Variable Metric Primal-Dual Method for Convex Optimization Problems with Changing Constraints

Igor Konnov

Kazan, Russia, e-mail:konn-igor@ya.ru

Abstract. We propose a modified primal-dual method for general convex optimization problems with changing affine constraints. We establish convergence of the method that uses variable metric matrices at each iteration. This approach yields new opportunities for control of the parameters according to the constraints changes. In case of the multi-agent optimization problems the method can be adjusted to the changing communication topology and enables the agents to choose the parameters separately of each other.

Keywords: Convex optimization, changing constraints, primal-dual method, variable metric, multi-agent optimization, changing topology

1 Introduction

The standard optimization problem consists in finding the minimal value of some goal function $\tilde{f}$ on a feasible set $\tilde{D}$. For brevity, we write this problem as

$$\min_{v \in \tilde{D}} \to \tilde{f}(v).$$

The feasible set $\tilde{D}$ is usually determined by equality and inequality constraints involving some functions. Usually, all these functions are fixed, but may contain some uncertain parameters. At the same time, there exist examples of applied large-scale problems that contain superfluous constraints and variables, but only some of them can be utilized at a given iterate. Various decentralized multi-agent optimization problems can serve as examples of such systems; see e.g. [1–3] and the references therein.

In [4], we presented a primal-dual method for finding a solution of these problems. That method is a modification of the method from [5] for usual constrained optimization problems. In this paper we propose a further modification of the primal-dual method, which enables us to enhance its convergence and implementation properties. More precisely we now admit variable metric matrices instead of the usual unit matrices. This approach opens new opportunities for control of the parameters according to the constraints changes. In case of the multi-agent optimization problems the method can be adjusted to the changing communication topology and enables the agents to choose the parameters separately of each other.

2 The general problem with changing constraints and its properties

Let us consider first a general optimization problem of the form

$$\min_{x \in D} \to f(x) \tag{1}$$

for some function $f : \mathbb{E} \to \mathbb{R}$ and set $D \subseteq \mathbb{E}$ in a finite-dimensional space $\mathbb{E}$. The set of its solutions is denoted by D^*, and the optimal function value by f^*, i.e.

$$f^* = \inf_{x \in D} f(x).$$

It will be suitable for us to specialize this problem as follows. For each $x \in \mathbb{E}$, let $x = (x_i)_{i=1,\dots,m}$, i.e. $x^\top = (x_1^\top, \dots, x_m^\top)$, where $x_i = (x_{i1}, \dots, x_{in})^\top$ for $i = 1, \dots, m$, hence $\mathbb{E} = \mathbb{R}^{mn}$. This means that each vector x is divided into m subvectors $x_i \in \mathbb{R}^n$. In case $n = 1$ we obtain the custom coordinates of x. Next, we suppose that

$$D = \{x \in X \mid Ax = b\}, \tag{2}$$

where X is a subset of $\mathbb{R}^{mn}$, the matrix A has ln rows and mn columns, so that $b = (b_i)_{i=1,\dots,l}$, $b_i \in \mathbb{R}^n$ for $i = 1, \dots, l$, and $b \in \mathbb{R}^{ln}$.

In what follows, we will use the following basic assumptions.

(A1) The function $f : \mathbb{R}^{mn} \to \mathbb{R}$ is convex, X is a convex and closed set in $\mathbb{R}^{mn}$.
(A2) The set D^* is nonempty, either X is a polyhedral set or $Ax' = b$ for some $x' \in \mathrm{ri} X$.

For brevity, we set $M = \{1, \dots, m\}$ and $L = \{1, \dots, l\}$. It is clear that the matrix A is represented as follows:

$$A = \begin{pmatrix} A_1 \\ A_2 \\ \dots \\ A_l \end{pmatrix} = \left(A_1^\top A_2^\top \dots A_l^\top\right)^\top,$$

where A_i is the corresponding $n \times mn$ sub-matrix of A for $i \in L$. We will write this briefly

$$A = \left(\{A_i^\top\}_{i \in L}\right)^\top.$$

Similarly, we can determine some other submatrices

$$A_I = \left(\{A_i^\top\}_{i \in I}\right)^\top$$

for any $I \subseteq L$, hence $A = A_L$. Setting

$$F_I = \{x \in \mathbb{R}^{mn} \mid A_I x = b_I\} \text{ and } D_I = \{x \in X \mid A_I x = b_I\} = X \bigcap F_I, \tag{3}$$

where $b_I = (b_i)_{i \in I}$, we obtain a family of optimization problems

$$\min_{x \in D_I} \to f(x). \tag{4}$$

As above, we denote the solution set of problem (3)–(4) by D_I^*, and the optimal function value by f_I^*, so that $D_L^* = D^*$ and $f_L^* = f^*$. We intend to recall some properties related to superfluous constraints. We will denote by F^* the solution set of the optimization problem

$$\min_{x \in X} \to f(x),$$

and its optimal function value by f^{**}. If the set $F^* \bigcap F_I$ is nonempty for some $I \subseteq L$, then $f^{**} = f_I^*$ and $F^* \bigcap F_I = D_I^*$.

Definition 1. *We say that $I \subseteq L$ is a basic index set if*

$$A_I x = b_I \implies Ax = b.$$

Clearly, if I is a basic index set, then $f_I^* = f^*$, $D_I = D$, and $D_I^* = D^*$. For each problem (3)–(4) associated with an index set $I \subseteq L$ we can define its Lagrange function

$$\mathcal{L}_I(x, y) = f(x) + \langle y_I, A_I x - b_I \rangle$$

and the corresponding saddle point problem. It appears more suitable to utilize the general Lagrange function

$$\mathcal{L}(x, y) = f(x) + \langle y, Ax - b \rangle,$$

with the modified dual feasible set. Namely, we say that $w^* = (x^*, y^*) \in X \times Y_I$ is a saddle point for problem (3)–(4) if

$$\forall y \in Y_I, \quad \mathcal{L}(x^*, y) \le \mathcal{L}(x^*, y^*) \le \mathcal{L}(x, y^*) \quad \forall x \in X, \tag{5}$$

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 $W_I^* = D_I^* \times Y_I^*$ the set of saddle points in (5) since D_I^* is precisely the solution set of problem (3)–(4), whereas Y_I^* is the set of its Lagrange multipliers. Since $D_L^* = D^*$, we also set $Y^* = Y_L^*$, i.e. $W^* = D^* \times Y^*$ is the set of saddle points for the initial problem (1)–(2). Properties of Lagrange functions and their saddle points are described in detail e.g. in [6].

