EFFICIENT LEVEL SET METHODS FOR SPARSE OPTIMIZATION PROBLEMS WITH LEAST SQUARES CONSTRAINTS

LI QIAN

PhD

The Hong Kong Polytechnic University

2025

The Hong Kong Polytechnic University

Department of Applied Mathematics

Efficient level set methods for sparse optimization problems with least squares constraints

LI Qian

A thesis submitted in partial fulfilment of the requirements

for the degree of Doctor of Philosophy

December 2024

Certificate of Originality

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma, except where due acknowledgement has been made in the text.

________________________(Signed)

LI Qian (Name of student)

Dedicate to my parents.

Abstract

This thesis is concerned with an important class of sparse optimization problems with least squares constraints. The motivation for this work arises from the well established model known as Basis Pursuit Denoising (BPDN), which aims to find a sparse representation (or approximation) of a solution to an underdetermined least squares problem. This sparse representation problem has a wide range of applications in various fields such as signal processing and statistics. In this thesis, inspired by the recently developed level set approaches, we will propose an efficient sieving based secant method to address the challenges posed by the targeted problems.

Firstly, an efficient dimension reduction technique, called adaptive sieving, will be introduced for solving unconstrained sparse optimization problems. This technique addresses the original problem by solving a series of reduced problems that have substantially lower dimensionality. Extensive numerical experiments demonstrate the high efficiency and promising performance of this technique in solving sparse optimization problems, especially in high dimensional settings. The significance of this technique is its integration into the secant method discussed later, which will efficiently reduce the dimension of the level set subproblems for computing the value function.

Secondly, we will develop the sieving based secant method to solve the sparse optimization problems with least squares constraints. In the literature, people use the bisection method to find a root of a univariate nonsmooth equation $\varphi(\lambda) = \varrho$

for some $\varrho > 0$, where $\varphi(\cdot)$ is the value function computed by a solution of the corresponding regularized least squares optimization problem. When the objective function in the constrained problem is a polyhedral gauge function, we prove that (a) for any positive integer k, $\varphi(\cdot)$ is piecewise C^k in an open interval containing the solution λ^* to the equation $\varphi(\lambda) = \varrho$; (b) the Clarke Jacobian of $\varphi(\cdot)$ is always positive. These results allow us to establish the essential ingredients of the fast convergence rates of the secant method.

Finally, with all the preparations completed, we will then develop our package, called SMOP, as it is a root finding based secant method for solving the sparse optimization problems. The high efficiency of SMOP is demonstrated by extensive numerical results on high dimensional real applications. We point out that, in the special case where the objective function is the ℓ_1 norm, our numerical results show that the secant method and the semismooth Newton method are comparable in terms of the number of iterations, which also demonstrates the high efficiency of the secant method. Note that, different from the semismooth Newton method, the secant method does not need to compute the generalized Jacobian of the value function. This paves the way for using the introduced secant method to solve more complicated sparse optimization problems with least-squares constraints, where the computation of the generalized Jacobian of the value function is impractical.

Acknowledgements

I sincerely thank my Ph.D. chief supervisor, Prof. Sun Defeng, for his patient guidance and profound wisdom, which have had a lasting and positive impact on my academic and intellectual growth. His wealth of academic experience and exceptional perspectives on different issues have significantly expanded my academic horizons and deeply impacted my capacity to engage in critical thinking and independently address problems. It would be an understatement to say that this work would not have been possible without his knowledge and guidance. Moreover, his valuable support and encouragement helped me a lot during my doctoral studies, making me more confident in completing my studies.

I am deeply grateful to my Ph.D. co-supervisor, Dr. Jiang Binyan, for providing me with statistical knowledge and insights. This has greatly enhanced my understanding of the problem from different viewpoints. Additionally, I would like to especially acknowledge Dr. Yuan Yancheng for our joint work on sparse optimization problems, as well as for his friendship and support along the way. I would also like to offer special thanks to Ma Jiaming for his valuable suggestions and to Ma Chiyu for the discussion on the secant method in my thesis.

I would also like to express my sincere appreciation for my friends and the members of our optimization group: Wang Yue, Wu Can, Yan Hui, Zhang Guojun, Chen Kaihuang, Zheng Yangzi, Qiu Zicheng and Fan Xinyu. Their friendship has brought genuine joy and happiness to my studies, and together we have created many wonderful, unforgettable moments during the past years.

Last but not least, I would like to thank my parents for all the care and support they provided during my studies. It was their strong support that helped me through all the challenging times.

LI Qian

Contents

Chapter 1

Introduction

In this thesis, we focus on designing efficient algorithms for solving sparse optimization problems with least squares constraints. In particular, we are interested in the following problem

$$\min_{x \in \mathbb{R}^n} \; \{p(x) \mid \|Ax - b\| \leq \varrho\}\,, \qquad \text{(CP}(\varrho))$$

where $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$ are given data, ϱ is a given parameter satisfying $0 < \varrho < \|b\|$, and $p : \mathbb{R}^n \to (-\infty, +\infty]$ is a proper closed convex function with $p(0) = 0$ that possesses the property of promoting sparsity. Without loss of generality, we assume that (CP(ϱ)) admits active solutions here. Owing to the rapid development of data engineering and technology, modern datasets often have much higher dimensions than before, making it increasingly challenging to solve (CP(ϱ)). In this thesis, we aim to design highly efficient algorithms to address (CP(ϱ)) with a given function $p(\cdot)$, such as the ℓ_1 norm function.

1.1 Motivations and related methods

The motivation for solving the general sparse optimization problems with least squares constraints (CP(ϱ)) arises from the following problem:

$$\min_{x \in \mathbb{R}^n} \; \{\|x\|_1 \mid \|Ax - b\| \leq \varrho\}\,, \qquad (1.1)$$

where $\|\cdot\|_1$ denotes the ℓ_1 norm function. This problem was introduced in (Donoho et al., 2005; Donoho and Elad, 2006; Candés et al., 2006; Tropp, 2006) and is commonly known as the basis pursuit denoising problem (BPDN), which is a relaxation of the basis pursuit problem (BP) proposed in (Chen and Donoho, 1994; Chen et al., 2001; Donoho and Elad, 2003). Originally, BP was designed to find a representation of the signal that minimizes the ℓ_1 norm of its coefficients with a given dictionary A. Subsequently, BPDN was proposed to accommodate the presence of noise in the data, where the only difference between BPDN and BP is that BPDN relaxes the constraint in BP from $Ax = b$ to $\|Ax - b\| \leq \varrho$. Note that when the noise exists, BPDN is not the only form for combining the ℓ_1 norm and least squares functions for sparse approximation. For a positive number λ, Chen et al. (2001) suggested the regularized least squares problem

$$\min_{x\in\mathbb{R}^n} \left\{ \frac{1}{2}\|Ax - b\|^2 + \lambda\|x\|_1 \right\}, \tag{1.2}$$

which is the so called Lagrangian form of the following problem proposed by Tibshirani (1996):

$$\min_{x\in\mathbb{R}^n} \left\{ \|Ax - b\| \mid \|x\|_1 \leq \vartheta \right\}, \tag{1.3}$$

where $\vartheta > 0$.

The relationship between (1.1), (1.2) and (1.3) is that, with appropriate choices of the parameters ϱ, λ and ϑ, the solutions to these problems coincide, indicating that the three problems are equivalent in some sense (Van den Berg and Friedlander, 2008). However, this equivalent relationship is impractical, as the parameters that establish the equivalence are often unknown in most cases, except for special situations, such as when A is orthogonal. Although (1.3) is originally named Lasso, the equivalence between (1.2) and (1.3), along with the widespread use of (1.2), makes that some also refer to problem (1.2) as Lasso. In this thesis, we will refer to the unconstrained version as Lasso. We are particularly interested in the scenario where the noise level ϱ can be estimated approximately. In this case, the model (1.1) is preferred. Although Chen et al. (2001) suggested that when the noise level is known, one can set $\lambda = \varrho\sqrt{2\log n}$ to achieve certain optimal properties related to the mean squared error, this result normally holds under the condition that A is orthogonal.

