Scientific Research

Neural Network Approach for Solving Singular Convex Optimization with Bounded Variables[*]

Rendong Ge[#], Lijun Liu, Yi Xu

School of Science, Dalian Nationalities University, Dalian, China
Email: [#]bgrbgg@163.com, liulijun@dlnu.edu.cn, xuyi@dlnu.edu.cn

ABSTRACT

Although frequently encountered in many practical applications, singular nonlinear optimization has been always recognized as a difficult problem. In the last decades, classical numerical techniques have been proposed to deal with the singular problem. However, the issue of numerical instability and high computational complexity has not found a satisfactory solution so far. In this paper, we consider the singular optimization problem with bounded variables constraint rather than the common unconstraint model. A novel neural network model was proposed for solving the problem of singular convex optimization with bounded variables. Under the assumption of rank one defect, the original difficult problem is transformed into nonsingular constrained optimization problem by enforcing a tensor term. By using the augmented Lagrangian method and the projection technique, it is proven that the proposed continuous model is convergent to the solution of the singular optimization problem. Numerical simulation further confirmed the effectiveness of the proposed neural network approach.

Keywords: Neural Networks; Singular Nonlinear Optimization; Stationary Point; Augmented Lagrangian Function; Convergence; LaSalle's Invariance Principle Plain

1. Introduction

The nonlinear model with rank one defect is of great importance for its singular nature. Follwing works of Schnabel and Dan Feng [1-3], we have made great improvement for such problem by applying Tensor methods by numerical solution [4,5]. For large-scale computational problems, the computation of the classical numerical method is still far from satisfactory.

In recent years, neural network approaches were proposed to deal with classical nonlinear optimization problems. Xia and Wang [6] presented neural networks for solving nonlinear convex optimization with bounded constraints and box constraints, respectively. Xia [7,8], Xia and Wang [9,10] developed several neural networks for solving linear and quadratic convex programming problems, monotone linear complementary problems, and a class of monotone variational inequality problems. Recently, projection neural networks for solving monotone variational inequality problems are developed in

[11-13] and recurrent neural networks for solving nonconvex optimization problem have been also studied [14,15]. It is regrettable that the study of singular nonlinear optimization problems in the neural network method have not been involved .

Recently, more attention were paid to the singular optimization problems due to real applications. For example, in the problem of singular optimal control, assume the state equation is depicted as

$$\frac{\mathrm{d}x}{\mathrm{d}t} = F(x,u,t)$$

where x is an n-dimensional state vector and u is an m-dimensional control vector, and the control piecewise functions satisfy that $|u_j| \le M_j, j = 1,2,\cdots,m$. The cost functional is given as

$$J = \Xi\left(x\left(t_f\right),t_f\right) + \int_{t_0}^{t_f} L(x,u,t)\mathrm{d}t$$

for which the Hamiltonian function is

$$H = L(x,u,t) + \lambda^{\mathrm{T}} F(x,u,t).$$

According to the maximum's principle, when the control variables change in the constrain boundary, the

minimum conditions of the optimal control function H are derived as

$$\frac{\partial H}{\partial u} = 0, \frac{\partial^2 H}{\partial^2 u} \geq 0$$

If on a given time interval $[t_1, t_2] \subseteq [t_0, t_f]$, there exists that $\det\left(\frac{\partial^2 H}{\partial^2 u}\right) = 0$, then this becomes the so-called singular optimal control problem. The optimal trajectory corresponding to the segment called singular arc. The numerical methods for solving such singular control problem can be referred to [16,17].

Due to the inherent Parallel mechanism and high-speed of hardware implementations, efforts to tackle such problems by using neural systems are promising and creative. Attempt was made for the first time in our recent paper [18] to solve unconstraint singular optimization problem, which turned out to be feasible and effecttive. In this paper, we further improve such result to the case of singular optimization problem with bounded variables constraint.

This paper is organized as follows. In Section 2, the nonlinear singular convex optimization problem and its equivalent formulations are described. In Section 3, a recurrent neural network model is proposed to solve such singular nonlinear optimization problems. Global convergence of the proposed neural network is analyzed. Finally, in Section 4, several illustrative examples are presented to evaluate the effectiveness of the proposed neural network method.

2. Problem Formulation and Neural Design

Let $\Omega = \{x \in R^n \mid l \leq x \leq h\}$. Assume $f(x): \Omega \to R$, is a continuous differentiable convex function. Consider the following unconstrained convex programming problem,

$$\begin{aligned}\min \quad & f(x) \\ \text{s.t.} \quad & l \leq x \leq h\end{aligned} \tag{1}$$

which can be easily transformed to equivalent non-negative bounded convex programming problem by using the such transformation as $u = x - l$,

$$\begin{aligned}\min \quad & f(x) \\ \text{s.t.} \quad & 0 \leq x \leq c.\end{aligned} \tag{2}$$

Let x^* is the unique optimal solution to (2). We will discuss the solution of (2) under the following assumptions.