Observe that (5) is rewritten equivalently as follows:

$$A_I x^* = b_I, \quad \mathcal{L}(x^*, y^*) \le \mathcal{L}(x, y^*) \quad \forall x \in X.$$

Besides, if we take $I = \varnothing$, then $Y_I = \{\mathbf{0}\}$, hence we can write $D_I^* = F^*$ and $Y_I^* = \{\mathbf{0}\}$. If assumptions (A1)–(A2) are fulfilled and I is a basic index set, then $D_I^* = D^*$ and $Y_I^* \subseteq Y^*$. Also, if assumptions (A1)–(A2) are fulfilled, the set $F^* \bigcap F_I$ is nonempty for some $I \subseteq L$, then $F^* \bigcap F_I = D_I^*$ and $\mathbf{0} \in Y_I^*$.

Let Q be a $s \times s$ symmetric and positive definite matrix. Then we can define the corresponding scalar product of any points u', u'' and the norm of any point u in $\mathbb{R}^s$ as follows:

$$\langle u', u'' \rangle_Q = \langle Qu', u'' \rangle = (u'')^T Q u'$$

and

$$\|u\|_Q = \sqrt{\langle Qu, u\rangle}.$$

In case $Q = E$ where E is the $s \times s$ unit matrix we obtain the standard scalar product and norm. Let Q', Q'' be two $s \times s$ symmetric matrices and let Θ denote the $s \times s$ zero matrix. Then $Q' \succeq Q''$ ($Q' \succ Q''$) means that the matrix $Q' - Q''$ is positive semi-definite (definite), or equivalently, $Q' - Q'' \succeq \Theta$ ($Q' - Q'' \succ \Theta$).

We need two auxiliary properties of sequences of symmetric and positive definite matrices.

Lemma 1. *Suppose a sequence $\{V_k\}$ of $s \times s$ symmetric and positive definite matrices satisfies the conditions:*

$$V_k \succeq V_{k+1} \succeq V, \; k = 1, 2, \dots, \tag{6}$$

where V is a $s \times s$ symmetric and positive definite matrix. Then

$$\lim_{k\to\infty} V_k = \bar{V} \succeq V.$$

Proof. It follows from (6) that $\|V_k\| \geq \|V_{k+1}\| \geq \|V\|$. Hence, the sequence $\{V_k\}$ is bounded and has limit points. Let V' and V'' be two different limit points of $\{V_k\}$, i.e.

$$V' = \lim_{i\to\infty} V_{k_i}, \; V'' = \lim_{j\to\infty} V_{l_j}.$$

For each k_i there exists $l_j > k_i$ such that $V_{k_i} \succeq V_{l_j}$, therefore $V' \succeq V''$. Similarly we can obtain the reverse inequality $V'' \succeq V'$. It follows that

$$\langle (V'' - V')u, u\rangle = 0 \quad \forall u \in \mathbb{R}^s,$$

hence $\|V'' - V'\| = 0$ and $V'' = V' = \bar{V}$. □

Lemma 2. *Suppose a sequence $\{Q_k\}$ of $s \times s$ symmetric and positive definite matrices satisfies the conditions:*

$$(1 + \alpha_k)Q_k \succeq Q_{k+1} \succeq Q, \; k = 1, 2, \dots, \tag{7}$$

where Q is a $s \times s$ symmetric and positive definite matrix,

$$\alpha_k \geq 0, \quad \sum_{k=0}^{\infty} \alpha_k = \alpha' < \infty. \tag{8}$$

Then

$$\lim_{k\to\infty} Q_k = \bar{Q} \succeq Q.$$

Proof. Set

$$V_k = \prod_{i=k}^{\infty}(1 + \alpha_i)Q_k,$$

then, due to (7) and (8) we have $V_k \succeq Q_k \succeq Q$ and

$$V_{k+1} = \prod_{i=k+1}^{\infty}(1+\alpha_i)Q_{k+1} \leq \prod_{i=k}^{\infty}(1+\alpha_i)Q_k = V_k.$$

It follows that the sequence $\{V_k\}$ satisfies the conditions in (6). On account of Lemma 1 we have

$$\lim_{k\to\infty} V_k = \bar{V} \succeq Q.$$

However,

$$\lim_{k\to\infty} \prod_{i=0}^{\infty}(1+\alpha_i) < \infty$$

due to (8), hence

$$\lim_{k\to\infty} \prod_{i=k}^{\infty}(1+\alpha_i) = 1.$$

Since

$$Q_k = \left\{\prod_{i=k}^{\infty}(1+\alpha_i)\right\}^{-1} V_k,$$

we obtain

$$\lim_{k\to\infty} Q_k = \bar{V} \succeq Q,$$

therefore $\bar{V} = \bar{Q}$. □

3 Primal-dual method for the family of saddle point problems

We intend to find saddle points in (5) by a modification of the method proposed in [4]. First we note that the set of saddle points for the initial problem (1)–(2) is nonempty under the assumptions in (A1)–(A2); see e.g. Theorems 5.2–5.3 in [6, pp.57–59] and [7, Corollary 28.2.2]. Therefore, this is the case for each saddle point problem in (5) associated with a basic index set I. Denote by $\pi_U(u)$ the projection of u onto U. Also, for simplicity we will write $Y_{(k)} = Y_{I_k}$, $Y^*_{(k)} = Y^*_{I_k}$, etc. Then the method is described as follows.

