e., $\|(a,b)\|_p=\sqrt[p]{|a|^p+|b|^p}$. In other words, in the function ϕ_p, the 2-norm of (a,b) in the FB function is replaced by a more general p-norm. The function ϕ_p is still an NCP-function, which naturally induces another NCP-function $\psi_p:\mathbb{R}^2\to\mathbb{R}_+$ given by

$$\psi_p(a,b):=\frac{1}{2}|\phi_p(a,b)|^2. \tag{8}$$

For any given $p>1$, the function ψ_p is shown to possess all favorable properties of the FB function ψ; see [2–4]. It plays an important part in our study throughout the paper. Like Φ, the operator $\Phi_p:\mathbb{R}^n\to\mathbb{R}^n$ defined as

$$\Phi_p(x)=\begin{pmatrix}\phi_p(x_1,F_1(x))\\ \cdot\\ \cdot\\ \cdot\\ \phi_p(x_n,F_n(x))\end{pmatrix} \tag{9}$$

yields a family of merit functions $\Psi_p:\mathbb{R}^n\to\mathbb{R}_+$ for the NCP:

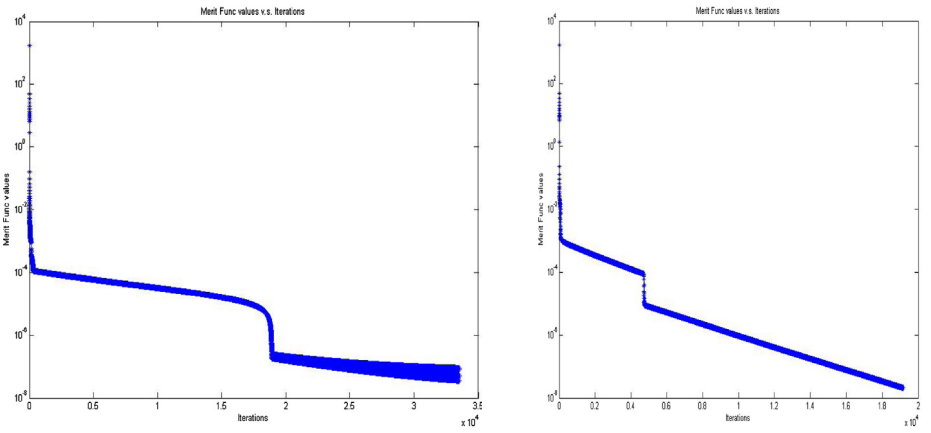
$$\Psi_p(x):=\frac{1}{2}\|\Phi_p(x)\|^2=\sum_{i=1}^{n}\psi_p(x_i,F_i(x)). \tag{10}$$

Analogously, the NCP is equivalent to an unconstrained minimization problem:

$$\min_{x\in\mathbb{R}^n}\Psi_p(x). \tag{11}$$

It was shown that if F is monotone [15] or an P_0-function [10], then any stationary point of Ψ is a global minima of the unconstrained minimization $\min_{x\in\mathbb{R}^n}\Psi(x)$, and hence solves the NCP. The similar results were generalized to Ψ_p-case in [4]. On the other hand, there are many classical iterative methods applied to this unconstrained minimization of the NCP. Derivative-free methods [29] are suitable for problems where the derivatives of F are not available or expansive. Some derivative-free algorithms with global convergence results were proposed to solve the NCP based on generalized Fischer–Burmeister merit function. For example, [4,5] pointed out that the performance of the algorithm is influenced by parameter p. In addition, there have been observed some phenomenon in the derivative-free algorithm studied in [5]. More specifically, there occurs kind of "cliff" in the convergent behavior depicted as Fig. 1.

During these years, we are frequently asked about what is the main factor causing this and how parameter p affects convergent behavior? These are what we are eager to know of. In light of our earlier numerical experience, we find that figuring out the geometric properties of ϕ_p and ψ_p may be a key way to answer the aforementioned puzzles. In view of this



(a) Convergent behavior in some descent algo-rithm ($p = 1.1$)

(b) Convergent behavior in some de-scent algorithm ($p = 1000$)

Fig. 1. "Cliff" phenomenon that appears in some derivative-free algorithm.

motivation, we aim to do analysis from geometric view in this paper. More specifically, the objective of this paper is to study the relation between convergent behavior and parameter p via aspect of geometry in which the graphs of ϕ_p and ψ_p can be regarded as families of surfaces embedded in $\mathbb{R}^3$.

This paper is organized as follows. In Section 2, we propose some geometric properties of ϕ_p and present its surface structure by figures. In Section 3, we study properties of ψ_p, and summarize the comparison between ϕ_p and ψ_p. In Section 4, we investigate a geometric visualization to see possible convergence behavior with different p by a few examples. Finally, we state the conclusion.

2. Geometric view of ϕ_p

In this section, we study some geometric properties of ϕ_p and interpret their meanings. We present the family of surfaces of $\phi_p(a,b)$ where $p \in (1,+\infty)$, see Figs. 2 and 3. When we fix a real number p with $1 < p < +\infty$, Fig. 3 gives us intuitive image that the surface shape is indeed influenced by the value of p. From the definition of p-norm, we know that $\|(a,b)\|_1 := |a| + |b|$, and $\|(a,b)\|_\infty := \max\{|a|,|b|\}$. It is trivial that $\phi_p(a,b) \to \phi_1(a,b) := |a| + |b| - (a+b)$ pointwisely, see Fig. 3(a) and (b). On the other hand $\phi_p(a,b) \to \phi_\infty(a,b) := \max\{|a|,|b|\} - (a+b)$ pointwisely, see Fig. 3(e) and (f). Note that $\phi_1(a,b)$ is not an NCP function because when $a > 0$ and $b > 0$, we have $\phi_1(a,b) = 0$ whereas $\phi_\infty(a,b)$ is an NCP function but not differentiable when $a = b$.

Next, we give some lemmas which will be used in subsequent analysis.

Lemma 2.1 [6, Lemma 3.1]. *If $a > 0$ and $b > 0$, then $(a+b)^p > a^p + b^p$ for all $p \in (1,+\infty)$.*

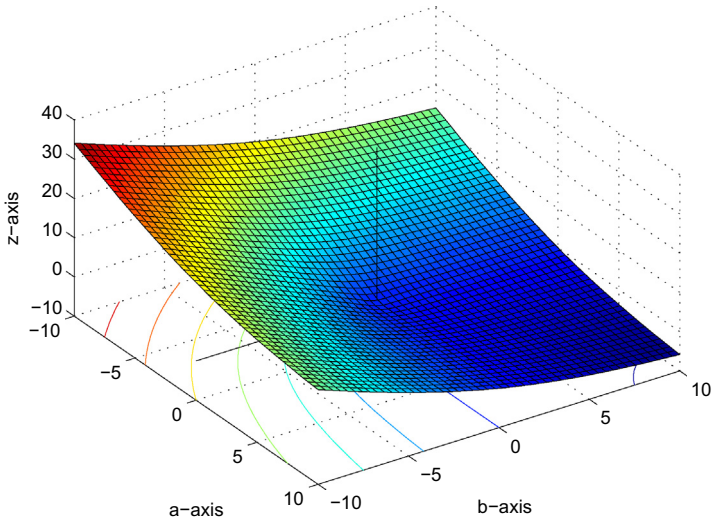
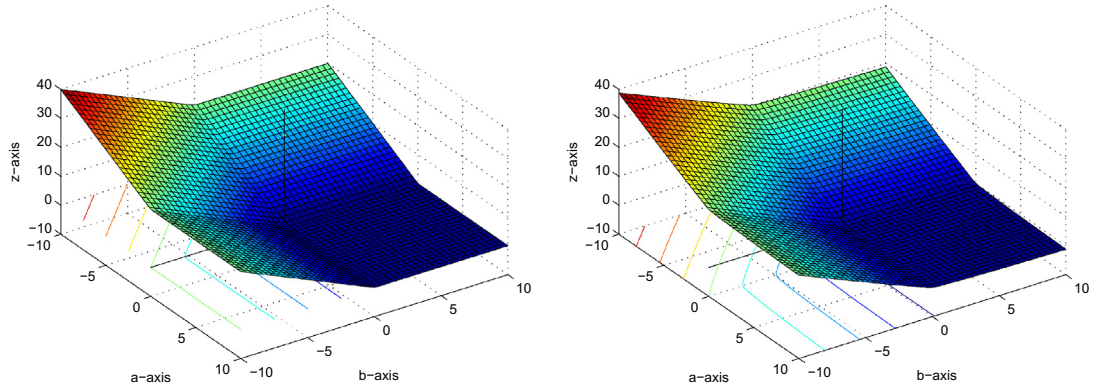
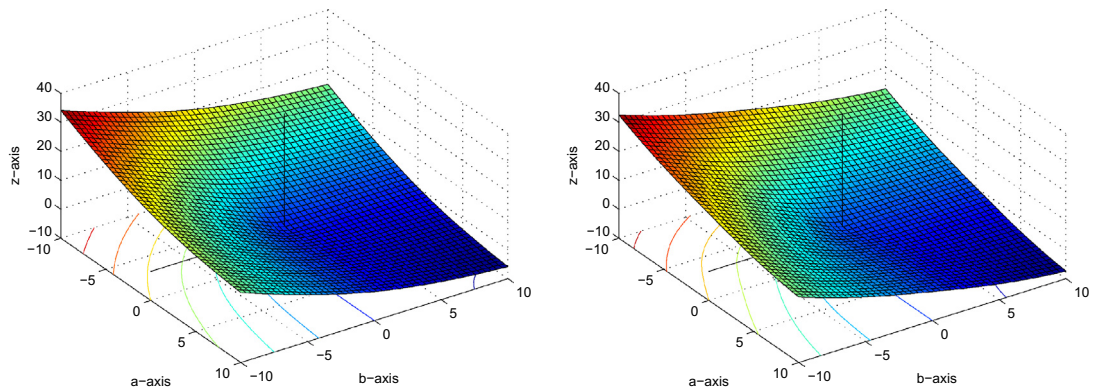


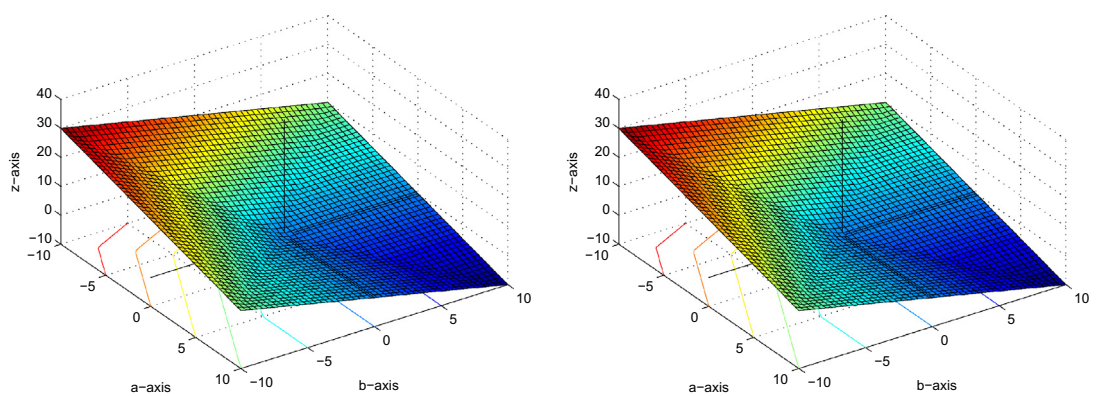
Fig. 2. The surface of $z = \phi_2(a,b)$ with $(a,b) \in [-10,10] \times [-10,10]$.



(a) $z = \phi_1(a, b)$ (b) $p = 1.1$



(c) $p = 2$ (d) $p = 3$



(e) $p = 100$ (f) $z = \phi_\infty(a, b)$

Fig. 3. The surface of $z = \phi_p(a, b)$ with different p.

Lemma 2.2 [17, Lemma 1.3]. *Let* $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ *and* $\|x\|_p := \left(\sum_{i=1}^n |x_i|^p\right)^{\frac{1}{p}}$. *If* $1 < p_1 < p_2$, *then* $\|x\|_{p_2} \leqslant \|x\|_{p_1} \leqslant n^{\left(\frac{1}{p_1} - \frac{1}{p_2}\right)} \|x\|_{p_2}$.

Lemma 2.3 [5, Lemma 3.2]. *Let* $\phi_p : \mathbb{R}^2 \to \mathbb{R}$ *be given as in (7) where* $p \in (1, +\infty)$. *Then,*

$$\left(2 - 2^{\frac{1}{p}}\right) |\min\{a, b\}| \leqslant |\phi_p(a, b)| \leqslant \left(2 + 2^{\frac{1}{p}}\right) |\min\{a, b\}|.$$

Proposition 2.1. *Let $\phi_p : \mathbb{R}^2 \to \mathbb{R}$ be given as in (7) where $p \in (1, +\infty)$. Then,*

(a) $(a > 0$ *and* $b > 0) \iff \phi_p(a, b) < 0$;
(b) $(a = 0$ *and* $b \geqslant 0)$ or $(b = 0$ *and* $a \geqslant 0) \iff \phi_p(a, b) = 0$;
(c) $b = 0$ *and* $a < 0 \Rightarrow \phi_p(a, b) = -2a > 0$;
(d) $a = 0$ and $b < 0 \Rightarrow \phi_p(a, b) = -2b > 0$.

Proof

(a) If $a > 0$ and $b > 0$, it is easy to see $\phi_p(a, b) < 0$ by Lemma 2.1. Conversely, because $\sqrt[p]{|a|^p + |b|^p} \geqslant |a|$ and $\sqrt[p]{|a|^p + |b|^p} \geqslant |b|$, we have $\sqrt[p]{|a|^p + |b|^p} \geqslant \max\{|a|, |b|\}$. Suppose $a \leqslant 0$ or $b \leqslant 0$, then we have $\max\{|a|, |b|\} \geqslant (a + b)$ which implies $\phi_p(a, b) \geqslant 0$. This is a contradiction.
(b) By definition of $\phi_p(a, b)$, we know

$$\phi_p(a, 0) = |a| - a = \begin{cases} 0 & a \geqslant 0, \\ -2a & a < 0, \end{cases} \quad \phi_p(0, b) = |b| - b = \begin{cases} 0 & b \geqslant 0, \\ -2b & b < 0, \end{cases}$$

which say that $(a = 0$ and $b \geqslant 0)$ or $(b = 0$ and $a \geqslant 0) \Rightarrow \phi_p(a, b) = 0$. Conversely, suppose $\phi_p(a, b) = 0$. If $a < 0$ or $b < 0$, mimicking the arguments of part (a) yields

$$\sqrt[p]{|a|^p + |b|^p} > \max\{|a|, |b|\} > a + b,$$

which implies $\phi_p(a, b) > 0$. Thus, there must hold $a \geqslant 0$ and $b \geqslant 0$. Furthermore, one of a and b must be 0 from part (a).

The proofs of (c) and (d) are direct from the proof of part (b). □

Proposition 2.1(a) shows that $\phi_p(a, b)$ is negative on the first quadrant of $\mathbb{R}^2$-plane, see Fig. 4, while Proposition 2.1(b) shows that $\phi_p(a, b) = 0$ can only happen on the nonnegative semiaxes (i.e., $a \geqslant 0,\ b = 0$ or $a = 0,\ b \geqslant 0$). In fact, this proposition is also equivalent to saying that $\phi_p(a, b)$ is an NCP-function. In addition, Proposition 2.1(b)–(d) indicate that the value of p does not affect the value $\phi_p(a, b)$ on the a-axis and b-axis.

Proposition 2.2. *Let $\phi_p : \mathbb{R}^2 \to \mathbb{R}$ be given as in (7) where $p \in (1, +\infty)$. Then,*

(a) $\phi_p(a, b) = \phi_p(b, a)$;
(b) ϕ_p *is convex, i.e.,*

$$\phi_p(\alpha w + (1 - \alpha)w') \leqslant \alpha\phi_p(w) + (1 - \alpha)\phi_p(w')$$

for all $w, w' \in \mathbb{R}^2$ and $\alpha \in [0, 1]$;
(c) *if $1 < p_1 < p_2$, then $\phi_{p_1}(a, b) \geqslant \phi_{p_2}(a, b)$.*

Proof. The verifications for part (a) and (b) are straightforward, we omit them. Part (c) is true by applying Lemma 2.2. □

Proposition 2.2(a) shows the symmetric property of $\phi_p(a, b)$ which means there have a couple of points on plane between line $a = b$ having the same height. In other words, surface $z = \phi_p(a, b)$ has the same structure on second and forth quadrant of the plane, see Figs. 4–6. Proposition 2.2(b) says that the shape of surface is convex because the function ϕ_p is convex while Proposition 2.2(c) implies that the value of ϕ_p is decreasing when the value of p is increasing. In summary, the value of p would affect geometric structure.