Assumption 1 $f(x)$ *is both strictly convex and four times continuous differentiable. For optimum point x^*, there exists $v \in R^n$ such that $\operatorname{rank}\left(\nabla^2 f(x^*)\right) = n - 1$ and $Null\left(\nabla^2 f(x^*)\right) = \{v\}$.*

Assumption 2 *For $x \neq x^*$, there exists $u^T \nabla^2 f(x) u > 0$*

for any nonzero $u \in R^n$. Moreover, $\left\|\nabla^2 f(x)\right\|$ and $\left\|\nabla^3 f(x)\right\|$ are all uniformly bounded.

Assumption 3 *For any $v \in Null\left(\nabla^2 f(x^*)\right)$, the quantity*

$$\nabla^4 f(x^*) v^4 \triangleq v^T \nabla^2 \left(v \nabla^2 f(x^*) v\right) v > 0.$$

(The reason for this assumption can be found, for example, in [4].)

Lemma 1 *For any $p \in R^n$ and $p^T v \neq 0$, $v \in Null\left(\nabla^2 f(x^*)\right)$, $\left(\nabla^2 f(x) + pp^T\right)$ is nonsingular at x^*.*

Proof. It is easy to verify this result, thus its proof is omitted here for the sake of saving space.

It is seen that when p in Lemma 2.1 take random values, the condition $p^T v \neq 0$ is satisfied with probability 1.

Define function $F(x)$ as follows

$$F(x) = f(x) + \lambda h(x),$$

where

$$h(x) = \mu(x) \nabla^2 f(x) \mu(x),$$

$$\mu(x) = \left(\nabla^2 f(x) + pp^T\right)^{-1} p$$

where $p^T v \neq 0$.

According to the definition of $F(x)$, we have the following important result,

Lemma 2 *For any $\lambda > 0$, the Hessian matrix $\nabla^2 F(x^*)$ is positive definite. Moreover, if λ is small enough, then $\nabla^2 F(x)$ is positive definite for any $x \in R^n$.*

Proof. This conclusion can be proved easily according to the results in [4] under Assumption A2. Thus the proof is omitted here.

Because the Hessian matrix of $f(x)$ is singular at x^* for (2), it is generally difficult to obtain ideal convergence results by conventional optimization algorithm (see [1-4]). In order to overcome this difficulty, alternatively we deal with equivalent unconstrained convex optimization problem as follows,

$$\begin{aligned}\min \quad & F(x) \\ \text{s.t.} \quad & 0 \leq x \leq c,\end{aligned} \tag{3}$$

for which we can establish the equivalent lemma as follows,

Lemma 3 *x^* is a solution of (2) if and only if x^* is a solution of (3).*

Further consider the difficulty caused by computing the matrix inverse, we turn optimization question (3) into the following equivalent constrained optimization problem.

$$\begin{aligned}\min \quad & g(x, y) = f(x) + \lambda y^T \nabla^2 f(x) y \\ \text{s.t.} \quad & \left(\nabla^2 f(x) + pp^T\right) y = p \quad 0 \leq x \leq c.\end{aligned} \tag{4}$$

Define Lagrange function of (4) as follows,

$$\tilde{L}(x,y,z) = f(x) + \lambda y^{\mathrm{T}} \nabla^2 f(x) y$$
$$+ z^{\mathrm{T}} \left[\left(\nabla^2 f(x) + pp^{\mathrm{T}} \right) y - p \right]$$

By Assumptions A2 and Lemma 2, it is easy to know that the function $g(x,y)$ is strictly convex. And based on the Karush-Kuhn-Tucker sufficient conditions, the KKT point $(\hat{x}, \hat{y})$ of the formula (4) is a unique optimal solution of the optimization question (4) and there exits $\hat{z} \in R^n$ satisfies the following condition:

$$\begin{cases} \left(\nabla_x \tilde{L}(x,y,z) \right)_i \begin{cases} \geq 0, & \text{if } x_i = 0, \\ \leq 0, & \text{if } x_i = c_i, \quad (i = 1,2,\cdots,n) \\ = 0, & \text{if } 0 < x_i < c_i, \end{cases} \\ \nabla_y \tilde{L}(x,y,z) = 0, \\ \nabla_z \tilde{L}(x,y,z) = 0, \\ x \in \Omega = \left\{ x \in R^n \middle| 0 \leq x \leq c \right\} \end{cases}$$

Equivalently, the point $(\hat{x}, \hat{y}, \hat{z})$ satisfies the following condition,

$$\begin{cases} (x - \hat{x})^{\mathrm{T}} \left(\nabla_x \tilde{L}(\hat{x}, \hat{y}, \hat{z}) \right) \geq 0, & x \in \Omega \\ \nabla_y \tilde{L}(\hat{x}, \hat{y}, \hat{z}) = 0, \\ \nabla_z \tilde{L}(\hat{x}, \hat{y}, \hat{z}) = 0. \end{cases} \quad (5)$$