We then generalize the penalty function in (1.1) from ℓ_1 norm function to a proper closed convex function $p(\cdot)$ that satisfies $p(0) = 0$ to encourage sparsity. This forms the main problem (CP(ϱ)) that we address in this thesis. Similar to what we discussed in the case where $p(\cdot) = \|\cdot\|_1$, the constrained optimization problem (CP(ϱ)) is usually preferred in practical modeling since we can regard ϱ as the noise level, which can be estimated in many applications. However, the optimization problem (CP(ϱ)) is perceived to be more challenging to solve in general due to the complicated geometry of the feasible set (Aravkin et al., 2019).

The sparse optimization problem (CP(ϱ)) has a wide range of applications in various fields such as signal processing and statistics. It plays a crucial role in significant areas such as the signal/image reconstruction: the sparse signal recovery (Candés and Romberg, 2006; Candés and Wakin, 2008; Cai et al., 2009, 2016) and the MRI reconstruction (Lustig et al., 2007, 2008), the image denoising (Papageorgiou et al., 2017; Baraldi et al., 2019) and the robust linear regression (Jin and Rao, 2010). In practical applications, datasets often have high dimensional characteristics, which drives us to develop efficient and robust algorithms to handle them and meet contemporary demands.

1.1.1 The dimension reduction techniques

Before discussing the algorithms for addressing (CP(ϱ)), we will first introduce dimension reduction techniques for the unconstrained sparse optimization problem. In this thesis, our algorithm solves (CP(ϱ)) by finding the root of a univariate nonlinear equation, where each iteration requires computing a value function by solving an unconstrained sparse optimization problem. Consequently, it is crucial to solve the unconstrained sparse optimization problem efficiently.

Let $\lambda > 0$ be a given positive parameter. We can also generalize (1.2) to the following regularized problem of the form:

$$\min_{x \in \mathbb{R}^n} \left\{ \frac{1}{2}\|Ax - b\|^2 + \lambda p(x) \right\}. \tag{$\mathrm{P_{LS}}(\lambda)$}$$

A dimension reduction technique is crucial for reducing the computational complexity of solving ($\mathrm{P_{LS}}(\lambda)$). Tibshirani et al. (2012) proposed strong screening rules (SSR) conditioned on a unit slope bound for regression problems with the ℓ_1 norm penalty function. However, this screening rule is not considered safe (the unit slope bound assumption may be violated). Therefore, the authors suggested using the Karush–Kuhn–Tucker (KKT) conditions for verification. In light of this limitation, several safe screening rules have been proposed. The first safe screening rule, proposed by El Ghaoui (2012), involves adding constraints to the dual problem and then checking the KKT conditions to eliminate variables when $p(\cdot) = \|\cdot\|_1$. Later, Wang et al. (2013) introduced a new safe screening rule based on the dual polytope projections (DPP) and an enhanced version of DPP, for the case that $p(\cdot)$ is the ℓ_1 norm and the grouped ℓ_1 norm, by carefully studying the geometry of the dual problem. Then, Liu et al. (2014) demonstrated that the DPP can be viewed as a relaxed version of their proposed safe screening rule using variational inequalities (Sasvi). This approach uses the variational inequality that offers the necessary and sufficient optimality conditions for the dual problem. For other safe screening rules, one may refer to the survey by Xiang et al. (2016) and the references therein. The safe screening rules can eliminate variables that are guaranteed to be zero. While this is advantageous because it does not discard nonzero elements, it also presents a limitation in practice, as it typically identifies only a small subset of these zero elements. Then, Zeng et al. (2021) proposed hybrid safe-strong rules (HSSR), which combines the SSR and the safe screen rules. The motivation for proposing HSSR is primarily to replace the unnecessary KKT checks required by SSR with safe screening rules. However, these screening rules are often specific to particular problems and can be challenging to apply to models with general regularizers. Moreover, it is necessary for the reduced problems to be solved exactly.

Recently, the adaptive sieving (Yuan et al., 2023, 2022) has been proposed for solving sparse optimization problems in the form of

$$\min_{x\in\mathbb{R}^n} \left\{\Phi(x) + P(x)\right\},$$

where $\Phi : \mathbb{R}^n \to \mathbb{R}$ is a continuously differentiable convex function, and $P : \mathbb{R}^n \to (-\infty, +\infty]$ is a closed proper convex function. This strategy sequentially identifies the index set of zeros using a proximal residual function defined by the KKT conditions, and the authors showed that this strategy stops in a finite number of iterations. Specifically, the adaptive sieving makes solving the original problem by addressing a sequence of reduced problems with dimensions much smaller than that of the original problem. The high efficiency of the adaptive sieving technique for addressing a broad range of sparse optimization problems has been demonstrated in (Yuan et al., 2023, 2022; Li et al., 2023; Wu et al., 2023). Note that, in numerical experiments, we observed that the adaptive sieving strategy can often reach a satisfactory solution in just a few iterations.

Apparently, ($\mathrm{P}_{\mathrm{LS}}(\lambda)$) fits within the general framework of the problem addressed by the adaptive sieving strategy. Now, let us consider how to solve the subproblems within it for ($\mathrm{P}_{\mathrm{LS}}(\lambda)$). The dual of ($\mathrm{P}_{\mathrm{LS}}(\lambda)$) can be written as

$$\max_{y\in\mathbb{R}^m, u\in\mathbb{R}^n} \left\{-\frac{1}{2}\|y\|^2 + \langle b, y\rangle - \lambda p^*(u) \mid A^T y - \lambda u = 0\right\}, \tag{$\mathrm{D}_{\mathrm{LS}}(\lambda)$}$$

where $p^*(\cdot)$ is the Fenchel conjugate function of $p(\cdot)$, i.e., $p^*(z) = \sup_{x\in\mathbb{R}^n} \{\langle z, x\rangle -$

$p(x)\}$, $z \in \mathbb{R}^n$. There are several dual based approaches that have demonstrated superior performance. The semismooth Newton augmented Lagrangian (SSNAL, Li et al. (2018b)) has demonstrated exceptional performance in solving the ℓ_1 penalized least squares problem (($\mathrm{P_{LS}}(\lambda)$) with $p(\cdot) = \|\cdot\|_1$). It addresses the dual problem ($\mathrm{D_{LS}}(\lambda)$) using the augmented Lagrangian method, with each subproblem being solved by a semismooth Newton method. Additionally, it exhibits an asymptotic superlinear convergence rate. Similarly, when $p(x) = \sum_{i=1}^{n} \gamma_i |x|_{(i)}$, $x \in \mathbb{R}^n$ with given parameters $\gamma_1 \geq \gamma_2 \geq \cdots \geq \gamma_n \geq 0$ and $\gamma_1 > 0$, where $|x|_{(1)} \geq |x|_{(2)} \geq \cdots \geq |x|_{(n)}$, Luo et al. (2019) introduced a semismooth Newton based augmented Lagrangian method (Newt-ALM) to solve ($\mathrm{D_{LS}}(\lambda)$). When the penalty function is the non-overlapping group Lasso regularization, i.e, $p(x) = \sum_{l=1}^{g} w_l \|x_{G_l}\|$, where for any $l = 1, 2, \cdots, g$, $w_l > 0$ and $G_l \subseteq \{1, 2, \cdots, n\}$ is the index set that includes all the features in the l-th group, Zhang et al. (2020) proposed a Hessian based algorithm that implements a superlinearly convergent inexact semismooth Newton method. Note that dual based methods are preferred due to their computational advantages, mainly because the dimension n of the problems being addressed is usually much larger than m (the number of rows of A). However, the situation we face may change now, as the adaptive sieving strategy may reduce the dimension of the subproblem to less than m. In this situation, a primal based algorithm may offer some computational advantages.

