Stable Convergence of a Primal-Dual Method for Multi-agent Optimization Problems

Igor Konnov

Kazan, Russia, konn-igor@ya.ru

Abstract. We describe a class of primal-dual methods for convex constrained multi-agent optimization problems. We show that these methods possess stable convergence properties under different assumptions. Significant examples of applications are also given.

Keywords: Convex optimization, primal-dual method, multi-agent optimization, stable convergence

1 Introduction

The custom computational method is based on the assumption that all the necessary data can be collected in one computer. Besides, there exist a number of methods whose implementation can be made within complex computational systems, however, under the control from a certain central unit. For many recent applications related to large systems it is standard that the private information of elements should not be moved. Moreover, constant transmission of information back and forth between the central unit and other elements is not suitable since this usually leads to increasing the data noise and mistakes and to very slow computational procedures due to various transmission data delays. In addition, the central unit capacity may be smaller essentially than the necessary total information volume. Then we can apply decentralized procedures within multi-agent systems, where the information is distributed among agents (units).

In [1, 2], we presented primal-dual methods for finding a solution of multi-agent convex optimization problems and established their convergence under different assumptions. In this paper we show that these methods possess stable convergence properties. More precisely, each limit point of the iteration sequence belongs to the fixed bounded set and under some additional assumptions coincides with the normal solution point.

2 Auxiliary properties

Let U^* and V^* be some convex and closed sets in the spaces $\mathbb{R}^{l_1}$ and $\mathbb{R}^{l_2}$, respectively, and let $W^* = U^* \times V^* \in \mathbb{R}^l$ where $l = l_1 + l_2$. We first consider convergence properties of sequences $\{w^k\}$ where $w^k = (u^k, v^k)$ to a point of the set W^*. A set $X \subseteq \mathbb{R}^n$ is called a linear manifold, if for each pair of points $x, y \in X$ and for all $\alpha \in \mathbb{R}$ we have $\alpha x + (1-\alpha)y \in X$. Then

$$X = \{x \in \mathbb{R}^n \mid Ax = b\},$$

where A is an $m \times n$ matrix and $b \in \mathbb{R}^m$. A set $X \subseteq \mathbb{R}^n$ is called a generalized linear manifold, if for each point $x \in X$ there exists a linear manifold $\tilde{X}$ such that $x \in \tilde{X} \subseteq X$.

Let us also define the matrix

$$P = \begin{pmatrix} B & \Theta \\ \Theta^\top & C \end{pmatrix}$$

where B and C are $l_1 \times l_1$ and $l_2 \times l_2$ symmetric and positive definite matrices, and Θ is the $l_1 \times l_2$ zero matrix. Hence, P is a symmetric and positive definite matrix and

$$\langle Pw, w\rangle = \langle Bu, u\rangle + \langle Cv, v\rangle,$$

i.e.

$$\|w\|_P^2 = \|u\|_B^2 + \|v\|_C^2$$

for any $w = (u, v)$. For the starting point $w^0 = (u^0, v^0)$ of the sequence $\{w^k\}$ we define

$$u^*_{(n)} = \arg\min\{\|u - u^0\|_B \mid u \in U^*\}$$

and

$$v^*_{(n)} = \arg\min\{\|v - v^0\|_C \mid v \in V^*\},$$

hence for $w^*_{(n)} = (u^*_{(n)}, v^*_{(n)})$ we have

$$w^*_{(n)} = \arg\min\{\|w - w^0\|_P \mid w \in W^*\}.$$

If W^* is a generalized linear manifold, then so are U^* and V^*. Given a point $\bar{w} = (\bar{u}, \bar{v}) \in W^*$, we set

$$\bar{w}' = \arg\min\{\|w - w^0\|_P \mid w \in W'\},$$

where $W' = U' \times V'$ is a linear manifold such that $\bar{w} \in W' \subseteq W^*$. Then U' is a linear manifold such that $\bar{u} \in U' \subseteq U^*$, V' is a linear manifold such that $\bar{v} \in V' \subseteq V^*$, and we have $\bar{w}' = (\bar{u}', \bar{v}') \in W^*$, where

$$\bar{u}' = \arg\min\{\|u - u^0\|_B \mid u \in U'\}$$

and

$$\bar{v}' = \arg\min\{\|v - v^0\|_C \mid v \in V'\}.$$

Lemma 1. *Suppose a sequence $\{w^k\}$ satisfies the conditions:*

(i)

$$\|w^{k+1} - w^*\|_P \le \|w^k - w^*\|_P, \quad k = 0, 1, \dots, \tag{1}$$

for any $w^* = (u^*, v^*) \in W^*$,

(ii)
$$\lim_{k\to\infty} w^k = \bar{w} = (\bar{u}, \bar{v}) \in W^*. \tag{2}$$

If V^* is a linear manifold, then $\bar{v} = v^*_{(n)}$. If W^* is a linear manifold, then $\bar{w} = w^*_{(n)}$.

