57º SBPO
SIMPÓSIO BRASILEIRO DE PESQUISA OPERACIONAL
Gramado • RS
DE 5 A 9 DE OUTUBRO DE 2025

Convergence Analysis of an Inexact Gradient Method for Smooth Optimization

Humberto Gimenes Macedo[a,b]
Luis Felipe Bueno[a]
[a]Universidade Federal de São Paulo (UNIFESP)
[b]Instituto Tecnlógico de Aeronáutica (ITA)
[a]Av. Cesare Mansueto Giulio Lattes, 1201 - Eugênio de Melo, São José dos Campos - SP
[b]Praça Marechal Eduardo Gomes, 50 - Vila das Acácias, São José dos Campos - SP
gimenes.humberto@unifesp.br, lfelipebueno@gmail.com

ABSTRACT

Gradient-based algorithms have been employed to solve numerous problems due to their efficiency, particularly in large-scale scenarios. The convergence theories of these algorithms rely on the assumption that exact first-order information is available. However, these algorithms frequently operate with inexact information regarding the optimization problem, such as the objective function and its gradient. In recent decades, studies have analyzed the convergence of these algorithms in inexact environments for solving both optimization problems and systems of nonlinear equations. This study presents a convergence analysis of an inexact gradient method for smooth and unconstrained optimization, considering three distinct step size strategies: constant, variable, and diminishing. The analysis demonstrates that the method converges to stationary points under the condition that the inexact gradient error is upper bounded by a quadratically summable sequence of error tolerances. The results obtained from a numerical experiment strongly support the proposed convergence theory.

KEYWORDS. Inexact Gradient Method. Smooth and Unconstrained Optimization. Inexact First-Order Information.

Paper topics (Nonlinear Optimization, Continuous Optimization)

RESUMO

Algoritmos baseados em gradiente têm sido empregados para resolver numerosos problemas devido a sua eficiência. A teoria de convergência desses algoritmos depende da disponibilidade de informações exatas de primeira ordem. Entretanto, esses algoritmos frequentemente operam com informações inexatas do problema de otimização, como a função objetivo e seu gradiente. Nas últimas décadas, estudos concentraram-se na análise de convergência desses algoritmos em ambientes inexatos visando resolver problemas de otimização e sistemas não lineares. Este estudo apresenta uma análise de convergência de um método do gradiente inexato para otimização suave e irrestrita, considerando três estratégias distintas de tamanho de passo: constante, variável e decrescente. A análise demonstra que o método converge para pontos estacionários, desde que o erro do gradiente inexato seja limitado superiormente por uma sequência de tolerâncias de erro quadraticamente somável. Os resultados obtidos por meio de um experimento numérico estão em conformidade com a teoria de convergência proposta.

PALAVRAS CHAVE. Método do Gradiente Inexato. Otimização Irrestrita e Suave. Informação Inexata de Primeira Ordem.

Tópicos (Otimização Não Linear, Otimização Contínua)

DOI: 10.59254/sbpo-2025-212676
https://proceedings.science/p/212676?lang=pt-br

1. Introduction

In general, the convergence analysis of gradient-based algorithms relies on the assumption that exact first-order information is available. However, certain situations may lead to errors, particularly when the function lacks an analytic form, is affected by noise, or is a smooth approximation of a non-smooth function [Khanh et al., 2023]. In fact, obtaining such information consistently results in errors due to floating-point arithmetic. Furthermore, errors may arise when solving auxiliary problems through iterative methods to acquire information about the problem. Such errors are particularly prevalent in optimization problems with partial differential equation constraints [Hinze et al., 2009]. In these problems, to compute the objective function and its gradient, it is necessary to perform one or more numerical simulations and solve auxiliary linear systems using iterative methods. As a result, both the objective function and its gradient are often inexact. Despite the inevitability of inexact information, it is essential to recognize that inexactness is not invariably detrimental; it can be purposefully utilized to reduce the computational cost of algorithms.

In recent decades, numerous studies have been conducted on the inexactness in gradient-based algorithms to address both optimization problems and nonlinear systems of equations. These studies generally aim to develop new convergence theories that ensure the convergence of this class of algorithms under inexact scenarios. In particular, gradient methods have been rigorously investigated within inexact environments. For instance, Polyak [1987] analyzed the impact of inexact gradient values on the gradient method for convex and smooth problems. Bertsekas and Tsitsiklis [2000] justified the convergence of the gradient method in the presence of errors in smooth problems, provided that these errors are smaller than the product of the step size and an expression dependent on the exact gradient. Later, d'Aspremont [2008] extended the accelerated gradient method introduced by Nesterov [1983] to incorporate inexact gradients.

Devolder et al. [2014] introduced the concept of inexact first-order oracles and analyzed, within this framework, the performance of various first-order methods for smooth and convex problems, such as the primal, dual, and accelerated gradient methods. It was observed that the accelerated gradient method accumulates errors linearly with the number of iterations, whereas the gradient method does not experience this issue. To address this problem, Devolder et al. [2013] proposed the intermediate gradient method. The primary characteristic of this method is its ability to combine the advantages of the previously mentioned methods. Specifically, the method demonstrates a convergence rate comparable to that of the accelerated gradient method and robustness similar to that of the gradient method in scenarios where the oracle is inexact. Later, Dvurechensky and Gasnikov [2016] expanded the intermediate gradient method to include problems with stochastic oracles. Recently, Khanh et al. [2023] introduced a reduced inexact gradient method for identifying stationary points of continuously differentiable and non-convex functions.

Due to the ubiquity of inexact first-order information in practical optimization problems, this study investigates the convergence of an inexact gradient method for unconstrained and smooth optimization. The analysis considers three step size strategies: constant, variable, and diminishing. The main objective is to show that every limit point (if any) of the sequence generated by the method is a stationary point, even when the gradient information is inexact. Although many works in the literature address convergence under inexact gradients with various step size rules, the novelty of this work lies in the simplicity and generality of the method analyzed. The algorithm is very close to classical gradient method, differing only in that the gradient is inexact and the error follows a general, quadratically summable sequence. Unlike prior works that focus on more complex or tailored algorithms, this study demonstrates that even a basic gradient method with minimal modifications can achieve convergence in inexact settings, making it a practical and theoretically sound approach for real-world problems where exact derivatives may be unavailable or costly to compute. This ar-

DOI: 10.59254/sbpo-2025-212676
https://proceedings.science/p/212676?lang=pt-br

ticle is organized as follows. Section 2 presents the theoretical preliminaries. Section 3 introduces the inexact gradient method, followed by its convergence analysis in Section 4. Section 5 presents a numerical experiment to support the theoretical findings. Final remarks are given in Section 6.

