Numer. Math. Theor. Meth. Appl.
doi: 10.4208/nmtma.OA-2024-0105

Vol. **18**, No. 2, pp. 463-486
May 2025

Convergence of a Generalized Primal-Dual Algorithm with an Improved Condition for Saddle Point Problems

Fan Jiang[1], Yueying Luo[2], Xingju Cai[2] and Tanxing Wang[2,*]

[1] *Department of Information and Computing Science, School of Mathematics and Statistics, Nanjing University of Information Science and Technology, Nanjing 210044, China*

[2] *School of Mathematical Sciences, Ministry of Education Key Laboratory for NSLSCS, Nanjing Normal University, Nanjing 210023, China*

Received 7 September 2024; Accepted (in revised version) 7 January 2025

Abstract. We consider a general convex-concave saddle point problem that frequently arises in large-scale image processing. First-order primal-dual algorithms have garnered significant attention due to their promising results in solving saddle point problems. Notably, these algorithms exhibit improved performance with larger step sizes. In a recent series of articles, the upper bound on step sizes has been increased, thereby relaxing the convergence-guaranteeing condition. This paper analyzes the generalized primal-dual method proposed in [B. He, F. Ma, S. Xu, X. Yuan, SIAM J. Imaging Sci. 15 (2022)] and introduces a better condition to ensure its convergence. This enhanced condition also encompasses the optimal upper bound of step sizes in the primal-dual hybrid gradient method. We establish both the global convergence of the iterates and the ergodic $\mathcal{O}(1/N)$ convergence rate for the objective function value in the generalized primal-dual algorithm under the enhanced condition.

AMS subject classifications: 94A08, 90C25, 65K10, 68U10

Key words: Generalized primal-dual algorithm, saddle point problem, convex programming, convergence rate.

1. Introduction

Consider the following fundamental convex-concave saddle point problem:

$$\min_{x\in\mathcal{X}} \max_{y\in\mathcal{Y}} \mathcal{L}(x,y) := f(x) + \langle Kx, y\rangle - g(y), \tag{1.1}$$

*Corresponding author. *Email addresses:* 15905154902@163.com (F. Jiang), 15244848891@163.com (Y.Y. Luo), caixingju@njnu.edu.cn (X.J. Cai), wangtanxing2021@163.com (T.X. Wang)

where $\mathcal{X}$ and $\mathcal{Y}$ are two finite-dimensional real Euclidean spaces, each equipped with the inner product $\langle\cdot,\cdot\rangle$ and its induced norm $\|\cdot\| = \sqrt{\langle\cdot,\cdot\rangle}$. $K : \mathcal{X} \to \mathcal{Y}$ is a bounded linear operator with the operator norm $L = \|K\|$, $f : \mathcal{X} \to (-\infty,\infty]$ and $g : \mathcal{Y} \to (-\infty,\infty]$ are proper closed convex functions. Numerous specific instances, including zero-sum games, convex programs with linear constraints, and a variety of variational models for image processing, fall under this problem (1.1) [11, 33–35].

It is well known that (1.1) is equivalent to the primal problem

$$\min_{x\in\mathcal{X}} f(x) + g^*(Kx), \tag{1.2}$$

and the dual problem

$$\min_{y\in\mathcal{Y}} f^*(-K^*y) + g(y), \tag{1.3}$$

where g^* is the Fenchel conjugate (see the definition in Section 2) of the function g.

For solving the convex-concave saddle point problem (1.1), one benchmark is the first-order primal-dual algorithm proposed by Chambolle and Pock [5, 7]. This algorithm can be expressed as

$$\begin{cases} x^{k+1} = \operatorname{prox}_{\tau f}\big(x^k - \tau K^* y^k\big), & \text{(1.4a)} \\ y^{k+1} = \operatorname{prox}_{\sigma g}\big(y^k + \sigma K(x^{k+1} + \theta(x^{k+1} - x^k))\big), & \text{(1.4b)} \end{cases}$$

where $\theta \in [0,1]$ is a combination parameter, $\tau > 0$ and $\sigma > 0$ are the proximal parameters (or step sizes) for the two subproblems (1.4a) and (1.4b), respectively. When $\theta = 0$, this algorithm (1.4) reduces to the Arrow-Hurwicz method in [1], which received significant attention due to its promising performance in solving the total variation image restoration problem [35]. To guarantee the convergence of the Arrow-Hurwicz method, additional restrictive assumptions should be posed, such as strong convexity of the objective functions in (1.1) or some requirements on step sizes τ and σ as studied in [16, 35]. In addition, one can take the iterate generated by (1.4) as a predictor and obtain the next iterate by a simple correction step to ensure convergence [17]. Note that the range of θ can be further extended to the real number field (see [3, 18]). The scheme (1.4) with $\theta = 1$, known as the primal-dual hybrid gradient method (PDHG) in [5, 6], can be described as

$$\begin{cases} x^{k+1} = \operatorname{prox}_{\tau f}\left(x^k - \tau K^* y^k\right), \\ y^{k+1} = \operatorname{prox}_{\sigma g}\left(y^k + \sigma K(2x^{k+1} - x^k)\right), \end{cases} \tag{1.5}$$

and is the most popular choice since it does not require strong convexity or an additional correction step. It is worth mentioning that He and Yuan [17] elegantly interpreted this scheme (1.5) as an instance of the generalized proximal point algorithm with the corresponding proximal measure (CP-PPA), and it is easy to deduce the convergence of (1.5) from a PPA point of view. The equivalence of the Douglas-Rachford splitting method (DRSM) and PDHG was recently demonstrated by [28]. Note that the alternating direction method of multipliers (ADMM) [13, 14] is also able