In order to discuss the constrained programming problem (4), first, we define a augmented Lagrangian function of (4) as follows

$$L(x,y,z,k) = f(x) + \lambda y^{\mathrm{T}} \nabla^2 f(x) y$$
$$+ z^{\mathrm{T}} \left[\left(\nabla^2 f(x) + pp^{\mathrm{T}} \right) y - p \right]$$
$$+ \frac{k}{2} \left\| \left(\nabla^2 f(x) + pp^{\mathrm{T}} \right) y - p \right\|^2, \quad x \in \Omega,$$

where $k > 0$ is a penalty parameter and z is an approximation of the Lagrange multiplier vector. Hence, the problem (4) can be solved by the stationary point of the following problem,

$$\min_{x \in \Omega, y, z \in R^n} L(x,y,z,k) \quad (6)$$

Then, the condition (5) can be written as

$$\begin{cases} (x - \hat{x})^{\mathrm{T}} \nabla_x L(\hat{x}, \hat{y}, \hat{z}, k) \geq 0, & x \in \Omega, \\ \nabla_y L(\hat{x}, \hat{y}, \hat{z}, k) = 0, \\ \nabla_z L(\hat{x}, \hat{y}, \hat{z}, k) = 0. \end{cases} \quad (7)$$

Now, we introduce the projection function P_Ω as follows,

$$R^n \to \Omega, P_\Omega(u) = \left(P_1(u), P_2(u), \cdots, P_n(u) \right)$$

where

$$P_i(u) = \begin{cases} 0, & u_i < 0, \\ u_i, & u_i \in [0, c_i], \\ c_i, & u_i > c_i. \end{cases} \quad (8)$$

From the projection conclusion as shown in [19], the first inequality of (7) can be equivalently represented as

$$P_\Omega \left(\hat{x} - \alpha \nabla_x L(\hat{x}, \hat{y}, \hat{z}, k) \right) - \hat{x} = 0, \forall \alpha > 0.$$

So the optimal solution of (4) and the stationary point of (6) meet with the conditions

$$\begin{cases} x = P_\Omega \left(x - \alpha \nabla_x L(x,y,z,k) \right), \forall \alpha > 0, \\ \nabla_y L(x,y,z,k) = 0, \\ \left(\nabla_x^2 f(x) + pp^{\mathrm{T}} \right) y = p. \end{cases} \quad (9)$$

3. Stability Analysis of the Neural Network Model

By Theorem 8 and Theorem 9 in [20], there exists a constant $k > 0$, such that if $c^* = (x^*, y^*, z^*)$ is an optimal solution of the problem (6), then (x^*, y^*) is an optimal solution of the problem (4) and

$$\min_{x \in \Omega, y, z \in R^n} L(x,y,z,k) = f(x^*).$$

Notice that $L(x,y,z) \equiv L(x,y,z,k)$. By the Lagrange function defined above, we can describe the neural network model by the following nonlinear dynamic system for solving (10). The logical graph is shown in **Figure 1**.

$$\begin{cases} \dfrac{\mathrm{d}x}{\mathrm{d}t} = P_\Omega \left(x - \alpha \nabla x L(x,y,z) \right) - x \\ \quad = P_\Omega \Big(x - \alpha \big(\nabla f(x) + \lambda \nabla^3 f(x) yy + \nabla^3 f(x) yz \\ \quad\quad + k \nabla^3 f(x) y \big(\left(\nabla^2 f(x) + pp^{\mathrm{T}} \right) y - p \big) \big) \Big) - x \\ \dfrac{\mathrm{d}v}{\mathrm{d}t} = -\beta \nabla y L(x,y,z) \\ \quad = -\beta \Big(2\lambda \nabla^2 f(x) y + \left(\nabla^2 f(x) + pp^{\mathrm{T}} \right) z \\ \quad\quad + k \left(\nabla^2 f(x) + pp^{\mathrm{T}} \right) \big(\left(\nabla^2 f(x) + pp^{\mathrm{T}} \right) y - p \big) \Big) \\ \dfrac{\mathrm{d}w}{\mathrm{d}t} = -\beta \nabla z L(x,y,z) = -\beta \big(\left(\nabla^2 f(x) + pp^{\mathrm{T}} \right) y - p \big) \\ y_j = s(v_j), j = 1,2,\cdots,n \\ z_k = s(w_k), k = 1,2,\cdots,n \end{cases}$$
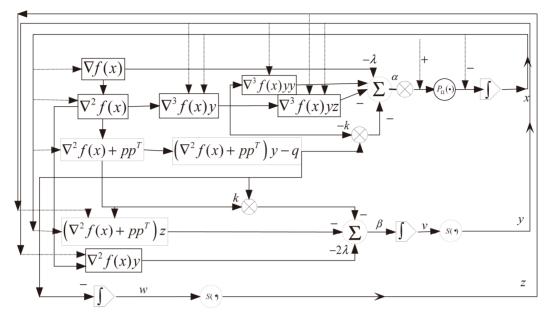
(10)

where

Figure 1. Logical graph of the proposed neural network model.

$$\nabla f(x) = \big(f_1(x), f_2(x), \cdots, f_n(x)\big)^{\mathrm{T}}, \alpha > 0, \beta > 0$$

$$\nabla^3 f(x) y = \big(\nabla^2 f_1(x) y, \nabla^2 f_2(x) y, \cdots, \nabla^2 f_n(x) y\big)^{\mathrm{T}},$$

and the activation function $s(\cdot)$ is continuously differentiable and satisfies that $s'(\cdot) > 0$.