2. Mathematical Preliminaries

This section introduces basic results concerning smooth functions. Smoothness represents a characteristic of functions often observed in convergence theorems of gradient-based algorithms, as it ensures that the gradient remains relatively stable across the domain.

Definition (Smoothness). Let $L \geq 0$. A continuously differentiable function $\widetilde{F} : \mathbb{R}^m \to \mathbb{R}$ is said to be L-smooth with respect to the Euclidean norm if its gradient $\boldsymbol{\nabla}\widetilde{F} : \mathbb{R}^m \to \mathbb{R}^m$ is Lipschitz continuous, as follows:

$$\left\|\boldsymbol{\nabla}\widetilde{F}(\mathbf{x}) - \boldsymbol{\nabla}\widetilde{F}(\mathbf{y})\right\|_2 \leq L\|\mathbf{x}-\mathbf{y}\|_2, \quad \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^m.$$

The constant L is referred to as the Lipschitz constant.

An important fact about smooth functions is that they can be upper bounded by a quadratic function at every point within their domains, as ensured by the following lemma.

Lemma 2.1 (Quadratic Upper Bound). *Let $\widetilde{F} : \mathbb{R}^m \to \mathbb{R}$ be an L-smooth function. Then,*

$$\widetilde{F}(\mathbf{y}) \leq \widetilde{F}(\mathbf{z}) + \boldsymbol{\nabla}\widetilde{F}(\mathbf{z})^T(\mathbf{y}-\mathbf{z}) + \frac{L}{2}\|\mathbf{y}-\mathbf{z}\|_2^2, \quad \forall \mathbf{y}, \mathbf{z} \in \mathbb{R}^m. \tag{1}$$

Proof. The proof can be found in Beck [2017]. □

Under moderate conditions, Lemma 2.1 is valid even when the gradient is inexact, as demonstrated by the subsequent lemma.

Lemma 2.2 (Inexact Quadratic Upper Bound). *Let $\widetilde{F} : \mathbb{R}^m \to \mathbb{R}$ be an L-smooth function. If the inexact gradient of $\widetilde{F}$, denoted by $\overline{\boldsymbol{\nabla}}\widetilde{F} : \mathbb{R}^m \to \mathbb{R}^m$, satisfies the condition*

$$\left\|\overline{\boldsymbol{\nabla}}\widetilde{F}(\mathbf{z}) - \boldsymbol{\nabla}\widetilde{F}(\mathbf{z})\right\|_2 \leq \delta_{\mathbf{z}}, \quad \forall \mathbf{z} \in \mathbb{R}^m, \tag{2}$$

where $\boldsymbol{\nabla}\widetilde{F} : \mathbb{R}^m \to \mathbb{R}^m$ is the exact gradient of $\widetilde{F}$ and $\delta_{\mathbf{z}} \in \mathbb{R}_{\geq 0}$ is the error tolerance, then

$$\widetilde{F}(\mathbf{y}) \leq \widetilde{F}(\mathbf{z}) + \overline{\boldsymbol{\nabla}}\widetilde{F}(\mathbf{z})^T(\mathbf{y}-\mathbf{z}) + \frac{L'}{2}\|\mathbf{y}-\mathbf{z}\|_2^2 + \delta'_{\mathbf{z}}, \quad \forall \mathbf{y}, \mathbf{z} \in \mathbb{R}^m,$$

wherein $L' = 2L$ and $\delta'_{\mathbf{z}} = \delta_{\mathbf{z}}^2/2L$.

Proof. As $\widetilde{F}$ is L-smooth, from Lemma 2.1 it follows that

$$\widetilde{F}(\mathbf{y}) \leq \widetilde{F}(\mathbf{z}) + \boldsymbol{\nabla}\widetilde{F}(\mathbf{z})^T(\mathbf{y}-\mathbf{z}) + \frac{L}{2}\|\mathbf{y}-\mathbf{z}\|_2^2, \quad \forall \mathbf{y} \in \mathbb{R}^m. \tag{3}$$

Note that the term with the gradient of $\widetilde{F}$ in (3) can be rewritten as

$$\boldsymbol{\nabla}\widetilde{F}(\mathbf{z})^T(\mathbf{y}-\mathbf{z}) = \overline{\boldsymbol{\nabla}}\widetilde{F}(\mathbf{z})^T(\mathbf{y}-\mathbf{z}) + \left[\boldsymbol{\nabla}\widetilde{F}(\mathbf{z}) - \overline{\boldsymbol{\nabla}}\widetilde{F}(\mathbf{z})\right]^T(\mathbf{y}-\mathbf{z}). \tag{4}$$

Furthermore, the second term in (4) can be upper bounded by utilizing Young-Fenchel's inequality, as stated in Vasin et al. [2023]. This leads to

$$\left[\boldsymbol{\nabla}\widetilde{F}(\mathbf{z}) - \overline{\boldsymbol{\nabla}}\widetilde{F}(\mathbf{z})\right]^T(\mathbf{y}-\mathbf{z}) \leq \frac{1}{2L}\left\|\boldsymbol{\nabla}\widetilde{F}(\mathbf{z}) - \overline{\boldsymbol{\nabla}}\widetilde{F}(\mathbf{z})\right\|_2^2 + \frac{L}{2}\|\mathbf{y}-\mathbf{z}\|_2^2. \tag{5}$$

DOI: 10.59254/sbpo-2025-212676
https://proceedings.science/p/212676?lang=pt-br

Observe that (4) can be effectively bounded above using (5), yielding the following expression:

$$\nabla \widetilde{F}(\mathbf{z})^T(\mathbf{y}-\mathbf{z}) \leq \overline{\nabla}\widetilde{F}(\mathbf{z})^T(\mathbf{y}-\mathbf{z}) + \frac{1}{2L}\left\|\nabla \widetilde{F}(\mathbf{z}) - \overline{\nabla}\widetilde{F}(\mathbf{z})\right\|_2^2 + \frac{L}{2}\|\mathbf{y}-\mathbf{z}\|_2^2. \tag{6}$$

Inequality (3) can be reformulated using (6), as follows:

$$\widetilde{F}(\mathbf{y}) \leq \widetilde{F}(\mathbf{z}) + \overline{\nabla}\widetilde{F}(\mathbf{z})^T(\mathbf{y}-\mathbf{z}) + \frac{1}{2L}\left\|\nabla \widetilde{F}(\mathbf{z}) - \overline{\nabla}\widetilde{F}(\mathbf{z})\right\|_2^2 + L\|\mathbf{y}-\mathbf{z}\|_2^2. \tag{7}$$