Proposition 2.3. *If $\{(a^k, b^k)\} \subseteq \mathbb{R}^2$ with $(a^k \to -\infty)$ or $(b^k \to -\infty)$ or $(a^k \to +\infty$ and $b^k \to +\infty)$, then $|\phi_p(a^k, b^k)| \to +\infty$ for $k \to +\infty$.*

Proof. This can be found in [26, p. 20]. □

Proposition 2.3 implies the increasing direction on surface. This can be seen from the contour graph of $z = \phi_p(a, b)$ which is plotted in Fig. 4, where the deep color presents the lower height. In order to understand the structure of the surface, it is nature to investigate special curves on the surface. We consider a family of curves $\alpha_{r,p} : \mathbb{R} \to \mathbb{R}^3$ defined as follows:

$$\alpha_{r,p}(t) := \big(r + t, r - t, \phi_p(r + t, r - t)\big) \tag{12}$$

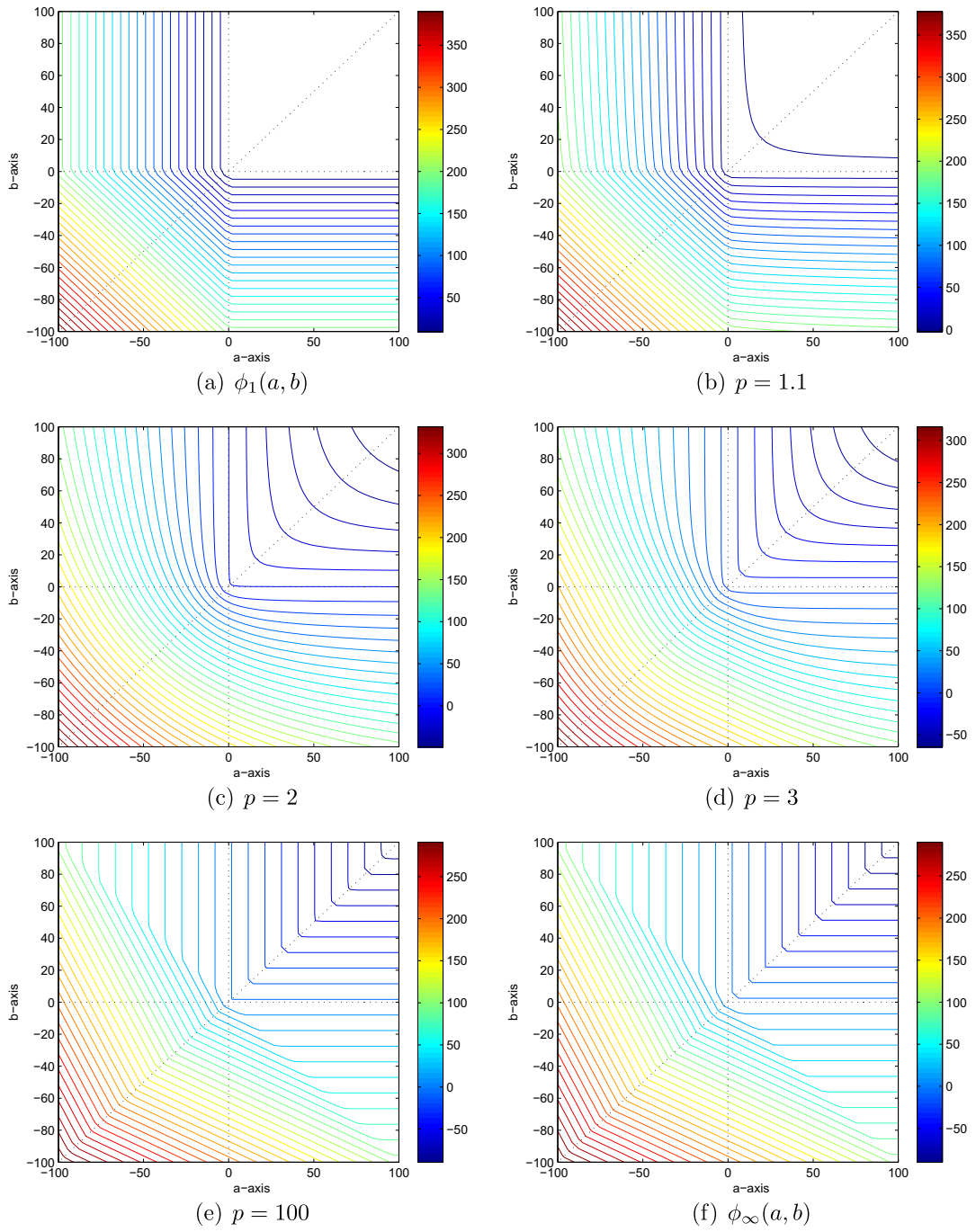


Fig. 4. Level curves of $z = \phi_p(a,b)$ with different p.

where $r \in \mathbb{R}$ and $p \in (1,+\infty)$ are two arbitrary fixed real number. These curves can be viewed as the intersection of surface $z = \phi_p(a,b)$ and plane $a + b = 2r$, see Fig. 6. We study some properties regarding these special curves.

Lemma 2.4. *Let $\phi_p : \mathbb{R}^2 \to \mathbb{R}$ be given as in (7) where $p \in (1,+\infty)$. Fix any $r \in \mathbb{R}$, we define $f : \mathbb{R} \to \mathbb{R}$ as $f(t) := \phi_p(r+t, r-t)$, then f is a convex function.*

Proof. We know that ϕ_p is a convex function by Proposition 2.3 and observe that f is a composition of ϕ_p and an affine function. Thus, f is convex since it is a composition of a convex function and an affine function (the composition of two convex functions is not necessarily convex, however, our case does guarantee the convexity because one of them is affine). □

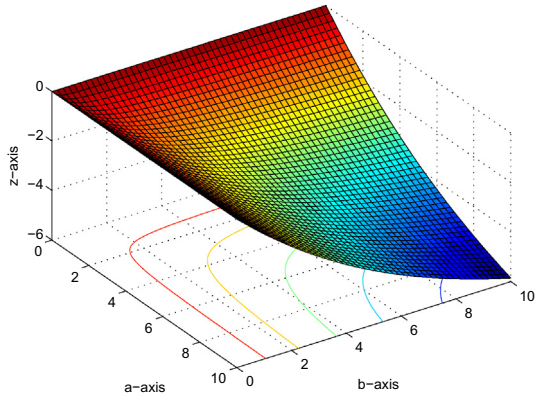
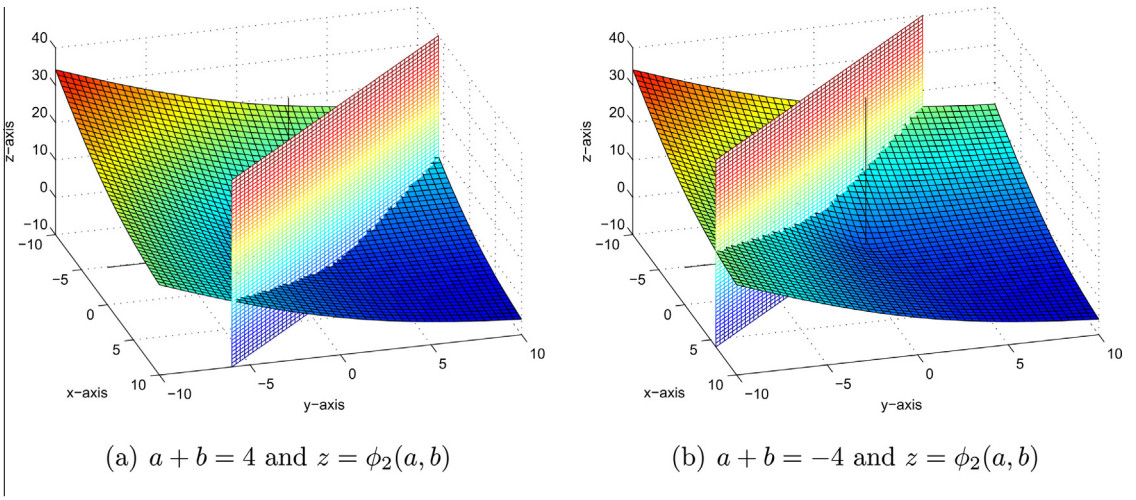
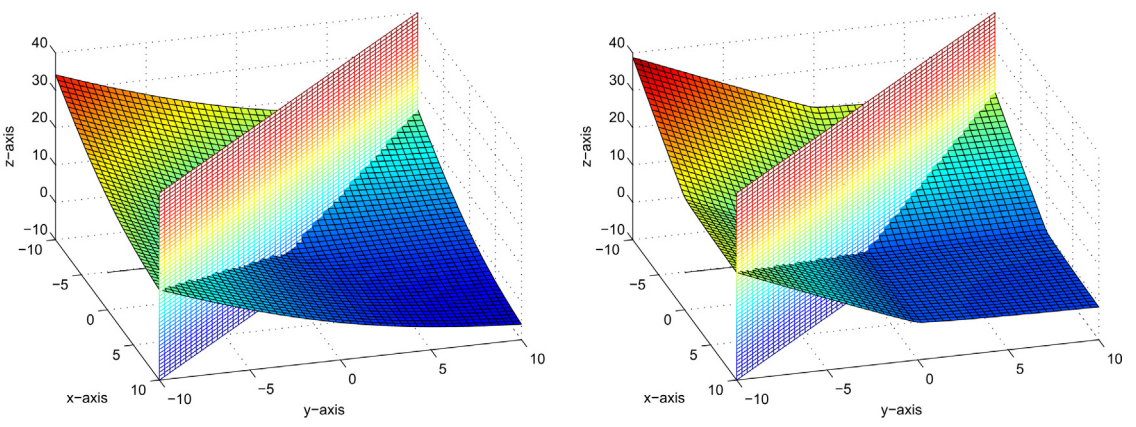


Fig. 5. The surface of $z = \phi_2(a, b)$ with $(a, b) \in [0, 10] \times [0, 10]$.



(a) $a + b = 4$ and $z = \phi_2(a, b)$ (b) $a + b = -4$ and $z = \phi_2(a, b)$



(c) $a + b = 0$ and $z = \phi_2(a, b)$ (d) $a + b = 0$ and $z = \phi_{1.1}(a, b)$

Fig. 6. The curve intersected by surface $z = \phi_p(a, b)$ and plane $a + b = 2r$.

Theorem 2.1. *Let $\phi_p : \mathbb{R}^2 \to \mathbb{R}$ be given as in (7) where $p \in (1, +\infty)$. Suppose a and b are constrained on the curve determined by $a + b = 2r$ $(r \in \mathbb{R})$ and the surface. Then, $\phi_p(a, b)$ attains its minima $\phi_p(r, r) = 2^{\frac{1}{p}}|r| - 2r$ along this curve at $(a, b) = (r, r)$.*

Proof. We know that $\phi_p(a, b)$ is differentiable except $(0, 0)$, therefore we discuss two cases as follows.

(i) Case (1): $r = 0$. Because $a + b = 0$, a and b have opposite sign to each other except $a = b = 0$, from Proposition 2.1, we know $\phi_p(a,b) \geqslant 0$ under this case. Thus, when $(a,b) = (0,0)$, $\phi_p(a,b)$ attains its minima zero.
(ii) Case (2): $r \neq 0$. Fix r and $p > 1$. Let $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ be respectively defined as

$$f(t) := \phi_p(r+t, r-t), \quad g(t) := |r+t|^p + |r-t|^p.$$

Then, we calculate that

$$f'(t) = \frac{g'(t)}{p(g(t))^{\frac{p-1}{p}}} \quad \text{and} \quad g'(t) = p\Big[\operatorname{sgn}(r+t)(r+t)^{p-1} - \operatorname{sgn}(r-t)(r-t)^{p-1}\Big].$$

We know $g(t) > 0$ for all $t \in \mathbb{R}$. It is clear $g'(0) = 0$, and hence $f'(0) = 0$. By Lemma 2.4, $f(t)$ is convex on $\mathbb{R}$. In addition, it is also continuous, therefore, $t = 0$ is a critical point of $f(t)$ which is also a global minimizer of $f(t)$. The proof is done since $a = b = r$ and $\phi_p(r,r) = 2^{\frac{1}{p}}|r| - 2r$ when $t = 0$. □

Lemma 2.4 and Theorem 2.1 show that the curve determined by the plane $a + b = 2r$ and the surface $z = \phi_p(a,b)$ is convex and attains minima when $a = b$, see Fig. 7. We now study curvature of the family of curves $\alpha_{r,p}$ defined as in (12) at point $(r, r, \phi_{r,p}(r,r))$. Because function ϕ_p is not differentiable at $(a,b) = (0,0)$ (i.e., $r = 0$), we choose two points $(-t_0, t_0, \phi_{0,p}(-t_0, t_0))$ and $(t_0, -t_0, \phi_{0,p}(t_0, -t_0))$ where $t_0 > 0$, and calculate the value of cosine function of the angle between $\alpha_{0,p}(-t_0), \alpha_{0,p}(t_0)$, see Fig. 8.

Proposition 2.4. *Let $\alpha_{r,p} : \mathbb{R} \to \mathbb{R}^3$ be defined as in (12), and $\cos_p(\theta)$ be cosine function of the angle between two vectors $\alpha_{0,p}(-t_0)$ and $\alpha_{0,p}(t_0)$ where $t_0 > 0$. Then,*

(a) $\cos_p(\theta) = \dfrac{2^{\frac{2}{p}} - 6}{\sqrt{\left(2^{\frac{2}{p}} - 2\right)^2 + 32}}$;
(b) $\cos_p(\theta) \to -\frac{1}{3}$ as $p \to 1$, *and* $\cos_p(\theta) \to -\frac{5}{33}$ as $p \to +\infty$;
(c) *if* $1 < p_1 < p_2$, *then* $\cos_{p_1}(\theta) < \cos_{p_2}(\theta)$.

Proof

(a) By direct computation, we obtain

$$\cos_p(\theta) = \frac{\alpha_{0,p}(-t_0) \cdot \alpha_{0,p}(t_0)}{\|\alpha_{0,p}(-t_0)\|\|\alpha_{0,p}(t_0)\|} = \frac{2^{\frac{2}{p}} - 6}{\sqrt{\left(2^{\frac{2}{p}} + 6\right) + 2^{\frac{1}{p}+2}}\sqrt{\left(2^{\frac{2}{p}} + 6\right) - 2^{\frac{1}{p}+2}}} = \frac{2^{\frac{2}{p}} - 6}{\sqrt{\left(2^{\frac{2}{p}} - 2\right)^2 + 32}}.$$

(b) From part (a), let $f : (1, +\infty) \to \mathbb{R}$ be $f(p) := \cos_p(\theta)$. Then $f(p)$ is continuous on $(1, +\infty)$. By taking the limit, we have $\cos_p(\theta) \to -\frac{1}{3}$ as $p \to 1$, and $\cos_p(\theta) \to -\frac{5}{33}$ as $p \to +\infty$.