It is easy to see that if $c^* = \big(x^*, y^*, z^*\big)$ is an optimal solution of the problem (6), then it is a equilibrium point of network (10). Conversely, if $\big(x^*, y^*, z^*\big)$ is a equilibrium point of network (10), it must be KKT point of original problem (4). To analyze the convergence of the neural network (10), the following lemmas are first introduced (see [19]).

Lemma 4 *Assume that the set* $\Omega \subset R^n$ *is a closed convex set, then the following two inequalities hold,*

$$\big(P_\Omega(x') - y'\big)^{\mathrm{T}}\big(x - P_\Omega(x')\big) \ge 0, \ \forall x' \in R^n, \ y' \in \Omega$$

$$\big\| P_\Omega(x') - P_\Omega(y') \big\| \le \big\| x' - y' \big\|, \ \forall x', y' \in R^n$$

where $P_\Omega : R^n \to \Omega$ is a projection operator defined as $P_\Omega(\gamma) = \min_{\varsigma \in \Omega} \| \gamma - \varsigma \|$.

Lemma 5 *For any initial point* $\big(x(t_0), v(t_0), w(t_0)\big) \in R^{3n}$, *there exists a unique continuous solution* $\big(x(t), v(t), w(t)\big) \in R^{3n}$ *for (10). Moreover,* $x(t) \in \Omega$ *provided that* $x(t_0) \in \Omega$. *The equilibrium point of* (10) *solves* (5).

Proof: By Assumption A1, $P_\Omega\big(x - \alpha\nabla_x L(x, y, z, k)\big) - x$, $\nabla_y L(x, y, z, k)$ and $\big(\nabla^2 f(x) + pp^{\mathrm{T}}\big)y - p$ are locally Lipschitz continuous. According to local existence theorem of ordinary differential equation, there exists a unique continuous solution $\big(x(t), v(t), w(t)\big)$ of (10) for (t_0, T).

Next, let initial point $x(t_0) \in \Omega$. Since

$$\frac{\mathrm{d}x}{\mathrm{d}t} + x = P_\Omega\big(x - \alpha\nabla x L(x, y, z)\big),$$

we have

$$\int_{t_0}^{t}\bigg(\frac{\mathrm{d}x}{\mathrm{d}t} + x\bigg)\mathrm{e}^s \mathrm{d}s = \int_{t_0}^{t} \mathrm{e}^s P_\Omega\big(x - \alpha\nabla x L(x, y, z)\big)\mathrm{d}s$$

Or equivalently,

$$x(t) = \mathrm{e}^{-(t-t_0)}x(t_0) + \mathrm{e}^{-t}\int_{t_0}^{t} \mathrm{e}^s P_\Omega\big(x - \alpha\nabla x L(x, y, z)\big)\mathrm{d}s.$$

So, $x(t) \ge 0$ provided that $x(t_0) \ge 0$, $P_\Omega\big(x - \alpha\nabla_x L(x, y, z, k)\big) \ge 0$, and since

$$x(t) = \mathrm{e}^{-(t-t_0)}x(t_0) + \mathrm{e}^{-t}\int_{t_0}^{t} \mathrm{e}^s P_\Omega\big(x - \alpha\nabla x L(x, y, z)\big)\mathrm{d}s$$

$$\le \mathrm{e}^{-(t-t_0)}x(t_0) + \mathrm{e}^{-t}\big(\mathrm{e}^t - \mathrm{e}^{t_0}\big)c$$

$$= c - \big(c - x(t_0)\big)\mathrm{e}^{-(t-t_0)}$$

$$\le c$$

Thus, $x(t) \in \Omega$ provided that $x(t_0) \in \Omega$.

Before establishing the convergence theorem, we need the property of the following augmented Lagrangian function.

Assumption 4 $L(x, y, z)$ *satisfies the local monotone property of following definition about* x.

$$\big(x - x^*\big)^{\mathrm{T}}\big(\nabla L_x(x, y, z) - \nabla_x L\big(x^*, y^*, z^*\big)\big) \ge 0.$$

Now we are ready to establish the stability and the convergence results of network (10).

Theorem 6 Assume that $f(x) : R^n \to R$ is strictly convex and the fourth differentiable, and $c^* = \big(x^*, y^*, z^*\big)$ is a global optimal solution of the problem (6), if the initial point $\big(x(t_0), y(t_0), z(t_0)\big)$

with $x(t_0) \in \Omega$ is chosen in a small neighborhood of the equilibrium point, then the proposed neural network of (10) is stable in the sense of Lyapunov and globally convergent to the stationary point (x^*, y^*, z^*), where x^* is the optimal solution of (2).