Since (2) holds, one can rewrite (7) as

$$\widetilde{F}(\mathbf{y}) = \widetilde{F}(\mathbf{z}) + \overline{\nabla}\widetilde{F}(\mathbf{z})^T(\mathbf{y}-\mathbf{z}) + \frac{L'}{2}\|\mathbf{y}-\mathbf{z}\|_2^2 + \delta'_{\mathbf{z}}, \quad \forall \mathbf{y} \in \mathbb{R}^m,$$

where $L' = 2L$ and $\delta'_{\mathbf{z}} = \delta_{\mathbf{z}}^2/2L$. □

3. Inexact Gradient Method (IGM)

Given an L-smooth function $\widetilde{F} : \mathbb{R}^m \to \mathbb{R}$, for which there exist $\mathbf{z}_\star \in \mathbb{R}^m$ such that $\widetilde{F}(\mathbf{z}_\star) = \min_{\mathbf{z}\in\mathbb{R}^m} \widetilde{F}(\mathbf{z})$, the purpose is to solve unconstrained optimization problems of the form

$$\min_{\mathbf{z}\,\in\,\mathbb{R}^m} \quad \widetilde{F}(\mathbf{z})$$

using an IGM. This method is defined by the following iterative updating rule:

$$\mathbf{z}_{k+1} = \mathbf{z}_k - t_k \overline{\nabla}\widetilde{F}(\mathbf{z}_k), \quad \forall k \in \mathbb{N}_0, \tag{8}$$

where $\mathbf{z}_0 \in \mathbb{R}^m$ is a given initial point, $(t_k)_{k\in\mathbb{N}_0}$ is a non-negative sequence of step sizes. The inexact gradient $\overline{\nabla}\widetilde{F} : \mathbb{R}^m \to \mathbb{R}^m$ is an approximation to the exact gradient $\nabla\widetilde{F} : \mathbb{R}^m \to \mathbb{R}^m$ that satisfies the following condition:

$$\left\|\overline{\nabla}\widetilde{F}(\mathbf{z}_k) - \nabla\widetilde{F}(\mathbf{z}_k)\right\|_2 \leq \delta_k, \quad \forall k \in \mathbb{N}_0, \tag{9}$$

wherein $(\delta_k)_{k\in\mathbb{N}_0}$ is a non-negative sequence of error tolerances. Similar to the gradient method (GM), the IGM also has a descent lemma that is essential for formulating its convergence theory.

Lemma 3.1 (**Inexact Descent Lemma**). *Let $\widetilde{F} : \mathbb{R}^m \to \mathbb{R}$ be an L-smooth function. Given an initial point $\mathbf{z}_0 \in \mathbb{R}^m$, consider the sequence $(\mathbf{z}_k)_{k\in\mathbb{N}}$ generated by equation* (8)*. If* (9) *holds, then*

$$\widetilde{F}(\mathbf{z}_{k+1}) \leq \widetilde{F}(\mathbf{z}_k) - t_k\left(1 - \frac{L't_k}{2}\right)\left\|\overline{\nabla}\widetilde{F}(\mathbf{z}_k)\right\|_2^2 + \delta'_k, \quad \forall k \in \mathbb{N}_0, \tag{10}$$

where $\delta'_k = \delta_k^2/2L$ and $L' = 2L$.

Proof. As $\widetilde{F} : \mathbb{R}^m \to \mathbb{R}$ is an L-smooth function and (9) holds, it follows from Lemma 2.2 that

$$\widetilde{F}(\mathbf{y}) \leq \widetilde{F}(\mathbf{z}_k) + \overline{\nabla}\widetilde{F}(\mathbf{z}_k)^T(\mathbf{y}-\mathbf{z}_k) + \frac{L'}{2}\|\mathbf{y}-\mathbf{z}_k\|_2^2 + \delta'_k, \quad \forall \mathbf{y} \in \mathbb{R}^m, \tag{11}$$

Applying (11) with $\mathbf{y} = \mathbf{z}_{k+1}$ and utilizing (8) in the resulting expression yields

$$\begin{aligned}\widetilde{F}(\mathbf{z}_{k+1}) &\leq \widetilde{F}(\mathbf{z}_k) + \overline{\nabla}\widetilde{F}(\mathbf{z}_k)^T(\mathbf{z}_{k+1}-\mathbf{z}_k) + \frac{L'}{2}\|\mathbf{z}_{k+1}-\mathbf{z}_k\|_2^2 + \delta'_k \\ &\leq \widetilde{F}(\mathbf{z}_k) - t_k\left(1 - \frac{L't_k}{2}\right)\left\|\overline{\nabla}\widetilde{F}(\mathbf{z}_k)\right\|_2^2 + \delta'_k.\end{aligned}$$

□

DOI: 10.59254/sbpo-2025-212676
https://proceedings.science/p/212676?lang=pt-br

According to (9), if the inexact gradient approaches zero, the exact gradient will also tend to zero.

Lemma 3.2 (Exact Stationarity). *Let $\widetilde{F} : \mathbb{R}^m \to \mathbb{R}^m$ be a differentiable function and $(\mathbf{z}_k)_{k\in\mathbb{N}_0}$ be the sequence generated by* (8)*, where $\mathbf{z}_0 \in \mathbb{R}^m$ is a given initial point. Suppose that* (9) *holds for a non-negative and quadratically summable sequence $(\delta_k)_{k\in\mathbb{N}_0}$. If*

$$\lim_{k\to\infty} \overline{\boldsymbol{\nabla}}\widetilde{F}(\mathbf{z}_k) = \mathbf{0}_{\mathbb{R}^m},$$

then

$$\lim_{k\to\infty} \boldsymbol{\nabla}\widetilde{F}(\mathbf{z}_k) = \mathbf{0}_{\mathbb{R}^m}. \tag{12}$$

Proof. From the triangle inequality and (9), one obtains

$$\begin{aligned}\left\|\boldsymbol{\nabla}\widetilde{F}(\mathbf{z}_k)\right\|_2 &= \left\|\overline{\boldsymbol{\nabla}}\widetilde{F}(\mathbf{z}_k) + \boldsymbol{\nabla}\widetilde{F}(\mathbf{z}_k) - \overline{\boldsymbol{\nabla}}\widetilde{F}(\mathbf{z}_k)\right\|_2 \\ &\leq \left\|\overline{\boldsymbol{\nabla}}\widetilde{F}(\mathbf{z}_k)\right\|_2 + \left\|\boldsymbol{\nabla}\widetilde{F}(\mathbf{z}_k) - \overline{\boldsymbol{\nabla}}\widetilde{F}(\mathbf{z}_k)\right\|_2 \\ &\leq \left\|\overline{\boldsymbol{\nabla}}\widetilde{F}(\mathbf{z}_k)\right\|_2 + \delta_k. \end{aligned} \tag{13}$$

As both terms on the right-hand side of (13) approach zero as $k \to \infty$, (12) is obtained. □

4. Convergence Analysis

This section develops the convergence theory of the IGM for smooth functions using three distinct step size strategies: constant, variable, and diminishing. It will be demonstrated that every limit point (if any) of the sequence generated by (8) is a stationary point under each of these strategies.