Method (PDM). *Step 0:* Choose an index set $I_0 \subseteq L$, a point $w^0 = (x^0, y^0) \in X \times Y_{(0)}$. Define a sequence $\{\lambda_k\}$ of positive numbers and a sequence $\{B_k\}$ of $mn \times mn$ symmetric and positive definite matrices. Set $k = 1$.

Step 1: Choose an index set $I_k \subseteq L$, take and a matrix B_k and a number $\lambda_k > 0$.

Step 2: Take $p^k = \pi_{Y_{(k)}}[y^{k-1} + \lambda_k(Ax^{k-1} - b)]$.

Step 3: Take $x^k = \text{argmin}\{f(x) + \langle p^k, Ax - b\rangle + 0.5\lambda_k^{-1}\|x - x^{k-1}\|^2_{B_k} \mid x \in X\}$.

Step 4: Take $y^k = \pi_{Y_{(k)}}[y^{k-1} + \lambda_k(Ax^k - b)]$. Set $k = k+1$ and go to Step 1.

First we observe that

$$p^k = \mathrm{argmin}\{-\mathcal{L}(x^{k-1}, p) + 0.5\lambda_k^{-1}\|p - y^{k-1}\|^2 \mid p \in Y_{(k)}\}$$

and

$$y^k = \mathrm{argmin}\{-\mathcal{L}(x^k, y) + 0.5\lambda_k^{-1}\|y - y^{k-1}\|^2 \mid y \in Y_{(k)}\}.$$

Therefore, each iteration of (PDM) involves two projection (proximal) steps in the dual variable y and one proximal step in the primal variable x. The point $w^k = (x^k, y^k)$ belongs to $X \times Y_{(k)}$.

Lemma 3. *Suppose U is a closed convex set in the space $\mathbb{R}^s$, $\varphi : \mathbb{R}^s \to \mathbb{R}$ is a convex function, u is a point in $\mathbb{E}$, and Q is a $s \times s$ symmetric and positive definite matrix. If*

$$\mu(z) = \varphi(z) + (2\lambda)^{-1}\|z - u\|_Q^2, \ \lambda > 0,$$

and

$$v = \mathrm{argmin}\{\mu(z) \mid z \in U\},$$

then

$$\varphi(v) - \varphi(z) \leq (2\lambda)^{-1}\{\|z - u\|_Q^2 - \|z - v\|_Q^2 - \|v - u\|_Q^2\} \quad \forall z \in U. \tag{9}$$

Proof. Since the function φ is convex, using the squared Q-norm properties gives

$$\mu(z) - \mu(v) \geq (2\lambda)^{-1}\|z - v\|_Q^2 \quad \forall z \in U.$$

This inequality is equivalent to (9). □

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, P_k is a symmetric and positive definite matrix and

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

i.e.

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

Proposition 1. *Suppose that assumptions (A1)–(A2) are fulfilled. For any pair $w^* = (x^*, y^*) \in D^*_{(k)} \times Y^*_{(k)}$ we have*

$$\begin{aligned} \|w^k - w^*\|_{P_k}^2 \ &\leq \|w^{k-1} - w^*\|_{P_k}^2 - \|p^k - y^k\|^2 - \|p^k - y^{k-1}\|^2 - \|x^k - x^{k-1}\|_{B_k}^2 \\ &+2\lambda_k\langle y^k - p^k, A(x^k - x^{k-1})\rangle \\ &= \|w^{k-1} - w^*\|_{P_k}^2 - \|p^k - y^k\|^2 - \|p^k - y^{k-1}\|^2 - \|x^k - x^{k-1}\|_{B_k}^2 \\ &+2\lambda_k^2\|A_{(k)}(x^k - x^{k-1})\|^2. \end{aligned} \tag{10}$$

Proof. Choose any $w^* = (x^*, y^*) \in D^*_{(k)} \times Y^*_{(k)}$. Setting $\varphi(z) = \mathcal{L}(z, p^k)$, $\lambda = \lambda_k$, $Q = B_k$, $U = X$, $u = x^{k-1}$, $v = x^k$, and $z = x^*$ in (9) gives

$$2\lambda_k\{\mathcal{L}(x^k, p^k) - \mathcal{L}(x^*, p^k)\} \le \|x^* - x^{k-1}\|^2_{B_k} - \|x^* - x^k\|^2_{B_k} - \|x^k - x^{k-1}\|^2_{B_k}.$$

Also, using (5) with $I = I_k$, $x = x^k$, and $y = p^k$ gives

$$2\lambda_k\{\mathcal{L}(x^*, p^k) - \mathcal{L}(x^k, y^*)\} \le 0.$$

Adding these inequalities, we obtain

$$\|x^k - x^*\|^2_{B_k} \le \|x^{k-1} - x^*\|^2_{B_k} - \|x^k - x^{k-1}\|^2_{B_k} + 2\lambda_k\langle p^k - y^*, Ax^k - b\rangle. \quad (11)$$

On the other hand, setting $\varphi(z) = -\mathcal{L}(x^{k-1}, z)$, $\lambda = \lambda_k$, $Q = E$, $U = Y_{(k)}$, $u = y^{k-1}$, $v = p^k$, and $z = y^k$ in (9) gives

$$2\lambda_k\{\mathcal{L}(x^{k-1}, y^k) - \mathcal{L}(x^{k-1}, p^k)\} \le \|y^k - y^{k-1}\|^2 - \|p^k - y^k\|^2 - \|p^k - y^{k-1}\|^2.$$

Next, setting $\varphi(z) = -\mathcal{L}(x^k, z)$, $\lambda = \lambda_k$, $Q = E$, $U = Y_{(k)}$, $u = y^{k-1}$, $v = y^k$, and $z = y^*$ in (9) gives

$$2\lambda_k\{\mathcal{L}(x^k, y^*) - \mathcal{L}(x^k, y^k)\} \le \|y^* - y^{k-1}\|^2 - \|y^* - y^k\|^2 - \|y^k - y^{k-1}\|^2.$$

Adding these inequalities, we obtain

$$\begin{aligned}\|y^k - y^*\|^2 \ &\le \|y^{k-1} - y^*\|^2 - \|p^k - y^k\|^2 - \|p^k - y^{k-1}\|^2 \\ &\quad -2\lambda_k\{\langle y^* - y^k, Ax^k - b\rangle + \langle y^k - p^k, Ax^{k-1} - b\rangle\}. \qquad (12)\end{aligned}$$

Now adding (11) and (12) gives the first inequality in (10). Since

$$\langle y^k - p^k, A(x^k - x^{k-1})\rangle = \lambda_k\|A_{(k)}(x^k - x^{k-1})\|^2,$$

we conclude also that the second relation in (10) holds true. □

Now we can indicate conditions that provide basic convergence properties.