Proof. We define $V: \Omega \to R$ as follows:

$$V(x(t), y(t), z(t))$$
$$= L(x(t), y(t), z(t)) - f(x^*) + \frac{1}{2\alpha}\|x(t) - x^*\|^2,$$

We want to show that V is a suitable Lypunov function for dynamic system (10), it is evident that

$$V(x(t), y(t), z(t)) > \frac{1}{2\alpha}\|x(t) - x^*\|^2,$$
$$(x(t), y(t), z(t)) \neq (x^*, y^*, z^*),$$
$$V(x^*, y^*, z^*) = 0.$$

Then

$$\frac{\mathrm{d}V}{\mathrm{d}t} = \sum_{i=1}^{n}\left(\frac{\partial L}{\partial u_i} \cdot \frac{\mathrm{d}x_i}{\mathrm{d}t} + \frac{\partial L}{\partial y_i} \cdot \frac{\mathrm{d}y_i}{\mathrm{d}t} + \frac{\partial L}{\partial z_i} \cdot \frac{\mathrm{d}z_i}{\mathrm{d}t}\right)$$
$$+ (x - x^*)^{\mathrm{T}} P_\Omega\left((x - \alpha\nabla_x L(x, y, z)) - x\right)\big/\alpha$$
$$= \sum_{i=1}^{n}\left(\frac{\partial L}{\partial x_i} \cdot \frac{\mathrm{d}x_i}{\mathrm{d}t} + \frac{\partial L}{\partial y_i} \cdot \frac{\mathrm{d}y_i}{\mathrm{d}v_i} \cdot \frac{\mathrm{d}v_i}{\mathrm{d}t} + \frac{\partial L}{\partial z_i} \cdot \frac{\mathrm{d}z_i}{\mathrm{d}w_i} \cdot \frac{\mathrm{d}w_i}{\mathrm{d}t}\right) \quad (11)$$
$$+ (x - x^*)^{\mathrm{T}} P_\Omega\left((x - \alpha\nabla_x L(x, y, z)) - x\right)\big/\alpha.$$

Consequently,

$$\left(P_\Omega(x - \nabla_x L(x, y, z)) - x\right)^{\mathrm{T}} \nabla_x L(x, y, z)$$
$$\leq -\left\|P_\Omega(x - \nabla_x L(x, y, z))\right.$$
$$\left. - x\right\|^2\big/a - (x - x^*)^{\mathrm{T}} \nabla_x L(x, y, z) \quad (13)$$
$$+ (x - x^*)^{\mathrm{T}}\left(x - P_\Omega(x - \nabla_x L(x, y, z))\right)\big/\alpha.$$

Then, we obtain from (12), (13) and Assumption A4

$$\frac{\mathrm{d}V}{\mathrm{d}t}(x, y, z) \leq -\left\|P_\Omega(x - \nabla_x L(x, y, z)) - x\right\|^2\big/\alpha$$
$$- (x - x^*)^{\mathrm{T}} \nabla_x L(x, y, z)$$
$$- \beta\left[\nabla_y L(x, y, z)\right]^{\mathrm{T}} G_v' \nabla_y L(x, y, z) \quad (14)$$
$$- \beta\left[\nabla_z L(x, y, z)\right]^{\mathrm{T}} G_w' \nabla_z L(x, y, z)$$
$$\leq 0,$$

$$\frac{\mathrm{d}V}{\mathrm{d}t}(x^*, y^*, z^*) = 0 \quad (15)$$

And denote that

$$G_v' = diag\left(s'(v_1), s'(v_2), \cdots, s'(v_n)\right),$$
$$G_w' = diag\left(s'(w_1), s'(w_2), \cdots, s'(w_n)\right)$$

we have

$$\frac{\mathrm{d}V}{\mathrm{d}t} = \sum_{i=1}^{n}\left(\frac{\partial L}{\partial x_i} \cdot \frac{\mathrm{d}x_i}{\mathrm{d}t} + \frac{\partial L}{\partial y_i} \cdot s'(v_i) \cdot \frac{\mathrm{d}v_i}{\mathrm{d}t} + \frac{\partial L}{\partial z_i} \cdot s'(w_i) \cdot \frac{\mathrm{d}w_i}{\mathrm{d}t}\right)$$
$$+ (x - x^*)^{\mathrm{T}} P_\Omega\left((x - \alpha\nabla_x L(x, y, z)) - x\right)\big/\alpha$$
$$= \left[\nabla_x L(x, y, z)\right]^{\mathrm{T}} \frac{\mathrm{d}x}{\mathrm{d}t} + \left[\nabla_y L(x, y, z)\right]^{\mathrm{T}} G_v' \frac{\mathrm{d}v}{\mathrm{d}t}$$
$$+ \left[\nabla_z L(x, y, z)\right]^{\mathrm{T}} G_w' \frac{\mathrm{d}w}{\mathrm{d}t}$$
$$+ (x - x^*)^{\mathrm{T}} P_\Omega\left((x - \alpha\nabla_x L(x, y, z)) - x\right)\big/\alpha$$
$$= \left[\nabla_x L(x, y, z)\right]^{\mathrm{T}} P_\Omega\left((x - \alpha\nabla_x L(x, y, z)) - x\right)$$
$$+ (x - x^*)^{\mathrm{T}} P_\Omega\left((x - \alpha\nabla_x L(x, y, z)) - x\right)\big/\alpha$$
$$- \beta\left[\nabla_y L(x, y, z)\right]^{\mathrm{T}} G_v' \nabla_y L(x, y, z)$$
$$- \beta\left[\nabla_z L(x, y, z)\right]^{\mathrm{T}} G_w' \nabla_z L(x, y, z).$$
$$(12)$$

