Neural Networks 61 (2015) 10–21

Contents lists available at ScienceDirect

Neural Networks

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

A one-layer recurrent neural network for constrained nonconvex optimization☆

Guocheng Li [a], Zheng Yan [b], Jun Wang [b,c,*]

[a] Department of Mathematics, Beijing Information Science and Technology University, Beijing, China
[b] Department of Mechanical and Automation Engineering, The Chinese University of Hong Kong, Shatin, New Territories, Hong Kong
[c] School of Control Science and Engineering, Dalian University of Technology, Dalian, Liaoning, China

ARTICLE INFO

Article history:
Received 21 May 2014
Received in revised form 22 August 2014
Accepted 18 September 2014
Available online 28 September 2014

Keywords:
Nonconvex optimization
Exact penalty function
Recurrent neural network
Finite time convergence

ABSTRACT

In this paper, a one-layer recurrent neural network is proposed for solving nonconvex optimization problems subject to general inequality constraints, designed based on an exact penalty function method. It is proved herein that any neuron state of the proposed neural network is convergent to the feasible region in finite time and stays there thereafter, provided that the penalty parameter is sufficiently large. The lower bounds of the penalty parameter and convergence time are also estimated. In addition, any neural state of the proposed neural network is convergent to its equilibrium point set which satisfies the Karush–Kuhn–Tucker conditions of the optimization problem. Moreover, the equilibrium point set is equivalent to the optimal solution to the nonconvex optimization problem if the objective function and constraints satisfy given conditions. Four numerical examples are provided to illustrate the performances of the proposed neural network.

1. Introduction

In this paper, the following constrained nonconvex minimization problem is considered:

$$\begin{array}{ll}\text{minimize} & f(x)\\ \text{subject to} & g_i(x) \leq 0, \quad i \in I = \{1, 2, \dots, m\},\end{array} \tag{1}$$

where $x \in R^n$ is the decision vector; f and $g_i, : R^n \to R\ (i \in I)$ are continuously differentiable functions, but not necessarily convex. The feasible region

$$\mathcal{F} = \{x \in R^n : g_i(x) \leq 0,\ i \in I\}$$

is assumed to be a nonempty set. We denote by $\mathcal{G}$ the set of global solutions of problem (1) as,

$$\mathcal{G} = \{x \in \mathcal{F} : f(y) \geq f(x),\ \forall\, y \in \mathcal{F}\}.$$

Many problems in engineering applications can be formulated as dynamic optimization problems such as kinematic control of redundant robot manipulators (Wang, Hu, & Jiang, 1999), nonlinear model predictive control (Piche, Sayyar-Rodsari, Johnson, & Gerules, 2000; Yan & Wang, 2012), hierarchical control of interconnected dynamic systems (Hou, Gupta, Nikiforuk, Tan, & Cheng, 2007), compressed sensing in adaptive signal processing (Balavoine, Romberg, & Rozell, 2012), and so on. For example, real-time motion planning and control of redundant robot manipulators can be formulated as constrained dynamic optimization problems with nonconvex objective functions for simultaneously minimizing kinetic energy and maximizing manipulability. Similarly, in nonlinear and robust model predictive control, optimal control commands have to be computed with a moving time window by repetitively solving constrained optimization problems with nonconvex objective functions for error and control variation minimization, and robustness maximization. The difficulty of dynamic optimization is significantly amplified when the optimal solutions have to be obtained in real time, especially in the presence of uncertainty. In such applications, compared with traditional numerical optimization algorithms, neurodynamic optimization approaches based on recurrent neural networks have several unique advantages. Recurrent neural networks can be physically implemented in designated hardware/firmware, such as very-large-scale integration (VLSI) reconfigurable analog chips, optical chips,

☆ The work described in the paper was supported by the Research Grants Council of the Hong Kong Special Administrative Region, China under Grant CUHK416812E; and by the National Natural Science Foundation of China under Grant 61273307.

* Corresponding author at: Department of Mechanical and Automation Engineering, The Chinese University of Hong Kong, Shatin, New Territories, Hong Kong. Tel.: +852 39438472.