Theorem 1. *Suppose that assumptions (A1)–(A2) are fulfilled,*

$$\bigcap_{k=j}^{\infty} W^*_{(k)} \neq \varnothing \ \textit{for some } j \ge 1, \quad (13)$$

the sequence $\{\lambda_k\}$ and the matrix sequence $\{B_k\}$ satisfy the conditions:

$$(1 + \alpha_k)B_k \succeq B_{k+1} \succeq B, \ k = 1, 2, \dots, \quad (14)$$

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

$$B_k - 2\lambda_k^2 A^\top_{(k)} A_{(k)} \succeq \tau E, \ \lambda_k \ge \lambda' > 0, \ k = 1, 2, \dots, \quad (15)$$

for some $\tau > 0$ where E is the $mn \times mn$ unit matrix, the sequence $\{\alpha_k\}$ satisfies the conditions in (8). Then:

(i) the sequence $\{w^k\}$ has limit points,
(ii) each of these limit points is a solution of problem (5) for some $I \subseteq L$,
(iii) for any limit point $\bar{w}$ of $\{w^k\}$ such that

$$\bar{w} \in \bigcap_{k=j}^{\infty} W^*_{(k)} \text{ for some } j \geq 1,$$

it holds that

$$\lim_{k\to\infty} w^k = \bar{w}. \tag{16}$$

Proof. Take any point

$$w^* \in \bigcap_{k=j}^{\infty} W^*_{(k)}.$$

Then from (10) and (15) we have

$$\|w^k - w^*\|^2_{P_k} \leq \|w^{k-1} - w^*\|^2_{P_k} - \|p^k - y^k\|^2 - \|p^k - y^{k-1}\|^2 - \tau\|x^k - x^{k-1}\|^2 \tag{17}$$

for $k = j, j+1, \ldots$ Due to Lemma 2 it follows from (15) and (8) that

$$\lim_{k\to\infty} B_k = \bar{B} \succeq B, \tag{18}$$

hence

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

where

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

and

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

It also follows from (17) and (14) that

$$\begin{aligned} \|w^k - w^*\|^2_{P_k} &\leq (1+\alpha_{k-1})\|w^{k-1} - w^*\|^2_{P_{k-1}} - \|p^k - y^k\|^2 - \|p^k - y^{k-1}\|^2 \\ &-\tau\|x^k - x^{k-1}\|^2 \end{aligned} \tag{19}$$

for $k = j, j+1, \ldots$ This inequality gives

$$\|w^k - w^*\|^2_{P} \leq \|w^k - w^*\|^2_{P_k} \leq C\|w^1 - w^*\|^2_{P_1}$$

where $C < \infty$. Hence, the sequence $\{w^k\}$ is bounded and has limit points, i.e. part (i) is true. Besides, (19) gives

$$\lim_{k\to\infty} \|w^k - w^*\|_{P_k} = \sigma \geq 0 \tag{20}$$

(see e.g. Lemma 2 in [8, p.50]) and

$$\lim_{k\to\infty} \|p^k - y^k\| = \lim_{k\to\infty} \|p^k - y^{k-1}\| = \lim_{k\to\infty} \|x^k - x^{k-1}\| = 0, \tag{21}$$

hence

$$\lim_{k\to\infty} \|y^k - y^{k-1}\| = 0. \tag{22}$$

Let $\bar{w} = (\bar{x}, \bar{y})$ be an arbitrary limit point of $\{w^k\}$, i.e.

$$\bar{w} = \lim_{s\to\infty} w^{k_s}.$$

Then there exists $J \subseteq L$ such that $J = I_{k_s}$ for infinitely many times. Without loss of generality we can suppose that $J = I_{k_s}$ for any s. Then $w^{k_s} = (x^{k_s}, y^{k_s}) \in X \times Y_J$ for any s, hence $\bar{w} = (\bar{x}, \bar{y}) \in X \times Y_J$. Setting $\varphi(z) = \mathcal{L}(z, p^k)$, $\lambda = \lambda_k$, $Q = B_k$, $U = X$, $u = x^{k-1}$, $v = x^k$, and $z = x \in X$ in (9) gives

$$\begin{aligned} 2\lambda_k\{\mathcal{L}(x^k, p^k) - \mathcal{L}(x, p^k)\} &\le \|x - x^{k-1}\|^2_{B_k} - \|x - x^k\|^2_{B_k} - \|x^k - x^{k-1}\|^2_{B_k} \\ &= 2\langle B_k(x^k - x^{k-1}), x - x^k\rangle. \end{aligned}$$

Taking the limit $k = k_s \to \infty$ due to (21)–(22) and (18) gives

$$\mathcal{L}(\bar{x}, \bar{y}) - \mathcal{L}(x, \bar{y}) \le 0. \tag{23}$$

Also, setting $\varphi(z) = -\mathcal{L}(x^k, z)$, $\lambda = \lambda_k$, $Q = E$, $U = Y_J$, $u = y^{k-1}$, $v = y^k$, and $z = y \in Y_J$ in (9) gives

$$\begin{aligned} 2\lambda_k\{\mathcal{L}(x^k, y) - \mathcal{L}(x^k, y^k)\} &\le \|y^{k-1} - y\|^2 - \|y^k - y\|^2 - \|y^k - y^{k-1}\|^2 \\ &= 2\langle y^k - y^{k-1}, y - y^k\rangle. \end{aligned}$$

Taking the limit $k = k_s \to \infty$ due to (21)–(22) gives

$$\mathcal{L}(\bar{x}, y) - \mathcal{L}(\bar{x}, \bar{y}) \le 0. \tag{24}$$

It follows from (23) and (24) that $\bar{w} = (\bar{x}, \bar{y}) \in W^*_J = D^*_J \times Y^*_J$. Hence, part (ii) is also true.