Proof. It follows from (1) and (2) that

$$\|\bar{w} - w^*\|_P \le \|w^k - w^*\|_P \le \|w^0 - w^*\|_P, \quad k = 0, 1, \ldots, \tag{3}$$

for any $w^* = (u^*, v^*) \in W^*$. Choose any $\alpha \in \mathbb{R}$ and set $w(\alpha) = (\bar{u}, \bar{v} + \alpha(v^*_{(n)} - \bar{v}))$. Since V^* is a linear manifold, $\bar{v} + \alpha(v^*_{(n)} - \bar{v}) \in V^*$, hence $w(\alpha) \in W^*$. By using the definitions and (3), we have

$$\begin{aligned}
&\|w^0 - w(\alpha)\|_P^2 = \|u^0 - \bar{u}\|_B^2 + \|v^0 - (\bar{v} + \alpha(v^*_{(n)} - \bar{v})\|_C^2 \\
&\ge \|\bar{u} - \bar{u}\|_B^2 + \|\bar{v} - (\bar{v} + \alpha(v^*_{(n)} - \bar{v}))\|_C^2 = \|\alpha(v^*_{(n)} - \bar{v})\|_C^2 \\
&= \|\bar{v} - v^0 + v^0 - (\bar{v} + \alpha(v^*_{(n)} - \bar{v}))\|_C^2 \\
&= \|v^0 - (\bar{v} + \alpha(v^*_{(n)} - \bar{v}))\|_C^2 + \|\bar{v} - v^0\|_C^2 + 2\langle v^0 - (\bar{v} + \alpha(v^*_{(n)} - \bar{v})), C(\bar{v} - v^0)\rangle \\
&= \|v^0 - (\bar{v} + \alpha(v^*_{(n)} - \bar{v}))\|_C^2 - \|\bar{v} - v^0\|_C^2 - 2\alpha\langle v^*_{(n)} - \bar{v}, C(\bar{v} - v^0)\rangle.
\end{aligned}$$

It follows that

$$\begin{aligned}
\|\bar{v} - v^0\|_C^2 &\ge -\|u^0 - \bar{u}\|_B^2 - 2\alpha\langle v^*_{(n)} - \bar{v}, C(\bar{v} - v^0)\rangle \\
&= -\|u^0 - \bar{u}\|_B^2 + 2\alpha\|\bar{v} - v^*_{(n)}\|_C^2 + 2\alpha\langle v^*_{(n)} - \bar{v}, C(v^0 - v^*_{(n)})\rangle.
\end{aligned}$$

By definition,
$$\langle v^*_{(n)} - \bar{v}, C(v^0 - v^*_{(n)})\rangle = 0,$$
hence
$$\|\bar{v} - v^*_{(n)}\|_C^2 \le (2\alpha)^{-1}\left\{\|u^0 - \bar{u}\|_B^2 + \|\bar{v} - v^0\|_C^2\right\}$$
for any $\alpha > 0$. Setting $\alpha \to +\infty$, we obtain $\bar{v} = v^*_{(n)}$. The second assertion is proved similarly. □

We extend the previous results to the case of generalized linear manifolds.

Proposition 1. *Suppose a sequence $\{w^k\}$ satisfies conditions (i) and (ii) of Lemma 1. If V^* is a generalized linear manifold, then $\bar{v} = \bar{v}'$ for any linear manifold V' such that $\bar{v} \in V' \subseteq V^*$. If W^* is a generalized linear manifold, then $\bar{w} = \bar{w}'$ for any linear manifold W' such that $\bar{w} \in W' \subseteq W^*$.*

The proof follows the lines of Lemma 1.

It should be observed that the above properties extend and modify those from [3, pp.283–285].

In the general case we can evaluate the location of $\bar{w}$. Set

$$H(z, w) = \left\{\tilde{w} \in \mathbb{R}^l \mid \|z - \tilde{w}\|_P \le \|z - w\|_P\right\}.$$

Lemma 2. *Suppose a sequence $\{w^k\}$ satisfies conditions (i) and (ii) of Lemma 1. Then*

$$\bar{w} \in \bigcap_{z \in W^*} H(z, w^0) \subseteq H(w^*_{(n)}, w^0). \tag{4}$$

Clearly, the estimates in (4) follow from (3).

Let us now take a convex optimization problem of the form

$$\min_{x \in X} \to \mu(x) \tag{5}$$

where X is a convex and closed set in $\mathbb{R}^s$, $\mu : \mathbb{R}^s \to \mathbb{R}$ is a convex function. The set of its solutions is denoted by X^*. It is well known that the optimization problem (5) is equivalent to the variational inequality: find $x^* \in X$ such that

$$\exists g^* \in \partial\mu(x^*), \ \langle g^*, x - x^* \rangle \geq 0 \quad \forall x \in X, \tag{6}$$

where $\partial\mu(x)$ is the sub-differential of the function μ at x. Let

$$K^\star = \{p \in \mathbb{R}^s \ | \langle p, q \rangle \geq 0 \quad \forall q \in K\}$$

denote the conjugate cone for the set K and let

$$K(X, x) = \{q \in \mathbb{R}^s \mid \exists \bar{\lambda} > 0, x + \lambda q \in X, \forall \lambda \in (0, \bar{\lambda})\}$$

denote the cone of feasible directions for the set X at x. We then can observe that (6) can be equivalently rewritten as

$$S(x^*) \neq \varnothing,$$

where

$$S(x) = \partial\mu(x) \bigcap [K(X, x)]^\star.$$

We can obtain a useful property of the sets $S(x^*)$ under additional assumptions.