In the first inequality of Lemma 4, let $x' = x - \alpha\nabla_x L(x, y, z)$ and $y' = x^*$, then

$$\left(P_\Omega(x - \alpha\nabla_x L(x, y, z)) - x^*\right)^{\mathrm{T}}\left(x - \alpha\nabla_x L(x, y, z) - P_\Omega(x - \alpha\nabla_x L(x, y, z))\right) \geq 0.$$

consequently, $V\left(P(x(t)), y(t), z(t)\right)$ is Lypunov function, and by (14), it is evident that

$$\frac{\mathrm{d}V}{\mathrm{d}t} = 0 \Leftrightarrow \frac{\mathrm{d}x}{\mathrm{d}t} = 0, \frac{\mathrm{d}v}{\mathrm{d}t} = 0, \frac{\mathrm{d}w}{\mathrm{d}t} = 0.$$

By the Lypunov stability theory, systems (10) is asymptotically stable. Therefore, when the initial point (x_0, y_0, x_0) is obtain near to the equilibrium point, the set $\left\{(x(t), y(t), z(t))\big| t \geq t_0\right\}$ is bounded. By also using LaSall's invariant principle, the trajectory of the neural network (10) $\left\{(u(t), y(t), z(t))\right\}$ will converge to the maximum invariant subset of the following set

$$E = \left\{(x, y, z) \in S\Big|\frac{\mathrm{d}V}{\mathrm{d}t} = 0\right\}.$$

Assume again that $X^* = \left\{x\big|(x, y, z) \in E\right\}$, then we have

$$\lim_{t \to \infty} dist(x(t), X^*) = 0.$$

Specially, when $X^* = \{x^*\}$, we have

$$\lim_{t \to \infty} x(t) = x^*.$$

The proof is completed.

4. Numerical Examples

In order to verify the effectiveness of the presented algorithm in this paper, three examples were selected from the literature [21]. These Examples has been used to check the effectiveness of the new algorithms (see [1-5]).

For the first example, it is easy to verify that the Hessian matrix of the object function $f_p(x) = \frac{1}{2}\|F(x)\|^2$ is rank one deficiency at the minimizer x^*. For the last two examples, the corresponding Hessian matrix is nonsingular at the minimizer. In order to adapt them to the singular case, we have adopt the same procedure as proposed in [1] by introducing function transformation as follows,

$$F(x) := F(x) - \nabla F(x^*) A (A^T A)^{-1} A^T (x - x^*) \quad (16)$$

where x^* is the root of $F(x) = 0$ and

$$F(x) : R^n \to R^m, A \in R^{n \times 1}, A^T = (1,1,\cdots,1).$$

Now we can construct the relevant objection function $f_p(x)$

$$f_p(x) = \frac{1}{2}\|F(x)\|^2$$

for which its Hessian matrix being rank one deficiency and if the Interval $[l, h]$ includes the root of the original $F(x)$, it can be checked that the root of the original $F(x)$ is the minimizer of the optimization problem $\min f_p(x), l \le x \le h$.

Ex. 1: Modified Powell Singular Function:
a) $n = 4, m = 4$
b) $f_1(x) = x_1 + 10x_2 + x_3$

$$f_2(x) = 5^{1/2}(x_3 - x_4)$$

$$f_3(x) = x_2 + (x_2 - 2x_3)^2$$

$$f_4(x) = 10^{1/2}(x_1 - x_4)^2$$

c) $f = 0$ at the minimizer $x^* = (0,0,0,0)$.

Ex. 2: Beale Function:
a) $n = 2, m = 3$
b) $f_1(x) = y_1 - x_1(1 - x_2)$

$$f_2(x) = y_2 - x_1(1 - x_2^2)$$

$$f_3(x) = y_3 - x_1 * (1 - x_2^3)$$

$$y_1 = 1.5; y_2 = 2.25; y_3 = 2.625$$

c) $f = 0$ at the minimizer $x^* = (3, 0.5)$.

Ex. 3: Modified Broyden Tridiagonal Function:

a) $n = 4, m = 4$
b) $f_1(x) = (3 - 2x_1) x_1 - 2x_2 + 1$

$$f_2(x) = (3 - 2x_2) x_2 - x_1 - 2x_3 + 2$$

$$f_3(x) = (3 - 2x_3) x_3 - x_2 - 2x_4 + 2$$

$$f_4(x) = (3 - 2x_4) x_4 - x_3$$

c) $f = 0$ at the minimizer $x^* = (1,1,1,1)$.

Meanwhile, we compare the dynamic behavior of the proposed model with the classical projection gradient system [11] as follows,

$$\frac{dx}{dt} = P_\Omega(x - \alpha \nabla f_p(x)) - x, \alpha > 0 \quad (17)$$

for which the Hessian matrix at the minimizer is generally assumed to be nonsingular.

We use Matlab 7.0 to simulate the dynamics of the corresponding systems. The integral curves are obtained by using the ODE function ode15s for the numerical integration. For the proposed system (10) (**PR**) and the classical projection gradient system (10) (**PG**), we have chosen the same initial point to numerically solve the ODEs.