For addressing the primal problem ($\mathrm{P_{LS}}(\lambda)$), Khanh et al. (2023) proposed the generalized damped Newton algorithms, assuming that $A^T A$ is positive definite. However, this assumption is quite stringent. We may then employ a smoothing method to address the primal problem directly. Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be a locally Lipschitz continuous function. Consider the equation $F(x) = 0$, $x \in \mathbb{R}^n$. Generally, a smoothing method begins by constructing a smoothing approximation function $G : \mathbb{R}_{++} \times \mathbb{R}^n \to \mathbb{R}^n$ for $F(\cdot)$. This function is designed such that for any $\epsilon > 0$,

$G'(\epsilon, \cdot)$ is continuously differentiable on $\mathbb{R}^n$. Furthermore, for any $x \in \mathbb{R}^n$, it holds that $\|F(x) - G(\epsilon, x)\| \to 0$ as $\epsilon \downarrow 0$. Then, a smoothing method addresses the equation $F(x) = 0$ by (inexactly) solving the following problems for a given positive sequence $\{\epsilon^k\}$, $k = 0, 1, 2, \cdots$: $G(\epsilon^k, x) = 0$. In (Chen et al., 1998), the authors introduced a smoothing Newton method that incorporates the derivative of $G(\cdot, \cdot)$ with respect to the variable x in the Newton iteration. The update rule is given by:

$$x^{k+1} = x^k - t_k G'_x(\epsilon^k, x^k)^{-1} F(x^k), \tag{1.4}$$

where the parameter $\epsilon^k > 0$, the stepsize $t_k > 0$, and $G'_x(\epsilon^k, x^k)$ denotes the derivative of $G(\cdot, \cdot)$ with respect to x at (ϵ^k, x^k). The analysis of the convergence of this smoothing method relies on two key assumptions: (a) the Jacobian consistency property, as defined in (Chen et al., 1998, Definition 2.1), is satisfied, and (b) there exists a constant $\mu > 0$ such that for any $\epsilon \in \mathbb{R}_{++}$ and $x \in \mathbb{R}^n$,

$$\|G(\epsilon, x) - F(x)\| \leq \mu\epsilon. \tag{1.5}$$

It has been verified in (Chen et al., 1998) that many smoothing functions satisfy the Jacobian consistency property, however, some smoothing functions do not meet the condition in (1.5). To circumvent such challenges, a class of squared smoothing Newton methods was introduced in (Qi et al., 2000; Qi and Sun, 2002). Subsequently, Gao and Sun (2009) presented an inexact smoothing Newton method for reducing the computational cost.

Then, drawing inspiration from the work of (Gao and Sun, 2009), we will develop a smoothing Newton method to directly solve ($\mathrm{P}_{\mathrm{LS}}(\lambda)$) by finding a solution of

$$F(x) = x - \mathrm{Prox}_{\lambda p}(x - A^T Ax + A^T b) = 0, \quad x \in \mathbb{R}^n,$$

where $\mathrm{Prox}_{\lambda p}(\cdot)$ is the proximal mapping associated with $\lambda p(\cdot)$. Let $G : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ be a locally Lipschitz continuous function such that

$$G(\epsilon, \tilde{x}) \to F(x) \quad \text{as} \quad (\epsilon, \tilde{x}) \to (0, x), \tag{1.6}$$

and the function $G(\cdot,\cdot)$ is continuously differentiable around $(\epsilon, x) \in \mathbb{R} \times \mathbb{R}^n$ except when $\epsilon = 0$. Denote $E : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R} \times \mathbb{R}^n$ as

$$E(\epsilon, x) := \begin{pmatrix} \epsilon \\ G(\epsilon, x) \end{pmatrix}, \quad \forall\, (\epsilon, x) \in \mathbb{R} \times \mathbb{R}^n. \tag{1.7}$$

Then, we will solve $F(x) = 0$, $x \in \mathbb{R}^n$ by solving

$$E(\epsilon, x) = 0. \tag{1.8}$$

So far, we have only provided details of the smoothing Newton method for the case that $p(\cdot) = \|\cdot\|_1$, while other cases will be addressed in the future work. Specifically, we use the Huber function to approximate

$$\mathrm{Prox}_{\lambda\|\cdot\|_1}(x - A^T Ax + A^T b), \quad x \in \mathbb{R}^n,$$

and then use the inexact Newton method to solve the corresponding smooth equation (1.8). Furthermore, we will show that the smoothing Newton method exhibits quadratic convergence under the Linear independence constraint qualification (LICQ) condition to the dual problem.

1.1.2 Related algorithms

Some algorithms such as the alternating direction method of multipliers (ADMM) (Glowinski and Marroco, 1975; Gabay and Mercier, 1976) are applicable to solve (CP(ϱ)). Nevertheless, to obtain an acceptable solution remains challenging for these algorithms. In particular, when applying the ADMM for solving (CP(ϱ)), it is computationally expensive to form the matrix AA^T or to solve the linear systems involved in the subproblems. Even though dimension reduction techniques exist, it is still unclear how to apply dimension reduction techniques to (CP(ϱ)) due to the potential infeasibility issue for reduced problems.

A popular approach for solving (CP(ϱ)) and the more general convex constrained optimization problems is the level set method (Van den Berg and Friedlander, 2008, 2011; Aravkin et al., 2019), which has been widely used in many interesting applications (Van den Berg and Friedlander, 2008, 2011; Aravkin et al., 2019; Li et al., 2018b). The idea of exchanging the role of the objective function and the constraints, which is the key for the level set method, has a long history and can date back to Queen Dido's problem (see (Richard and George, 2001, Page 548)). Readers can refer to (Aravkin et al., 2019, Section 1.3) and the references therein for a discussion of the history of the level set method. In particular, the level set method developed in (Van den Berg and Friedlander, 2008, 2011) solves the optimization problem (CP(ϱ)) by finding a root of the following univariate nonlinear equation

$$\phi(\tau) = \varrho, \qquad (E_\phi)$$

where $\phi(\cdot)$ is the value function of the following level set problem

$$\phi(\tau) := \min_{x \in \mathbb{R}^n} \{\|Ax - b\| \,|\, p(x) \leq \tau\}, \quad \tau \geq 0. \qquad (1.9)$$

Therefore, by executing a root finding procedure for (E_ϕ) (e.g., the bisection method), one can obtain a solution to (CP(ϱ)) by solving a sequence of problems in the form of (1.9) parameterized by τ. In implementations, one needs an efficient procedure to compute the metric projection of given vectors onto the constraint set $\mathcal{F}_p(\tau) := \{x \in \mathbb{R}^n \,|\, p(x) \leq \tau, \tau > 0\}$. However, such an efficient computation procedure may not be available. One example can be found in (Li et al., 2018b), where $p(\cdot)$ is the fused Lasso regularizer (Tibshirani et al., 2005). Moreover, it is still not clear to us how to deal with the infeasibility issue when a dimension reduction technique is applied to (1.9).

Recently, Li et al. (2018b) proposed a level set method for solving (CP(ϱ)) via solving a sequence of ($\mathrm{P_{LS}}(\lambda)$). Let $\Omega(\lambda)$ be the solution set to ($\mathrm{P_{LS}}(\lambda)$). Define the gauge $\Upsilon(\cdot \mid C)$ of a nonempty convex set $C \subseteq \mathbb{R}^n$ as $\Upsilon(x \mid C) := \inf\{\nu \geq 0 \mid x \in \nu C\}$, $x \in \mathbb{R}^n$. Denote $\partial p(0)$ as the subdifferential of $p(\cdot)$ at the origin. We assume

$$\lambda_\infty := \Upsilon(A^T b \,|\, \partial p(0)) \in (0, +\infty) \tag{1.10}$$

and that for any $\lambda' > 0$, there exists $(y(\lambda'), u(\lambda'), x(\lambda')) \in \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^n$ satisfying the following KKT system

$$x \in \partial p^*(u), \quad y - b + Ax = 0, \quad A^T y - \lambda' u = 0, \tag{KKT}$$

where $\partial p^*(\cdot)$ is the subdifferential of $p^*(\cdot)$. Consequently, the solution set $\Omega(\lambda)$ to $(\mathrm{P_{LS}}(\lambda))$ is nonempty, and $b - Ax(\lambda)$ is invariant for all $x(\lambda) \in \Omega(\lambda)$ since the solution $(y(\lambda), u(\lambda))$ to $(\mathrm{D_{LS}}(\lambda))$ is unique. Based on this fact, Li et al. (2018b) proposed to solve $(\mathrm{CP}(\varrho))$ by finding the root of the following equation with the bisection method:

$$\varphi(\lambda) := \|Ax(\lambda) - b\| = \varrho, \tag{E_φ}$$

where $x(\lambda) \in \Omega(\lambda)$ is any solution to $(\mathrm{P_{LS}}(\lambda))$. We assume that (E_φ) has at least one solution $\lambda^* > 0$. We then know that any $x(\lambda^*) \in \Omega(\lambda^*)$ is a solution to $(\mathrm{CP}(\varrho))$ (Li et al., 2018b; Friedlander and Tseng, 2007). There are several advantages to this approach. Firstly, it requires computing the proximal mapping of $p(\cdot)$, which is normally easier than computing the projection over the constraint set of (1.9). Secondly, efficient algorithms are available to solve the regularized least squares problem $(\mathrm{P_{LS}}(\lambda))$ for a wide class of functions $p(\cdot)$ (Li et al., 2018a,b; Luo et al., 2019; Zhang et al., 2020; Beck and Teboulle, 2009; Glowinski and Marroco, 1975). More importantly, this approach is well suited for applying dimension reduction techniques to solve $(\mathrm{P_{LS}}(\lambda))$ as can be seen in subsequent sections.