(c) From part (b), we know $f'(p) = \dfrac{6 - \left(1 - \frac{\ln 2}{p}\right)2^{\frac{2}{p}}}{\sqrt{(2^{\frac{2}{p}} - 2)^2 + 32}}$ which implies $f'(p) > 0$ for all $p > 1$. Therefore, $f(p)$ is a strictly increasing function on $(1, +\infty)$. □

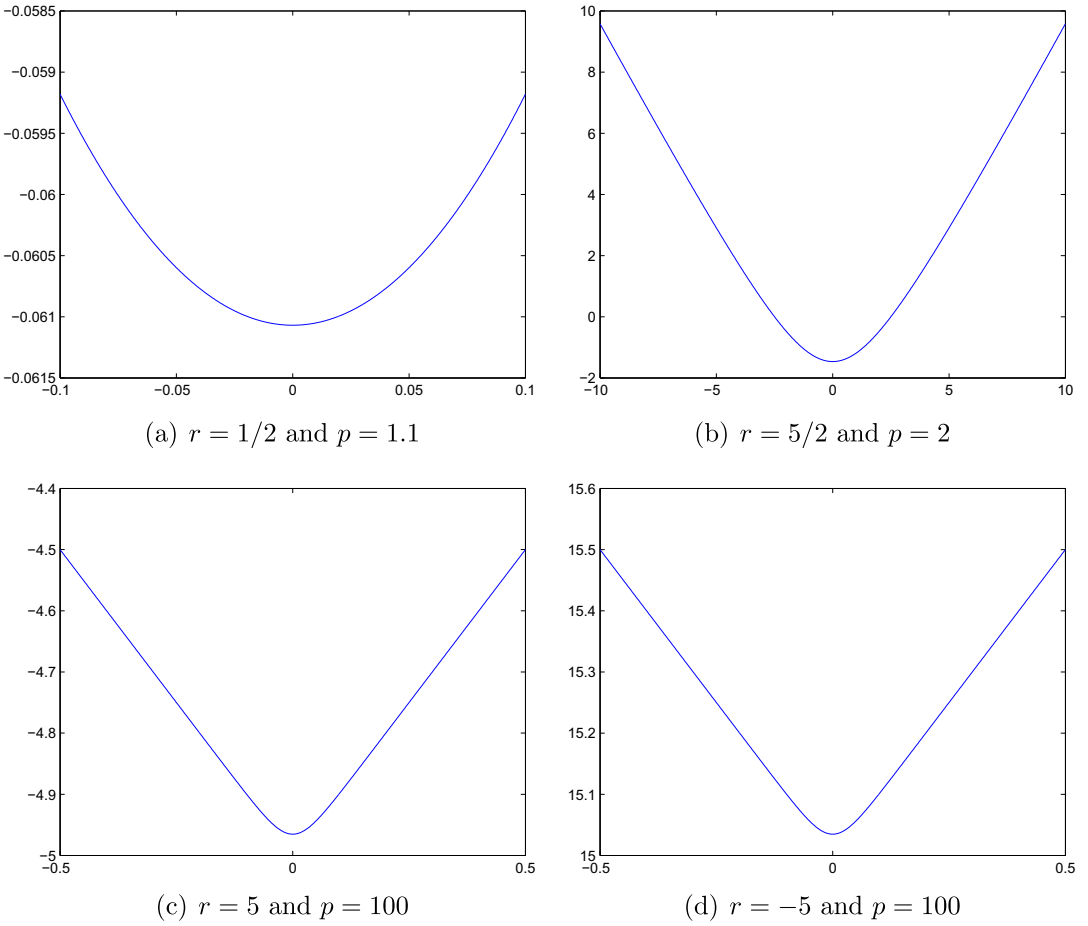
Proposition 2.5. *Let $\alpha_{r,p} : \mathbb{R} \to \mathbb{R}^3$ be defined as in (12). Then the following hold.*

(a) The curvature at point $\alpha_{r,p}(0) = (r, r, \phi_p(r,r))$ is $\kappa_p(0) = \frac{(p-1)2^{\frac{1}{p}-1}}{|r|}$.
(b) $\kappa_p(0) \to 0$ as $p \to 1$ and $\kappa_p(0) \to +\infty$ as $p \to +\infty$.
(c) If $1 < p_1 < p_2$, then $\kappa_{p_1}(0) < \kappa_{p_2}(0)$.

Proof

(a) Because $\alpha_{r,p}(t) = (r + t, r - t, \phi_p(r + t, r - t))$, we know

$$\alpha'_{r,p}(0) = (1, -1, 0) \quad \text{and} \quad \alpha''_{r,p}(0) = \left(0, 0, \frac{(p-1)2^{\frac{1}{p}}}{|r|}\right).$$



(a) $r = 1/2$ and $p = 1.1$ (b) $r = 5/2$ and $p = 2$

(c) $r = 5$ and $p = 100$ (d) $r = -5$ and $p = 100$

Fig. 7. The curve $f(t) = \phi_p(r + t, r - t)$.

Recall the formulation of curvature

$$\kappa_p(t) = \frac{|\alpha'_{r,p}(t) \wedge \alpha''_{r,p}(t)|}{|\alpha'_{r,p}(t)|^3},$$

where wage operator means the outer product of two vectors. Thus, we have

$$\kappa_p(0) = \frac{|\alpha'_{r,p}(0) \wedge \alpha''_{r,p}(0)|}{|\alpha'_{r,p}(0)|^3} = \frac{(p-1)2^{\frac{1}{p}-1}}{|r|}.$$

(b) Let $f : (1, +\infty) \to \mathbb{R}$ be defined as

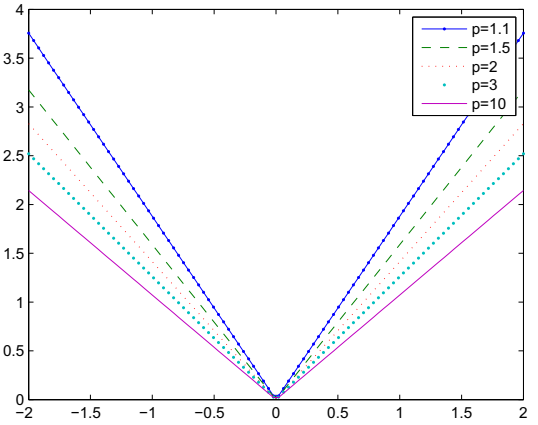
$$f(p) := \kappa_p(0) = \frac{(p-1)2^{\frac{1}{p}-1}}{|r|},$$

then obviously $f(p)$ is continuous on $\mathbb{R}$. Thus, the desired result follows by taking the limit directly.

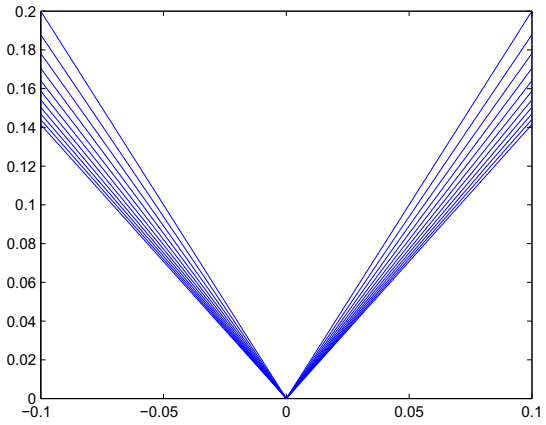
(c) From part (b), we compute that

$$f'(p) = \frac{2^{\frac{1}{p}-1}}{|r|}\left(1 - \frac{\ln 2}{p} + \frac{\ln 2}{p^2}\right),$$

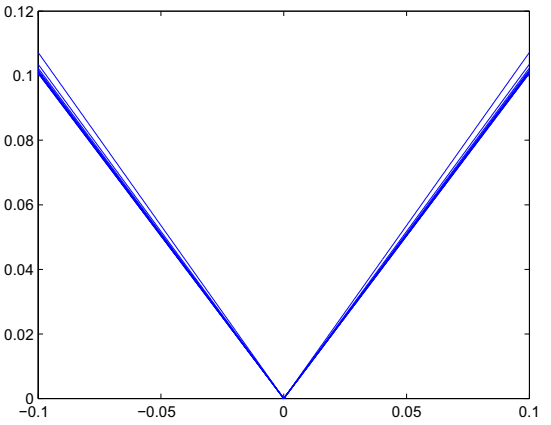
which implies $f'(p) > 0$ for all $p \in (1, +\infty)$. Then $f(p)$ is strictly increasing on $(1, +\infty)$. □



(a) Angle with different p



(b) The change of angle as $p \to 1$.



(c) The change of angle as $p \to +\infty$.

Fig. 8. Angle between vectors $\alpha_{0,p}(-t_0)$ and $\alpha_{0,p}(t_0)$.

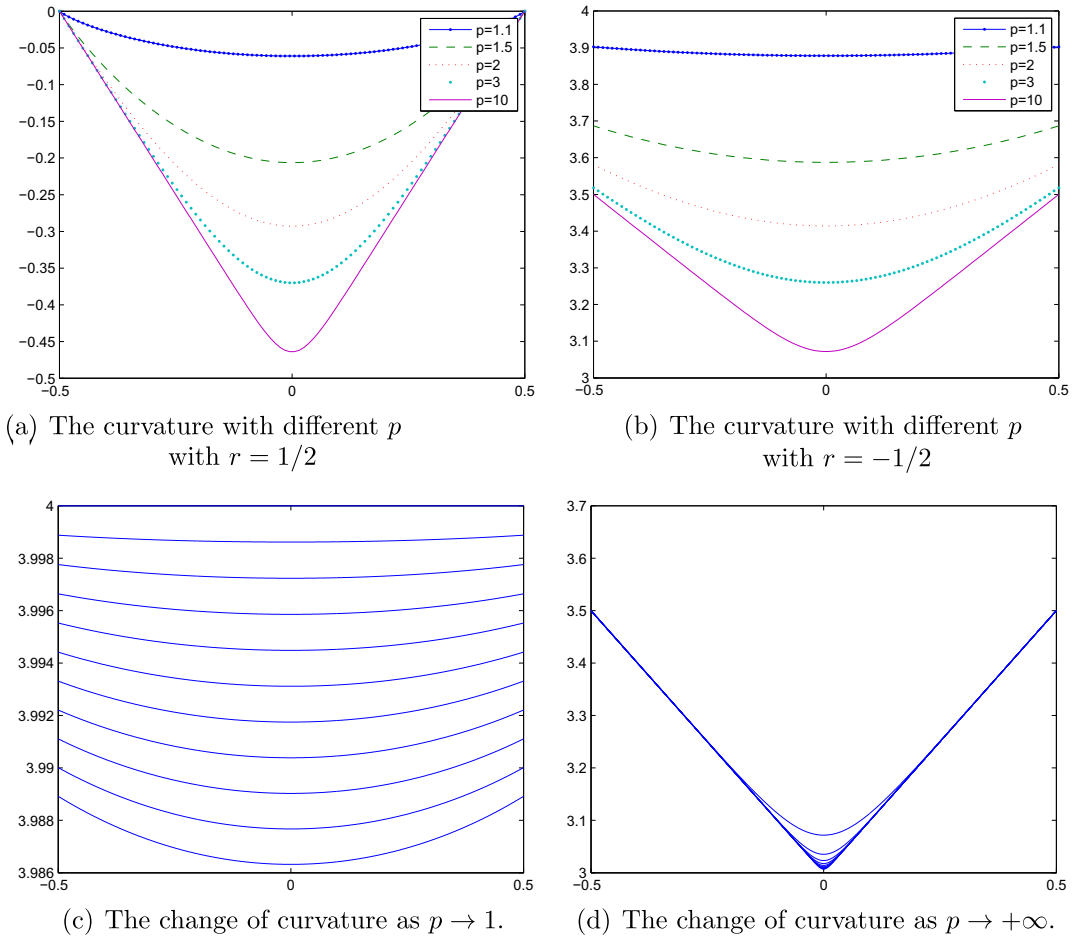
The above two propositions shows how p affect the geometric structure, see Fig. 9(a) and (b). Proposition 2.5(b) says that when $p \to 1$ the curve becomes a straight line, see Fig. 9(c). Note that when $p \to +\infty$ the curve becomes more and more sharp at the point. This curve is not differentiable when $t = 0$, see Fig. 9(d). To sum up, from all properties we presented in this section we realize that p indeed affect the geometric behavior of surface $z = \phi_p(a, b)$ both locally and globally.

3. Geometric view of ψ_p

In previous section, we see that generalized FB function ϕ_p is convex and differentiable everywhere except $(0, 0)$. To the contrast, the function $\psi_p(a, b)$ defined as in (8) is non-convex, but continuously differentiable everywhere. Nonetheless, ϕ_p and ψ_p have many similar geometric properties as will be seen later. In this section, we study some properties like what we have done in Section 2 and compare the difference between ψ_p and ϕ_p (see Figs. 10 and 11).

Proposition 3.1. *Let $\psi_p : \mathbb{R}^2 \to \mathbb{R}$ be given as in (8) where $p \in (1, +\infty)$. Then,*

(a) $\psi_p(a, b) \geqslant 0, \ \forall (a, b) \in \mathbb{R}^2$;
(b) $\psi_p(a, b) = \psi_p(b, a), \ \forall (a, b) \in \mathbb{R}^2$;
(c) $(a = 0$ *and* $b \geqslant 0)$ *or* $(b = 0$ *and* $a \geqslant 0) \iff \psi_p(a, b) = 0$;
(d) $b = 0$ *and* $a < 0 \Rightarrow \psi_p(a, b) = 2a^2 > 0$;
(e) $a = 0$ *and* $b < 0 \Rightarrow \psi_p(a, b) = 2b^2 > 0$;
(f) ψ_p *is continuously differentiable everywhere.*



(a) The curvature with different p with $r = 1/2$

(b) The curvature with different p with $r = -1/2$

(c) The change of curvature as $p \to 1$.

(d) The change of curvature as $p \to +\infty$.

Fig. 9. The curvature $\kappa_p(0)$ at point $\alpha_{r,p}(0)$.

Proof. Parts (d) and (e) come from Propositions 2.5(c) and 2.1(d), please see [2–4] for the rest. □

Proposition 2.2(c) says that the value of ϕ_p is decreasing with respect to p. To the contrast, ψ_p does not have such property. More specifically, it is true for ψ_p to hold such property only on certain quadrants.

Proposition 3.2. *Suppose* $1 < p_1 < p_2$ *and* $(a,b) \in \mathbb{R}^2$. *Then,*

(a) *if* $a < 0$ *or* $b < 0$, *then* $\psi_{p_1}(a,b) \geqslant \psi_{p_2}(a,b)$;
(b) *if* $a > 0$ *and* $b > 0$, *then* $\psi_{p_1}(a,b) \leqslant \psi_{p_2}(a,b)$.

Proof

(a) This is clear from Proposition 2.2(c).
(b) Suppose $a > 0$ and $b > 0$, from Proposition 2.1(a), we have $\phi_p(a,b) < 0$. Then Proposition 2.2(c) yields $\phi_{p_1}(a,b) \geqslant \phi_{p_2}(a,b)$, and hence $\phi^2_{p_1}(a,b) \leqslant \phi^2_{p_2}(a,b)$. □

Since ψ_p is not convex in general. The counterpart of Theorem 2.1 is as below.

Theorem 3.1. *Let* $\psi_p(a,b)$ *be defined as* (8) *with* $a + b = 2r$. *Then, the following hold.*

(a) *If* $r \in \mathbb{R}^+$ *and* $a > 0, b > 0$, *then* $\psi_p(a,b)$ *attains maxima* $\left(2^{\frac{2}{p}-1} - 2^{\frac{1}{p}+1} + 2\right)r^2$ *when* $(a,b) = (r,r)$.
(b) *If* $r \in \mathbb{R}^- \cup \{0\}$, *then* $\psi_p(a,b)$ *attains minima* $\left(2^{\frac{2}{p}-1} + 2^{\frac{1}{p}+1} + 2\right)r^2$ *when* $(a,b) = (r,r)$.

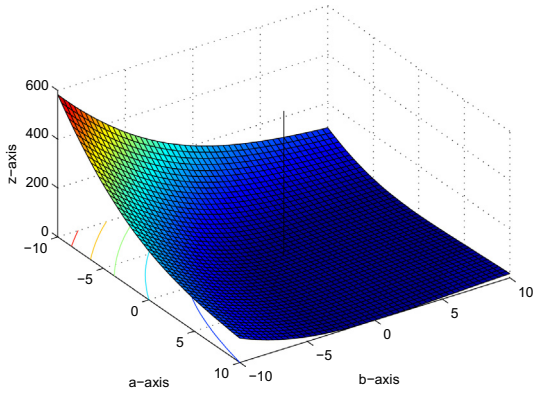


Fig. 10. The surface of $z = \psi_2(a,b)$ with $(a,b) \in [-10,10] \times [-10,10]$.

Proof

(a) When $a > 0$ and $b > 0$, Proposition 2.1(a) says that $\phi_p(a,b) < 0$. Since $\phi_p^2(a,b) > 0$, by Theorem 2.1, the minima of $\phi_p(a,b)$ becomes maxima of $\psi_p(a,b)$.
(b) This is a consequence of Theorem 2.1. □

The aforementioned results show ψ_p has many similar properties like ϕ_p hold, see Figs. 11 and 12, where we denote $\psi_1(a,b) := \frac{1}{2}|\phi_1(a,b)|^2$ and $\psi_\infty(a,b) = \frac{1}{2}|\phi_\infty(a,b)|^2$. However, there still are some differences between ϕ_p and ψ_p. For example, ψ_p is not convex whereas ϕ_p is. Fig. 13 depicts the increasing direction of ψ_p. Note that $\psi_p(a,b)$ is nonnegative and has different properties when $a > 0$ and $b > 0$, see Fig. 11.