4.1. Constant Step Size

Consider the IGM with a constant step size such that $t_k \equiv t > 0$ for all $k \in \mathbb{N}_0$. The subsequent theorem establishes the convergence of this method to a stationary point under the specified condition.

Theorem 4.1 (Convergence for a Constant Step Size). *Let $\widetilde{F} : \mathbb{R}^m \to \mathbb{R}$ be an L-smooth function. Given an initial point $\mathbf{z}_0 \in \mathbb{R}^m$, consider the sequence $(\mathbf{z}_k)_{k\in\mathbb{N}}$ generated by* (8) *with a constant step size $t \in \left(0, \frac{2}{L'}\right)$. If* (9) *holds for a non-negative and quadratically summable sequence $(\delta_k)_{k\in\mathbb{N}_0}$, then*

$$\lim_{k\to\infty} \overline{\boldsymbol{\nabla}}\widetilde{F}(\mathbf{z}_k) = \mathbf{0}_{\mathbb{R}^m} \quad \textit{and} \quad \lim_{k\to\infty} \boldsymbol{\nabla}\widetilde{F}(\mathbf{z}_k) = \mathbf{0}_{\mathbb{R}^m}.$$

Proof. As $\widetilde{F}$ is an L-smooth function and (9) holds, Lemma 3.1 assures that

$$\widetilde{F}(\mathbf{z}_{k+1}) \leq \widetilde{F}(\mathbf{z}_k) - t\left(1 - \frac{L't}{2}\right)\left\|\overline{\boldsymbol{\nabla}}\widetilde{F}(\mathbf{z}_k)\right\|_2^2 + \delta_k', \quad \forall k \in \mathbb{N}_0, \tag{14}$$

Summing (14) from 0 to $K-1$ (where $K \geq 1$), one obtains

$$\sum_{k=0}^{K-1}\left[\widetilde{F}(\mathbf{z}_{k+1}) - \widetilde{F}(\mathbf{z}_k)\right] \leq -t\left(1 - \frac{L't}{2}\right)\sum_{k=0}^{K-1}\left\|\overline{\boldsymbol{\nabla}}\widetilde{F}(\mathbf{z}_k)\right\|_2^2 + \frac{1}{2L}\sum_{k=0}^{K-1}\delta_k^2. \tag{15}$$

Note that the left-hand side of (15) is a telescoping sum. Thus, this inequality can be rewritten as

$$\widetilde{F}(\mathbf{z}_K) \leq \widetilde{F}(\mathbf{z}_0) + \frac{1}{2L}\sum_{k=0}^{K-1}\delta_k^2 - t\left(1 - \frac{L't}{2}\right)\sum_{k=0}^{K-1}\left\|\overline{\boldsymbol{\nabla}}\widetilde{F}(\mathbf{z}_k)\right\|_2^2. \tag{16}$$

DOI: 10.59254/sbpo-2025-212676
https://proceedings.science/p/212676?lang=pt-br

Since $\widetilde{F}$ is lower bounded and $(\delta_k)_{k\in\mathbb{N}_0}$ is quadratically summable, it follows from (16) that the series $\sum_{k=0}^{\infty}\left\|\overline{\boldsymbol{\nabla}}\widetilde{F}(\mathbf{z}_k)\right\|_2^2$ must converge. This implies that $\left\|\overline{\boldsymbol{\nabla}}\widetilde{F}(\mathbf{z}_k)\right\|_2 \to 0$ as $k\to\infty$. Consequently, $\lim_{k\to\infty}\overline{\boldsymbol{\nabla}}\widetilde{F}(\mathbf{z}_k)=\mathbf{0}_{\mathbb{R}^m}$. Therefore, Lemma 3.2 ensures that $\lim_{k\to\infty}\boldsymbol{\nabla}\widetilde{F}(\mathbf{z}_k)=\mathbf{0}_{\mathbb{R}^m}$. □

4.2. Variable Step Size

The flexibility of the IGM can be enhanced by allowing a variable step size. In this context, suppose that there exists $\bar{t},\widetilde{t}>0$ such that $\bar{t}\le\widetilde{t}<\frac{2}{L'}$. Then, $\widetilde{t}$ can be represented as follows:

$$\widetilde{t}=\frac{2-\gamma}{L'},\quad \text{with}\quad 0<\gamma\le 2-\bar{t}L'.$$

Consequently, instead of a constant step size, one can adopt a sequence $(t_k)_{k\in\mathbb{N}_0}$ such that

$$t_k\in\left[\bar{t},\frac{2-\gamma}{L'}\right],\quad \forall k\in\mathbb{N}_0. \tag{17}$$

The following theorem establishes the convergence of the IGM to a stationary point under these conditions.

Theorem 4.2 (Convergence for a Variable Step Size). *Let $\widetilde{F}:\mathbb{R}^m\to\mathbb{R}$ be an L-smooth function. Given an initial point $\mathbf{z}_0\in\mathbb{R}^m$, consider the sequence $(\mathbf{z}_k)_{k\in\mathbb{N}}$ generated by (8) with a step size sequence $(t_k)_{k\in\mathbb{N}_0}$ such that $t_k\in\left[\bar{t},\frac{2-\gamma}{L'}\right]$, where $\bar{t}>0$ and $0<\gamma\le 2-\bar{t}L'$. If (9) holds for a non-negative and quadratically summable sequence $(\delta_k)_{k\in\mathbb{N}_0}$, then*

$$\lim_{k\to\infty}\overline{\boldsymbol{\nabla}}\widetilde{F}(\mathbf{z}_k)=\mathbf{0}_{\mathbb{R}^m}\quad \textit{and}\quad \lim_{k\to\infty}\boldsymbol{\nabla}\widetilde{F}(\mathbf{z}_k)=\mathbf{0}_{\mathbb{R}^m}.$$

Proof. Given that $\widetilde{F}$ is L-smooth and (9) holds, it follows from Lemma 3.1 that

$$\widetilde{F}(\mathbf{z}_{k+1})\le\widetilde{F}(\mathbf{z}_k)-t_k\left(1-\frac{L't_k}{2}\right)\left\|\overline{\boldsymbol{\nabla}}\widetilde{F}(\mathbf{z}_k)\right\|_2^2+\delta_k',\quad \forall k\in\mathbb{N}_0. \tag{18}$$