Next, if

$$\bar{w} \in \bigcap_{k=j}^{\infty} W^*_{(k)} \text{ for some } j \ge 1,$$

we can set $w^* = \bar{w}$ in (20). It follows now from (18) that

$$0 \le \sigma^2 = \lim_{k_s\to\infty} \|w^{k_s} - \bar{w}\|^2_{P_{k_s}} = \lim_{k_s\to\infty} \{\|w^{k_s} - \bar{w}\|^2_{B_{k_s}} + \|y^{k_s} - \bar{y}\|^2\} = 0,$$

i.e. $\sigma = 0$. However, we have

$$\|w^k - \bar{w}\|_P \le \|w^k - \bar{w}\|_{P_k},$$

therefore,

$$0 \le \lim_{k\to\infty} \|w^k - \bar{w}\|_P = 0,$$

which gives (16) and part (iii) is true. □

Let us discuss the matrix inequality in (15). Define by $\mu(B)$ the minimal eigenvalue of B. Since

$$\|A_{(k)}^{\top} A_{(k)}\| \leq \|A^{\top} A\|$$

and

$$\mu(B) \leq \mu(B_k),$$

it suffices to take $\tau < \mu(B)$ and

$$\lambda_k < \sqrt{0.5(\mu(B) - \tau)/\|A^{\top} A\|}. \tag{25}$$

However, taking into account peculiarities of the problem, we can utilize some more efficient choice of the parameters. The above properties enable us to establish convergence to a solution under suitable conditions.

Definition 2. *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*

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

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

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

then clearly $J^s \subseteq J^*$ if $\mathcal{P}^* \neq \varnothing$.

Theorem 2. *Suppose that assumptions (A1)–(A2) are fulfilled, the sequences $\{\lambda_k\}$ and $\{B_k\}$ satisfy conditions (14), (15), and (8) for some $\tau > 0$.*

(i) If J^s is a basic index set, then the sequence $\{w^k\}$ has limit points and each of these limit points belongs to W^*.
(ii) If J^s is a basic index set and $J^s \in \mathcal{P}$ or $J^s = J^*$, then

$$\lim_{k \to \infty} w^k = w^* \in W^*. \tag{26}$$

Proof. Let $J = J^s$ be a basic index set. By assumption, the sets W_J^* and W^* are now nonempty. Besides, $W_J^* \subseteq W_{(k)}^*$ for k large enough, hence condition (13) holds. Then the sequence $\{w^k\}$ has limit points due to Theorem 1 (i). Also, there exists $I \subseteq L$ such that $J \subseteq I = I_{k_s}$ for infinitely many times. But now I is a nonempty basic index set, hence $W_I^* \subseteq W^*$. Following the lines of part (ii) of Theorem 1, we obtain that any limit point of $\{w^{k_s}\}$ will belong to $W_I^* \subseteq W^*$. Therefore, part (i) is true.

In case (ii) we first take the case where $J \in \mathcal{P}$. It follows that $J = I_{k_s}$ for infinitely many times. Then we have similarly that any limit point w^* of $\{w^{k_s}\}$ will belong to W_J^*, but

$$w^* \in \bigcap_{k=j}^{\infty} W_{(k)}^* \text{ for some } j \geq 1, \tag{27}$$

and (26) follows from Theorem 1 (iii).

Now we take the case where $J = J^s = J^*$. Let $w^* = (x^*, y^*)$ be a limit point of $\{w^k\}$, i.e.

$$w^* = \lim_{s\to\infty} w^{k_s}$$

and let $I \in \mathcal{P}^*$. By definition, for each k_s there exists a number l_s such that $I = I_{l_s}$ and $k_s \leq l_s \leq k_s + d$. Due to (21)–(22) we have

$$w^* = \lim_{s\to\infty} w^{l_s},$$

but $y_i^{l_s} = \mathbf{0}$ for any $i \notin I$, hence $y_i^* = \mathbf{0}$ for any $i \notin I$. It follows that $w^* = (x^*, y^*) \in D_J^* \times Y_J^*$ and (27) holds. Then (26) also follows from Theorem 1 (iii). □

Theorem 3. *Suppose that assumptions (A1)–(A2) are fulfilled, the sequences* $\{\lambda_k\}$ *and* $\{B_k\}$ *satisfy conditions (14), (15), and (8) for some* $\tau > 0$.

(i) If $F^* \bigcap F_L \neq \varnothing$, *then the sequence* $\{w^k\}$ *has limit points.*
(ii) If $F^* \bigcap F_L \neq \varnothing$ *and each* $I \in \mathcal{P}$ *is a basic index set, then all the limit points of* $\{w^k\}$ *belong to* W^*.
(iii) If $F^* \bigcap F_L \neq \varnothing$, J^s *is a basic index set and* $J^s \in \mathcal{P}$ *or* $J^s = J^*$, *then the sequence* $\{w^k\}$ *converges to a point of* W^*.

Proof. By assumption, we now have $F^* \bigcap F_I = D_I^*$, $D_I^* \neq \varnothing$, and $\mathbf{0} \in Y_I^*$ for any $I \subseteq L$. It follows that

$$\left\{F^* \bigcap F_L\right\} \times \{\mathbf{0}\} \subseteq \bigcap_{k=1}^{\infty} W_{(k)}^*.$$

Therefore, (13) holds and assertion (i) follows from Theorem 1 (i). Following the lines of part (ii) of Theorem 1, we obtain that any limit point of $\{w^{k_s}\}$ will belong to $W_I^* \subseteq W^*$ where I is a nonempty basic index set. Therefore, assertion (ii) is also true. Assertion (iii) clearly follows from Theorem 2. □

The conditions of part (ii) of Theorem 2 are satisfied if for instance we take the rule $I_k \subseteq I_{k+1}$ or $I_{k+1} \subseteq I_k$ for index sets. These rules can be also applied in part (iii) of Theorem 3. It is possible to utilize more general symmetric and positive definite matrices P_k, i.e. apply variable metric matrices instead of the unit matrices with respect to y as well. However, computational preferences of this version are not clear, moreover, the method becomes very complicated.