In order to further understand the geometric properties, we define a family of curves as follows:

$$\beta_{r,p}(t) := \big(r+t, r-t, \psi_p(r+t, r-t)\big), \tag{13}$$

where r is a fixed real number, and $t \in \mathbb{R}$. This family of curves can be regarded as intersection of plane $a + b = 2r$ and surface $z = \psi_p(a,b)$, see Fig. 14.

Proposition 3.3. *Let $\beta_{r,p} : \mathbb{R} \to \mathbb{R}^3$ be defined as in (13). Then the following hold.*

(a) *The curvature at point $\beta_{r,p}(0) = (r, r, \psi_p(r,r))$ is $\bar{\kappa}_p(0) = (p-1)2^{\frac{1}{p}}\left(1 - 2^{\frac{1}{p}-1}\right)$.*
(b) *$\bar{\kappa}_p(0) \to 0$ as $p \to 1$ and $\bar{\kappa}_p(0) \to +\infty$ as $p \to +\infty$.*
(c) *If $1 < p_1 < p_2$, then $\bar{\kappa}_{p_1}(0) < \bar{\kappa}_{p_2}(0)$.*

Proof

(a) From $\beta_{r,p}(t) = \big(r+t, r-t, \psi_p(r+t, r-t)\big)$, we know

$$\beta'_{r,p}(0) = (1,-1,0) \quad \text{and} \quad \beta'_{r,p}(0) = \left(0, 0, (p-1)2^{\frac{2}{p}} - \operatorname{sgn}(r)(p-1)2^{\frac{1}{p}+1}\right),$$

which yields

$$\bar{\kappa}_p(r) = \frac{|\beta'_{r,p}(0) \wedge \beta'_{r,p}(0)|}{|\beta'_{r,p}(0)|^3} = (p-1)2^{\frac{1}{p}}\left(1 - 2^{\frac{1}{p}-1}\right).$$

(b) Let $f : (1,+\infty) \to \mathbb{R}$ be defined as $f(p) := \bar{\kappa}_p(0) = (p-1)2^{\frac{1}{p}}\left(1 - 2^{\frac{1}{p}-1}\right)$.
Then the result follows by taking the limit directly.
(c) From part (b), it can be verified that $f'(p) > 0$ for all $p \in (1,+\infty)$. Thus, $f(p)$ is strictly increasing on $(1,+\infty)$. □

Fig. 14 depicts the change of the curve when we have different value of p in which we can see the change of curvature when p is close to one or infinity. We state an addendum to part (a) here: the curvature at another two special points $\beta_{r,p}(-r) = (0, 2r, 0)$, $\beta_{r,p}(r) = (2r, 0, 0)$ is the same, namely, $\bar{\kappa}_p(r) = \bar{\kappa}_p(-r) = \frac{1}{2}$. Note that although ψ_p is differentiable everywhere, the mean curvature at $(0,0)$ does not exist. To end up this section, we summarize the similarity and difference between ϕ_p and ψ_p as below.

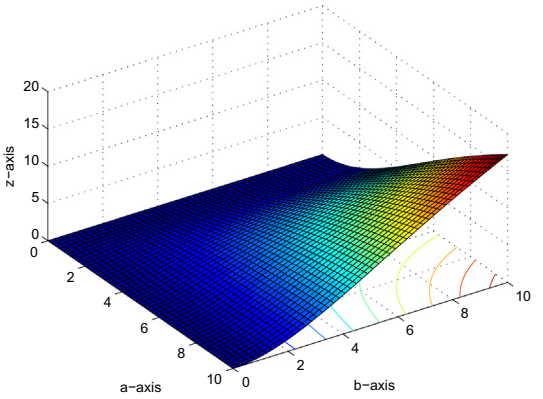


Fig. 11. The surface of $z = \psi_2(a, b)$ with $(a, b) \in [0, 10] \times [0, 10]$.

	$\phi_p(a, b)$	$\psi_p(a, b)$
Difference	Convex differentiable everywhere except $(0, 0)\phi_p(a, b) < 0$ when $a > 0$ and $b > 0$	Nonconvex differentiable everywhere $\psi_p(a, b) \geqslant 0, \forall (a, b) \in \mathbb{R}^2$
Similarity	(1) NCP-function (2) Symmetry (i.e. $\phi_p(a, b) = \phi_p(b, a)$ and $\psi_p(a, b) = \psi_p(b, a)$) (3) The function is not affected by p on axes (4) When $(a^k \to -\infty)$ or $(b^k \to -\infty)$ or $(a^k, b^k \to +\infty)$ there have $\lvert\phi_p(a^k, b^k)\rvert \to \infty$ and $\lvert\psi_p(a^k, b^k)\rvert \to \infty$ (5) Non-coercive	

4. Geometric analysis of merit function in descent algorithms

In this section, we employ derivative-free descent algorithms presented in [4,5] to solve the unconstrained minimization problem (11) by using the merit function (10). We then compare two algorithms and study their convergent behavior by investigating an intuitive visualization. We first list these two algorithms as below.

Algorithm 4.1 [4, Algorithm 4.1].

(Step 0) Given real numbers $p > 1$ and a starting point $x^0 \in \mathbb{R}^n$. Choose the parameters $\sigma \in (0, 1), \beta \in (0, 1)$ and $\varepsilon \geqslant 0$. Set $k := 0$.
(Step 1) If $\Psi_p(x^k) \leqslant \varepsilon$, then stop.
(Step 2) Let m_k be the smallest nonnegative integer m satisfying

$$\Psi_p(x^k + \beta^m d^k) \leqslant (1 - \sigma\beta^{2m})\Psi_p(x^k),$$

where

$$d^k := -\nabla_b \psi_p(x^k, F(x^k))$$

and

$$\nabla_b \psi_p(x, F(x)) := \left(\nabla_b \psi_p(x_1, F_1(x)), \ldots, \nabla_b \psi_p(x_n, F_n(x))\right)^T.$$

(Step 3) Set $x^{k+1} := x^k + \beta^{m_k} d^k$, $k := k + 1$ and go to Step 1.

Algorithm 4.2 [5, Algorithm 4.1].

(Step 0) Given real numbers $p > 1$ and $\alpha \geqslant 0$ and a starting point $x^0 \in \mathbb{R}^n$. Choose the parameters $\sigma \in (0, 1)$, $\beta \in (0, 1)$, $\gamma \in (0, 1)$ and $\varepsilon \geqslant 0$. Set $k := 0$.
(Step 1) If $\Psi_{\alpha,p}(x^k) \leqslant \varepsilon$, then stop.
(Step 2) Let m_k be the smallest nonnegative integer m satisfying

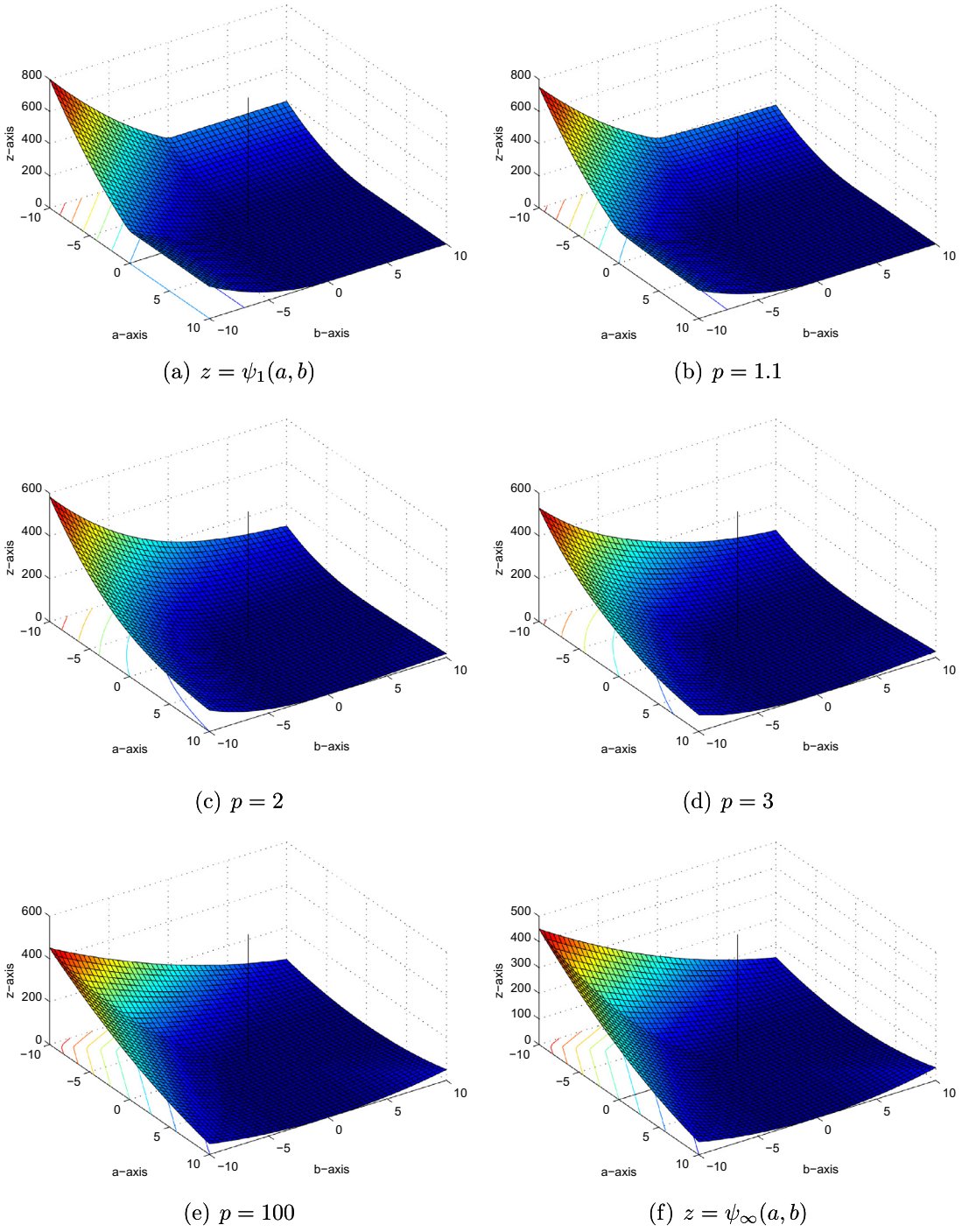


Fig. 12. The surface of $z = \phi_p(a, b)$ with different p.

$$\Psi_{\alpha,p}(x^k + \beta^m d^k(\gamma^m)) \leqslant (1 - \sigma\beta^{2m})\Psi_{\alpha,p}(x^k),$$

where

$$d^k(\gamma^m) := -\nabla_b \psi_{\alpha,p}(x^k, F(x^k)) - \gamma^m \nabla_a \psi_{\alpha,p}(x^k, F(x^k))$$

and

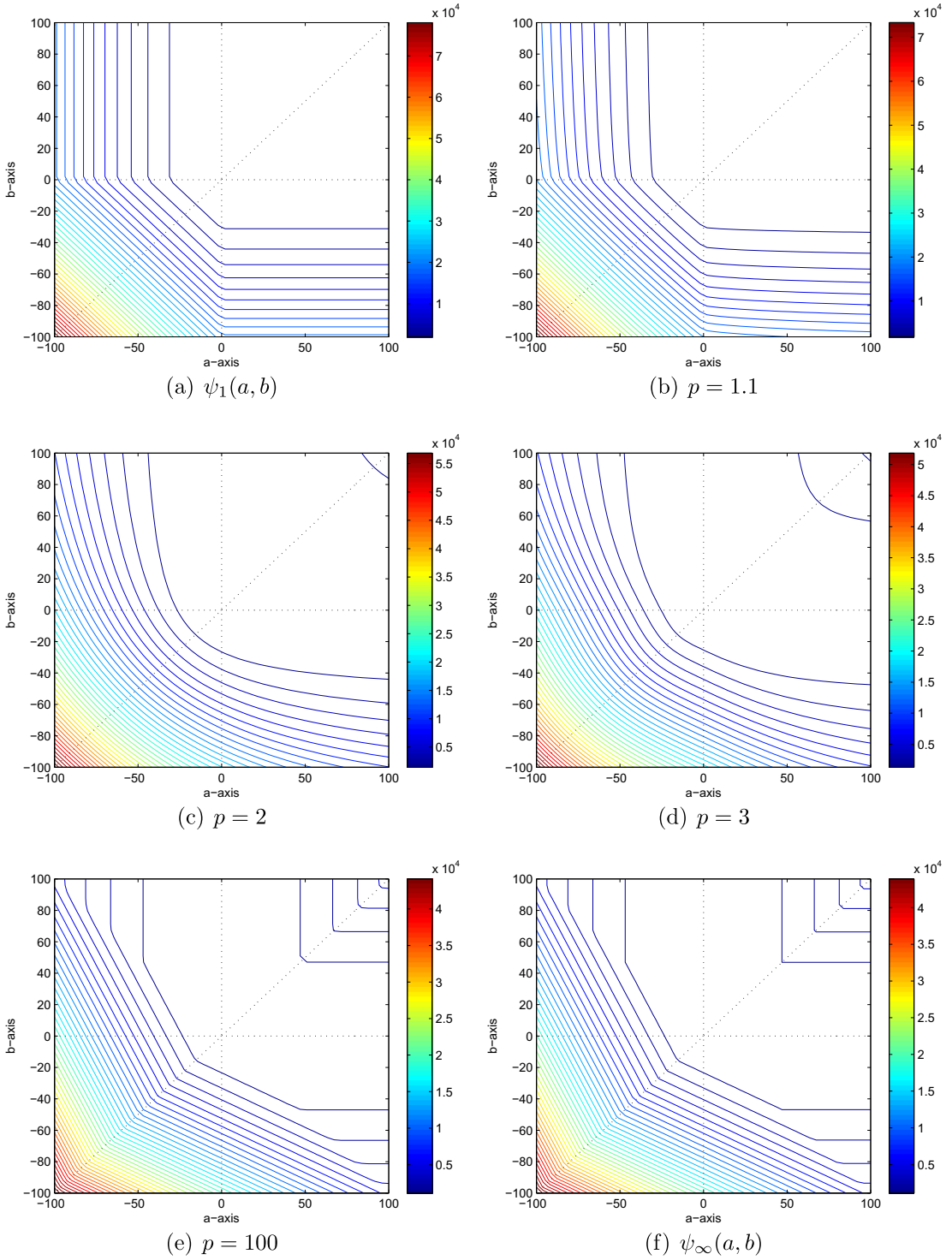
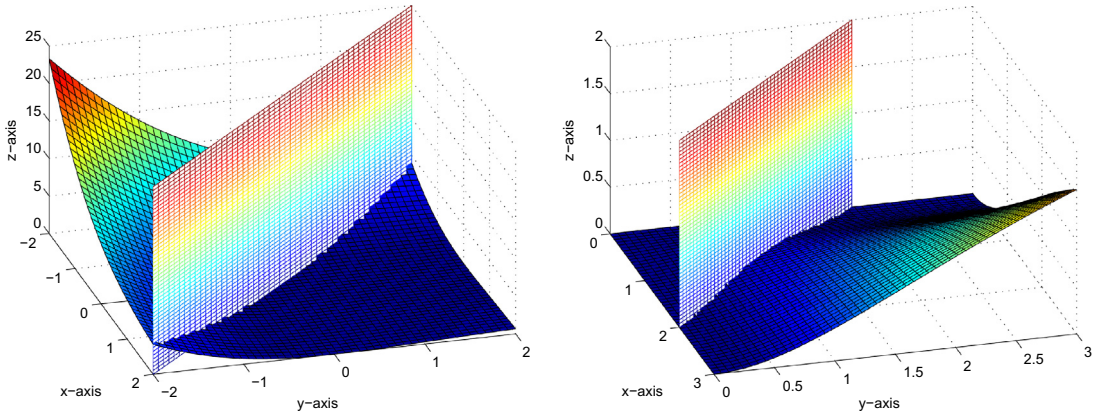


Fig. 13. Level curves of $z = \psi_p(a,b)$ with different p.

$$\nabla_a \psi_{\alpha,p}(x, F(x)) := \left(\nabla_a \psi_{\alpha,p}(x_1, F_1(x)), \ldots, \nabla_a \psi_{\alpha,p}(x_n, F_n(x))\right)^T,$$
$$\nabla_b \psi_{\alpha,p}(x, F(x)) := \left(\nabla_b \psi_{\alpha,p}(x_1, F_1(x)), \ldots, \nabla_b \psi_{\alpha,p}(x_n, F_n(x))\right)^T.$$

(Step 3) Set $x^{k+1} := x^k + \beta^{m_k} d^k(\gamma^{m_k})$, $k := k+1$ and go to Step 1.