For Ex.1, we choose $x_0 = (01, 0.9, 1.5238, 10.9605)$, $\alpha = \beta = 1$ for both **PR** and **PG** and let y_0 and z_0 be some random values between 0 and 12. The other parameters **PR** are chosen as $l = 0$, $h = 20$, $\lambda = 0.0000001$, $k = 3000$. The results are shown in **Figures 2** and **3**. It can be seen, in **Figure 2**, that the integral curves response of **PR** converge to f_p's minimizer $x^* = (0,0,0,0)$. On the contrary, as shown in **Figure 3**, the curves of **PG** failed to converge to f_p's minimizer with the same initial point.

For Ex. 2, we choose $l = 0, h = 3, \lambda = 0.0000001$; $x_0 = rand(1,6)$; $k = 5000, \alpha = \beta = 1$. Similar results are obtained, *i.e.*, the proposed system **PR** successfully
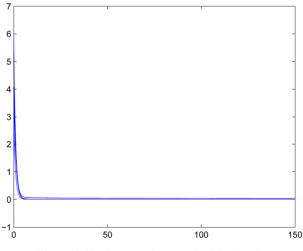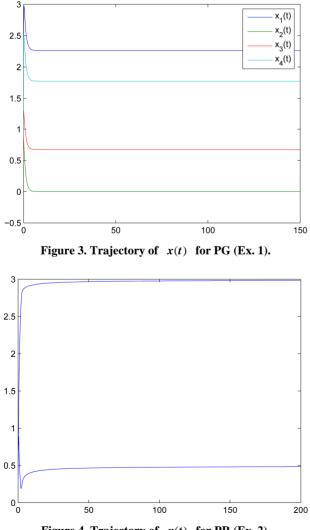

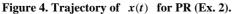**Figure 2. Trajectory of $x(t)$ for PR (Ex. 1).**

found the minimizer $x^* = (3, 0.5)$ while the classical system **PG** failed. The results are shown in **Figures 4** and **5** respectively.

Figures 6 and 7 show the corresponding results for Ex. 3 with initial conditions chosen as $x_0 = (-1, -1, -1, -1)$, $l = 0, h = 1$, $\lambda = 0.0000001$, $k = 5000$, $\alpha = 1$, $\beta = 100$. The proposed system **PR** finally got the minimizer $x^* = (1, 1, 1, 1)$, while the system **PR** got stuck all the time.

5. Concluding Remarks

Singular nonlinear convex optimization problems have been traditionally studied by classical numerical methods. In this paper, a novel neural network model was established to solve such a difficult problem. Under some mild assumptions, the unconstrained nonlinear optimization problem is turned into a constrained optimization problem. By establishing the relationship between KKT points and the augmented Lagrange function, a neural
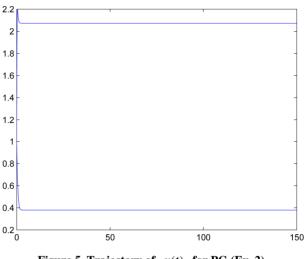
Figure 5. Trajectory of $x(t)$ for PG (Ex. 2).

Figure 6. Trajectory of $x(t)$ for PR (Ex. 3).

Figure 3. Trajectory of $x(t)$ for PG (Ex. 1).

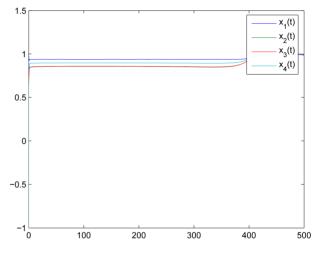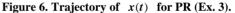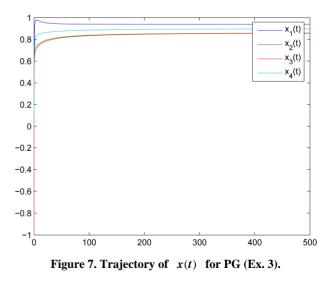**Figure 7. Trajectory of $x(t)$ for PG (Ex. 3).**

Figure 4. Trajectory of $x(t)$ for PR (Ex. 2).

network model is successfully obtained. Global analysis with illustrative examples supports the presented results.

REFERENCES

[1] R. B. Schnabel and T.-T. Chow, "Tensor Methods for Unconstrained Optimization Using Second Derivatives," *SIAM*: *SIAM Journal on Optimization*, Vol. 1, No. 3, 1991, pp. 293-315. [doi:10.1137/0801020](doi:10.1137/0801020)

[2] D. Feng and R. B. Schnabel, "Tensor Methods for Equality Constrained Optimization," *SIAM*: *SIAM Journal on Optimization*, Vol. 6, No. 3, 1996, pp. 653-673. [doi:10.1137/S1052623494270790](doi:10.1137/S1052623494270790)

[3] A. Bouaricha, "Tensor Methods for Large Sparse Unconstrained Optimization," *SIAM*: *SIAM Journal on Optimization*, Vol. 7, No. 3, 1997, pp. 732-756. [doi:10.1137/S1052623494267723](doi:10.1137/S1052623494267723)

[4] R. Ge and Z. Xia, "Solving a Type of Modified BFGS Algorithm with Any Rank Defects and the Local Q-Superliner Convergence Properties," *Journal of Computational and Applied Mathematics*, Vol. 22, No. 1-2, 2006, pp. 193-208.