In this thesis, we propose an efficient sieving based secant method for solving $(\mathrm{CP}(\varrho))$ by finding the root of (E_φ). We call our algorithm **SMOP** as it is a root finding based **S**ecant **M**ethod for solving the **O**ptimization **P**roblem $(\mathrm{CP}(\varrho))$. We focus on the case where $p(\cdot)$ is a gauge function (see (Rockafellar, 1970, Section 15)), i.e., $p(\cdot)$ is a nonnegative positively homogeneous convex function with $p(0) = 0$. We start by studying the properties of the value function $\varphi(\cdot)$ and the convergence rates of the secant method for solving (E_φ). To address the computational challenges for solving ($\mathrm{P}_{\mathrm{LS}}(\lambda)$) and computing the function value of $\varphi(\cdot)$, we incorporate the adaptive sieving technique into the secant method to effectively reduce the dimension of ($\mathrm{P}_{\mathrm{LS}}(\lambda)$). We point out that the adaptive sieving technique is naturally implemented in the level set method. This is because, in the level set method, we need to solve problems with a sequence of penalty parameters. As a result, each adaptive sieving iteration (except the first) is warm started. Extensive numerical results will be presented in this thesis to demonstrate the superior performance of the proposed algorithm in solving ($\mathrm{CP}(\varrho)$). Moreover, in the special case that $p(\cdot)$ is the ℓ_1 norm function, we also present numerical results of the semismooth Newton method for (E_φ). The results indicate that the secant method and the semismooth Newton method are comparable in terms of the number of iterations. However, unlike the semismooth Newton method, the secant method does not require the computation of the generalized Jacobian of the value function. This makes it convenient to use the secant method for more complex sparse optimization problems with least squares constraints, where calculating the generalized Jacobian is impractical.

1.2 Contributions

To efficiently solve the sparse optimization problem ($\mathrm{CP}(\varrho)$), we introduce a sieving based secant method for finding the root of (E_φ). Previously, the level set method introduced by Li et al. (2018b) solves ($\mathrm{CP}(\varrho)$) by finding the root of (E_φ) using the bisection method, with each subproblem addressed by SSNAL (Li et al., 2018a). We then improve the efficiency of this level set method in two key ways. Firstly, we improve the efficiency of solving subproblems in the level set method by implementing the adaptive sieving strategy (Yuan et al., 2023, 2022). Additionally, we introduce a primal based algorithm that works with SSNAL to address the reduced problems generated by the adaptive sieving strategy. Secondly, we propose a secant method for finding the root of (E_φ), which significantly reduces the number of iterations needed compared to the bisection method.

In each iteration of the level set method, we need to compute the value function, which is derived from solving an unconstrained sparse optimization problem. There are several dual based algorithms that can be applied to solve unconstrained sparse optimization problems with a specified $p(\cdot)$. For example, SSNAL (Li et al., 2018a) is suited for $p(\cdot) = \|\cdot\|_1$ and Newt-ALM (Luo et al., 2019) is applicable in cases where $p(\cdot)$ is the sorted ℓ_1 norm function. However, directly using such algorithms is not sufficiently efficient. Our first improvement involves implementing the adaptive sieving technique, which allows us to solve the original problem by addressing several reduced problems with significantly smaller dimensions. The second improvement we made is the introduction of the smoothing Newton method for directly solving the primal problem. When combined with the dual based algorithms, this approach allows for higher efficiency in solving the reduced problems in the adaptive sieving strategy. This is because that we can choose to use either the primal or dual based algorithm depending on the relationship between the dimensionality of the reduced problem and the number of rows in A. So far, we have only provided details of the smoothing Newton method for the case where $p(\cdot) = \|\cdot\|_1$, while other cases will be addressed in the future work. In this case, we show that the algorithm converges quadratically to a solution under the assumption that the LICQ condition to the dual problem holds, which is less stringent than the condition presented in (Khanh et al., 2023) for the generalized damped Newton algorithm. At last, we point out that the adaptive sieving technique is naturally applicable in the level set method, as it involves solving subproblems with a sequence of penalty parameters. This allows each problem solved in the adaptive sieving strategy to be warm started, except for the first one.

To find the root of (E_φ), we develop the secant method by carefully analyzing the properties of the value function. Specifically, when $p(\cdot)$ is a gauge function, we prove that $\varphi(\cdot)$ is (strongly) semismooth for a wide class of instances of $p(\cdot)$ via connecting $(\mathrm{D}_{\mathrm{LS}}(\lambda))$ to a metric projection problem. More importantly, when $p(\cdot)$ is a polyhedral gauge function, we show that $\varphi(\cdot)$ is locally piecewise C^k on $(0, \lambda_\infty)$ for any integer $k \geq 1$; and for any $\bar{\lambda} \in (0, \lambda_\infty)$, $v > 0$ for any $v \in \partial\varphi(\bar{\lambda})$. Then, under the assumption that $p(\cdot)$ is a polyhedral gauge function, we show that the secant method converges at least 3-step Q-quadratically for solving (E_φ), and if $\partial_{\mathrm{B}}\varphi(\lambda^*)$ is a singleton, the secant method converges superlinearly with Q-order at least $(1+\sqrt{5})/2$. Furthermore, for a general strongly semismooth function $\varphi(\cdot)$, if $\partial\varphi(\lambda^*)$ is a singleton and nondegenerate, the secant method converges superlinearly with R-order of at least $(1+\sqrt{5})/2$.

With all the preparations completed, we develop a package, named SMOP, for addressing $(\mathrm{CP}(\varrho))$. This incorporates a fast convergent secant method for root finding of (E_φ), along with an adaptive sieving technique for efficiently reducing the dimensionality of subproblems in the form of $(\mathrm{P}_{\mathrm{LS}}(\lambda))$, where each subproblem is solved by a fast primal or dual based algorithm. The efficiency of the proposed algorithm for solving $(\mathrm{CP}(\varrho))$ will be demonstrated by extensive numerical experiments.

1.3 Thesis organization

The remainder of the thesis is organized as follows. In Chapter 2, we will present some preliminaries relevant to the following chapters. This includes useful properties associated with the Moreau-Yosida regularization and semismooth functions. Then, we will review the existing results of the secant method for semismooth equations and present the details of the ADMM for (CP(ϱ)). In Chapter 3, we will introduce the adaptive sieving technique for the unconstrained sparse optimization problems, and develop a smoothing Newton method to directly address the subproblems (1.1) in the adaptive sieving strategy with quadratic convergence under the assumption that the LICQ to the dual problem holds. Additionally, we present a warm-started path-following adaptive sieving approach for extremely large unconstrained sparse optimization problems. In Chapter 4, we will develop the secant based method for solving (CP(ϱ)). We begin by analyzing the properties of the value function $\varphi(\cdot)$ and constructing the HS-Jacobian of $\varphi(\cdot)$ for some $p(\cdot)$. Next, we will demonstrate the fast convergence of the secant method for solving (E_φ) and design a global version of the algorithm. The extensive numerical experiments on real datasets are presented in Chapter 5, where $p(\cdot)$ is the ℓ_1 norm function, sorted ℓ_1 norm function, and an interesting example in which $p(\cdot)$ is a non-polyhedral function. In Chapter 6, we will present the final conclusions of this thesis and explore several potential directions for the future research.