Proposition 2. *Suppose that the set X in (5) is polyhedral. Then*

$$S(x^*) = S \quad \forall x^* \in X^*. \tag{7}$$

Proof. Since the set X is polyhedral, it can be defined as follows:

$$X = \{x \in \mathbb{R}^s \mid Ax \geq b\},$$

where A is an $m \times s$ matrix, $b \in \mathbb{R}^m$. If $X^* \neq \varnothing$, then the Lagrange function

$$L(x, y) = \mu(x) + \langle y, b - Ax \rangle$$

has a saddle point $w^* = (x^*, y^*) \in \mathbb{R}^s \times \mathbb{R}^m_+$, i.e. it holds that

$$\forall y \in \mathbb{R}^m_+, \quad L(x^*, y) \leq L(x^*, y^*) \leq L(x, y^*) \quad \forall x \in \mathbb{R}^s, \tag{8}$$

moreover, the set of these saddle points is of the form $X^* \times Y^*$, where Y^* is the set of solutions of the dual problem to (5). Observe that the inequalities in (8) can be replaced with the following system of relations:

$$\begin{array}{l} \langle Ax^* - b, y - y^* \rangle \geq 0 \quad \forall y \in \mathbb{R}^m_+, \\ \exists g^* \in \partial\mu(x^*),\ g^* - A^\top y^* = \mathbf{0}. \end{array}$$

Taking into account the equivalence of (5) and (6) and using (8) we now obtain

$$S(x^*) = \{p \in \mathbb{R}^s \mid p = A^\top y^*,\ \exists y^* \in Y^*\}.$$

It follows that the set $S(x^*)$ is fixed for any $x^* \in X^*$ and (7) holds true. □

Corollary 1. *Suppose that the set X in (5) is polyhedral and the function μ is differentiable. Then*

$$\mu'(x^*) = c \quad \forall x^* \in X^*. \tag{9}$$

In fact, now $\partial\mu(x) = \{\mu'(x)\}$ and (7) gives (9).

Proposition 3. *Suppose that the set X in (5) is a linear manifold, i.e.*

$$X = \{x \in \mathbb{R}^s \mid Ax = b\},$$

where A is an $m \times s$ matrix, $b \in \mathbb{R}^m$. Then the set Y^* of solutions of the dual problem to (5) is a generalized linear manifold. If in addition the function μ is differentiable, then the set Y^* is a linear manifold.

Proof. If $X^* \neq \varnothing$, then the Lagrange function

$$L(x, y) = \mu(x) + \langle y, b - Ax \rangle$$

must now have a saddle point $w^* = (x^*, y^*) \in \mathbb{R}^s \times \mathbb{R}^m$, i.e. it holds that

$$\forall y \in \mathbb{R}^m, \quad L(x^*, y) \leq L(x^*, y^*) \leq L(x, y^*) \quad \forall x \in \mathbb{R}^s,$$

and the set of these saddle points is of the form $X^* \times Y^*$. The above inequalities can be replaced with the following system of relations:

$$Ax^* = b,\ \exists g^* \in S(x^*),\ g^* - A^\top y^* = \mathbf{0}.$$

On account of (7) we now obtain

$$Y^* = \{y^* \in \mathbb{R}^m \mid g^* = A^\top y^*,\ \exists g^* \in S\}.$$

It follows that the set Y^* is a generalized linear manifold. If in addition the function μ is differentiable, then $S(x^*) = \{\mu'(x^*)\}$, but $\mu'(x^*) = c$ for any $x^* \in X^*$ due to Corollary 1. We now obtain

$$Y^* = \{y^* \in \mathbb{R}^m \mid c = A^\top y^*\},$$

hence Y^* is a linear manifold. □

3 Primal-dual method for multi-agent optimization problems

We now describe the primal-dual method proposed in [2] for the multi-agent optimization problems and its general convergence properties. Usually, the multi-agent optimization problem is defined as follows:

$$\min \to \left\{ \sum_{i=1}^{m} f_i(v) \;\middle|\; \bigcap_{i=1}^{m} X_i \right\}, \tag{10}$$

where m is the number of agents (units) in the system. That is, the information about any function f_i and set X_i is known only to the i-th agent and may be unknown even to its neighbours. The agents are joined by some transmission links for limited information exchange, and the topology of the communication network may vary from time to time. This decentralized system has to find a concordant solution defined by (10).

For this reason, we replace (10) with the family of optimization problems of the form

$$\min_{x \in X_I} \to f(x) = \sum_{i=1}^{m} f_i(x_i), \tag{11}$$

where $x = (x_i)_{i=1,\dots,m} \in \mathbb{R}^{mn}$, i.e. $x^\top = (x_1^\top, \dots, x_m^\top)$, $x_i = (x_{i1}, \dots, x_{in})^\top$ for $i = 1, \dots, m$,

$$X_I = X' \bigcap X_I'', \; X' = X_1 \times \ldots \times X_m = \prod_{i=1}^{m} X_i, \; X_i \subseteq \mathbb{R}^n, \; i = 1, \dots, m; \tag{12}$$

the set X_I'' describes the information exchange scheme within the current topology of the communication network, and I is the index set of arcs of the corresponding oriented graph, i.e.

$$X_I'' = \{x \in \mathbb{R}^{mn} \mid x_s - x_t = \mathbf{0}, \; i = (s,t) \in I\}. \tag{13}$$

Taking index sets I, we obtain various constraint sets X_I'' corresponding to the oriented graphs $\mathcal{G}_I$. This means that the formulation of the problem associates each pair of vertices (agents) (s,t) to one oriented arc $i \in I$ whose direction is fixed in any communication network. At the same time, the agents can in principle use both the directions of each arc for communication. This non-oriented graph of the real communication network is denoted by $\mathcal{F}_I$. Let $L = \{1, \dots, l\}$ denote the index set of all the possible oriented arcs for the m agents, so that $I \subseteq L$. We will call $I \subseteq L$ a basic index set if the graph $\mathcal{F}_I$ is connected. Next, for each arc $i = (s,t)$ we can define the $n \times mn$ sub-matrix

$$F_i = (F_{i1} \cdots F_{im}), \tag{14}$$

where

$$F_{ij} = \begin{cases} E, \text{ if } j = s, \\ -E, \text{ if } j = t, \\ \Theta, \text{ otherwise}, \end{cases} \tag{15}$$