Variable metric matrices were applied in primal-dual alternating directions methods for custom convex optimization problems; see [9, 10]. Their replacement depended only on the behaviour of the method along the trajectory. Hence, these methods did not take into account possible changes of constraints. Nevertheless, we observe that the rules in (14) and (8) are weaker than those in [9, 10].

4 Primal-dual method for multi-agent optimization problems

We now describe a specialization of the proposed method to the multi-agent optimization problem

$$\min \to \left\{ \sum_{i=1}^{m} f_i(v) \;\middle|\; \bigcap_{i=1}^{m} X_i \right\}, \tag{28}$$

where m is the number of agents (units) in the system. That is, the information about the function f_i and set X_i is known only to the i-th agent and may be unknown even to its neighbours. Besides, it is usually supposed that the agents are joined by some transmission links for information exchange so that the system is usually a connected network, whose topology may vary from time to time. This decentralized system has to find a concordant solution defined by (28).

For this reason, we replace (28) with the family of optimization problems of the form

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

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 $i = 1, \dots, m$,

$$D_I = X \bigcap F_I, \ X = X_1 \times \ldots \times X_m = \prod_{i=1}^{m} X_i, \ X_i \subseteq \mathbb{R}^n, \ i = 1, \dots, m; \tag{30}$$

the set F_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. More precisely, the maximal (full) communication network with non-oriented edges denoted by $\mathcal{F}$ corresponds to the set

$$\tilde{F} = \{x \in \mathbb{R}^{mn} \mid x_s = x_t, \ s, t = 1, \dots, m, \ s \neq t\},$$

i.e. each edge is associated with two directions or equations ($x_s = x_t$ and $x_t = x_s$). However, this definition of topology is superfluous. It seems more suitable to associate each pair of vertices (agents) (s, t) to one oriented arc i, so that $L = \{1, \dots, l\}$ is the index set of all these arcs, hence $l = m(m-1)/2$. In other words, we fix only one direction of each arc and obtain the oriented graph

$\mathcal{G}$. At the same time, the agents will use both the directions of each arc for communication. Taking subsets $I \subseteq L$, we obtain various constraint sets

$$F_I = \{x \in \mathbb{R}^{mn} \mid x_s - x_t = \mathbf{0},\ i = (s,t) \in I\}\,, \tag{31}$$

corresponding to the oriented graphs $\mathcal{G}_I$. Using non-oriented edges instead of the arcs in $\mathcal{G}_I$, we obtain the communication network $\mathcal{F}_I$. It follows that $\mathcal{F} = \mathcal{F}_L$, $\mathcal{G} = \mathcal{G}_L$, and $F = F_L$. Next, for each arc $i = (s,t)$ we can define the $n \times mn$ sub-matrix

$$A_i = (A_{i1} \cdots A_{im})\,,$$

where

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

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

$$F_I = \{x \in \mathbb{R}^{mn} \mid A_I x = \mathbf{0}\}\,,$$

where

$$A_I = \left(\{A_i^\top\}_{i \in I}\right)^\top,$$

which corresponds to the definition in (3) for $b_I = \mathbf{0}$ and any $I \subseteq L$, hence we can set $A = A_L$. Therefore, our problem (29)–(31) corresponds to (3)–(4).

In what follows, we will use the following basic assumptions.

(B1) 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.
(B2) The set $D^* = D^*_L$ is nonempty.

These assumptions imply (A1)–(A2). If the graph $\mathcal{F}_I$ for some $I \subseteq L$ is connected, then I a basic index set. Now we present an implementation of Method (PDM) for the multi-agent optimization problem (29)–(31), 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. Since many oriented graphs $\mathcal{G}_I$ are associated with the same graph $\mathcal{F}_I$, we 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 $i \in \mathcal{N}_I^-(j)$. That is, we will fix the oriented graph $\mathcal{G}$ and its subgraphs $\mathcal{G}_I$ such that agent j is associated with all the outgoing arcs for vertex j. The general Lagrange function for problems (29)–(31) is written as follows:

$$\begin{aligned} \mathcal{L}(x,y) = f(x) + \langle y, Ax\rangle &= \sum_{j\in M} f_j(x_j) + \sum_{i \in L}\langle y_i, A_i x\rangle \qquad (32) \\ &= \sum_{j \in M}\left\{ f_j(x_j) + \sum_{i \in \mathcal{N}_L^-(j)} \langle y_i, x_j\rangle - \sum_{i \in \mathcal{N}_L^+(j)} \langle y_i, x_j\rangle \right\}. \end{aligned}$$

The saddle point problems are defined in (5). Next, the method involves the auxiliary matrix B_k at each k-th iteration. We will take it to be block-diagonal, i.e.

$$B_k = \begin{pmatrix} B_k^1 & \Theta & \dots & \Theta \\ \Theta & B_k^2 & \dots & \Theta \\ \dots & \dots & \dots & \dots \\ \Theta & \Theta & \dots & B_k^m \end{pmatrix} \tag{33}$$

where B_k^s is an $n \times n$ symmetric and positive definite matrix for $s \in M$ and Θ is the $n \times n$ zero matrix. As in Section 3, for simplicity we will write $Y_{(k)} = Y_{I_k}$, $Y_{(k)}^* = Y_{I_k}^*$, etc.

Method (PDMI). 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_s^0 and y_i^0 for $i \in \mathcal{N}_{(0)}^-(s)$ and reports these values to its neighbours. This means that $y_i^0 = \mathbf{0}$ for $i \notin I_0$. The agents define a common sequence $\{\lambda_k\}$ of positive numbers. Separately, each s-th agent chooses a sequence $\{B_k^s\}$ 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_s^{k-1} and y_i^{k-1}, $i \in \mathcal{N}_{(k-1)}^-(s)$, 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 $y_i^k = \mathbf{0}$ for $i \notin I_k$.