E-mail addresses: xyliguocheng@sohu.com (G. Li), zyan@mae.cuhk.edu.hk (Z. Yan), jwang@mae.cuhk.edu.hk (J. Wang).

http://dx.doi.org/10.1016/j.neunet.2014.09.009
0893-6080/

graphic processing units (GPU), field programmable gate array (FPGA), digit signal processor (DSP), and so on. Recent technological advances make the design and implementation of neural networks more feasible at a more reasonable cost (Asai, Kanazawa, & Amemiya, 2003).

Since the pioneering work of Hopfield neural networks (Hopfield & Tank, 1985; Tank & Hopfield, 1986), neurodynamic optimization has achieved great success in the past three decades. For example, a deterministic annealing neural network was proposed for solving convex programming problems (Wang, 1994), a Lagrangian network was developed for solving convex optimization problems with linear equality constraints based on the Lagrangian optimality conditions (Xia, 2003), the primal–dual network (Xia, 1996), the dual network (Xia, Feng, & Wang, 2004), and the simplified dual network (Liu & Wang, 2006) were developed for solving convex optimization problems based on the Karush–Kuhn–Tucker optimality conditions, projection neural networks were developed for constrained optimization problems based on the projection method (Gao, 2004; Hu & Wang, 2007; Liu, Cao, & Chen, 2010; Xia, Leung, & Wang, 2002). In recent years, neurodynamic optimization approaches have been extended to nonconvex and generalized convex optimization problems. For example, a Lagrangian neural network was proposed for nonsmooth convex optimization by using the Lagrangian saddle-point theorem (Cheng et al., 2011), a recurrent neural network with global attractivity was proposed for solving nonsmooth convex optimization problems (Bian & Xue, 2013), several neural networks were developed for nonsmooth pseudoconvex or quasiconvex optimization using the Clarke's generalized gradient (Guo, Liu, & Wang, 2011; Hosseini, Wang, & Hosseini, 2013; Hu & Wang, 2006; Liu, Guo, & Wang, 2012; Liu & Wang, 2013). In addition, various neural networks with finite-time convergence property were developed (Bian & Xue, 2009; Forti, Nistri, & Quincampoix, 2004, 2006; Xue & Bian, 2008).

Despite the enormous success, neurodynamic optimization approaches would reach their solvability limits at constrained optimization problems with unimodal objective functions and are important for global optimization with general nonconvex objective functions. Little progress has been made on nonconvex optimization in the neural network community. Instead of seeking global optimal solutions, a more attainable and meaningful goal is to design neural networks for searching critical points (e.g., Karush–Kuhn–Tucker points) of nonconvex optimization problems. Xia, Feng, and Wang (2008) proposed a neural network for solving nonconvex optimization problems with inequality constraints, whose equilibrium points correspond to the KKT points. But the condition that the Hessian matrix of the associated Lagrangian function is positive semidefinite for the global convergence is too strong. In this paper, a one-layer recurrent neural network based on an exact penalty function method is proposed for searching KKT points of nonconvex optimization problems with inequality constraints. The contribution of this paper can be summarized as follows. (1) State of the proposed neural network is convergent to the feasible region in finite time and stays there thereafter, with a sufficiently large penalty parameter; (2) the proposed neural network is convergent to its equilibrium point set; (3) any equilibrium point x^* of the proposed neural network corresponds to a KKT twofold (x^*, λ^*) of the nonconvex problem and vice versa; (4) if the objective function and the constraint functions meet one of the following conditions: (a) the objective function and the constraint functions are convex; (b) the objective function is pseudoconvex and the constraint functions are quasiconvex, then the state of the proposed network converges to the global optimal solution. If the objective function and the constraint functions are invex with respect to the same kernel, then the state of the proposed network converges to optimal solution set. Hence, the results presented in Li, Yan, and Wang (2014) can be viewed as special cases of this paper.