1.4 Notation

Let $\mathcal{X}$, $\mathcal{Y}$ and $\mathcal{Z}$ be real finite dimensional Euclidean spaces, each equipped with an inner product $\langle \cdot, \cdot \rangle$ and the induced norm $\|\cdot\|$. Let $n \geq 1$ be any given integer. Denote the nonnegative orthant and the positive orthant of $\mathbb{R}^n$ as $\mathbb{R}^n_+$ and $\mathbb{R}^n_{++}$, respectively. We denote $[n] := \{1, 2, \ldots, n\}$. We denote the subvector generated by $x \in \mathbb{R}^n$ indexed by $K \subseteq [n]$ as x_K and submatrix generated by the columns (rows) of $A \in \mathbb{R}^{m \times n}$ indexed by $K \subseteq [n]$ ($K \subseteq [m]$) as $A_{:K}$ ($A_{K:}$). For any $x \in \mathbb{R}^n$ and any integer $q \geq 1$, the ℓ_q norm of x is defined as $\|x\|_q := \sqrt[q]{|x_1|^q + \cdots + |x_n|^q}$. We denote $\|\cdot\| = \|\cdot\|_2$. Let $U \subseteq \mathbb{R}^n$ be an open set. We say that a function $f : U \to \mathbb{R}$

is C^k for some integer $k \geq 1$ if $f(\cdot)$ is k-times continuously differentiable on U. Let $p : \mathbb{R}^n \to (-\infty, +\infty]$ be a proper closed convex function. The conjugate $p^*(\cdot)$ of $p(\cdot)$ is defined by

$$p^*(x^*) = \sup_{x \in \mathbb{R}^n} \{\langle x, x^* \rangle - p(x)\}, \quad x^* \in \mathbb{R}^n.$$

Let $p(\cdot)$ be a nonnegative positively homogeneous convex function such that $p(0) = 0$, i.e. a gauge function (Rockafellar, 1970, Section 15), then the polar of $p(\cdot)$ is defined by

$$p^\circ(y) := \inf\{\nu \geq 0 \,|\, \langle y, x \rangle \leq \nu p(x), \ \forall\, x \in \mathbb{R}^n\}, \quad y \in \mathbb{R}^n.$$

Let $C \subseteq \mathbb{R}^n$ be a convex set. The interior $\operatorname{int} C$ of C and the relative interior $\operatorname{ri} C$ of C are defined as

$$\operatorname{int} C = \{x \mid \exists \epsilon > 0, \, x + \epsilon \mathbb{B} \subset C\},$$

$$\operatorname{ri} C = \{x \in \operatorname{aff} C \mid \exists\, \epsilon > 0, (x + \epsilon \mathbb{B}) \cap \operatorname{aff} C \subset C\},$$

where $\operatorname{aff} C$ denotes the affine hull of C and $\mathbb{B}$ denotes the Euclidean unit ball. The effective domain $\operatorname{dom} f$ of a convex function $f : \mathbb{R}^n \to (-\infty, +\infty]$ is defined by $\operatorname{dom} f = \{x \in \mathbb{R}^n \mid f(x) < +\infty\}$.

Chapter 2

Preliminaries

In this chapter, we will begin by introducing the Moreau-Yosida regularization, a useful technique for approximating nonsmooth functions and facilitating efficient optimization methods. Then, we will introduce semismooth functions, an important subclass of Lipschitz continuous functions. The property of semismoothness plays a crucial role in analyzing the convergence rates of various algorithms, including Newton's method for nonsmooth equations. Since we will be using the secant method in our algorithm, in this chapter, we will summarize the existing results on the convergence properties of the secant method for solving semismooth equations. Besides, we will also introduce the well known alternating direction method of multipliers (ADMM) to solve equation (CP(ϱ)). Its numerical performance will be compared with that of our algorithm in Section 5.

We start this chapter with some basic definitions. Let $g : \mathcal{X} \to \mathcal{Y}$ be a mapping. Various definitions of directional differentiability are available. For example, one can find some of these definitions in (Shapiro, 1990; Bonnans and Shapiro, 2000) and the references therein. We adopted the following definition from (Bonnans and Shapiro, 2000, Definition 2.44).

Definition 2.1 (Directionally differentiability)**.** *We say that $g(\cdot)$ is directionally* *differentiable at a point $x \in \mathcal{X}$ in a direction $h \in \mathcal{X}$ if the limit*

$$g'(x;h) := \lim_{t \downarrow 0} \frac{g(x+th) - g(x)}{t} \tag{2.1}$$

exists. If $g(\cdot)$ is directionally differentiable at x in every direction $h \in \mathcal{X}$, then $g(\cdot)$ is said to be directionally differentiable at x.

A concept, Bouligand differentiability (B-differentiability), is related to the directional differentiability, and the definition is given as follows.

Definition 2.2 (B-differentiability)**.** *(Pang, 1990, Definition 1) We say that $g(\cdot)$ is B-differentiable at a point $x \in \mathcal{X}$ if there is a function $BD(x) : \mathcal{X} \to \mathcal{Y}$, which is positively homogeneous of degree 1 (i.e., $BD(x)(th) = tBD(x)(h)$, for all $t \geq 0$ and for all $h \in \mathcal{X}$) such that*

$$g(x+h) - g(x) - BD(x)(h) = o(\|h\|). \tag{2.2}$$

If $g(\cdot)$ is B-differentiable at all points in a set S, we then say $g(\cdot)$ is B-differentiable in S.

Shapiro (1990) showed that if $g(\cdot)$ is further assumed to be locally Lipschitz continuous, then $g(\cdot)$ is B-differentiable at $x \in \mathcal{X}$ if and only if $g(\cdot)$ is directionally differentiable, and $BH(x)(h) = g'(x;h)$ for all $h \in \mathcal{X}$. We say that $g(\cdot)$ is directionally differentiable at a point $x \in \mathcal{X}$ of degree $\gamma \in (0, +\infty)$ if $g(\cdot)$ is directionally differentiable at x and for all $h \to 0$,

$$g(x+h) - g(x) - g'(x;h) = O(\|h\|^{1+\gamma}). \tag{2.3}$$

If $g(\cdot)$ is directionally differentiable at x of degree γ in every direction $h \in \mathcal{X}$, then $g(\cdot)$ is said to be directionally differentiable at x of degree γ. Therefore, if $g(\cdot)$ is locally Lipschitz continuous and directionally differentiable at a point $x \in \mathcal{X}$ of degree 1, it is also calmly B-differentiable at x (Ding et al., 2014; Ye and Zhou, 2017).

Fréchet differentiability is a commonly used concept, and the definition is given below.

Definition 2.3 (Fréchet differentiability)**.** *(Shapiro, 1990) Let $\mathcal{O}$ be an open subset of $\mathcal{X}$. We say that $g(\cdot)$ is Fréchet differentiable at a point $x \in \mathcal{O}$ if there is a continuous linear operator $FD(x) : \mathcal{X} \to \mathcal{Y}$ such that*

$$\lim_{h \to 0} \frac{\|g(x+h) - g(x) - FD(x)(h)\|}{\|h\|} = 0.$$

The continuous linear operator $FD(x)$ is called the Fréchet derivative of $g(\cdot)$ at x.

The following theorem is crucial for the generalized Jacobian defined in Section 2.6 of (Clarke, 1983).

Theorem 2.1 (Rademacher's theorem)**.** *Let $\mathcal{O}$ be an open subset of $\mathcal{X}$. If $g(\cdot)$ is locally Lipschitz continuous on $\mathcal{O}$, then $g(\cdot)$ is almost everywhere (Fréchet) differentiable in $\mathcal{O}$.*

Henceforth, we make the assumption that $g(\cdot)$ is locally Lipschitzian. According to Rademacher's theorem, we know that $g(\cdot)$ is almost everywhere (Fréchet) differentiable. Let D_g be the set of points at which $g(\cdot)$ is differenable and $g'(x)$ be the Jacobian of $g(\cdot)$ at $x \in D_g$. The Bouligand subdifferential (B-subdifferential) of g at $x \in \mathcal{X}$ is defined as

$$\partial_B g(x) = \left\{ \lim_{x^k \to x} g'(x^k), \; x^k \in D_g \right\}, \tag{2.4}$$

and the Clark generalized Jacobian of $g(\cdot)$ at $x \in \mathcal{X}$ is defined as the convex hull of $\partial_B g(x)$, i.e.,

$$\partial g(x) = \operatorname{conv} \left\{ \partial_B g(x) \right\}. \tag{2.5}$$

For finitely valued convex functions, the Clarke generalized Jacobian is the same as the subdifferential in convex analysis (Clarke, 1983, Proposition 2.27). Summarizing from (Clarke, 1983, Proposition 2.6.2) and (Clarke, 1983, Proposition 2.6.5), we have the following properties about $\partial g(\cdot)$.