Step 1: Each s-th agent sets

$$p_i^k = y_i^{k-1} + \lambda_k (x_s^{k-1} - x_t^{k-1}) \quad \forall i = (s, t), \ i \in \mathcal{N}_{(k)}^-(s). \tag{34}$$

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

Step 2: Each s-th agent calculates

$$v_s^k = \sum_{i \in \mathcal{N}_{(k)}^-(s)} p_i^k - \sum_{i \in \mathcal{N}_{(k)}^+(s)} p_i^k$$

and

$$x_s^k = \arg\min_{x_s \in X_s} \left\{ f_s(x_s) + \langle v_s^k, x_s \rangle + 0.5 \lambda_k^{-1} \|x_s - x_s^{k-1}\|_{B_k^s}^2 \right\} \tag{35}$$

and reports this value to its neighbours.

Step 3: Each s-th agent sets

$$y_i^k = y_i^{k-1} + \lambda_k (x_s^k - x_t^k) \quad \forall i = (s, t), \ i \in \mathcal{N}_{(k)}^-(s). \tag{36}$$

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

We observe that the agents do not store the dual variables related to the inactive arcs, i.e. $y_i^k = \mathbf{0}$ for $i \notin I_k$. If some arc $i = (s, t) \notin I_{k-1}$ becomes active at the k-th iteration, i.e. $i \in I_k$, then agent s simply sets $y_i^{k-1} = \mathbf{0}$. Due to (32), relations (34)–(36) correspond to Steps 2–4 of (PDM), respectively. Hence, the convergence properties of (PDMI) will follow directly from Theorems 2 and 3.

Corollary 1. *Suppose that assumptions (B1)–(B2) are fulfilled, the sequences* $\{\lambda_k\}$ *and* $\{B_k\}$ *satisfy conditions (14), (15), and (8) for some* $\tau > 0$.

(i) If J^s *is a basic index set, then the sequence* $\{w^k\}$, $w^k = (x^k, y^k)$, *generated by (PDMI) has limit points and each of these limit points belongs to* W^*.
(ii) If J^s *is a basic index set and* $J^s \in \mathcal{P}$ *or* $J^s = J^*$, *then (26) holds.*

Corollary 2. *Suppose that assumptions (B1)–(B2) are fulfilled, the sequences* $\{\lambda_k\}$ *and* $\{B_k\}$ *satisfy conditions (14), (15), and (8) for some* $\tau > 0$.

(i) If $F^* \bigcap F_L \neq \varnothing$, *then the sequence* $\{w^k\}$, $w^k = (x^k, y^k)$, *generated by (PDMI) has limit points.*
(ii) If $F^* \bigcap F_L \neq \varnothing$ *and each* $I \in \mathcal{P}$ *is a basic index set, then all the limit points of* $\{w^k\}$ *belong to* W^*.
(iii) If $F^* \bigcap F_L \neq \varnothing$, J^s *is a basic index set and* $J^s \in \mathcal{P}$ *or* $J^s = J^*$, *then the sequence* $\{w^k\}$ *converges to a point of* W^*.

Convergence of (PDMI) requires for all the agents to choose the common stepsize λ_k and the matrix B_k in accordance with (14), (15), and (8) at the k-th iteration. The usual choice $B_k = E$ was taken in [4], then the stepsize λ_k must satisfy the condition

$$\lambda_k \in \left[\tau, 0.5\sqrt{(1-\tau)/v(\mathcal{F}_{I_k})}\right] \tag{37}$$

for some $\tau \in (0, 1)$, where $v(\mathcal{F}_{I_k})$ is the maximal vertex degree of the graph $\mathcal{F}_{I_k}$; see (25). In case of varying communication network topology the agents may meet difficulties in evaluation of $v(\mathcal{F}_{I_k})$ and should take some its upper bound $\tilde{u} = v(\mathcal{F}_I)$ in (37) so that $\mathcal{F}_I \supseteq \mathcal{F}_{(k)}$. For instance, they can simply take $\tilde{u} = v(\mathcal{F})$ as in the fixed full graph case. However, setting this value in (37) may lead to the very small stepsize λ_k and slow convergence.

We intend to show that the agents are able to provide conditions (14), (15), and (8) independently with respect to the variable metric matrices. First of all we note that

$$A_{I_k}^\top A_{I_k} = H_{I_k} \otimes E = \begin{pmatrix} G_k^1 \\ G_k^2 \\ \cdots \\ G_k^m \end{pmatrix},$$

where H_{I_k} is the Kirchhoff matrix of the graph $\mathcal{F}_{I_k}$, E is the $n \times n$ unit matrix, $\otimes$ denotes the Kronecker product of matrices, $G_k^s = \left(G_k^{s1} \cdots G_k^{sm}\right)$ is the $n \times mn$ sub-matrix such that

$$G_k^{st} = \begin{cases} v_k^s E, & \text{if } s = t, \\ -E, & \text{if } (s,t) \in \mathcal{F}_{I_k}, \\ \Theta, & \text{otherwise,} \end{cases}$$

v_k^s is the degree of vertex s in the graph $\mathcal{F}_{I_k}$, Θ is the $n \times n$ zero matrix. Due to (33), relation (15) can be fulfilled from the separate comparison of the matrices B_k^s and G_k^s by each s-th agent. For instance, one can take

$$B_k^s = (1 + \beta_{k,s})\tilde{v}_k^s E, \;\; \beta_{k,s} > 0 \text{ for } s \in M, \tag{38}$$

where $\tilde{v}_k^s$ is some estimate of v_k^s. By using the Gershgorin theorem (see Theorem 5 in [11, Chapter XIV]), we obtain that (15) holds if

$$(1+\beta_{k,s})\tilde{v}_k^s - \tau > 4\lambda_k^2 v_k^s \text{ for } s \in M. \tag{39}$$