As $t_k\le(2-\gamma)/L'$ for all $k\in\mathbb{N}_0$, it can be verified that

$$-\frac{\gamma}{2}\ge-\left(1-\frac{L't_k}{2}\right),\quad \forall k\in\mathbb{N}_0. \tag{19}$$

Using (19), it is possible to rewrite (18) as follows

$$\widetilde{F}(\mathbf{z}_{k+1})\le\widetilde{F}(\mathbf{z}_k)-\frac{\gamma}{2}t_k\left\|\overline{\boldsymbol{\nabla}}\widetilde{F}(\mathbf{z}_k)\right\|_2^2+\delta_k'. \tag{20}$$

Summing inequality (20) from 0 to $K-1$ (where $K\ge1$) yields

$$\widetilde{F}(\mathbf{z}_K)\le\widetilde{F}(\mathbf{z}_0)+\frac{1}{2L}\sum_{k=0}^{K-1}\delta_k^2-\frac{\gamma}{2}\sum_{k=0}^{K-1}t_k\left\|\overline{\boldsymbol{\nabla}}\widetilde{F}(\mathbf{z}_k)\right\|_2^2. \tag{21}$$

Since $\widetilde{F}$ is lower bounded and $(\delta_k)_{k\in\mathbb{N}}$ is quadratically summable, (21) implies that

$$\sum_{k=0}^{\infty}t_k\left\|\overline{\boldsymbol{\nabla}}\widetilde{F}(\mathbf{z}_k)\right\|_2^2<\infty. \tag{22}$$

DOI: 10.59254/sbpo-2025-212676
https://proceedings.science/p/212676?lang=pt-br

Note that $t_k \geq \bar{t} > 0$ for all $k \in \mathbb{N}_0$. This means that the sequence of step sizes is bounded away from zero. Thus, from (22) one can obtain the following result:

$$\bar{t} \sum_{k=0}^{\infty} \left\| \overline{\nabla} \widetilde{F}(\mathbf{z}_k) \right\|_2^2 \leq \sum_{k=0}^{\infty} t_k \left\| \overline{\nabla} \widetilde{F}(\mathbf{z}_k) \right\|_2^2 < \infty.$$

Consequently,

$$\sum_{k=0}^{\infty} \left\| \overline{\nabla} \widetilde{F}(\mathbf{z}_k) \right\|_2^2 < \infty.$$

This ensures that $\lim_{k\to\infty} \overline{\nabla} \widetilde{F}(\mathbf{z}_k) = \mathbf{0}_{\mathbb{R}^m}$. Therefore, Lemma 3.2 implies that $\lim_{k\to\infty} \nabla \widetilde{F}(\mathbf{z}_k) = \mathbf{0}_{\mathbb{R}^m}$. □

4.3. Diminishing Step Size

Instead of considering a constant or variable step size, one approach involves utilizing a diminishing step size, which satisfies the following conditions:

$$\sum_{k=0}^{\infty} t_k^2 < \infty, \tag{23a}$$

and

$$\sum_{k=0}^{\infty} t_k = \infty. \tag{23b}$$

To establish the convergence of the IGM with a diminishing step size, the subsequent lemma is necessary.

Lemma 4.3 (Stationarity Lemma). *Let $\widetilde{F} : \mathbb{R}^m \to \mathbb{R}$ be an L-smooth function. Given an initial point $\mathbf{z}_0 \in \mathbb{R}^m$, consider the sequence $(\mathbf{z}_k)_{k\in\mathbb{N}}$ generated by* (8) *with a step size sequence $(t_k)_{k\in\mathbb{N}_0}$. Suppose that* (9) *holds for a non-negative and summable sequence $(\delta_k)_{k\in\mathbb{N}_0}$. If*

$$\sum_{k=0}^{\infty} t_k \left\| \overline{\nabla} \widetilde{F}(\mathbf{z}_k) \right\|_2^2 < \infty \tag{24}$$

and

$$\liminf_{k\to\infty} \left\| \overline{\nabla} \widetilde{F}(\mathbf{z}_k) \right\|_2 = 0, \tag{25}$$

then

$$\lim_{k\to\infty} \left\| \overline{\nabla} \widetilde{F}(\mathbf{z}_k) \right\|_2 = 0. \tag{26}$$

Proof. Suppose, in anticipation of deriving a contradiction, that (26) fails to hold. Consequently, there exists $\epsilon > 0$ such that

$$\limsup_{k\to\infty} \left\| \overline{\nabla} \widetilde{F}(\mathbf{z}_k) \right\|_2 \geq \epsilon. \tag{27}$$

Given the validity of (25) and (27), the sequence of inexact gradient norms demonstrates an oscillatory pattern, enabling the division of the sequence into segments characterized by relatively small or large norms. Accordingly, define $(m_j)_{j\in\mathbb{N}_0}$ and $(n_j)_{j\in\mathbb{N}_0}$ as sequences of indices such that

$$m_j < n_j < m_{j+1}, \quad \forall j \in \mathbb{N}_0,$$

DOI: 10.59254/sbpo-2025-212676
https://proceedings.science/p/212676?lang=pt-br

and

$$\left\|\overline{\nabla}\widetilde{F}(\mathbf{z}_k)\right\|_2 > \frac{\epsilon}{3}, \quad \text{for } m_j \leq k < n_j, \tag{28}$$

$$\left\|\overline{\nabla}\widetilde{F}(\mathbf{z}_k)\right\|_2 \leq \frac{\epsilon}{3}, \quad \text{for } n_j \leq k < m_{j+1}. \tag{29}$$

Since $(\delta_k)_{k\in\mathbb{N}_0}$ is summable, given $\frac{\epsilon}{12} > 0$ there exist a sufficiently large constant $\overline{j}_1 \in \mathbb{N}_0$ such that

$$\sum_{k=m_{\overline{j}_1}}^{\infty} \delta_k < \frac{\epsilon}{12}. \tag{30}$$

Similarly, as a result of (24), for any $\epsilon^2/18L > 0$ a sufficiently large constant $\overline{j}_2 \in \mathbb{N}_0$ can be identified such that

$$\sum_{k=m_{\overline{j}_2}}^{\infty} t_k \left\|\overline{\nabla}\widetilde{F}(\mathbf{z}_k)\right\|_2^2 < \frac{\epsilon^2}{18L}. \tag{31}$$