Proposition 2.1. *Let $g(\cdot)$ be locally Lipschitz on an open subset $\mathcal{O}$ of $\mathcal{X}$. For any $x \in \mathcal{O}$, the following properties holds:*

(a) $\partial g(x)$ is a nonempty convex compact subset of $\mathbb{R}^{|\mathcal{X}|\times|\mathcal{Y}|}$.

(b) $\partial g(\cdot)$ is closed at x; that is, if $x^k \to x$, $V^k \in \partial g(x^k)$, and $V^k \to V$, then $V \in \partial g(x)$.

(c) $\partial g(\cdot)$ is upper semicontinuous at x: for any $\epsilon > 0$, there is $\delta > 0$ such that, for all z in $x + \delta\mathbb{B}$,

$$\partial g(z) \subset \partial g(x) + \epsilon\mathbb{B}_{|\mathcal{X}|\times|\mathcal{Y}|}, \tag{2.6}$$

where $\mathbb{B}_{|\mathcal{X}|\times|\mathcal{Y}|}$ denotes the open unit ball in $\mathbb{R}^{|\mathcal{X}|\times|\mathcal{Y}|}$.

(d) For any $z \in \mathcal{O}$, one has

$$g(z) - g(x) \in conv\{\partial g(s)(z - x) \mid s \in [z, x]\}, \tag{2.7}$$

where $[z, x]$ is the line segment connecting z and x.

2.1 The Moreau-Yosida regularization

In this section, we will introduce an important tool, the Moreau-Yosida regularization, which can be seen as a possible way to smooth a nonsmooth function in some sense. Let $f(\cdot)$ be a proper closed convex function from $\mathcal{X}$ to $(-\infty, +\infty]$. Then, the Moreau envelope function $\Phi_f(\cdot)$ of $f(\cdot)$ is defined by

$$\Phi_f(x) := \min_{y\in\mathcal{X}} \left\{ f(y) + \frac{1}{2}\|y - x\|^2 \right\}, \quad \forall\, x \in \mathcal{X}. \tag{2.8}$$

It can be found in (Moreau, 1965; Yosida, 1964) that the following property holds, demonstrating that (2.8) is well defined.

Proposition 2.2. *For any $x \in \mathcal{X}$, problem* (2.8) *has a unique optimal solution.*

Definition 2.4 (Proximal mapping)**.** *The unique optimal solution of* (2.8) *is called the proximal mapping* $\mathrm{Prox}_f(\cdot)$ *associated with $f(\cdot)$, that is*

$$\mathrm{Prox}_f(x) := \arg\min_{y \in \mathcal{X}} \left\{ f(y) + \frac{1}{2}\|y - x\|^2 \right\}, \quad \forall\, x \in \mathcal{X}. \tag{2.9}$$

Note that, for any $x \in \mathcal{X}$, $\mathrm{Prox}_f(x)$ is usually referred to as the proximal point of x associated with $f(\cdot)$. The following are two well known proximal mappings associated with the ℓ_1 norm and the ℓ_2 norm. For any given $\varsigma > 0$ and $v \in \mathbb{R}^n$,

$$\begin{aligned} \mathrm{Prox}_{\varsigma\|\cdot\|_1}(v) &= \mathrm{sign}(v) \odot \max\{|v| - \varsigma e, 0\}, \\ \mathrm{Prox}_{\varsigma\|\cdot\|_2}(v) &= \begin{cases} \frac{v}{\|v\|}\max\{\|v\| - \varsigma, 0\}, & \text{if } v \neq 0, \\ 0, & \text{otherwise,} \end{cases} \end{aligned} \tag{2.10}$$

where $\odot$ denotes the pointwise multiplication.

The following proposition is important for providing useful properties of the Moreau-Yosida regularization, and the proof can be found in (Hiriart-Urruty and Lemaréchal, 1993, Theorem 4.1.4).

Proposition 2.3. *Let $f(\cdot)$ be a proper closed convex function from $\mathcal{X}$ to $(-\infty, +\infty]$, and $\mathrm{Prox}_f(\cdot)$ be the associated proximal point mapping. Then, the following properties hold.*

(a) *Both $\mathrm{Prox}_f(\cdot)$ and $(I - \mathrm{Prox}_f)(\cdot)$ are firmly non-expansive, i.e., $\forall\, x, y \in \mathcal{X}$,*

$$\|\mathrm{Prox}_f(x) - \mathrm{Prox}_f(y)\|^2 \leq \langle \mathrm{Prox}_f(x) - \mathrm{Prox}_f(y), x - y\rangle, \tag{2.11}$$

$$\|(x - \mathrm{Prox}_f(x)) - (y - \mathrm{Prox}_f(y))\|^2 \leq \langle (x - \mathrm{Prox}_f(x)) - (y - \mathrm{Prox}_f(y)), x - y\rangle. \tag{2.12}$$

(b) *The Moreau envelope* $\Phi_f(\cdot)$ *is continuously differentiable, and furthermore, it holds that*

$$\nabla \Phi_f(x) = x - \mathrm{Prox}_f(x), \quad \forall\, x \in \mathcal{X}.$$

A notably elegant and valuable property of the Moreau-Yosida regularization is the following theorem.

Theorem 2.2 (Moreau's decomposition theorem)**.** *Let* $f(\cdot)$ *be a proper closed convex function from* $\mathcal{X}$ *to* $(-\infty, +\infty]$ *and* $f^*(\cdot)$ *be its conjugate. Let* τ *be a positive scalar. Then one has the following decomposition*

$$x = \mathrm{Prox}_{\tau f}(x) + \tau \mathrm{Prox}_{\tau^{-1} f^*}(\tau^{-1} x), \quad \forall\, x \in \mathcal{X}. \tag{2.13}$$

Proof. For any $x \in \mathcal{X}$, let $z = \mathrm{Prox}_{\tau f}(x)$. From (Rockafellar, 1970, Theorem 23.5), we have $x - z \in \tau \partial f(z)$ and then $z \in \partial f^*(\tau^{-1} x - \tau^{-1} z)$. Thus $\tau^{-1} x - \tau^{-1} z = \mathrm{Prox}_{\tau^{-1} f^*}(\tau^{-1} x)$. Consequently, $x = z + \tau \mathrm{Prox}_{\tau^{-1} f^*}(\tau^{-1} x)$. This completes the proof. □

Now, we consider a special application of the Moreau-Yosida regularization. Let $C \subseteq \mathcal{X}$ be a closed convex set and define the indicator function of C as

$$\delta_C(x) = \begin{cases} 0, & \text{if } x \in C, \\ +\infty, & \text{otherwise,} \end{cases} \quad \forall\, x \in \mathcal{X}.$$

Then, for any $x \in \mathcal{X}$, the proximal point of x associated with $\delta_C(\cdot)$ simplifies to the metric projection of x onto C, due to the observation that

$$\min_{y \in \mathcal{X}} \left\{ \delta_C(y) + \frac{1}{2}\|y - x\|^2 \right\} \quad \Longleftrightarrow \quad \min_{y \in C} \frac{1}{2}\|y - x\|^2.$$

Thus, we denote

$$\Pi_C(x) := \mathrm{Prox}_{\delta_C}(x), \quad \forall\, x \in \mathcal{X}.$$

When $f(\cdot) = \delta_C(\cdot)$, we have from Proposition 2.3 (a) that $\Pi_C(\cdot)$ is globally Lipschitz continuous with modulus 1. Then, by Theorem 2.1, we know that it is differentiable almost everywhere on $\mathcal{X}$. Therefore, the B-subdifferential $\partial_B \Pi_C(\cdot)$ and the Clarke generalized Jacobian $\partial \Pi_C(\cdot)$ of $\Pi_C(\cdot)$ are well defined. Then, we introduce the following lemma (Meng et al., 2005, Proposition 1), which provides several important properties of $\partial \Pi_C(\cdot)$.

Proposition 2.4. *Let $C \subseteq \mathcal{X}$ be a closed convex set. Then, for any $x \in \mathcal{X}$ and $V \in \partial \Pi_C(\cdot)$, the following properties hold.*

(a) *V is self-adjoint, i.e., $\langle Vy, z \rangle = \langle y, Vz \rangle$ for any y, $z \in \mathcal{X}$.*

(b) *$\langle d, Vd \rangle \geq 0$ for all $d \in \mathcal{X}$.*

(c) *$\langle Vd,\, d \rangle \geq \|Vd\|^2$ for all $d \in \mathcal{X}$.*

2.2 Semismooth functions

The concept of semismoothness was introduced in (Mifflin, 1977) to define an important subclass of Lipschitz functions. Later, Qi and Sun (1993) extended the definition of semismoothness to vector valued functions in order to analyze the convergence of Newton's method for solving nondifferentiable equations.