This means that we can take the suitable fixed stepsize value $\lambda_k = \lambda$ for all the agents in order to avoid additional communications. For instance, if $\tilde{v}_k^s = v_k^s$ and we set $\lambda_k = 1$ for all the agents, then we should choose $\beta_{k,s} > 3$, then (15) holds for

$$\tau < \min_{s \in M} \{(\beta_{k,s} - 3)v_k^s\}.$$

Also, if $\tilde{v}_k^s = v_k^s$ and we set $\lambda_k = 0.5$, then it suffices to choose any $\beta_{k,s} > 0$, then (15) holds for

$$\tau < \min_{s \in M} \{\beta_{k,s} v_k^s\}.$$

This rule together with (39) has clear preference over (37) even in the case of the fixed communication network topology or centralized systems since the matrices B_k take into account the possible non-regularity of the graph. At the same time, each s-th agent can choose his/her own value $\beta_{k,s}$ without any information from the other agents, as indicated above. In comparison with (37), this approach seems much more efficient just for the multi-agent systems.

In order to satisfy (14) and (38), the choice of parameters must be subordinated to the conditions

$$0 < (1+\beta_s)v^s \leq (1+\beta_{k+1,s})\tilde{v}_{k+1}^s \leq (1+\alpha_k)(1+\beta_{k,s})\tilde{v}_k^s,$$

where the sequence $\{\alpha_k\}$ satisfies the conditions in (8). In case of stationary or non-increasing topology it suffices to take $\tilde{v}_k^s = v_k^s$ and the constant values of $\beta_{k,s} = \beta_s$, then the metric matrices B_k will correspond to the network topology. In case of non-monotone varying topology the separate agents can apply different strategies to satisfy the conditions in (14) and (38) by proper choice of $\beta_{k,s}$ and $\tilde{v}_k^s$ instead of the exact values v_k^s if necessary. It is supposed that each agent after some finite number of iterations can evaluate the behaviour of joining edges, in particular, the further upper and lower bounds of v_k^s. Afterwards, he/she can simply take the upper bound $u^s = \tilde{v}_k^s$ in (38) instead of the exact current values v_k^s and fixed values of $\beta_{k,s}$. Afterwards, the upper bound $u^s = \tilde{v}_k^s$ can be decreased if some edges become inactive. The other way consists in applying the exact current values v_k^s, but the evaluation of both the possible upper and lower bounds of v_k^s and changing $\beta_{k,s}$. It is suitable if the behaviour of the network topology is cyclic or if the difference between the upper and lower bounds is rather small. This approach makes the control process more smooth [1]. Next, we have to choose the fixed basic topology that corresponds to an arc index set

[1] After obtaining the main results of the paper, the author found that variable metric matrices were applied in [12] in similar primal-dual methods for multi-agent optimization problems. However, these matrices are used in [12] as a general tool for enhancing convergence properties, since they are used in the same manner both for the stationary and varying topology cases. Moreover, the methods in [12] are based

$J_0 \subset L$ so that it gives the connected graph $\mathcal{F}_{J_0}$ and $J_0 \subseteq I_k$ for any k. This means that all the arcs in J_0 remain always active. The status of the other arcs may vary from time to time. This approach provides convergence properties. In general, most known iterative methods for multi-agent optimization problems utilize some other models of the changing communication topology; see e.g. [13–15].

References

1. Khan, M., Pandurangan, G., Kumar, V.: Distributed algorithms for constructing approximate minimum spanning trees in wireless sensor networks. IEEE Trans. Paral. Distrib. Syst. **20**, 124–139 (2009)
2. Lobel, I., Ozdaglar, A., Feijer, D.: Distributed multi-agent optimization with state-dependent communication. Math. Program. **129**, 255–284 (2011)
3. Peng, Z., Yan, M., Yin, W.: Parallel and distributed sparse optimization. The 47th Asilomar Conference on Signals, Systems and Computers, pp.646–659. Pacific Grove, IEEE (2013)
4. 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)
5. Antipin, A.S.: On non-gradient methods for optimization of saddle functions. In: Karmanov, V.G. (ed.) Problems of Cybernetics. Methods and Algorithms for the Analysis of Large Systems, pp.4–13. Nauchn. Sovet po Probleme "Kibernetika", Moscow (1988) [In Russian]
6. Gol'shtein, E.G., Tret'yakov, N.V.: Modified Lagrange functions. Nauka, Moscow (1989) [Engl. transl. in John Wiley and Sons, New York (1996)].
7. Rockafellar, R.T.: Convex analysis. Princeton University Press, Princeton (1970)
8. Polyak, B.T.: Introduction to optimization. Nauka, Moscow (1983) [Engl. transl. in Optimization Software, New York (1987)]
9. Kontogiorgis, S., Meyer, R.: A variable-penalty alternating directions method for convex optimization. Math. Program. **83**, 29–53 (1998)
10. He, B., Wang, S., Yang, H.: A modified variable-penalty alternating directions method for monotone variational inequalities. J. Comput. Mathem. **21**, 495–504 (2003)
11. Gantmacher, F.R.: The theory of matrices. Nauka, Moscow, 1966) [In Russian]
12. Hamedani, E.Y., Aybat, N.S.: A decentralized primal-dual method for constrained minimization of a strongly convex function. Preprint arXiv:1908.11835v4 on 21 Feb. 2022. – 27 pp.
13. Blondel, V.D., Hendrickx, J.M., Olshevsky, A., Tsitsiklis, J.N.: Convergence in multiagent coordination, consensus, and flocking. In: Proceedings of the 44-th IEEE Conference on Decision and Control, Seville (2005) – 5 pp.
14. Nedić, A., Olshevsky, A.: Distributed optimization over time-varying directed graphs. IEEE Trans. Autom. Control **60**, 601–615 (2015)

---

on the different stochastic approach to treating the varying topology, hence, the parameters of the metric matrices have no such direct relations with the communication topology as in this paper.

15. Aybat, N.S., Hamedani, E.Y.: A primal-dual method for conic constrained distributed optimization problems. In: Advances in Neural Information Processing Systems, pp.5049–5057. Neural Information Processing Systems Foundation, Barcelona (2016)