(a) $a+b=0$ and $z=\psi_2(a,b)$ (b) $a+b=2$ and $z=\psi_2(a,b)$

Fig. 14. The curve intersected by surface $z=\psi_p(a,b)$ and plane $a+b=2r$.

In Algorithm 4.2, $\psi_{\alpha,p}:\mathbb{R}^2\to\mathbb{R}_+$ is an NCP-function defined by

$$\psi_{\alpha,p}(a,b):=\frac{\alpha}{2}(\max\{0,ab\})^2+\psi_p(a,b)=\frac{\alpha}{2}(ab)_+^2+\frac{1}{2}(\|(a,b)\|_p-(a+b))^2$$

with $\alpha\geqslant 0$ being a real parameter. When $\alpha=0$, the function $\psi_{\alpha,p}$ reduces to ψ_p. For comparing these two algorithms, we take $\alpha=0$ when we use Algorithm 4.2 in this section. Note that the descent direction in Algorithm 4.1 is lack of a certain symmetry whereas Algorithm 4.2 adopts a symmetric search direction. Under the assumption of monotonicity, i.e.,

$$\langle x-y,F(x)-F(y)\rangle\geqslant 0\quad\text{for all } x,y\in\mathbb{R}^n,$$

the error bound is proposed and Algorithm 4.2 is shown to have locally R-linear convergence rate in [5]. In other words, there exists a positive constant κ_2 such that

$$\|x^k-x^*\|\leqslant\kappa_2\left(\max\left\{\Psi_{\alpha,p}(x^k),\sqrt{\Psi_{\alpha,p}(x^k)}\right\}\right)^{\frac{1}{2}}\quad\text{when } \alpha=0.$$

Furthermore, the convergence rate of Algorithm 4.2 has a close relation with the constant

$$\left\lceil\log_{\frac{\gamma}{\beta}}\left(L_1+\frac{\sigma}{C(B,\alpha,p)}\right)\right\rceil\quad\text{where}\quad C(B,\alpha,p)=\frac{\left(2-2^{\frac{1}{p}}\right)^4}{\alpha B^2+\left(2+2^{\frac{1}{p}}\right)^2}.$$

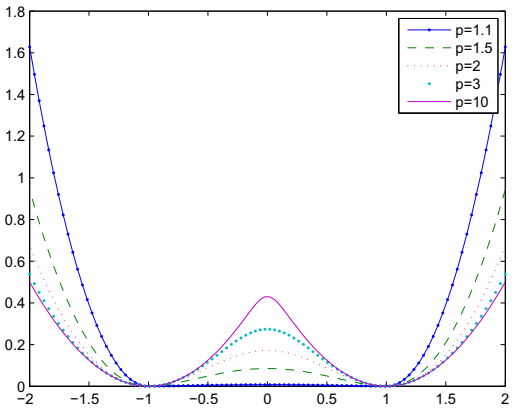
Therefore, when the value of p decreases, the convergence rate of Algorithm 4.2 becomes worse and worse, see Remark 4.1 in [5].

Recall that merit function $\Psi_p(x)$ is sum of n nonnegative functions ψ_p, i.e.,

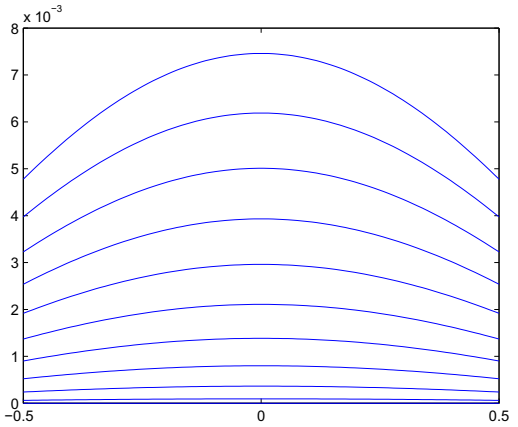
$$\Psi_p(x)=\sum_{i=1}^n\psi_p(x_i,F_i(x)).$$

This encourages us to view each component $\psi_p(x_i^k,F_i(x^k))$ for $i=1,2\ldots,n$ as the motion with different velocity on the same surface $z=\psi_p(a,b)$ at each iteration. Due to our study in Sections 2 and 3, we observe a visualization that help us understand the convergent behavior in details. Fig. 20 depicts the visualization in a four-dimensional NCP in Example 4.3. The merit function of this NCP is $\Psi_p(x)=\sum_{i=1}^4\psi_p(x_i,F_i(x))$. We plot point sequences $\{(x_i^k,F_i(x^k))\}$ for $i=1,2,3,4$ together with different color and level curve of surface $\psi_{1.1}(a,b)$ in Fig. 20(a). Vertical line represents value of x_i, horizontal line represents value of $F_i(x)$ and skew line means $x_i=F_i(x)$. We take initial point $x^0=(0,0,0,0)$ which implies $F(x^0)=(-6,-2,-1,-3)$, and observe convergent behavior separately with different i from initial point to the solution $x^*=(\sqrt{6}/2,0,0,1/2)$ which is on the horizontal line in this figure. Furthermore, we observe the position of point sequence on the surface in Fig. 20(a) and merit function which is the sum of their height at each iteration shown as in Fig. 20(b).

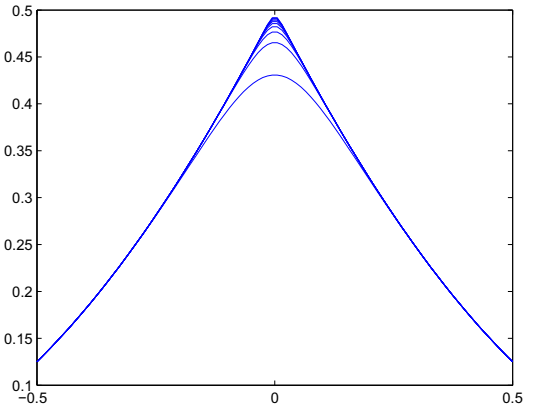
In one-dimensional NCP, F is continuously differentiable and there is only one variable x in F, so $(x,F(x))$ is continuous curve on $\mathbb{R}^2$ and merit function $\Psi_p(x)=\psi_p(x,F(x))$ is obviously a curve on the surface $z=\psi_p(a,b)$, see Fig. 16(a) and (b). Therefore, point sequence in one-dimensional problem can only lie on the curve $(x,F(x),\psi_p(x,F(x)))$.



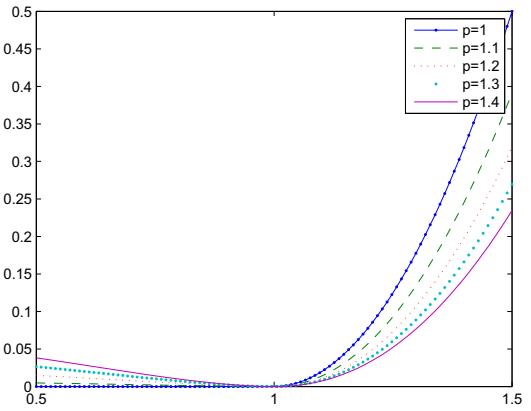
(a) The curvature with different p with $r = 1$



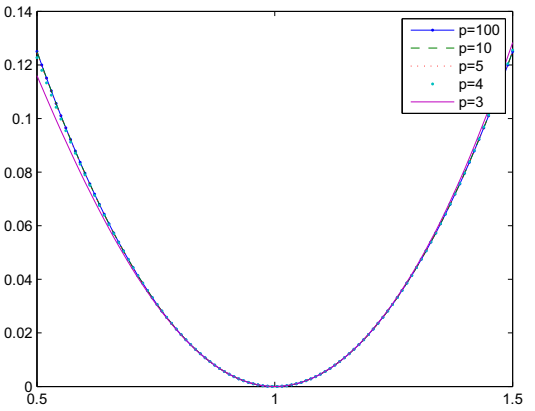
(b) The change of curvature as $p \to 1$ at $\beta_{1,p}(0)$



(c) The change of curvature as $p \to +\infty$ at $\beta_{1,p}(0)$.



(d) The change of curvature as $p \to 1$ at $\beta_{1,p}(1)$.

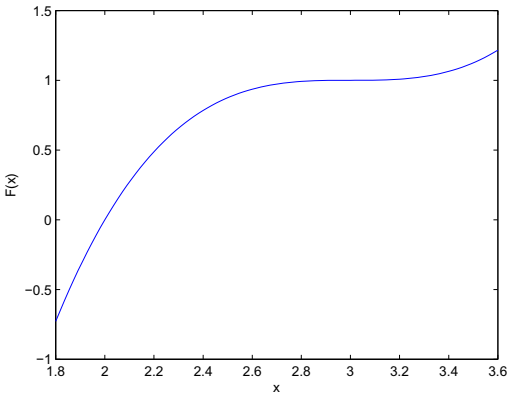


(e) The change of curvature as $p \to +\infty$ at $\beta_{1,p}(1)$.

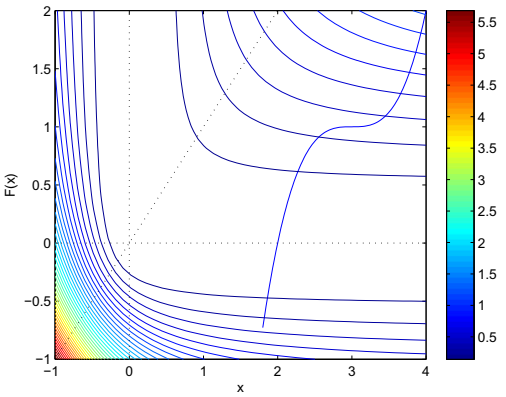
Fig. 15. The curvature $\bar{\kappa}_p(0)$ at point $\beta_{r,p}(0)$.

Example 4.1. Consider the NCP, where $F : \mathbb{R} \to \mathbb{R}$ is given by

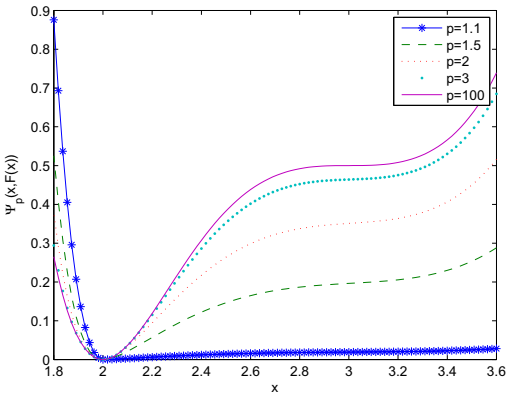
$$F(x) = (x-3)^3 + 1.$$



(a) Graph $F(x) = (x-3)^3 + 1$.



(b) Plane curve in (a) and level curve of surface $z = \psi_2(a,b)$ near the solution $(x^*, F(x^*)) = (2,0)$.



(c) The merit function with different p near the solution.

Fig. 16. Geometric view of NCP in Example 4.1.

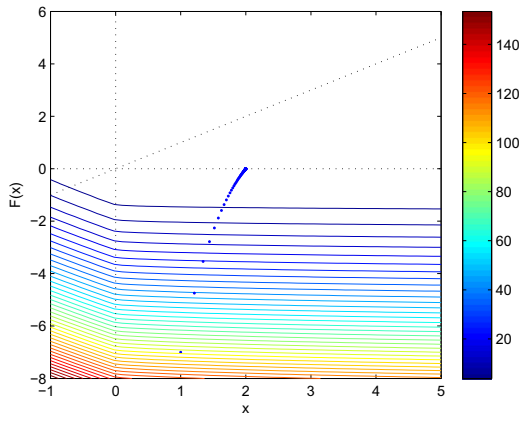
The unique solution of this NCP is $x^* = 2$. Note that F is strictly monotone, see geometric view of this NCP problem in Fig. 16. The value of merit function with each iteration is plotted in Fig. 16(c) which presents the different behavior of the functions with different value p near by the solution. Fig. 17(a)–(d) depict convergent behavior in Algorithm 4.1 from two direction with two different initial points, and Fig. 17(e) and (f) show convergent behavior with different p. Fig. 19(a)–(d) depict convergent behavior in Algorithm 4.2 from two direction with two different initial points. We found that Algorithm 4.2 always produce point sequence in or close to the boundary of feasible set, i.e., $\{(x, F(x)) : x \geqslant 0 \text{ and } F(x) \geqslant 0\}$. Based on Proposition 3.2, the speed of the decreasing of merit function with different initial point in Algorithm 4.1 is different when we increase p. But it is similar with different initial point in Algorithm 4.2. This phenomena is consistent with geometric properties studied in Section 3.

To show the importance of inflection point, we give an extreme example as follows:

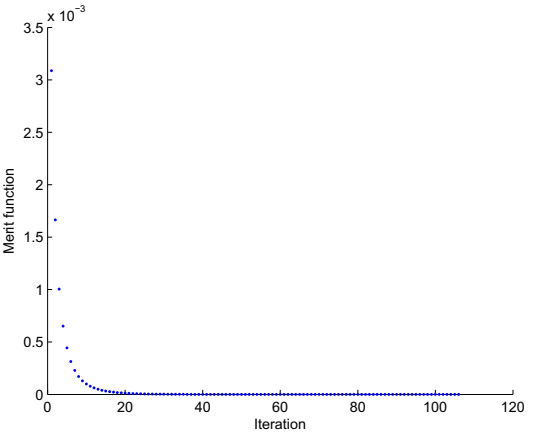
Example 4.2. Consider the NCP, where $F : \mathbb{R} \to \mathbb{R}$ is given by

$$F(x) = 1.$$

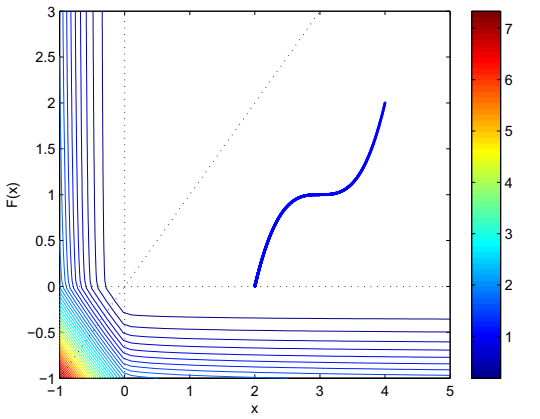
The unique solution of this NCP is $x^* = 0$. From above discussion, we know that point sequence is on the curve $(x, 1, \psi_p(x,1))$, see Fig. 18(a). Fig. 18(c) shows there is rapid decreasing of merit function form the 80th to 120th iteration. Fig. 18(b) shows the behavior during 80th to 120th iteration. Observing the width of the level curve in Fig. 18(b), we found that rapid decreasing may arise from the existence of inflection point on the surface. Figs. 18(c)–(f) and Fig. 19(e) and (f) show that the position of inflection point may change with different p.



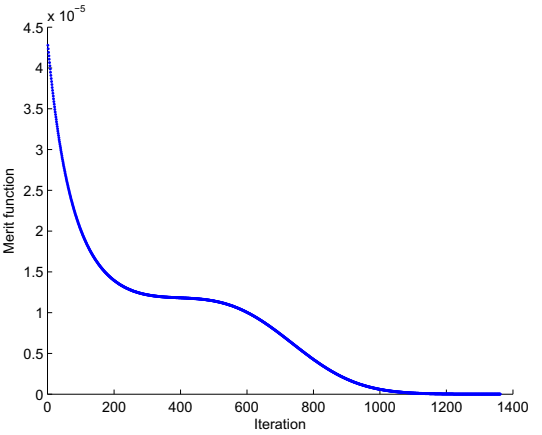
(a) Initial point $x^0 = 1$ and $p = 1.1$.



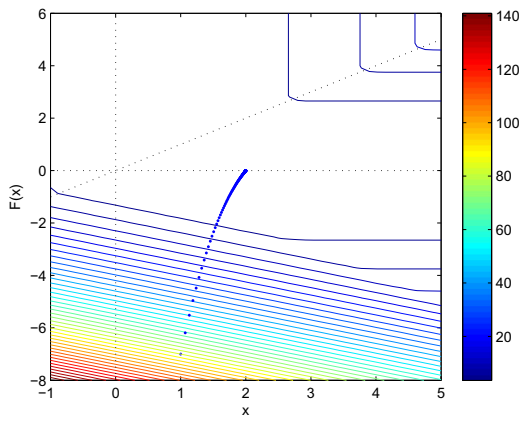
(b) Initial point $x^0 = 1$ and $p = 1.1$.