E is the $n \times n$ unit matrix, Θ is the $n \times n$ zero matrix. Then clearly

$$X_I'' = \{x \in \mathbb{R}^{mn} \mid F_I x = \mathbf{0}\}, \tag{16}$$

where

$$F_I = \left(\{F_i^\top\}_{i\in I}\right)^\top. \tag{17}$$

Therefore, we can set

$$X_I = \{x \in X' \mid F_I x = \mathbf{0}\}.$$

We denote by X_I^* the solution set of problem (11)–(13). In what follows, we will use the following basic assumptions.

(A1) For each $i = 1, \dots, m$, X_i is a convex and closed set in $\mathbb{R}^n$, $f_i : \mathbb{R}^n \to \mathbb{R}$ is a convex function. Either X' is a polyhedral set or $\mathrm{ri}X' \cap X_L''$ is nonempty.

(A2) The set $X^* = X_L^*$ is nonempty.

Now we present a solution method for the multi-agent optimization problem (11)–(13), where each agent (or unit) receives information only from its neighbours. Given an oriented graph $\mathcal{G}_I$ and an agent j, we denote by $\mathcal{N}_I^+(j)$ and $\mathcal{N}_I^-(j)$ the sets of incoming and outgoing arcs at j and suppose that agent j is responsible for calculation of the current values of the primal variable x_j and all the dual variables y_i and p_i such that $i \in \mathcal{N}_I^-(j)$.

The Lagrange function for problem (11)–(13) is written as follows:

$$\begin{aligned}\mathcal{L}(x,y) = f(x) + \langle y, Ax\rangle &= \sum_{j\in M} f_j(x_j) + \sum_{i\in L}\langle y_i, F_i x\rangle \\ &= \sum_{j\in M}\left\{ f_j(x_j) + \sum_{i\in\mathcal{N}_L^-(j)}\langle y_i, x_j\rangle - \sum_{i\in\mathcal{N}_L^+(j)}\langle y_i, x_j\rangle \right\}.\end{aligned}$$

A pair $w^* = (x^*, y^*) \in X' \times Y_I$ is a Lagrangian saddle point for problem (11)–(13) if

$$\forall y \in Y_I, \quad \mathcal{L}(x^*, y) \le \mathcal{L}(x^*, y^*) \le \mathcal{L}(x, y^*) \quad \forall x \in X, \tag{18}$$

where

$$Y_I = \left\{ y = (y_i)_{i\in L} \in \mathbb{R}^{ln} \mid y_i = \mathbf{0} \in \mathbb{R}^n \text{ for } i \notin I \right\}.$$

We denote by $W_I^* = X_I^* \times Y_I^*$ the set of saddle points in (18) since X_I^* is precisely the solution set of problem (11)–(13), whereas Y_I^* is the set of its Lagrange multipliers. Since $X_L^* = X^*$, we also set $Y^* = Y_L^*$ and $W^* = X^* \times Y^*$. We note that the set of saddle points for the initial problem $W_I^* = X_I^* \times Y_I^*$ is nonempty under the assumptions in (A1)–(A2) if I is a basic index set. Then also we have $X_I^* = X^*$.

Next, the method involves an auxiliary matrix B_k at each k-th iteration. We will take it to be block-diagonal, i.e.

$$B_k = \begin{pmatrix} B_k^1 & \Theta & \dots & \Theta \\ \Theta & B_k^2 & \dots & \Theta \\ \dots & \dots & \dots & \dots \\ \Theta & \Theta & \dots & B_k^m \end{pmatrix}$$

where B^s_k is an $n \times n$ symmetric and positive definite matrix for $s \in M$ and Θ is the $n \times n$ zero matrix. For the sake of simplicity we will write $Y_{(k)} = Y_{I_k}$, $F_{(k)} = F_{I_k}$, etc.

Method (PDM). At the beginning, the agents choose the communication topology by choosing the active arc index set $I_0 \subseteq L$. Next, each s-th agent chooses x^0_s and y^0_i for $i \in \mathcal{N}^-_{(0)}(s)$ and reports these values to its neighbours. This means that $y^0_i = \mathbf{0}$ for $i \notin I_0$. The agents define a common sequence $\{\lambda_k\}$ of positive numbers. Separately, each s-th agent chooses a sequence $\{B^s_k\}$ of $n\times n$ symmetric and positive definite matrices.

At the k-th iteration, $k = 1, 2, \dots$, each s-th agent has the values x^{k-1}_s and y^{k-1}_i, $i \in \mathcal{N}^-_{(k-1)}(s)$, and the same values of its neighbours. The agents choose the current communication topology by choosing the active arc index set $I_k \subseteq L$ and determine the stepsize λ_k. This means that they set $y^k_i = \mathbf{0}$ for $i \notin I_k$.

Step 1: Each s-th agent sets

$$p^k_i = y^{k-1}_i + \lambda_k(x^{k-1}_s - x^{k-1}_t) \quad \forall i = (s,t),\ i \in \mathcal{N}^-_{(k)}(s).$$

Then each s-th agent reports these values to its neighbours.