[5] R. Ge and Z. Xia, "A Type of Modified BFGS Algorithm with Rank Defects and Its Global Convergence in Convex Minimization," *Journal of Pure and Applied Mathematics: Advances and Applications*, Vol. 3, No. 1, 2010, pp. 17-35.

[6] Y. S. Xia and J. Wang, "On the Stability of Globally Projected Dynamical Systems," *Journal of Optimization Theory and Applications*, Vol. 106, No. 1, 2000, pp. 129-150. [doi:10.1023/A:1004611224835](doi:10.1023/A:1004611224835)

[7] Y. S. Xia and J. Wang, "A New Neural Network for Solving Linear Programming Problems and Its Applications," *IEEE Transactions on Neural Networks*, Vol. 7, No. 2, 1996, pp. 525-529. [doi:10.1109/72.485686](doi:10.1109/72.485686)

[8] Y. S. Xia, "A New Neural Network for Solving Linear and Quadratic Programming Problems," *IEEE Transactions on Neural Networks*, Vol. 7, No. 6, 1996, pp. 1544-1547. [doi:10.1109/72.548188](doi:10.1109/72.548188)

[9] Y. S. Xia, J. Wang, "A General Methodology for Designing Globally Convergent Optimization Neural Networks," *IEEE Transactions on Neural Networks*, Vol. 9, No. 6, 1998, pp. 1331-1343. [doi:10.1109/72.728383](doi:10.1109/72.728383)

[10] Y. S. Xia, "A Recurrent Neural Network for Solving Linear Projection Equations," *IEEE Transactions on Neural Networks* Vol. 13, No. 3, 2000, pp. 337-350. [doi:10.1016/S0893-6080(00)00019-8](doi:10.1016/S0893-6080(00)00019-8)

[11] Y. S. Xia, H. Leung and J. Wang, "A Projection Neural Network and Its Application to Constrained Optimization Problems," *IEEE Transactions on Circuits and Systems I*, Vol. 49, No. 4, 2002, pp. 447-458. [doi:10.1109/81.995659](doi:10.1109/81.995659)

[12] Y. S. Xia and J. Wang, "A General Projection Neural Network for Solving Monotone Variational Inequality and Related Optimization Problems," *IEEE Transactions on Neural Networks*, Vol. 15, No. 2, 2004, pp. 318-328. [doi:10.1109/TNN.2004.824252](doi:10.1109/TNN.2004.824252)

[13] X. Gao, L. Z. Liao and W. Xue, "A Neural Network for a Class of Convex Quadratic Minimax Problems with Constraints," *IEEE Transactions on Neural Networks*, Vol. 15, No. 3, 2004, pp. 622-628. [doi:10.1109/TNN.2004.824405](doi:10.1109/TNN.2004.824405)

[14] C. Y. Sun and C. B. Feng, "Neural Networks for Nonconvex Nonlinear Programming Problems: A Switching Control Approach," *Lecture Notes in Computer Science*, Vol. 3496, 2005, pp. 694-699. [doi:10.1007/11427391_111](doi:10.1007/11427391_111)

[15] Q. Tao, X. Liu and M. S. Xue, "A Dynamic Genetic Algorithm Based on Continuous Neural Networks for a Kind of Non-Convex Optimization Problems," *Applied Mathematics and Computation*, Vol. 150, No. 3, 2004, pp. 811-820. [doi:10.1016/S0096-3003(03)00309-6](doi:10.1016/S0096-3003(03)00309-6)

[16] D. J. Bell and D. H. Jacobson, "Singular Optimal Control Problems," Academic Press, New York, 1975.

[17] F. Lamnabhi-Lagarrigue and G. Stefani, "Singular Optimal Control Problem: On the Necessary Conditions of Optimality," *SIAM*: *SIAM Journal on Control and Optimization*, Vol. 28, No. 4, 1990, pp. 823-840. [doi:10.1137/0328047](doi:10.1137/0328047)

[18] L. Liu, R. Ge and P. Gao, "A Novel Neural Network for Solving Singular Nonlinear Convex Optimization Problems," *Lecture Notes in Computer Science*, Vol. 7063, 2011, pp. 554-561. [doi:10.1007/978-3-642-24958-7_64](doi:10.1007/978-3-642-24958-7_64)

[19] D. Kinderlehrer and G. Stampacchia, "An Introduction to Variational Inequalities and Their Applications," Academic Press, New York, 1980.

[20] X. Du, Y. Yang and M. Li, "Further Studies on the Hestenes-Powell augmented Lagrangian Function for Equality Constraints in Nonlinear Programming Problems," *OR Transactions*, Vol. 10, No. 1, 2006, pp. 38-46.

[21] J. J. More, B. S. Garbow and K. E. Hillstrom, "Testing Unconstrained Optimization Software," *ACM Transactions on Mathematical Software*, Vol. 7, No. 1, 1981, pp. 19-31. [doi:10.1145/355934.355936](doi:10.1145/355934.355936)