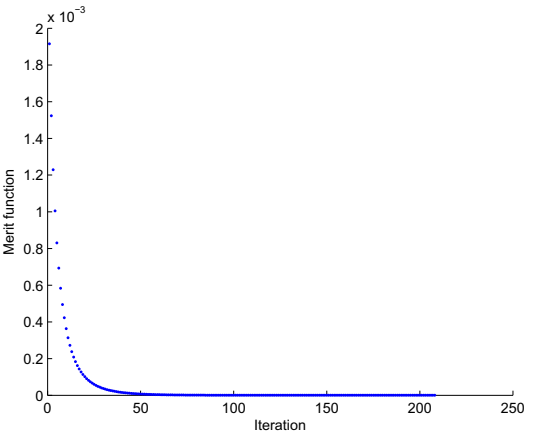
(c) Initial point $x^0 = 4$ and $p = 1.1$.



(d) Initial point $x^0 = 4$ and $p = 1.1$.



(e) Initial point $x^0 = 1$ and $p = 100$.

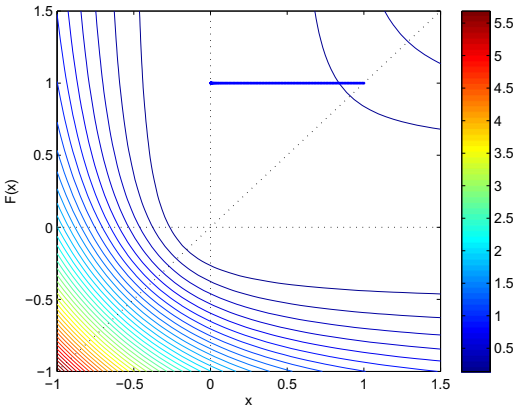


(f) Initial point $x^0 = 1$ and $p = 100$.

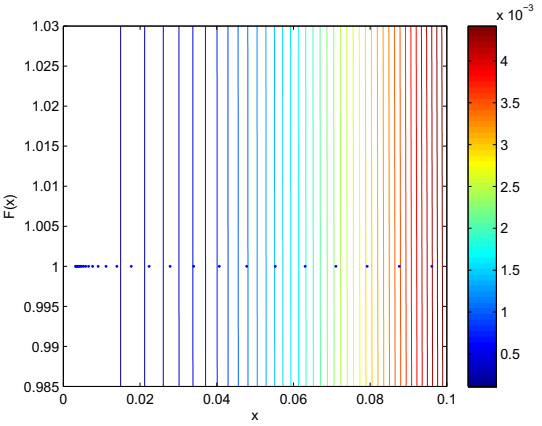
Fig. 17. Convergent behavior of Algorithm 4.1 and the value of merit function in Example 4.1.

Example 4.3. Consider the NCP, where $F : \mathbb{R}^4 \to \mathbb{R}^4$ is given by

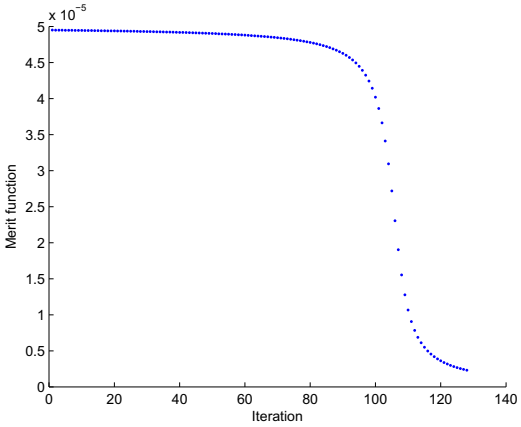
$$F(x) = \begin{pmatrix} 3x_1^2 + 2x_1x_2 + 2x_2^2 + x_3 + 3x_4 - 6 \\ 2x_1^2 + x_1 + x_2^2 + 3x_3 + 2x_4 - 2 \\ 3x_1^2 + x_1x_2 + 2x_2^2 + 2x_3 + 3x_4 - 1 \\ x_1^2 + 3x_2^2 + 2x_3 + 3x_4 - 3 \end{pmatrix}.$$



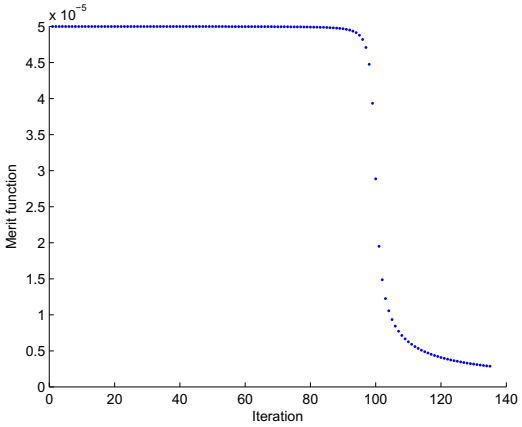
(a) $p = 2$.



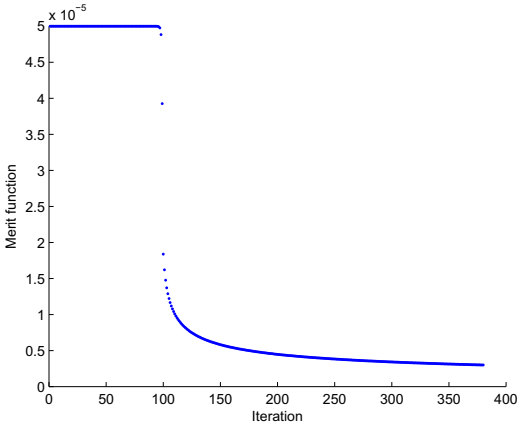
(b) Local convergent behavior of (a) with the 80th to 120th iteration of (c).



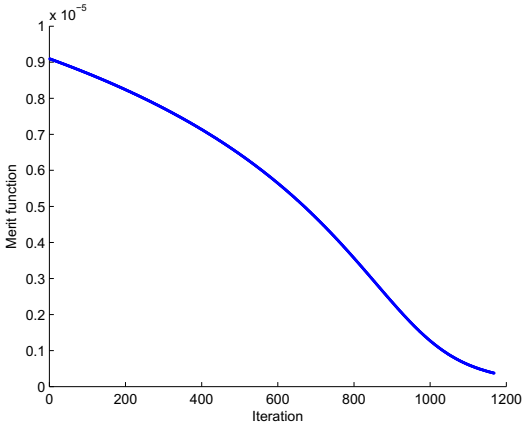
(c) $p = 2$.



(d) $p = 3$.



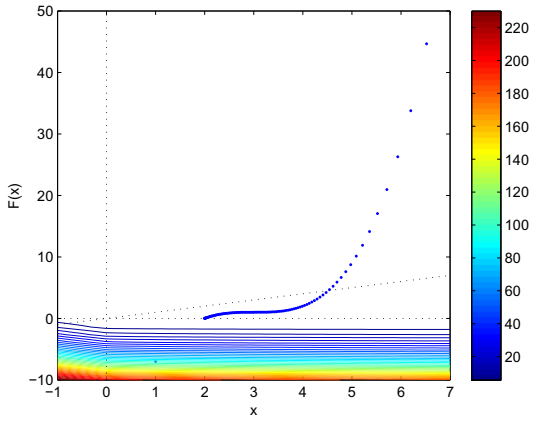
(e) $p = 5$.



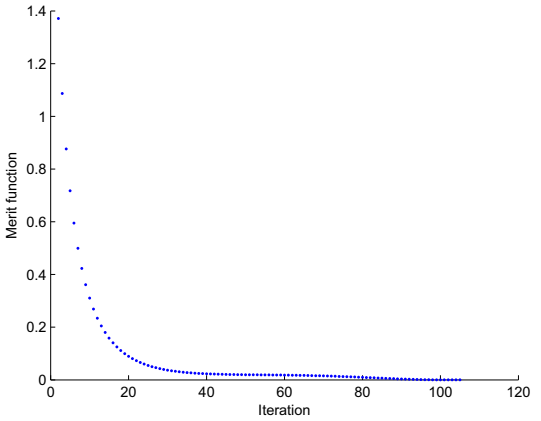
(f) $p = 1.1$.

Fig. 18. Convergent behavior of Algorithm 4.1 and merit function with initial point $x^0 = 1$ in Example 4.2.

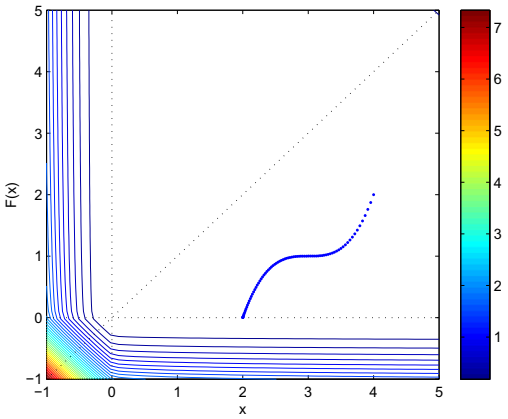
This is non-degenerate NCP and the solution is $x^* = \left(\sqrt{6}/2, 0, 0, 1/2\right)$. Fig. 20 shows that the behavior of merit function is consistent with Proposition 3.2(b) in Algorithm 4.1. Fig. 20 shows that the convergent behaviors are different by various initial points in Algorithm 4.1. Fig. 22 says that the convergent behavior is also different with different initial point in Algorithm 4.2. But we see that the behavior is still different between two algorithms even with the same initial point from Figs. 21 and



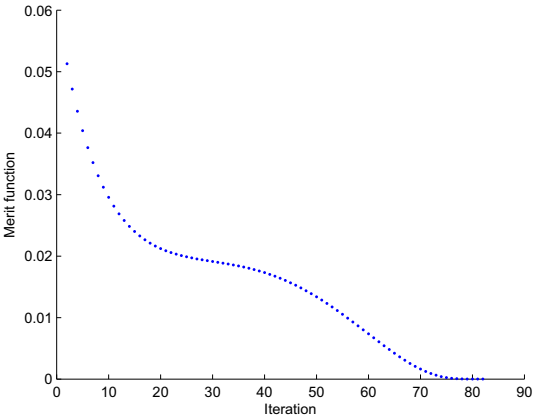
(a) Initial point $x^0 = 1$ and $p = 1.1$ in Example 4.1.



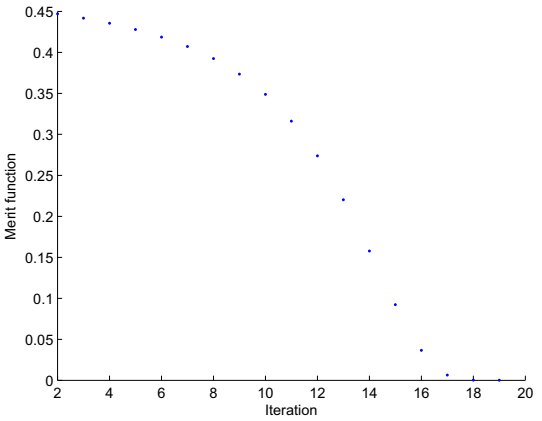
(b) Initial point $x^0 = 1$ and $p = 1.1$ in Example 4.1.



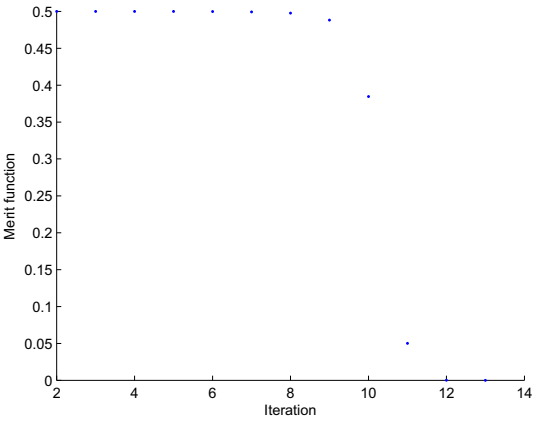
(c) Initial point $x^0 = 4$ and $p = 1.1$ in Example 4.1.



(d) Initial point $x^0 = 4$ and $p = 1.1$ in Example 4.1.



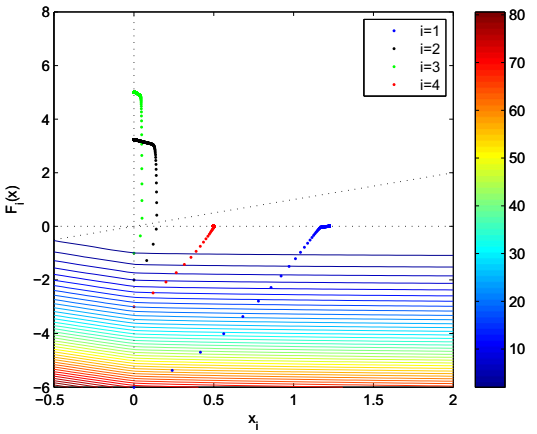
(e) Merit function in Example 4.2 with $p = 1.1$ and $x^0 = 10$.



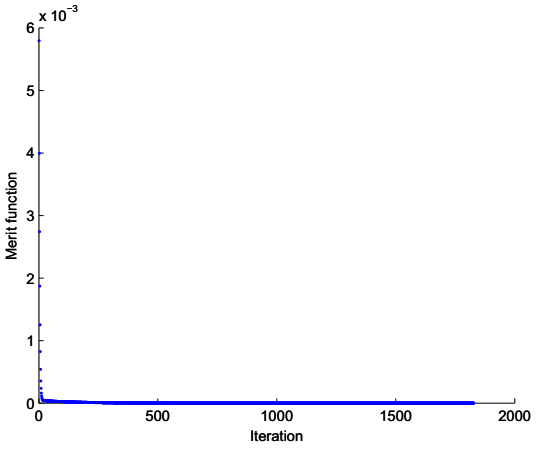
(f) Merit function in Example 4.2 with $p = 5$ and $x^0 = 10$.

Fig. 19. Convergent behavior of Algorithm 4.2 and the value of merit function in Examples 4.1 and 4.2.

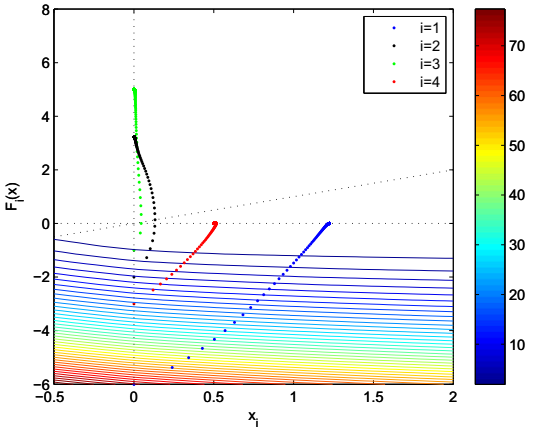
22. Fig. 23(e) shows that merit function decreases more and more quickly when p is smaller. However, Fig. 24(e) shows that merit function decreases more and more quickly when p is bigger.



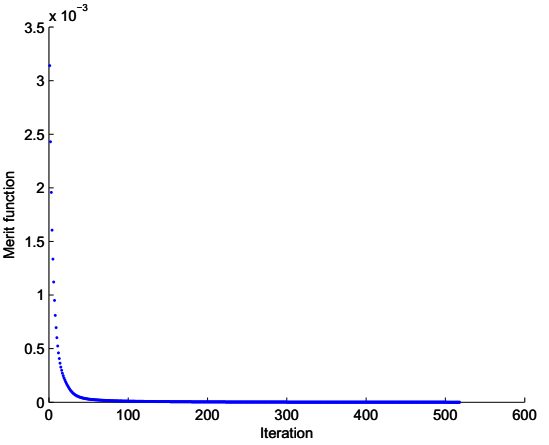
(a) $p = 1.1$.



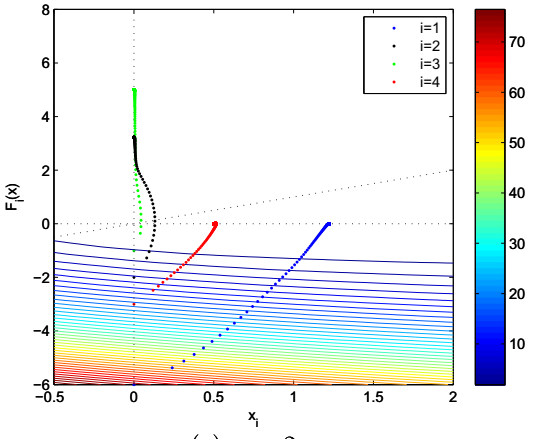
(b) Iteration and merit function $p = 1.1$.