Define $\overline{j} \in \mathbb{N}_0$ as

$$\overline{j} = \max(\overline{j}_1, \overline{j}_2).$$

Also, note that

$$\begin{aligned}\left\|\overline{\nabla}\widetilde{F}(\mathbf{z}_{k+1}) - \overline{\nabla}\widetilde{F}(\mathbf{z}_k)\right\|_2 &\leq \left\|\overline{\nabla}\widetilde{F}(\mathbf{z}_{k+1}) - \nabla\widetilde{F}(\mathbf{z}_{k+1})\right\|_2 + \left\|\nabla\widetilde{F}(\mathbf{z}_{k+1}) - \nabla\widetilde{F}(\mathbf{z}_k)\right\|_2 \\ &\quad + \left\|\overline{\nabla}\widetilde{F}(\mathbf{z}_k) - \nabla\widetilde{F}(\mathbf{z}_k)\right\|_2 \\ &\leq \delta_{k+1} + L\|\mathbf{z}_{k+1} - \mathbf{z}_k\|_2 + \delta_k. \end{aligned} \tag{32}$$

Consequently, for every $j \in \mathbb{N}_0$ where $j \geq \overline{j}$ and each $m \in \mathbb{N}_0$ satisfying $m_j \leq m \leq n_j - 1$, the implications of (30) and (32) are as follows:

$$\begin{aligned}\left\|\overline{\nabla}\widetilde{F}(\mathbf{z}_{n_j}) - \overline{\nabla}\widetilde{F}(\mathbf{z}_m)\right\|_2 &\leq \sum_{k=m}^{n_j-1} \left\|\overline{\nabla}\widetilde{F}(\mathbf{z}_{k+1}) - \overline{\nabla}\widetilde{F}(\mathbf{z}_k)\right\|_2 \\ &\leq \sum_{k=m}^{n_j-1} \delta_{k+1} + L\sum_{k=m}^{n_j-1} \|\mathbf{z}_{k+1} - \mathbf{z}_k\|_2 + \sum_{k=m}^{n_j-1} \delta_k \\ &= \frac{\epsilon}{6} + L\sum_{k=m}^{n_j-1} t_k \left\|\overline{\nabla}\widetilde{F}(\mathbf{z}_k)\right\|_2. \end{aligned} \tag{33}$$

By employing (28) and (31), one can express (33) in the following form:

$$\begin{aligned}\left\|\overline{\nabla}\widetilde{F}(\mathbf{z}_{n_j}) - \overline{\nabla}\widetilde{F}(\mathbf{z}_m)\right\|_2 &\leq \frac{\epsilon}{6} + \frac{3L}{\epsilon}\sum_{k=m}^{n_j-1} t_k \left\|\overline{\nabla}\widetilde{F}(\mathbf{z}_k)\right\|_2^2 \\ &\leq \frac{\epsilon}{6} + \frac{3L}{\epsilon}\frac{\epsilon^2}{18L} \\ &= \frac{\epsilon}{3}. \end{aligned} \tag{34}$$

DOI: 10.59254/sbpo-2025-212676
https://proceedings.science/p/212676?lang=pt-br

Based on (29) and (34), it can be established that

$$\begin{aligned}\left\|\overline{\nabla}\widetilde{F}(\mathbf{z}_m)\right\|_2 &= \left\|\overline{\nabla}\widetilde{F}(\mathbf{z}_m) - \overline{\nabla}\widetilde{F}(\mathbf{z}_{n_j}) + \overline{\nabla}\widetilde{F}(\mathbf{z}_{n_j})\right\|_2 \\ &\leq \left\|\overline{\nabla}\widetilde{F}(\mathbf{z}_m) - \overline{\nabla}\widetilde{F}(\mathbf{z}_{n_j})\right\|_2 + \left\|\overline{\nabla}\widetilde{F}(\mathbf{z}_{n_j})\right\|_2 \\ &\leq \frac{\epsilon}{3} + \frac{\epsilon}{3} \\ &= \frac{2\epsilon}{3}, \quad \forall j \geq \overline{j}, \quad m_j \leq m \leq n_j - 1. \end{aligned} \tag{35}$$

From (29) and (35), it follows that

$$\left\|\overline{\nabla}\widetilde{F}(\mathbf{z}_m)\right\|_2 \leq \frac{2\epsilon}{3}, \quad \forall j \geq \overline{j}, \quad m_j \leq m < m_{j+1}.$$

Thus,

$$\left\|\overline{\nabla}\widetilde{F}(\mathbf{z}_m)\right\|_2 \leq \frac{2\epsilon}{3}, \quad \forall m \geq m_{\overline{j}}.$$

However, this contradicts (27). Therefore, (26) holds. □

Based on Lemma 4.3, the subsequent theorem establishes the convergence of the IGM with a diminishing step size to a stationary point.

Theorem 4.4 (**Convergence for a Diminishing Step Size**). *Let $\widetilde{F} : \mathbb{R}^m \to \mathbb{R}$ be an L-smooth function. Given an initial point $\mathbf{z}_0 \in \mathbb{R}^m$, consider the sequence $(\mathbf{z}_k)_{k\in\mathbb{N}}$ generated by* (8) *with a step size sequence $(t_k)_{k\in\mathbb{N}_0}$ satisfying* (23a) *and* (23b)*. If* (9) *holds for a non-negative and summable sequence $(\delta_k)_{k\in\mathbb{N}_0}$, then*

$$\lim_{k\to\infty} \overline{\nabla}\widetilde{F}(\mathbf{z}_k) = \mathbf{0}_{\mathbb{R}^m} \quad \textit{and} \quad \lim_{k\to\infty} \nabla\widetilde{F}(\mathbf{z}_k) = \mathbf{0}_{\mathbb{R}^m}. \tag{36}$$

Proof. As a consequence of the smoothness of $\widetilde{F}$ and the validity of (9), Lemma 3.1 ensures that

$$\widetilde{F}(\mathbf{z}_{k+1}) \leq \widetilde{F}(\mathbf{z}_k) - t_k\left(1 - \frac{L't_k}{2}\right)\left\|\overline{\nabla}\widetilde{F}(\mathbf{z}_k)\right\|_2^2 + \delta'_k, \quad \forall k \in \mathbb{N}_0. \tag{37}$$

Since $(t_k)_{k\in\mathbb{N}_0}$ satisfies (23a), it converges to zero. Consequently, there exists a constant $\overline{k} \in \mathbb{N}$ such that

$$t_k \leq \frac{2(1-c)}{L'} \iff -c \geq -\left(1 - \frac{L't_k}{2}\right), \quad \forall k \geq \overline{k}, \tag{38}$$

where $c \in (0, 1)$. Using (38) in (37), one obtains

$$\widetilde{F}(\mathbf{z}_{k+1}) \leq \widetilde{F}(\mathbf{z}_k) - ct_k\left\|\overline{\nabla}\widetilde{F}(\mathbf{z}_k)\right\|_2^2 + \delta'_k, \quad \forall k \geq \overline{k}. \tag{39}$$