When a vector valued function $g : \mathcal{X} \to \mathcal{Y}$ is locally Lipschitzian (but not differentiable), the concept of semismoothness, as defined in Mifflin (1977) originally, was extended by Qi and Sun (1993).

Definition 2.5 (Semismoothness)**.** *Let $g : \mathcal{X} \to \mathcal{Y}$ be a locally Lipschitz continuous function. The function $g(\cdot)$ is said to be semismooth at $x \in \mathcal{X}$ if $g(\cdot)$ is directionally differentiable at x and for any $V \in \partial g(x+h)$ and $h \to 0$,*

$$g(x+h) - g(x) - V(h) = o(\|h\|). \tag{2.14}$$

The function $g(\cdot)$ is said to be γ-order ($0 < \gamma < \infty$) semismooth at $x \in \mathcal{X}$ if $g(\cdot)$ is semismooth and directional differentiable of degree γ at x and

$$g(x+h) - g(x) - V(h) = O(\|h\|^{1+\gamma}). \tag{2.15}$$

In particular, if $g(\cdot)$ is first-order semismooth, we call it strongly semismooth.

We know from (Qi and Sun, 1993) that the above definition of semismoothness coincides with that in (Potra et al., 1998). According to (Mifflin, 1977), we have that the convex functions are semismooth. It is well established in the nonsmooth optimization that twice continuously differentiable functions are strongly semismooth (indicated by (Facchinei and Pang, 2003a, Proposition 7.4.5)). Next, we introduce an important class of functions that is (strongly) semismooth.

Definition 2.6 (Piecewise affine/linear functions)**.** *(Scholtes, 2012) A function $g(\cdot)$ is called piecewise affine if it is continuous and there is a finite set of affine functions $g_i(\cdot)$, $i = 1, \cdots, l$ such that*

$$g(x) \in \{g_1(x), \cdots, g_l(x)\}, \quad \forall x \in \mathcal{X}.$$

If all $g_i(\cdot)$, $i = 1, \cdots, l$ are linear, then $g(\cdot)$ is called piecewise linear.

A more general class of functions is the piecewise C^k function, where $k \geq 1$. It is defined as follows: a function $g(\cdot) : \mathcal{O} \subseteq \mathcal{X} \to \mathcal{Y}$ is called piecewise C^k on the open set $\mathcal{O}$ if it is continuous and, for any $\bar{x} \in \mathcal{O}$, there is a neighborhood $\mathcal{N}_{\bar{x}} \subset \mathcal{O}$ of $\bar{x}$ and a finite set of C^k functions $g_i(\cdot)$, $i = 1, \cdots, l$ such that $g(x) \in \{g_1(x), \cdots, g_l(x)\}$, $\forall x \in \mathcal{V}_{\bar{x}}$. It follows from (Ulbrich, 2011, Proposition 2.26) that piecewise C^1 (also called piecewise smooth) functions are semismooth and piecewise C^2 functions are strongly semismooth. Apparently, a piecewise affine/linear function is also piecewise C^k for all $k \geq 1$. Therefore, it is strongly semismooth.

The following result on semismoothness is derived in (Mifflin, 1977, Theorem 5), while the result on γ-order semismoothness is given in (Fischer, 1997, Theorem 19).

Proposition 2.5. *Let $g(\cdot) : \mathcal{X} \to \mathcal{Y}$ be (γ-order) semismooth at $x \in \mathcal{X}$ and $f(\cdot) : \mathcal{Y} \to \mathcal{Z}$ be (γ-order) semismooth at $g(x)$. Then the composite function $(f \circ g)(\cdot)$ is (γ-order) semismooth at x.*

This results indicate that the (strong) semismoothness is preserved under operations such as the sum and difference of two (strongly) semismooth functions. Besides, in (Qi and Sun, 1993, Corollary 2.4), it has been established that if each component of the function $g(\cdot)$ is (strongly) semismooth, then $g(\cdot)$ itself is (strongly) semismooth. In particular, the projection operator $\Pi_K(\cdot)$ is a piecewise affine function when K is a nonempty closed convex polyhedron (Rockafellar and Wets, 2009, Example 12.31), which implies that $\Pi_K(\cdot)$ is strongly semismooth. The property of strong semismoothness in the projection onto a set is applicable not only to closed convex polyhedral sets but also to some non-polyhedral sets, such as the positive semidefinite cone (Sun and Sun, 2002), the second-order cone (Chen et al., 2003) and the ℓ_2 norm ball (Zhang et al., 2020). Additionally, we know that the norm function $\|\cdot\|_c : \mathcal{X} \to \mathbb{R}_+$, $c \in [1, +\infty]$ is strongly semismooth (Facchinei and Pang, 2003a, Proposition 7.4.8).

There is another type of function that has been proven to be strongly semismooth. Let us start with the following definition given in (Coste, 2000, Definition 1.4 and 1.5) and (Bolte et al., 2009, Definition 1).

Definition 2.7 (o-minimal structure)**.** *A structure expanding the real closed field is a collection $\mathcal{S} = \{\mathcal{S}_n\}$, where each $\mathcal{S}_n$ is a set of subsets of the affine space $\mathbb{R}^n$, satisfying the following statements:*

(a) All algebraic subsets of $\mathbb{R}^n$ are in $\mathcal{S}_n$, i.e., $\mathcal{S}_n$ contains every subset in the form

$$\{x \in \mathbb{R}^n \mid s_1(x) = ... = s_k(x) = 0\},$$

where $s_i : \mathbb{R}^n \to \mathbb{R}, i \in [k]$ are polynomials.

(b) For every n, $\mathcal{S}_n$ is a Boolean subalgebra of subsets of $\mathbb{R}^n$, i.e., $\mathcal{S}_n$ is closed under finite unions and intersections and taking complement.

(c) If $U \in \mathcal{S}_m$ and $V \in \mathcal{S}_n$, then $U \times V \in \mathcal{S}_{m+n}$.

(d) If $\Pi : \mathbb{R}^{n+1} \to \mathbb{R}^n$ is the canonical projection on the first n coordinates and $Q \in \mathcal{S}_{n+1}$, then $\Pi(Q) \in \mathcal{S}_n$.

The elements of $\mathcal{S}_n$ are called the definable subsets of $\mathbb{R}^n$. Moreover, we say that the structure $\mathcal{S}$ is o-minimal if it further satisfies

(e) The elements of $\mathcal{S}_1$ are the finite unions of points and intervals.

A map $f : V \subset \mathbb{R}^n \to \mathbb{R}^m$ is called definable in $\mathcal{S}$ if its graph is a definable subset of $\mathbb{R}^n \times \mathbb{R}^m$.

Subsequently, we can present the definition of the tame function as follows.

Definition 2.8 (Tame function)**.** *(Bolte et al., 2009, Definition 2) A set $U \subset \mathcal{R}^n$ is called tame if for every $r > 0$ there exists an o-minimal structure $\mathcal{S}$ over $\mathbb{R}$ such that $U \cap [-r, r]^n$ is definable in $\mathcal{S}$. A mapping $f : V \subset \mathbb{R}^n \to \mathbb{R}^m$ is tame if its graph is a tame subset of $\mathbb{R}^n \times \mathbb{R}^m$.*

We know from (Bolte et al., 2009, Theorem 1) that any locally Lipschitz tame mapping $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$ is semismooth. Furthermore, in (Bolte et al., 2009), it is pointed that function $f(\cdot)$ achieves γ-order ($\gamma > 0$) semismoothness if $f(\cdot)$ is (globally) semialgebraic or subanalytic and Lipschitz continuous (definitions are given in (Bolte et al., 2009, Example 1 and Example 2 (a))).