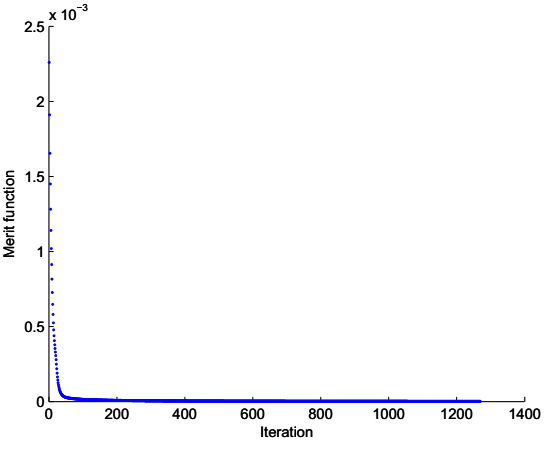
(c) $p = 1.5$.



(d) Iteration and merit function $p = 1.5$.

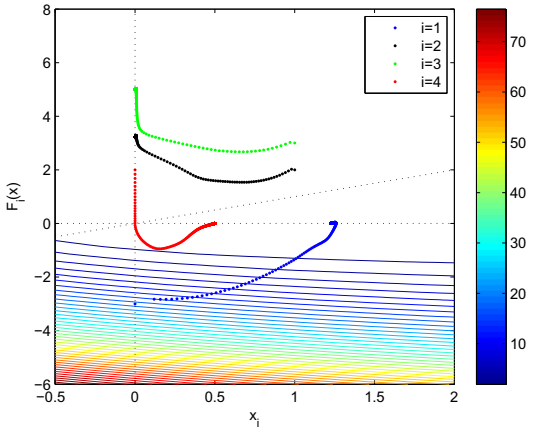


(e) $p = 2$.

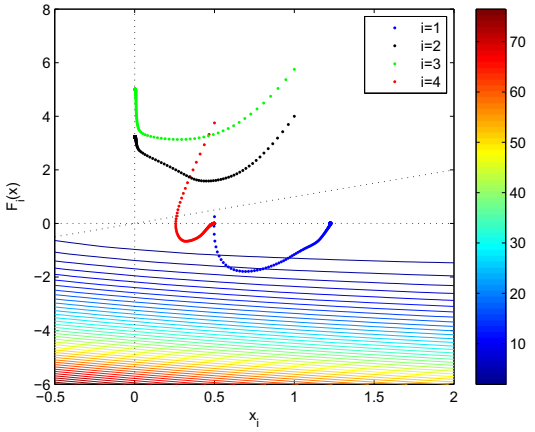


(f) Iteration and merit function $p = 2$.

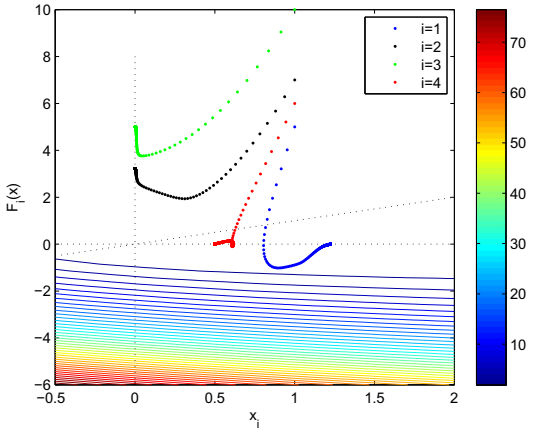
Fig. 20. Convergent behavior of Algorithm 4.1 and merit function with initial point $x^0 = (0, 0, 0, 0)$ in Example 4.3.



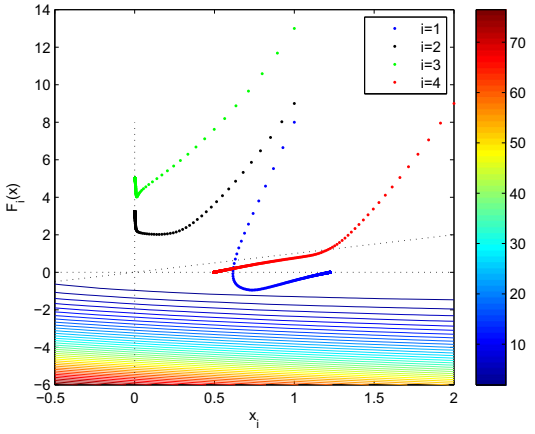
(a) $x^0 = (0, 1, 1, 0)$.



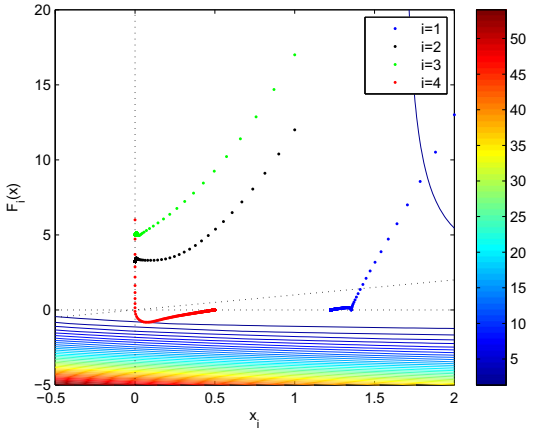
(b) $x^0 = (1/2, 1, 1, 1/2)$.



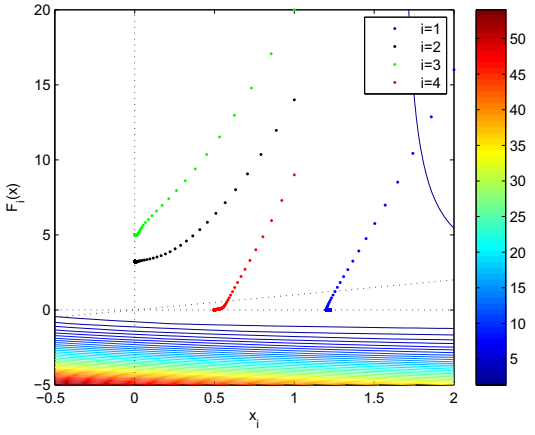
(c) $x^0 = (1, 1, 1, 1)$.



(d) $x^0 = (1, 1, 1, 2)$.

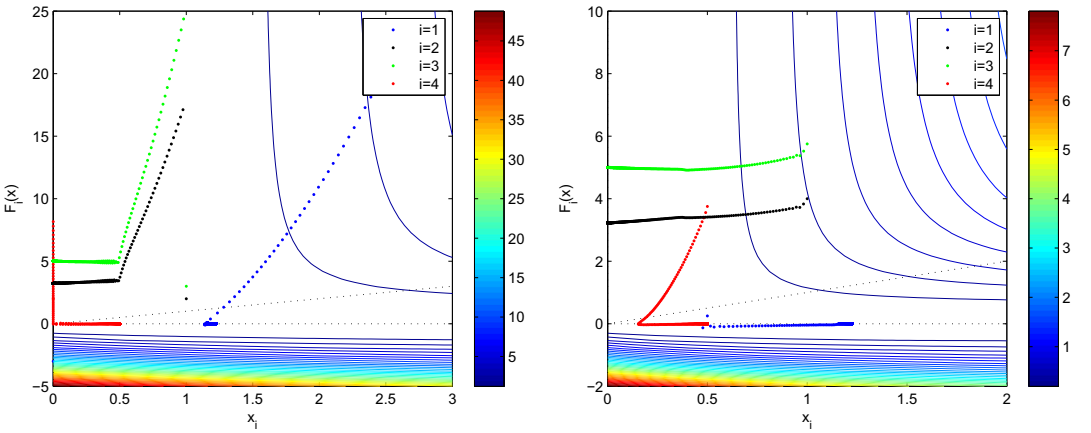


(e) $x^0 = (2, 1, 1, 0)$.

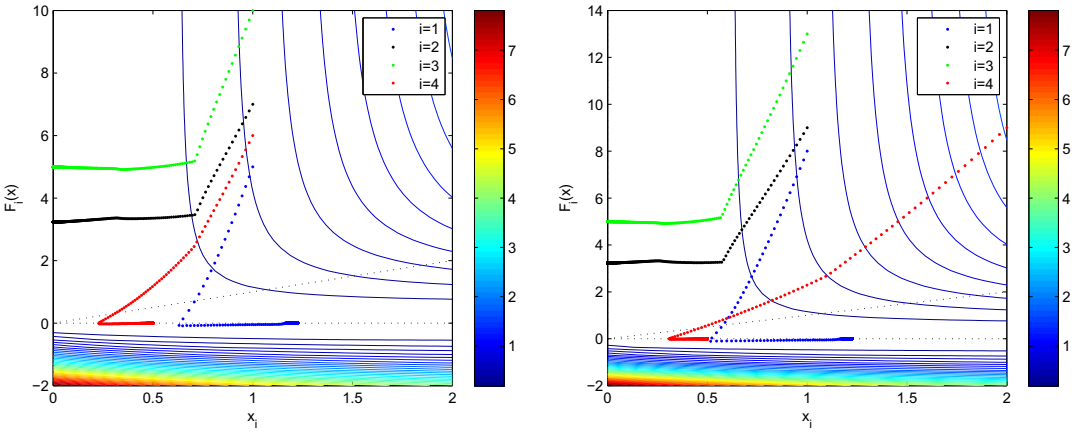


(f) $x^0 = (2, 1, 1, 1)$.

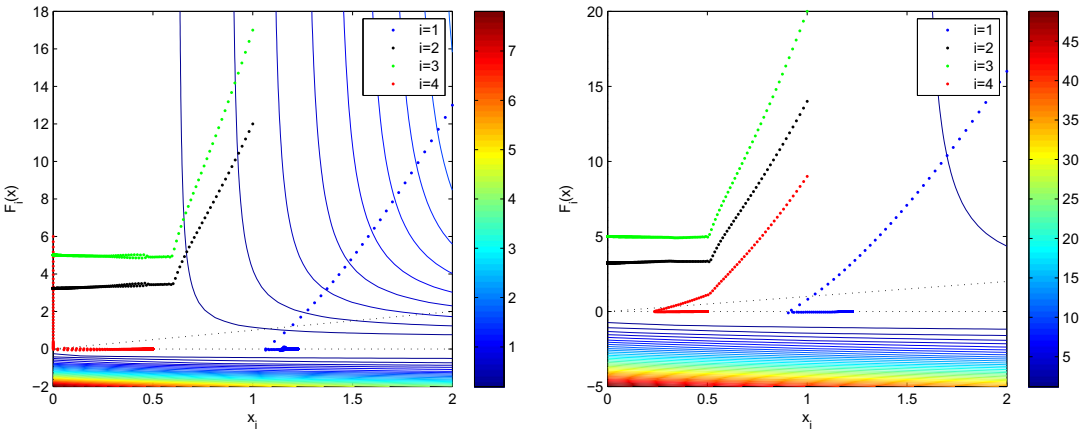
Fig. 21. Convergent behavior of Algorithm 4.1 with different initial point and $p = 2$ in Example 4.3.



(a) $x^0 = (0, 1, 1, 0)$. (b) $x^0 = (1/2, 1, 1, 1/2)$.

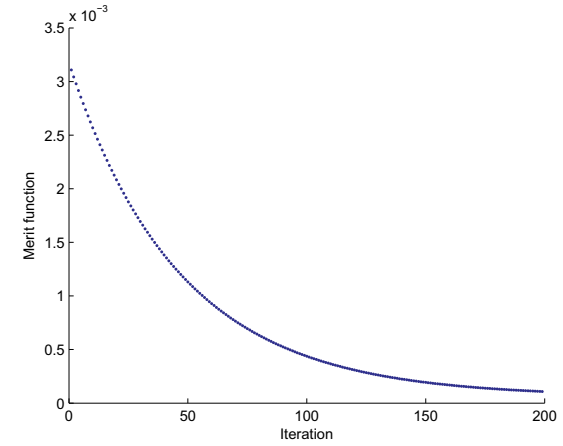


(c) $x^0 = (1, 1, 1, 1)$. (d) $x^0 = (1, 1, 1, 2)$.

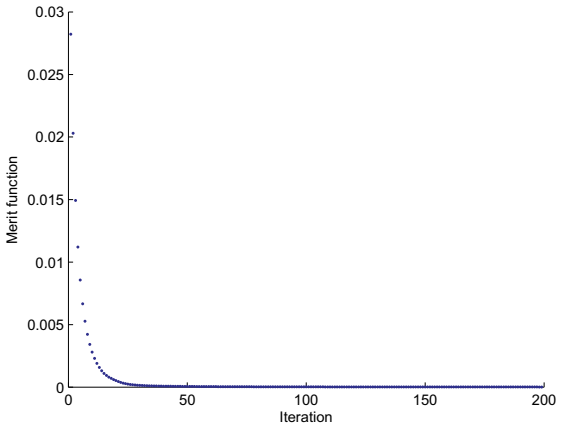


(e) $x^0 = (2, 1, 1, 0)$. (f) $x^0 = (2, 1, 1, 1)$.

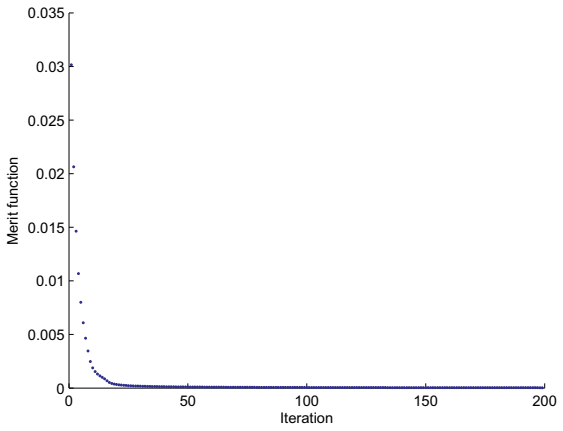
Fig. 22. Convergent behavior of Algorithm 4.2 with different initial point and $p = 2$ in Example 4.3.



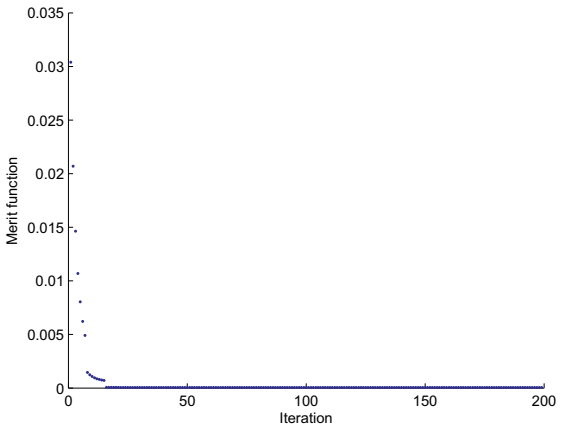
(a) $p = 1.1$.



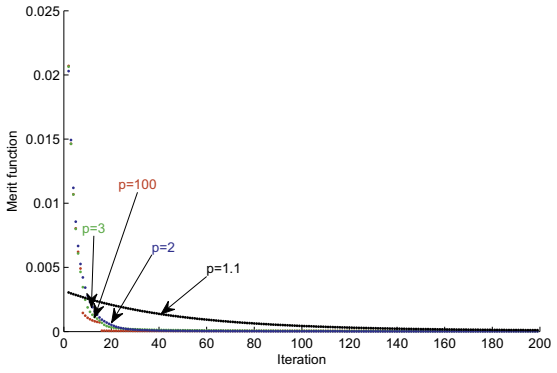
(b) $p = 2$.



(c) $p = 3$.

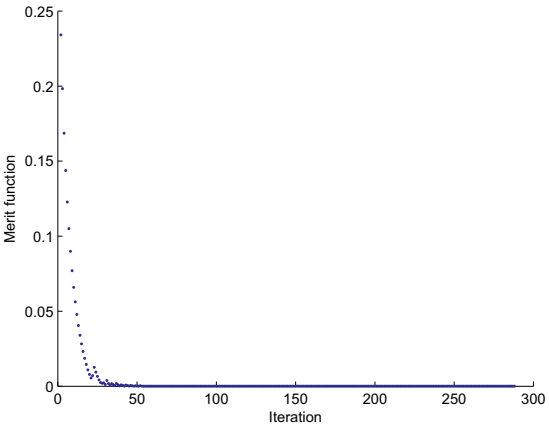


(d) $p = 100$.