Step 2: Each s-th agent calculates

$$v^k_s = \sum_{i \in \mathcal{N}^-_{(k)}(s)} p^k_i - \sum_{i \in \mathcal{N}^+_{(k)}(s)} p^k_i$$

and

$$x^k_s = \arg\min_{x_s \in X_s} \left\{ f_s(x_s) + \langle v^k_s, x_s\rangle + 0.5\lambda_k^{-1}\|x_s - x^{k-1}_s\|^2_{B^s_k} \right\}$$

and reports this value to its neighbours.

Step 3: Each s-th agent sets

$$y^k_i = y^{k-1}_i + \lambda_k(x^k_s - x^k_t) \quad \forall i = (s,t),\ i \in \mathcal{N}^-_{(k)}(s).$$

Then each s-th agent reports these values to its neighbours. The k-th iteration is complete.

Definition 1. *We say that $I \subseteq L$ is a support index set with respect to the sequence $\{w^k\}$ if $I = I_k$ for infinitely many k. We say that $I \subseteq L$ is a strongly support index set with respect to the sequence $\{w^k\}$ if it is a support index set and*

$$\inf_{I=I_j, k<j} \sup_{I=I_k} (j-k) \leq d < \infty.$$

We denote by $\mathcal{P}$ (respectively, by $\mathcal{P}^*$) the collection of all support (respectively, strongly support) index sets with respect to the sequence $\{w^k\}$. Also, we set

$$J = \bigcap_{I\in\mathcal{P}} I \text{ and } J^* = \bigcap_{I\in\mathcal{P}^*} I.$$

If $J \neq \varnothing$, then

$$W_J^* \subseteq \bigcap_{k=k_0}^{\infty} W_{(k)}^* \text{ for some } k_0 \geq 1. \tag{19}$$

In what follows we will define the matrices

$$P_k = \begin{pmatrix} B_k & \Theta \\ \Theta^\top & E \end{pmatrix}$$

where E is the $ln \times ln$ unit matrix and Θ is the $mn \times ln$ zero matrix. Hence,

$$\langle P_k w, w \rangle = \langle B_k x, x \rangle + \langle y, y \rangle,$$

i.e.

$$\|w\|_{P_k}^2 = \|x\|_{B_k}^2 + \|y\|^2.$$

We collect some convergence properties of (PDM) from Theorem 1 in [2].

Proposition 4. *Suppose that assumptions (A1)–(A2) are fulfilled, J is a basic index set, $J \in \mathcal{P}$ or $J = J^*$, the sequence $\{\lambda_k\}$ and the matrix sequence $\{B_k\}$ satisfy the conditions:*

$$(1 + \alpha_k) B_k \succeq B_{k+1} \succeq B, \;\; k = 1, 2, \dots, \tag{20}$$

for some $mn \times mn$ symmetric and positive definite matrix B and

$$B_k - 2\lambda_k^2 F_{(k)}^\top F_{(k)} \succeq \tau E, \;\; \lambda_k \geq \lambda' > 0, \;\; k = 1, 2, \dots, \tag{21}$$

for some $\tau > 0$ where E is the $mn \times mn$ unit matrix, the sequence $\{\alpha_k\}$ satisfies the conditions:

$$\alpha_k \geq 0, \quad \sum_{k=0}^{\infty} \alpha_k = \alpha' < \infty. \tag{22}$$

Then:

(i)

$$\|w^k - w^*\|_{P_k}^2 \leq (1 + \alpha_{k-1}) \|w^{k-1} - w^*\|_{P_{k-1}}^2 \text{ if } k \geq k_0,$$

(ii)

$$\lim_{k \to \infty} w^k = w^* \in W_J^*. \tag{23}$$

It was also proved in [2] that (20) and (22) imply

$$\lim_{k \to \infty} B_k = \bar{B} \succeq B,$$

hence

$$\lim_{k \to \infty} P_k = \bar{P} \succeq P,$$

where

$$\bar{P} = \begin{pmatrix} \bar{B} & \Theta \\ \Theta^\top & E \end{pmatrix} \text{ and } P = \begin{pmatrix} B & \Theta \\ \Theta^\top & E \end{pmatrix}.$$

Implementation of the conditions of Proposition 4 are discussed in detail in [2]. We note that these conditions allow for the agents to provide them independently even in the varying topology case. The agents can also apply different strategies if necessary.

4 Stable convergence in the stationary case

By using the results of the previous sections we can establish stability properties of (PDM). Let us first take the completely stationary case, where

$$I_k \equiv I,\ k = 0, 1, 2, \dots, \tag{24}$$

and

$$B_k \equiv B,\ k = 1, 2, \dots, \tag{25}$$

where B is an $mn \times mn$ symmetric and positive definite matrix, hence

$$P_k \equiv P,\ k = 1, 2, \dots$$

Relation (24) means that the network topology is fixed, whereas (25) means that we utilize the fixed metric matrix in (PDM). Then the assertion of Proposition 4 is simplified as follows.