Summing (39) from $\overline{k}$ to $\overline{K} - 1$ (where $\overline{K} > \overline{k} + 1$) yields

$$\widetilde{F}(\mathbf{z}_{\overline{K}}) \leq \widetilde{F}(\mathbf{z}_{\overline{k}}) + \frac{1}{2L}\sum_{k=\overline{k}}^{\overline{K}-1}\delta_k^2 - c\sum_{k=\overline{k}}^{\overline{K}-1} t_k\left\|\overline{\nabla}\widetilde{F}(\mathbf{z}_k)\right\|_2^2. \tag{40}$$

Rearranging the terms of (40), one arrives at

$$c\sum_{k=\overline{k}}^{\overline{K}-1} t_k\left\|\overline{\nabla}\widetilde{F}(\mathbf{z}_k)\right\|_2^2 \leq \widetilde{F}(\mathbf{z}_{\overline{k}}) - \widetilde{F}(\mathbf{z}_\star) + \frac{1}{2L}\sum_{k=\overline{k}}^{\overline{K}-1}\delta_k^2. \tag{41}$$

DOI: 10.59254/sbpo-2025-212676
https://proceedings.science/p/212676?lang=pt-br

As $(\delta_k)_{k\in\mathbb{N}_0}$ is a non-negative and summable sequence, it follows that it is also quadratically summable. Thus, (41) implies that

$$\sum_{k=\overline{k}}^{\infty} t_k \left\| \overline{\nabla} \widetilde{F}(\mathbf{z}_k) \right\|_2^2 < \infty. \tag{42}$$

To establish a contradiction, assume that for a given constant $\epsilon > 0$, there is $\widetilde{k} \in \mathbb{N}$ such that

$$\left\| \overline{\nabla} \widetilde{F}(\mathbf{z}_k) \right\|_2 \geq \epsilon, \quad \forall k \geq \widetilde{k}.$$

Thus, employing (23b) leads to the subsequent condition:

$$\sum_{k=\overline{k}}^{\infty} t_k \left\| \overline{\nabla} \widetilde{F}(\mathbf{z}_k) \right\|_2^2 \geq \epsilon^2 \sum_{k=\overline{k}}^{\infty} t_k = \infty,$$

which contradicts (42). As a result, one has

$$\liminf_{k\to\infty} \left\| \overline{\nabla} \widetilde{F}(\mathbf{z}_k) \right\|_2 = 0. \tag{43}$$

Due to (42) and (43), (36) follows from Lemmas 4.3 and 3.2. □

5. Numerical Experiment

To provide numerical evidence supporting the convergence theory of the IGM, the following problem was considered:

$$\min_{\mathbf{z}\,\in\,\mathbb{R}^3} \quad \widetilde{F}(\mathbf{z}) = \|\mathbf{A}\mathbf{z} - \mathbf{b}\|_2^2, \tag{44}$$

where $\widetilde{F} : \mathbb{R}^3 \to \mathbb{R}$, $\mathbf{A} = \begin{bmatrix} 2 & 3 & 1 \\ 1 & 4 & 2 \\ 3 & 2 & 5 \end{bmatrix}$, and $\mathbf{b} = \begin{bmatrix} 5 & 6 & 7 \end{bmatrix}^T$. It is noted that $\widetilde{F}$ is lower bounded by 0 and is L-smooth, since its exact gradient $\boldsymbol{\nabla}\widetilde{F} : \mathbb{R}^3 \to \mathbb{R}^3$, given by $\boldsymbol{\nabla}\widetilde{F}(\mathbf{z}) = 2\mathbf{A}^T[\mathbf{A}\mathbf{z} - \mathbf{b}]$, satisfies

$$\left\| \boldsymbol{\nabla}\widetilde{F}(\mathbf{x}) - \boldsymbol{\nabla}\widetilde{F}(\mathbf{y}) \right\|_2 \leq \underbrace{2\left\|\mathbf{A}^T\mathbf{A}\right\|_2}_{L} \|\mathbf{x} - \mathbf{y}\|_2, \quad \forall \mathbf{x},\mathbf{y} \in \mathbb{R}^3.$$

By employing the forward finite difference formula to compute the inexact gradient $\overline{\nabla}\widetilde{F} : \mathbb{R}^3 \to \mathbb{R}^3$ at each iterate generated by (8), Berahas et al. [2022] demonstrated that

$$\left\| \overline{\nabla} \widetilde{F}(\mathbf{z}_k) - \boldsymbol{\nabla} \widetilde{F}(\mathbf{z}_k) \right\|_2 \leq \underbrace{\frac{L\sigma_k\sqrt{3}}{2}}_{\delta_k},$$

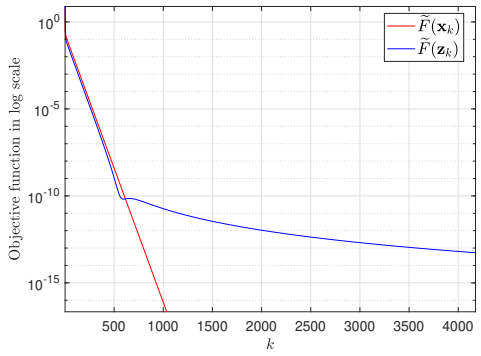
where σ_k is the finite difference perturbation parameter at the k-th iteration, and L denotes the Lipschitz constant. For the numerical experiment, $\sigma_k = 1/(k^2+1)$ and $L = 2\left\|\mathbf{A}^T\mathbf{A}\right\|_2$ were considered. The initial point was $\mathbf{z}_0 = \begin{bmatrix} 0.1 & 1 & 0.5 \end{bmatrix}^T$.

The problem (44) was addressed using both GM and IGM with either a constant step size, $t = 0.99/L$, or a diminishing step size, $t_k = 5/[L(k^{1/1.95}+1)]$, which fulfill (23a) and (23b). The stopping criterion for both methods was established based on the norm of the respective gradients (exact or inexact) falling below the tolerance $\epsilon = 10^{-8}$. It is important to note that this numerical

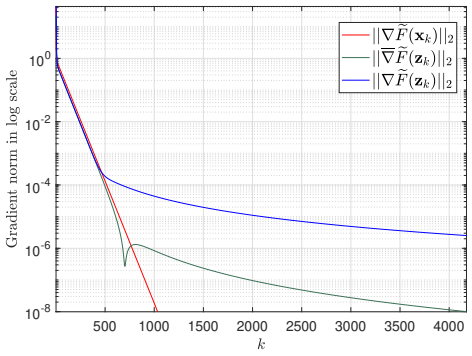
DOI: 10.59254/sbpo-2025-212676
https://proceedings.science/p/212676?lang=pt-br

experiment does not aim to demonstrate efficiency; rather, it seeks to illustrate that both exact and inexact gradients converge as prescribed by the theoretical analysis.

