Automatica 98 (2018) 34–41

Contents lists available at ScienceDirect

Automatica

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

Brief paper

Projected primal–dual gradient flow of augmented Lagrangian with application to distributed maximization of the algebraic connectivity of a network[☆]

Han Zhang[a,*], Jieqiang Wei[b], Peng Yi[c], Xiaoming Hu[a]

[a] Department of Mathematics, KTH Royal Institute of Technology, SE-100 44, Stockholm, Sweden
[b] Department of Automatic Control, KTH Royal Institute of Technology, SE-100 44, Stockholm, Sweden
[c] Department of Electrical and Systems Engineering, Washington University in St. Louis, USA

ARTICLE INFO

Article history:
Received 25 January 2017
Received in revised form 13 August 2018
Accepted 24 August 2018

Keywords:
Projected dynamical systems
Semi-definite programming
Distributed optimization

ABSTRACT

In this paper, a projected primal–dual gradient flow of augmented Lagrangian is presented to solve convex optimization problems that are not necessarily strictly convex. The optimization variables are restricted by a convex set with computable projection operation on its tangent cone as well as equality constraints. As a supplement of the analysis in Niederländer and Cortés (2016), we show that the projected dynamical system converges to one of the saddle points and hence finding an optimal solution. Moreover, the problem of distributedly maximizing the algebraic connectivity of an undirected network by optimizing the port gains of each nodes (base stations) is considered. The original semi-definite programming (SDP) problem is relaxed into a nonlinear programming (NP) problem that will be solved by the aforementioned projected dynamical system. Numerical examples show the convergence of the aforementioned algorithm to one of the optimal solutions. The effect of the relaxation is illustrated empirically with numerical examples. A methodology is presented so that the number of iterations needed to reach the equilibrium is suppressed. Complexity per iteration of the algorithm is illustrated with numerical examples.

1. Introduction

When solving a convex minimization problem with strong duality, it is well-known that the optimal solution is the saddle point of the Lagrangian. Hence it is natural to consider the gradient flow of Lagrangians (also known as saddle point dynamics) where the primal variable follows the negative gradient flow while the dual variable follows the gradient flow. Gradient flow of Lagrangians is first studied by Arrow et al. (1959), Kose (1956) and has been revisited by Feijer and Paganini (2010). Feijer and Paganini (2010) study the case of strictly convex problems and provides methodologies to transform non-strictly convex problems to strictly convex problems to fit the framework. The convergence is shown by employing the invariance principle for hybrid automata. Cherukuri, Mallada, and Cortés (2016) study the same strictly convex problem from the perspective of projected dynamical systems and are able to show the convergence by a LaSalle-like invariant principle for Carathéodory solutions. Instead of considering discontinuous dynamics, Dürr and Ebenbauer (2011) propose a smooth vector field for seeking the saddle points of strictly convex problems. Wang and Elia (2011) consider a strictly convex problem with equality constraints and with inequality constraints respectively. Saddle point dynamics is also used therein, however, it is worth noticing that their problem is still strictly convex. When they consider the problem with inequality constraints, logarithmic barrier function is used. Though considering nonsmooth problems, Zeng, Yi, and Hong (2017) use the projected saddle point dynamics of augmented Lagrangian whose equality constraint is the variable consensus constraint, and can be viewed as a special case of our problem. Instead of using the continuous-time saddle point dynamics, an iterative distributed augmented Lagrangian method is developed in Chatzipanagiotis, Dentcheva, and Zavlanos (2015). In a recent work (Niederländer and Cortés, 2016) and its conference version (Niederländer, Allgöwer, & Cortés, 2016), the authors consider the nonsmooth case of projected saddle point dynamics and the dynamics is the same as the ones in the current paper when the objective function is smooth.

In this paper, we will focus on maximizing network algebraic connectivity distributedly. In Simonetto, Keviczky, and Babuška (2013), the authors maximize the algebraic connectivity of a

☆ This work is supported by China Scholarship Council. The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Michael M. Zavlanos under the direction of Editor Christos G. Cassandras.