The remainder of this paper is organized as follows. Section 2 introduces some definitions and preliminary results. Section 3 discusses an exact penalty function. Section 4 presented a neural network model and analyzed its convergent properties. Section 5 provides simulation results. Finally, Section 6 concludes this paper.

2. Preliminaries

In this section, we present definitions and properties concerning the set-valued analysis, nonsmooth analysis, and the generalized convex function which are needed in the remainder of the paper. We refer readers to Aubin and Cellina (1984), Cambini and Martein (2009), Clarke (1969), Filippov (1988) and Pardalos (2008) for a more thorough research.

Let R^n be real Euclidean space with the scalar product $\langle x, y\rangle = \sum_{i=1}^n x_i y_i$, $x, y \in R^n$ and its related norm $\|x\| = [\sum_{i=1}^n x_i^2]^{\frac{1}{2}}$. Let $x \in R^n$ and $A \subset R^n$, $\mathrm{dist}(x, A) = \inf_{y\in A} \|x - y\|$ is the distance of x from A.

Definition 1. $F : R^n \hookrightarrow R^n$ is called a set-valued map, if to each point $x \in R^n$, there corresponds to a nonempty closed set $F(x) \subset R^n$.

Definition 2. Let F be a set-valued map. F is said to be upper semicontinuous at $x_0 \in R^n$ if $\forall \varepsilon > 0$, $\exists \delta > 0$ such that $\forall x \in (x_0 + \delta\mathcal{B})$, $F(x) \subset F(x_0) + \varepsilon\mathcal{B}$, where $\mathcal{B} = \mathcal{B}(0, 1)$ is the ball centered at the origin with radius 1. F is upper semicontinuous if it is so at every $x_0 \in R^n$.

A solution $x(t)$ of a differential inclusion is an absolutely continuous function, the derivative $\dot{x}(t)$ is only defined almost everywhere, so that its limit when $t \to \infty$ is not well defined. The concepts of limit and cluster points to a measurable function should be defined. Let $\mu(A)$ denote the Lebesgue measure of a measurable subset $A \subset R$.

Definition 3. Let $x : [0, \infty) \to R^n$ be a measurable function. $x^* \in R^n$ is the almost limit of $x(\cdot)$ if when $t \to \infty$ $\forall \varepsilon > 0, \exists T > 0$ such that

$$\mu\left\{t : \|x(t) - x^*\| > \varepsilon,\ t \in [0, \infty)\right\} < \varepsilon.$$

It can be written as $x^* = \mu - \lim_{t\to\infty} x(t)$. x^* is an almost cluster point of $x(\cdot)$ if when $t \to \infty$ $\forall \varepsilon > 0$,

$$\mu\left\{t : \|x(t) - x^*\| \le \varepsilon,\ t \in [0, \infty)\right\} = \infty.$$

The following propositions show that the usual concepts of limit and cluster are particular cases of almost limit and almost cluster point (Aubin & Cellina, 1984).

Proposition 1. *The limit x^* of $x : [0, \infty) \to R^n$ is an almost limit point. If $x(\cdot)$ is uniformly continuous, any cluster point x^* of $x(\cdot)$ is an almost cluster point.*

Proposition 2. *An almost limit x^* of a measurable function $x : [0, \infty) \to R^n$ is a unique almost cluster point. If $x(\cdot)$ has a unique almost cluster point x^* and $\{x(t) : t \in [0, \infty)\}$ is a bounded subset of R^n, $\mu - \lim_{t\to\infty} x(t) = x^*$.*

Proposition 3. *Let K be a compact subset of R^n and $x : [0, \infty) \to K$ be a measurable function, there exists an almost cluster $x^* \in K$ of $x(\cdot)$ when $t \to \infty$.*

Definition 4. Function $f : R^n \to R$ is said to be Lipschitz near $x \in R^n$ if there exist positive number k and ε such that $|f(x_2) - f(x_1)| \le k\|x_2 - x_1\|$, for all $x_1, x_2 \in x + \varepsilon\mathcal{B}$. If f is Lipschitz near any point of its domain, then it is said to be locally Lipschitz.