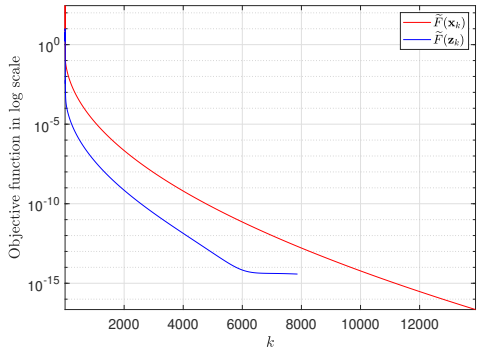
Figure 1 shows the objective function and gradient values for both methods with constant and diminishing step sizes. In the IGM, both exact and inexact gradient values were computed, as shown by the blue and green curves in Figures 1b and 1d. For the GM, only exact gradients were calculated, as shown by the red curves in the same figures. As can be seen, both gradients in each case reached the desired tolerance. The constant step size requires fewer iterations than the diminishing step size, as the latter's progressively smaller steps slow down convergence. Furthermore, notable differences in the objective function values deserve explanation. For the constant step size, the IGM gradients closely align with those from GM during the initial iterations, which accounts for the similar objective function values in Figure 1a. However, as the iterations progress, the IGM exact gradient decreases at a slower rate than the inexact gradient. Consequently, this results in inexact iterates that are far away from the solution, despite presenting a reduced inexact gradient norm. This fact explains the gradual decrease of the IGM objective function. In the case of the diminishing step size, the IGM gradients are consistently smaller than those of the GM, which elucidates why the IGM objective function values are lower than those of the GM (see Figure 1c).



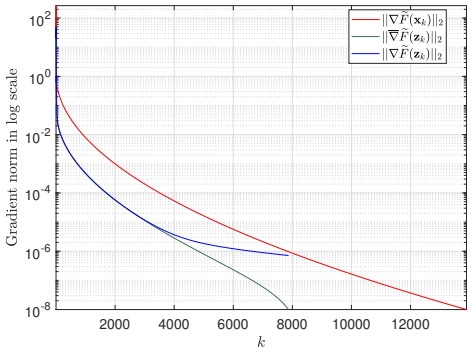
(a) Constant step size (objective function).



(b) Constant step size (gradient).



(c) Diminishing step size (objective function).



(d) Diminishing step size (gradient).

Figure 1: Objective function and gradient values produced by the GM and IGM using constant and diminishing step sizes. The sequence $(\mathbf{x}_k)_{k\in\mathbb{N}_0}$ in the legends is generated by the GM.

6. Conclusions and Future Research

This study establishes the convergence theory for an IGM applied to smooth functions, considering three step size strategies: constant, variable, and diminishing. Under certain reasonable

DOI: 10.59254/sbpo-2025-212676
https://proceedings.science/p/212676?lang=pt-br

assumptions, it was demonstrated that every limit point (if any) of the sequence generated by the method is a stationary point when utilizing any of these strategies, even within an inexact framework. A numerical experiment was conducted, and the results align with the proposed convergence theory. For future investigations, there are plans to extend the theoretical analysis to methods for constrained optimization and to apply more efficient sensitivity analysis techniques.

7. Acknowledgments

The authors acknowledge the financial support received from the Brazilian Federal Agency for Support and Evaluation of Graduate Education (CAPES) through grant number 88887.669708/2022-00, as well as from the São Paulo Research Foundation (FAPESP) through grant numbers 2013/07375-0, 2022/05803-3, and 2023/08706-1.

References

Beck, A. (2017). *First-Order Methods in Optimization*. Society for Industrial and Applied Mathematics, Philadelphia, PA. ISBN 978-1-61197-498-0.

Berahas, A. S., Cao, L., Choromanski, K., and Scheinberg, K. (2022). A Theoretical and Empirical Comparison of Gradient Approximations in Derivative-Free Optimization. *Foundations of Computational Mathematics*, 22(2):507–560. ISSN 1615-3383.

Bertsekas, D. P. and Tsitsiklis, J. N. (2000). Gradient Convergence in Gradient methods with Errors. *SIAM Journal on Optimization*, 10(3):627–642. ISSN 1052-6234. Publisher: Society for Industrial and Applied Mathematics.

d'Aspremont, A. (2008). Smooth Optimization with Approximate Gradient. *SIAM Journal on Optimization*, 19(3):1171–1183. ISSN 1052-6234. Publisher: Society for Industrial and Applied Mathematics.

Devolder, O., Glineur, F., and Nesterov, Y. (2013). Intermediate gradient methods for smooth convex problems with inexact oracle. Technical report, Technical report, CORE-2013017.

Devolder, O., Glineur, F., and Nesterov, Y. (2014). First-order methods of smooth convex optimization with inexact oracle. *Mathematical Programming*, 146(1):37–75. ISSN 1436-4646.

Dvurechensky, P. and Gasnikov, A. (2016). Stochastic Intermediate Gradient Method for Convex Problems with Stochastic Inexact Oracle. *Journal of Optimization Theory and Applications*, 171 (1):121–145. ISSN 1573-2878.

Hinze, M., Pinnau, R., Ulbrich, M., and Ulbrich, S. (2009). *Optimization with PDE Constraints*, volume 23 of *Mathematical Modelling: Theory and Applications*. Springer Netherlands, Dordrecht. ISBN 978-1-4020-8838-4 978-1-4020-8839-1. ISSN: 1386-2960.

Khanh, P. D., Mordukhovich, B. S., and Tran, D. B. (2023). Inexact Reduced Gradient Methods in Nonconvex Optimization. *Journal of Optimization Theory and Applications*. ISSN 1573-2878.

Nesterov, Y. (1983). A method for solving the convex programming problem with convergence rate O (1/k2). In *Dokl akad nauk Sssr*, volume 269, p. 543.

Polyak, B. T. (1987). *Introduction to Optimization*. Translations series in mathematics and engineering. Optimization Software, Publications Division, New York. ISBN 978-0-911575-14-9.

Vasin, A., Gasnikov, A., Dvurechensky, P., and Spokoiny, V. (2023). Accelerated gradient methods with absolute and relative noise in the gradient. *Optimization Methods and Software*, 38(6):1180–1229. ISSN 1055-6788. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/10556788.2023.2212503.

DOI: 10.59254/sbpo-2025-212676
https://proceedings.science/p/212676?lang=pt-br