Proposition 5. *Suppose that assumptions (A1)–(A2) are fulfilled, relations (24) and (25) hold, I is a basic index set, the sequence $\{\lambda_k\}$ satisfies the condition:*

$$B - 2\lambda_k^2 F_I^\top F_I \succeq \tau E,\ \lambda_k \geq \lambda' > 0,\ k = 1, 2, \dots, \tag{26}$$

for some $\tau > 0$. Then:

(i)

$$\|w^k - w^*\|_P^2 \leq \|w^{k-1} - w^*\|_P^2,\ k = 1, 2, \dots,$$

(ii)

$$\lim_{k\to\infty} w^k = \bar{w} \in W_I^*. \tag{27}$$

Clearly, rule (26) enables us to choose $\lambda_k > 0$ to be fixed as well. Now we can utilize the results of Section 2. For the starting point $w^0 = (x^0, y^0)$ we set

$$x^*_{(n)} = \arg\min\{\|x - x^0\|_B \mid x \in X^*\}$$

and

$$y^*_{(n)} = \arg\min\{\|y - y^0\| \mid y \in Y_I^*\},$$

hence for $w^*_{(n)} = (x^*_{(n)}, y^*_{(n)})$ we have

$$w^*_{(n)} = \arg\min\{\|w - w^0\|_P \mid w \in W_I^*\}.$$

If W_I^* is a generalized linear manifold, then so are X^* and Y_I^*. Given a point $\bar{w} = (\bar{x}, \bar{y}) \in W_I^*$, we set

$$\bar{w}' = \arg\min\{\|w - w^0\|_P \mid w \in W_I'\},$$

where $W_I' = X' \times Y_I'$ is a linear manifold such that $\bar{w} \in W_I' \subseteq W_I^*$. Then X' is a linear manifold such that $\bar{x} \in X' \subseteq X^*$, Y_I' is a linear manifold such that $\bar{y} \in Y_I' \subseteq Y_I^*$, and we have $\bar{w}' = (\bar{x}', \bar{y}') \in W^*$, where

$$\bar{x}' = \arg\min\{\|x - x^0\|_B \mid x \in X'\}$$

and

$$\bar{y}' = \arg\min\{\|y - y^0\| \mid y \in Y_I'\}.$$

Also, we now set

$$H(z, w) = \left\{ \tilde{w} \in \mathbb{R}^{(m+l)n} \mid \|z - \tilde{w}\|_P \le \|z - w\|_P \right\}.$$

Theorem 1. *Suppose that assumptions (A1)–(A2) are fulfilled, relations (24) and (25) hold, I is a basic index set, the sequence $\{\lambda_k\}$ satisfies condition (26) for some $\tau > 0$. Then:*

(i) Relation (27) holds.

(ii)

$$\bar{w} \in \bigcap_{z \in W_I^*} H(z, w^0) \subseteq H(w^*_{(n)}, w^0). \tag{28}$$

(iii) If X^* is a generalized linear manifold, then $\bar{x} = \bar{y}'$ for any linear manifold X' such that $\bar{x} \in X' \subseteq X^*$.

(iv) If Y_I^* is a generalized linear manifold, then $\bar{y} = \bar{y}'$ for any linear manifold Y_I' such that $\bar{y} \in Y_I' \subseteq Y_I^*$.

(v) if W_I^* is a generalized linear manifold, then $\bar{w} = \bar{w}'$ for any linear manifold W_I' such that $\bar{w} \in W_I' \subseteq W_I^*$.

The assertions follow directly from Lemma 2 and Propositions 1 and 5.

Corollary 2. *Suppose that the assumptions of Theorem 1 are fulfilled. Then:*

(i) Relations (27) and (28) hold.

(ii) If X^* is a linear manifold, then $\bar{x} = x^*_{(n)}$.

(iii) If Y_I^* is a linear manifold, then $\bar{y} = y^*_{(n)}$.

(iv) if W_I^* is a linear manifold, then $\bar{w} = w^*_{(n)}$.

The assertions follow directly from Lemma 1 and Theorem 1.

We conclude that (PDM) provides weak stability of limit points for its iteration sequence $\{w^k\}$ in the general case. That is, each limit point $\bar{w}$ belongs to the bounded set due to (28). This property is essential since the solution set W_I^* may be unbounded. Moreover, we can obtain stronger stability properties under certain additional assumptions.

5 Examples of applications

We now illustrate the stability properties of (PDM) obtained in the previous section. We take several significant examples of applications.

Example 1. (Distributed system of linear equations) We took a system of linear equations

$$Av = b,$$

where A is an $d \times n$ matrix, b is a vector in $\mathbb{R}^d$, or equivalently,

$$A_i v = b_i, \ i = 1, \dots, m;$$

where A_i is an $d_i \times n$ matrix, $b_i \in \mathbb{R}^{d_i}$, $i = 1, \dots, m$, so that

$$A^\top = \left(\{A_i^\top\}_{i=1,\dots,m} \right), \ b^\top = \left(\{b_i^\top\}_{i=1,\dots,m} \right), \text{ and } d = \sum_{i=1}^{m} d_i.$$

We can replace this system with its squared gap minimization problem

$$\min_{v \in \mathbb{R}^n} \to \tilde{f}(v) = 0.5 \sum_{i=1}^{m} \|A_i v - b_i\|^2,$$

which corresponds to (10). In the multi-agent setting, the information about any matrix A_i and vector b_i is known only to the i-th agent and may be unknown to its neighbours. Then we rewrite the above optimization problem in the format (11)–(13) as follows:

$$\min_{x \in X_I} \to f(x) = \sum_{i=1}^{m} f_i(x_i), \tag{29}$$

where

$$X_I = X_I'', \ f_i(x_i) = 0.5\|A_i x_i - b_i\|^2, \ i = 1, \dots, m, \tag{30}$$

$x = (x_i)_{i=1,\dots,m} \in \mathbb{R}^{mn}$, the set X_I'' describes the communication network topology, I is the index set of arcs, i.e.

$$X_I'' = \{x \in \mathbb{R}^{mn} \mid x_s - x_t = \mathbf{0}, \ i = (s,t) \in I\},$$

or briefly as in (16):

$$X_I'' = \{x \in \mathbb{R}^{mn} \mid F_I x = \mathbf{0}\}, \tag{31}$$

where the matrix F_I is defined in (14)–(15) and (17).