* Corresponding author.
E-mail addresses: hanzhang@kth.se (H. Zhang), jieqiang@kth.se (J. Wei), peng.yi@utoronto.ca (P. Yi), hu@kth.se (X. Hu).

https://doi.org/10.1016/j.automatica.2018.09.004
0005-1098/

mobile robot network distributedly. The authors use first-order Taylor expansion to approximate the original non-convex problem and get a convex problem. A more general linear dynamics is considered and a two-step algorithm is proposed to solve the problem distributedly. It is shown in Simonetto et al. (2013) that the algebraic connectivity is monotonically increasing with the algorithm, while the convergence to one optimal solution is not explicitly given. Schuresko and Cortés (2008), Yang et al. (2010) and Zavlanos and Pappas (2008) focus on assuring the connectivity distributedly, while the algebraic connectivity maximization is not considered.

The main contribution of this paper is as follows. As a supplement to Niederländer and Cortés (2016) and its conference version (Niederländer et al., 2016), we propose a novel analysis line regarding the convergence of the dynamical system to reach comparable results. Moreover, the problem of distributedly maximizing the algebraic connectivity of an undirected network by adjusting the "port gains" of each nodes (base stations) is considered. It is worth noticing that the problem motivates from a physical system and the goal is to enable each base station to compute its own optimal port gains only using its neighbours' information, the total number of nodes N and the information belonging to itself; one cannot "design" the communication network according to the structure of the problem or the algorithm. (For example, Pakazad, Hansson, Andersen, & Rantzer, 2015.) We solve the original problem, which is an SDP, by relaxing it into an NP problem. The NP problem is not strictly convex, hence we adapt the projected saddle point dynamics method proposed in this work to solve the aforementioned NP problem. Numerical examples show that the aforementioned algorithm converges to one of the optimal solutions.

2. Preliminaries and notations

We denote $\mathbb{1} = \mathbf{1}\mathbf{1}^T$ as an N dimensional all-one matrix, where $\mathbf{1}$ is an N dimensional all-one vector. The element located on the ith row and jth column of a matrix A is denoted as $[A]_{ij}$. If matrix $A_1 - A_2$ is positive semi-definite, then it will be denoted as $A_1 \succeq A_2$. We use $\|\cdot\|$ to denote 2-norm of vectors. $|S|$ denotes the cardinality of set S. And any notation with the superscript $*$ is denoted as the optimal solution to the corresponding optimization problem. $tr(\cdot)$ denotes the trace of a matrix. $\langle\cdot,\cdot\rangle_2$ is denoted as the inner-product in Euclidean space and $\langle A_1, A_2\rangle_M = tr(A_1A_2)$ denotes the inner-product in $\mathcal{S}^n$, which is the Hilbert space of $n\times n$ symmetric matrix.

Assume $K \subset \mathbb{R}^n$ is a closed and convex set, the projection of a point x to the set K is defined as $P_K(x) = \arg\min_{y\in K}\|x-y\|$. For $x \in K$, $v \in \mathbb{R}^n$, the projection of the vector v at x with respect to K is defined as: (see Brogliato, Daniilidis, Lemaréchal, & Acary, 2006; Nagurney & Zhang, 2012) $\Pi_K(x, v) = \lim_{\delta\to 0}\frac{P_K(x+\delta v)-x}{\delta} = P_{T_K(x)}(v)$, where $T_K(x)$ denotes the tangent cone of K at x. The interior, the boundary and the closure of K are denoted as $int(K)$, ∂K and $cl(K)$, respectively. The set of inward normals of K at x is defined as $n(x) = \{\gamma \mid \|\gamma\| = 1, \langle\gamma, x-y\rangle_2 \le 0, \forall y \in K\}$, and $\Pi_K(x, v)$ fulfils the following lemma:

Lemma 1 (*Nagurney & Zhang, 2012*)**.** *If $x \in int(K)$, then $\Pi_K(x, v) = v$; if $x \in \partial K$, then $\Pi_K(x, v) = v + \beta(x)n^*(x)$, where $n^*(x) = \arg\max_{n\in n(x)}\langle v, -n\rangle$ and $\beta(x) = \max\{0, \langle v, -n^*(x)\rangle\}$.*

Let F be a vector field such that $F : K \mapsto \mathbb{R}^n$, the projected dynamical system is given by $\dot{x} = \Pi_K(x, F(x))$. Note that the right hand side of above dynamics can be discontinuous on the ∂K. Hence given an initial value $x_0 \in K$, the system does not necessarily have a classical solution. However, if $F(x)$ is Lipschitz continuous, then it has a unique Carathéodory solution that continuously depends on the initial value (Nagurney & Zhang, 2012).