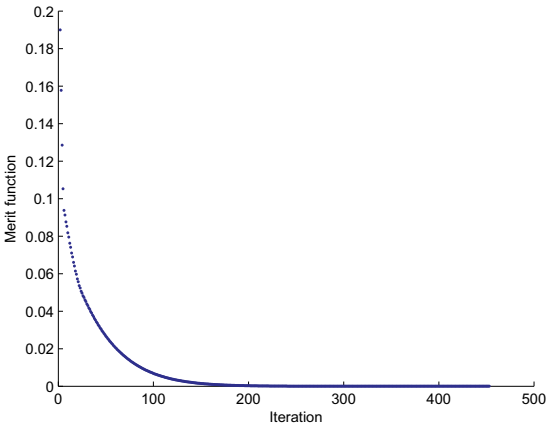


(e) All together.

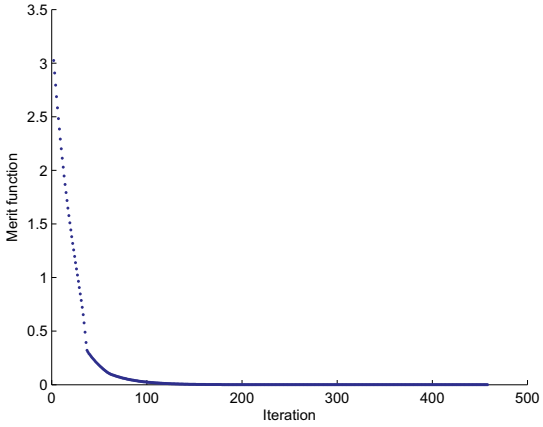
Fig. 23. Merit function in the first 200 iterations of Algorithm 4.1 with different p and initial point $x^0 = (2, 1, 1, 1)$ of Fig. 21(f).



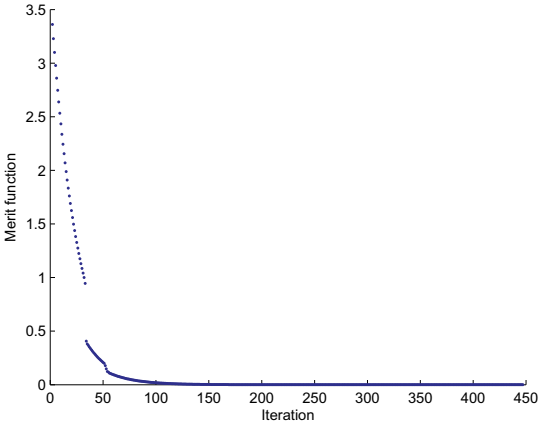
(a) $p = 1.1$.



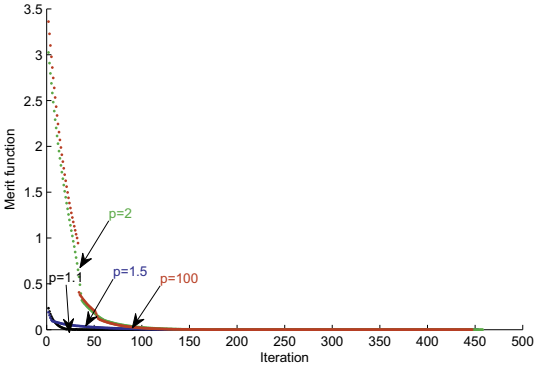
(b) $p = 1.5$.



(c) $p = 2$.



(d) $p = 100$.

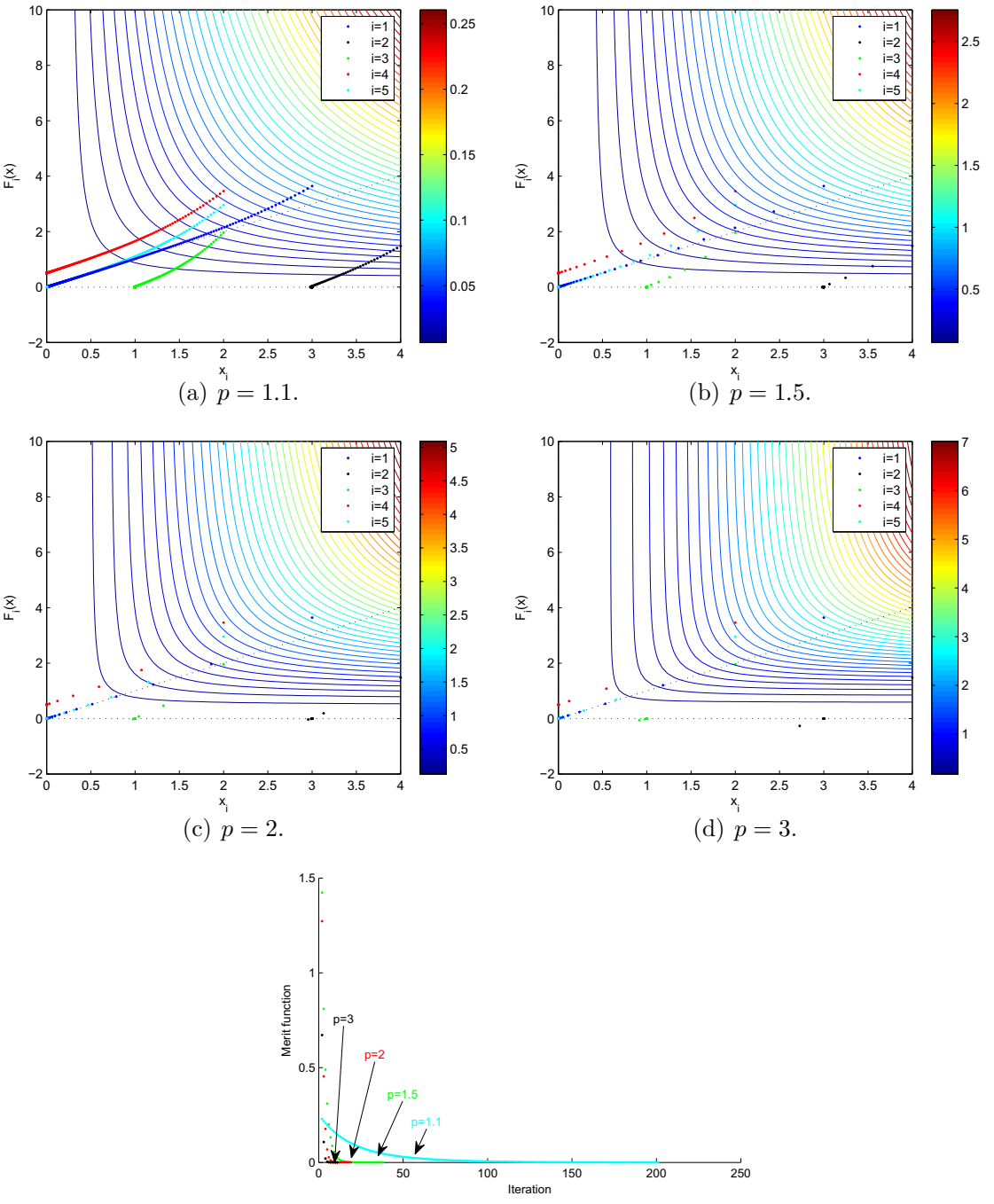


(e) All together.

Fig. 24. Merit function of Algorithm 4.2 with different p and initial point $x^0 = (2, 1, 1, 1)$ of Fig. 22(f).

Example 4.4. Consider the NCP, where $F : \mathbb{R}^5 \to \mathbb{R}^5$ is given by

$$F(x) = \begin{pmatrix} x_1 + x_2x_3x_4x_5/50 \\ x_2 + x_1x_3x_4x_5/50 - 3 \\ x_3 + x_1x_2x_4x_5/50 - 1 \\ x_4 + x_1x_2x_3x_5/50 + 1/2 \\ x_5 + x_1x_2x_3x_4/50 \end{pmatrix}.$$



(a) $p = 1.1$. (b) $p = 1.5$.

(c) $p = 2$. (d) $p = 3$.

(e) Merit function with different p.

Fig. 25. Convergent behavior with different p and merit function of Algorithm 4.2 with initial point $x^0 = (3, 4, 2, 2, 2)$ in Example 4.4.

This NCP has only one solution $x^* = (0, 3, 1, 0, 0)$. We choose initial point $x^0 = (3, 4, 2, 2, 2)$ in Algorithm 4.2, see Fig. 25. Figs. 25(e) and 24(e) show different result with different p.

The results of above examples suggest that the convergent behavior is influenced by the position of initial point, properties of $F(x)$, and the geometric structure of the NCP function ψ_p. Indeed, the convergent behavior of Algorithm 4.1 can be classified into three cases when starting from different initial points.

- Case 1: the point sequence $\{(x_i^k, F_i(x^k))\}$ $i = 1, 2, \ldots, n$ is almost located in or close to the boundary of surface $z = \psi_p(a, b)$ where $a > 0$ and $b > 0$, see Fig. 20(f).

- Case 2: the point sequence $\{(x_i^k, F_i(x^k))\}$ $i = 1, 2, \ldots, n$ is almost located in or close to the boundary of surface $z = \psi_p(a, b)$ where $a > 0$ and $b < 0$, see Fig. 16(a) and (e).
- Case 3: the point sequence does not belong to cases 1 and 2, see Fig. 20(a).

In addition, the value of merit function decreases more quickly when the value of p increases in case 1, see Fig. 23. The value of merit function decreases more quickly when the value of p increases in case 2, see Fig. 20. The value of merit function seems to depend on the slope of the surface, as visualization shows. Thus, the above two cases can be explained by Proposition 3.2 geometrically. Although the convergent behavior is complicated in case 3, if there exists some initial point x^0 such that $F_i(x^0) < 0$ for some i and p is small, for example, $p \in (1, 2)$, we can easily deduce that the value of merit function decreases more quickly when the value of p increases, like case 2 and see Fig. 17. This is because the surface $z = \psi_p(a, b)$ is much higher when $a > 0$ and $b < 0$ than when $a > 0$ and $b > 0$, see Figs. 13 and 15(d). Therefore the value of merit function is dominated by the component of $F(x)$ with initial point satisfy $F_i(x^0) < 0$ for some i. Such observations match up with those convergence results discussed in [4].

Convergent behavior in Algorithm 4.2 belongs to case 1, see Figs. 19(a) and 21. The behavior of merit function seems to depend on the height scale of surface with steady step length at each iteration, as visualization shows. Therefore surface in case 1 is closer to zero but becomes flatter when p is smaller. This is the geometrical reason why Algorithm 4.2 has a better a global convergence and a worse convergence rate when p decreases, see concluding remark in [5] and Fig. 24(e).

According to [4,5], Algorithm 4.2 has different global convergence result compared to Algorithm 4.1. However, this phenomena depends on initial point. In fact, we can choose suitable initial point by observing our visualization such that these two algorithms having similar global convergence result with different p, see Figs. 25(e) and 23(e). This is another benefit of analyzing visualization.

5. Conclusion

In this paper, we view graphs of function ϕ_p and function ψ_p as families of surfaces in $\mathbb{R}^3$ and study their geometric properties. It is shown that when using two descent methods to solve NCP with merit function Ψ_p, the convergent behaviors are influenced by descent directions and geometric structure of surface of ψ_p. By looking at the visualization, we observe that the existence of inflection point (in the graphs of ψ_p) is the main factor causes rapid decreasing ("cliff") of merit function. In addition, the convergent behavior of both algorithms is sensitive to initial points which also matches what is remarked in [5,7]. The proposed geometric view and visualization greatly help us realize that how convergent behavior is influenced by changing or adding parameters in NCP-functions.

References

[1] S.C. Billups, S.P. Dirkse, M.C. Soares, A comparison of algorithms for large scale mixed complementarity problems, Comput. Optim. Appl. 7 (1997) 3–25.
[2] J.-S. Chen, The semismooth-related properties of a merit function and a descent method for the nonlinear complementarity problem, J. Global Optim. 36 (2006) 565–580.
[3] J.-S. Chen, On some NCP-functions based on the generalized Fischer–Burmeister function, Asia-Pac. J. Oper. Res. 24 (2007) 401–420.
[4] J.-S. Chen, S.-H. Pan, A family of NCP-functions and a descent method for the nonlinear complementarity problem, Comput. Optim. Appl. 40 (2008) 389–404.
[5] J.-S. Chen, H.-T. Gao, S.-H. Pan, An R-linearly convergent derivative-free algorithm for the NCPs based on the generalized Fischer–Burmeister merit function, J. Comput. Appl. Math. 232 (2009) 455–471.
[6] J.-S. Chen, Z.-H. Huang, C.-Y. She, A new class of penalized NCP-functions and its properties, Comput. Optim. Appl. 50 (2011) 49–73.
[7] J.-S. Chen, C.-H. Ko, S.-H. Pan, A neural network based on generalized Fischer–Burmeister function for nonlinear complementarity problems, Inf. Sci. 180 (2010) 697–711.
[8] R.W. Cottle, J.-S. Pang, R.-E. Stone, The Linear Complementarity Problem, Academic Press, New York, 1992.
[9] S. Dafermos, An iterative scheme for variational inequalities, Math. Program. 26 (1983) 40–47.
[10] F. Facchinei, J. Soares, A new merit function for nonlinear complementarity problems and a related algorithm, SIAM J. Optim. 7 (1997) 225–247.
[11] F. Facchinei, J.S. Pang, Finite-Dimensional Variational Inequalities and Complementary Problems, vol. I and II, Springer, New York, 2003.
[12] A. Fischer, A special Newton-type optimization methods, Optimization 24 (1992) 269–284.
[13] A. Fischer, Solution of the monotone complementarity problem with locally Lipschitzian functions, Math. Program. 76 (1997) 513–532.
[14] M. Fukushima, Merit functions for varitional inequality and complementarity problem, in: G. Di Pillo, F. Giannessi (Eds.), Nonlinear Optimization and Applications, Pleneum Press, New York, 1996, pp. 155–170.
[15] C. Geiger, C. Kanzow, On the resolution of monotone complementarity problems, Comput. Optim. Appl. 5 (1996) 155–173.
[16] P.T. Harker, J.-S. Pang, Finite dimensional variational inequality and nonlinear complementarity problem: a survey of theory, algorithms and applications, Math. Program. 48 (1990) 161–220.
[17] N.J. Higham, Estimating the matrix p-norm, Numer. Math. 62 (1992) 539–555.
[18] H. Jiang, Unconstrained minimization approaches to nonlinear complementarity problems, J. Global Optim. 9 (1996) 169–181.
[19] C. Kanzow, Nonlinear complementarity as unconstrained optimization, J. Optim. Theory Appl. 88 (1996) 139–155.
[20] C. Kanzow, N. Yamashita, M. Fukushima, New NCP-functions and their properties, J. Optim. Theory Appl. 94 (1997) 115–135.
[21] O.L. Mangasarian, Equivalence of the complementarity problem to a system of nonlinear equations, SIAM J. Appl. Math. 31 (1976) 89–92.
[22] J.-S. Pang, Complementarity problems, in: R. Horst, P. Pardalos (Eds.), Handbook of Global Optimization, Kluwer Academic Publishers, Boston, MA, 1994, pp. 271–338.
[23] J.-S. Pang, Newton's method for B-differentiable equations, Math. Oper. Res. 15 (1990) 311–341.
[24] J.-S. Pang, D. Chan, Iterative methods for variational and complemantarity problems, Math. Program. 27 (99) (1982) 284–313.
[25] D. Sun, L.-Q. Qi, On NCP-functions, Comput. Optim. Appl. 13 (1999) 201–220.
[26] P. Tseng, Growth behavior of a class of merit functions for the nonlinear complementarity problem, J. Optim. Theory Appl. 89 (1996) 17–37.
[27] N. Yamashita, M. Fukushima, On stationary points of the implicit lagrangian for the nonlinear complementarity problems, J. Optim. Theory Appl. 84 (1995) 653–663.

[28] N. Yamashita, M. Fukushima, Modified newton methods for solving a semismooth reformulation of monotone complementarity problems, Math. Program. 76 (1997) 469–491.
[29] K. Yamada, N. Yamashita, M. Fukushima, A new derivative-free descent method for the nonlinear complementarity problems, in: G.D. Pillo, F. Giannessi (Eds.), Nonlinear Optimization and Related Topics, Kluwer Academic Publishers, Netherlands, 2000, pp. 463–487.