Let us apply (PDM) to problem (29)–(31). We observe that the Lagrangian saddle point problem (18) is now equivalent to the system of linear equations

$$F_I x^* = \mathbf{0}, \ A_j^\top (A x_j^* - b_j) + \sum_{i \in I} F_{ij}^\top y_i^* = \mathbf{0}, \ j = 1, \dots, m.$$

This means that the set W_I^* of Lagrangian saddle points is a linear manifold. Due to Corollary 2 the sequence $\{w^k\}$ generated by (PDM) converges to the point $\bar{w} = w^*_{(n)}$.

Example 2. (Distributed system of linear inequalities) We took a system of linear inequalities

$$Av \le b,$$

where A is an $d \times n$ matrix, b is a vector in $\mathbb{R}^d$, or equivalently,

$$A_i v \le b_i, \ i = 1, \dots, m;$$

where A_i is an $d_i \times n$ matrix, $b_i \in \mathbb{R}^{d_i}$, $i = 1, \dots, m$, so that

$$A^\top = \left(\{A_i^\top\}_{i=1,\dots,m}\right), \ b^\top = \left(\{b_i^\top\}_{i=1,\dots,m}\right), \ \text{and } d = \sum_{i=1}^m d_i.$$

We can replace this system with its squared gap minimization problem

$$\min_{v \in \mathbb{R}^n} \to \tilde{f}(v) = 0.5 \sum_{i=1}^m \|[A_i v - b_i]_+\|^2,$$

which corresponds to (10). Here and below $[u]_+$ denotes the projection of a point $u \in \mathbb{R}^s$ onto the non-negative orthant $\mathbb{R}^s_+$. For the multi-agent setting, we rewrite the above optimization problem as (29) where

$$X_I = X_I'', \ f_i(x_i) = 0.5\|[A_i x_i - b_i]_+\|^2, \ i = 1, \dots, m, \tag{32}$$

$x = (x_i)_{i=1,\dots,m} \in \mathbb{R}^{mn}$, the set X_I'' is defined in (31).

Let us apply (PDM) to problem (29),(31), and (32). We observe that the Lagrangian saddle point problem (18) is now equivalent to the system of equations

$$F_I x^* = \mathbf{0}, \ A_j^\top [A x_j^* - b_j]_+ + \sum_{i \in I} F_{ij}^\top y_i^* = \mathbf{0}, \ j = 1, \dots, m.$$

The set $W_I^* = X^* \times Y_I^*$ of Lagrangian saddle points need not be a linear manifold, but Y_I^* is a linear manifold due to Proposition 3. From Theorem 1 and Corollary 2 we obtain that the sequence $\{w^k\}$ generated by (PDM) converges to the point $\bar{w} = (\bar{x}, \bar{y}) \in W^*$, which belongs to the bounded set indicated in (28), besides, $\bar{y} = y^*_{(n)}$.

Example 3. (Decomposable system of linear equations) We took a system of linear equations

$$\sum_{i=1}^m A_i z_i = b,$$

where A_i is an $n \times d_i$ matrix, $z_i \in \mathbb{R}^{d_i}$, $i = 1, \dots, m$, b is a vector in $\mathbb{R}^n$. As above, we replace this system with its squared gap minimization problem

$$\min_{z \in \mathbb{R}^d} \to 0.5 \left\| \sum_{i=1}^m A_i z_i - b \right\|^2, \tag{33}$$

where

$$z = (z_i)_{i=1,\dots,m} \text{ and } d = \sum_{i=1}^m d_i.$$

However, this formulation is not suitable for the multi-agent setting, where the information about the sub-matrix A_i is known only to the i-th agent. For this reason, we take its dual problem

$$\min \to \{0.5\|x\|^2 - \langle b, x\rangle \mid A_i^\top x = \mathbf{0}, \ i = 1, \dots, m\}.$$

For the multi-agent setting, it can be rewritten in the format (11)–(13) as follows:

$$\min_{x \in X_I} \to f(x) = \sum_{i=1}^{m} f_i(x_i), \tag{34}$$

where

$$\begin{array}{l} X_I = X' \bigcap X''_I, \ X' = X_1 \times \ldots \times X_m, \\ X_i = \{x_i \in \mathbb{R}^n \mid A_i^\top x_i = \mathbf{0}\}, \\ f_i(x_i) = \left(0.5\|x_i\|^2 - \langle b, x_i \rangle\right)/m, \ i = 1, \ldots, m, \end{array} \tag{35}$$

$x = (x_i)_{i=1,\ldots,m} \in \mathbb{R}^{mn}$, and the set X''_I is defined in (31). Problem (34)–(35) clearly has a unique solution x^*. Let us apply (PDM) to this problem. We observe that the Lagrangian saddle point problem (18) is now equivalent to the system of relations

$$F_I x^* = \mathbf{0}, \ \exists \tilde{z}_j \in \mathbb{R}^{d_j}, \ (1/m)(x_j^* - b) + \sum_{i \in I} F_{ij}^\top y_i^* + A_j \tilde{z}_j = \mathbf{0}, \ j = 1, \ldots, m.$$