3. Problem formulation and projected saddle point dynamics

In this section, we consider the following optimization problem defined on $\mathbb{R}^n$:

$$\begin{aligned} \underset{x\in K}{\text{minimize}} \quad & f(x) \\ \text{subject to} \quad & Ax - b = 0, \end{aligned} \tag{1}$$

where $f : \mathbb{R}^n \mapsto \mathbb{R}$ and $A \in \mathbb{R}^{m\times n}$. K is a convex set such that calculating the projection on its tangent cone is computationally cheap. $f(x)$ is a convex function but not necessarily strictly convex. It is also assumed that the gradient of $f(x)$ is locally Lipschitz continuous and the Slater's condition holds for (1). Hence strong duality holds for (1).

The Lagrangian $\mathscr{L} : K \times \mathbb{R}^m \mapsto \mathbb{R}$ for the problem (1) is given by

$$\mathscr{L}(x, v) = f(x) + v^T(Ax - b), \tag{2}$$

where $v \in \mathbb{R}^m$ is the Lagrangian multiplier of the constraint $Ax - b = 0$. Since strong duality holds for (1), then (x^*, v^*) is a saddle point of $\mathscr{L}(x, v)$ if and only if x^* is an optimal solution to (1) and v^* is optimal solution to its dual problem. The augmented Lagrangian $\mathscr{L}_{\mathcal{A}} : K \times \mathbb{R}^m \mapsto \mathbb{R}$ for (1) is given by $\mathscr{L}_{\mathcal{A}}(x, v) = f(x) + v^T(Ax - b) + \frac{\rho}{2}(Ax - b)^T(Ax - b)$, where $\rho > 0$ is the damping parameter that will help to suppress the oscillation of x during optimization algorithms. Without loss of generality, we choose $\rho = 1$.

We propose to find the saddle point of (2) via the saddle point dynamics projected on the set K, i.e.,

$$\begin{aligned} \dot{x} &= \Pi_K(x, -\nabla f(x) - A^Tv - A^T(Ax - b)) \\ &= \Pi_K(x, -\frac{\partial \mathscr{L}_{\mathcal{A}}(x, v)}{\partial x}), \end{aligned} \tag{3a}$$

$$\dot{v} = Ax - b = \frac{\partial \mathscr{L}_{\mathcal{A}}(x, v)}{\partial v}. \tag{3b}$$

Note that it is assumed that $\nabla f(x)$ is locally Lipschitz continuous, therefore there is a unique Carathéodory solution for the dynamics (3).

4. Convergence analysis

In this section, we analyse the convergence for (3) and start with the analysis of the equilibrium point of (3). Niederländer and Cortés (2016) consider the nonsmooth case of projected saddle point dynamics and the dynamics are the same as the ones in the current paper when the objective function is smooth. As a supplement, we propose a novel analysis line regarding the stability of the dynamical system to reach comparable results.

Proposition 2. *(x^*, v^*) is a saddle point to (1) if and only if it is an equilibrium of (3).*

Proof. Since strong duality holds for (1), the optimality conditions become necessary and sufficient conditions. The optimality condition for (1) is given by $-\nabla f(x^*) - A^Tv^* \in N_K(x^*)$, $Ax^* - b = 0$, Eskelinen (2007), which implies $-\nabla f(x^*) - A^Tv^* + A^T(Ax^* - b) \in N_K(x^*)$, where $N_K(x^*)$ denotes the normal cone of K at x^*. This implies $\Pi_K(x^*, -\nabla f(x^*) - A^Tv^* - A^T(Ax^* - b)) = 0$, therefore, (x^*, v^*) is an equilibrium point of (3). On the other hand, if (x^*, v^*) is an equilibrium point of (3), it must have $-\nabla f(x^*) - A^Tv^* + A^T(Ax^* - b) \in N_K(x^*)$ and $Ax^* - b = 0$, which implies the optimality condition. □

Proposition 3. *Given an initial value $(x(0), v(0))$, where $x(0) \in K$, the trajectory of the projected dynamical system (3) asymptotically converges to one of the saddle points of (1).*