This means that the set Y_I^* of Lagrangian multipliers is a linear manifold. Due to Corollary 2 the sequence $\{w^k\}$ generated by (PDM) converges to the point $\bar{w} = w_{(n)}^*$, where $\bar{x} = x^*$ and $\bar{y} = y_{(n)}^*$. After calculating the point x^* we can find a solution of (33) by using the optimality conditions in (34)–(35) and taking the extended Lagrangian involving the constraints in X' with respect to the dual variables z_j:

$$(1/m)(x_j^* - b) + \sum_{i \in I} F_{ij}^\top y_i^* + A_j z_j^* = \mathbf{0}, \ j = 1, \ldots, m.$$

Example 4. (Fermat-Weber optimization problem) We took the well-known Fermat-Weber optimization problem:

$$\min_{v \in \mathbb{R}^n} \to \tilde{f}(v) = \sum_{i=1}^{m} \alpha_i \|v - a_i\|, \ \alpha_i > 0, \ i = 1, \ldots, m;$$

where a_i, $i = 1, \ldots, m$ are some given points (anchors). We suppose that the points a_i do not belong to the same straight line. Then the function $\tilde{f}$ is coercive, strictly convex, and non-smooth; see [4, Ch. V, §2]. Hence, the optimization problem has unique solution. It can be rewritten in the multi-agent format (11)–(13) as (29) where

$$X_I = X''_I, \ f_i(x_i) = \alpha_i \|x_i - a_i\|, \ i = 1, \ldots, m, \tag{36}$$

$x = (x_i)_{i=1,\ldots,m} \in \mathbb{R}^{mn}$, and the set X''_I is defined in (31). Problem (29), (31), and (36) also has a unique solution x^*. Let us apply (PDM) to this problem. From Theorem 1 and Corollary 2 we obtain that the sequence $\{w^k\}$ generated by (PDM) converges to the point $\bar{w} = (\bar{x}, \bar{y}) \in W^*$, which belongs to the bounded set indicated in (28). If $W_I^* = X^* \times Y_I^*$ is the set of its Lagrangian saddle points, then $X^* = \{x^*\}$, hence, $\bar{x} = x^*$. Due to Proposition 3, the set Y_I^* is a linear manifold if $x_i^* \neq a_i$, $i = 1, \ldots, m$ since the function f is differentiable at x^*. Then $\bar{y} = y_{(n)}^*$. Otherwise, Y_I^* is a generalized linear manifold. Then $\bar{y} = \bar{y}'$ for any linear manifold Y'_I such that $\bar{y} \in Y'_I \subseteq Y_I^*$.

6 Stable convergence in the non-stationary case

By using the results of Sections 2–3 we can establish somewhat weaker stability properties of (PDM) in the non-stationary case. Let us first discuss the case where (25) holds, but the network topology is not stationary, i.e. (24) is not fulfilled. Hence we still utilize the fixed metric matrix in (PDM). We recall that the i-th agent is associated with the i-th vertex of the communication graph. It follows from the results in [1, 2] that we can satisfy the conditions of Proposition 4 if each i-th agent will evaluate the maximal degree of the i-th vertex in the varying communication graphs $\mathcal{F}_{I_k}$. Then the assertions of Theorem 1 and Corollary 2 will be true, but we have to replace the point $w^0 = (x^0, y^0)$ with $w^{k_0} = (x^{k_0}, y^{k_0})$ where k_0 is defined in (19). Therefore, the stable convergence will be almost the same in this case.

Let us first discuss the general non-stationary case where neither (24) nor (25) will be fulfilled. This means that the network topology is non-stationary, and we utilize the variable metric matrices in (PDM). Then we can apply general convergence properties. Under the conditions of Proposition 4 we obtain that (23) holds, besides,

$$\|w^* - \tilde{w}\|_P^2 \leq \sigma \|w^{k_0} - \tilde{w}\|_{P_{k_0}}^2, \quad \forall \tilde{w} \in W_J^*,$$

where

$$\sigma = \prod_{i=k_0}^{\infty} (1 + \alpha_i) < +\infty$$

due to (22), and k_0 is defined in (19). Set

$$\tilde{H}(z, \bar{w}) = \left\{ w \in \mathbb{R}^{(m+l)n} \mid \|z - w\|_P \leq \sigma \|z - \bar{w}\|_{P_{k_0}} \right\},$$

then

$$w^* \in \bigcap_{z \in W_J^*} \tilde{H}(z, w^{k_0}) \subseteq \tilde{H}(w_{(n)}^*, w^{k_0}).$$

That is, each limit point w^* belongs to the bounded set.

References

1. Konnov, I.V.: Primal-dual method for optimization problems with changing constraints. In: P. Pardalos et al. (eds.) Mathematical Optimization Theory and Operations Research (MOTOR 2022), Lecture Notes in Computer Science **13367**, pp. 46–61. Springer, Cham (2022)
2. Konnov, I.V.: Variable metric primal-dual method for convex optimization problems with changing constraints. Preprint, stored on 18.08.2022 (Preprints.ru). https://doi.org/10.24108/preprints-3112463. – 18 pp.
3. Karmanov, V.G.: Mathematical programming. Nauka, Moscow (1986) [In Russian]
4. Polyak, B.T.: Introduction to optimization. Nauka, Moscow (1983) [Engl. transl. in Optimization Software, New York (1987)]