We will then introduce the implicit function theorem, which connects the semismoothness of a Lipschitz function and its corresponding implicit function. Let $H : \mathcal{X} \times \mathcal{Y} \to \mathcal{Y}$ be a locally Lipschitz continuous function in a neighborhood of $(\bar{x}, \bar{y})$, where $(\bar{x}, \bar{y}) \in \mathcal{X} \times \mathcal{Y}$ is a solution to $H(x, y) = 0$. Let $\pi_y \partial H(x, y)$ be the canonical projection of $\partial H(x, y)$ onto $\mathcal{Y}$. We say that $\pi_y \partial H(x, y)$ is of maximal rank if every $M \in \pi_y \partial H(x, y)$ has maximal rank. Now, we can state the following implicit function theorem due to (Clarke, 1976).

Theorem 2.3. *Let $H : \mathcal{X} \times \mathcal{Y} \to \mathcal{Y}$ be a locally Lipschitz continuous function in a neighborhood of $(\bar{x}, \bar{y})$, which is a solution to $H(x, y) = 0$. If $\pi_y \partial H(x, y)$ is of maximal rank, then there is an open neighborhood X of $\bar{x}$ and a Lipschitzian function $G(\cdot) : X \to \mathcal{Y}$ such that $G(\bar{x}) = \bar{y}$ and for every $x \in X$, $H(x, G(x)) = 0$.*

Based on the above Clarke's implicit function, Sun (2001) established the following result.

Theorem 2.4. *Suppose that all conditions in Theorem 2.3 hold. If $H(\cdot, \cdot)$ is (strongly) semismooth at $(\bar{x}, \bar{y})$, then $G(\cdot)$ is (strongly) semismooth at some point in the neighborhood X of $\bar{x}$.*

Note that the assumption of the maximal rank for $\pi_y \partial H(x, y)$ in Theorems 2.3 and 2.4 is equivalent to the following Clark's nonsingularity condition

$$\forall\, d \in \mathcal{Y}, \quad 0 \in \partial H(\bar{x}, \bar{y})(0, d) \Rightarrow d = 0. \tag{2.16}$$

This condition is stricter than

$$\forall\, d \in \mathcal{Y}, \quad 0 \in D_* H(\bar{x}, \bar{y})(0, d) \Rightarrow d = 0, \tag{2.17}$$

where $D_* H(\bar{x}, \bar{y})(0, d)$ denotes the strict derivative of $H(\cdot, \cdot)$ at $(\bar{x}, \bar{y})$ in the direction $(0, d)$ consisting of all points

$$\lim_{(x^k, y^k) \to (\bar{x}, \bar{y}),\, t_k \downarrow 0} \frac{H((x^k, y^k) + t_k(0, d)) - H(x^k, y^k)}{t_k}.$$

It follows from (Kummer, 1991b) that

$$\partial_B H(\bar{x}, \bar{y})(u, v) \subseteq D_* H(\bar{x}, \bar{y})(u, v) \subseteq \partial H(\bar{x}, \bar{y})(u, v), \quad \forall\, (u, v) \in \mathcal{X} \times \mathcal{Y}. \tag{2.18}$$

Note that, if we replace the assumption of maximal rank for $\pi_y \partial H(x, y)$ with (2.17), the results stated in Theorems 2.3 and 2.4 still hold according to (Kummer, 1991a, Theorem 1) and (Meng et al., 2005, Corollary 2).

2.3 The secant method for semismooth equations

In this section, we will introduce the secant method for solving semismooth equations. The classical secant method is one of the most efficient methods for solving smooth equations, and it is well known that the classical secant method converges superlinearly with Q-order of at least $(1+\sqrt{5})/2$ (Traub, 1964). Later, Potra et al. (1998) generalized the secant method for solving (strongly) semismooth equations and showed that the secant method converges 2-step or 3-step Q-superlinear (-quadratically).

Let $f : \mathbb{R} \to \mathbb{R}$ be a real valued functional. Denote for $x \neq y$ that

$$\delta_f(x, y) := (f(x) - f(y)) / (x - y). \tag{2.19}$$

For finding a zero of the function $f(\cdot)$, the secant method performs the following update iteration at the k-th iteration:

$$x^{k+1} = x^k - \delta_f(x^k, x^{k-1})^{-1} f(x^k).$$

The expression $\delta_f(x^k, x^{k-1})$ represents the divided difference approximation of the derivative $f'(x^k)$ (if it exists) using the previous two iterates x^k and x^{k-1} along with their values. If $f(\cdot)$ is smooth at a zero x^* of $f(x) = 0$, $x \in \mathbb{R}$ and the derivative $f'(x^*)$ is nonzero, the secant method is superlinearly convergent in the sense that $|x^{k+1} - x^*| = o(|x^k - x^*|)$. Furthermore, if the derivative $f'(\cdot)$ is Lipschitz continuous in the neighborhood of x^*, then the secant method converges superlinearly with a Q-order of at least $(1 + \sqrt{5})/2$, as shown by Traub (1964). For a specific class of nonsmooth equations known as (strongly) semismooth equations, Potra et al.

(1998) demonstrated that the secant method exhibits Q-superlinear (or quadratic) convergence in either 2 or 3 steps

Let $f : \mathbb{R} \to \mathbb{R}$ be a locally Lipschitz continuous function which is semismooth at a solution x^* to the equation

$$f(x) = 0. \tag{2.20}$$

The following lemma is useful for analyzing the convergence of the secant method. Part (a) of this lemma is from (Potra et al., 1998, Lemma 2.2), and Part (b) can be proved by following a similar procedure as in the proof of (Potra et al., 1998, Lemma 2.3).

Lemma 2.1. *Assume that $f : \mathbb{R} \to \mathbb{R}$ is semismooth at $\bar{x} \in \mathbb{R}$. Denote the lateral derivatives of f at $\bar{x}$ by*

$$\bar{d}^- := -f'(\bar{x}; -1) \quad \text{and} \quad \bar{d}^+ := f'(\bar{x}; 1). \tag{2.21}$$

(a) Then the lateral derivatives $\bar{d}^-$ and $\bar{d}^+$ exist and

$$\partial_{\mathrm{B}} f(\bar{x}) = \{\bar{d}^-, \bar{d}^+\};$$

(b) It holds that

$$\bar{d}^- - \delta_f(u, v) = o(1) \quad \text{for all } u \uparrow \bar{x},\ v \uparrow \bar{x}; \tag{2.22}$$

$$\bar{d}^+ - \delta_f(u, v) = o(1) \quad \text{for all } u \downarrow \bar{x},\ v \downarrow \bar{x}; \tag{2.23}$$

moreover, if $f(\cdot)$ is γ-order semismooth at $\bar{x}$ for some $\gamma > 0$, then

$$\bar{d}^- - \delta_f(u, v) = O(|u - \bar{x}|^\gamma + |v - \bar{x}|^\gamma) \quad \text{for all } u \uparrow \bar{x},\ v \uparrow \bar{x}; \tag{2.24}$$

$$\bar{d}^+ - \delta_f(u, v) = O(|u - \bar{x}|^\gamma + |v - \bar{x}|^\gamma) \quad \text{for all } u \downarrow \bar{x},\ v \downarrow \bar{x}. \tag{2.25}$$

We analyze the convergence of the secant method described in Algorithm 1 with two generic starting points x^{-1} and x^0. The convergence properties of the secant method for semismooth equations can be found in Theorem 2.5.

Algorithm 1 A secant method for solving (2.20)

Input: $x^{-1}, x^0 \in \mathbb{R}$.
Initialization: Set $k = 0$.
while $f(x^k) \neq 0$ **do**
  Step 1. Compute

$$x^{k+1} = x^k - (\delta_f(x^k, x^{k-1}))^{-1} f(x^k). \tag{2.26}$$

  Step 2. Set $k = k + 1$.
end while
Output: x^k.

Theorem 2.5 (Potra, Qi, and Sun (1998, Theorem 3.2)). *Suppose that $f : \mathbb{R} \to \mathbb{R}$ is semismooth at a solution x^* of (2.20). Let d^- and d^+ be the lateral derivatives of $f(\cdot)$ at x^* as defined in (2.21). If d^- and d^+ are both positive (or negative), then there are two neighborhoods $\mathcal{U}$ and $\mathcal{N}$ of x^*, $\mathcal{U} \subseteq \mathcal{N}$, such that for $x^{-1}, x^0 \in \mathcal{U}$, Algorithm 1 is well defined and produces a sequence of iterates $\{x^k\}$ such that $\{x^k\} \subseteq \mathcal{N}$. The sequence $\{x^k\}$ converges to x^* 3-step Q-superlinearly, i.e., $|x^{k+3} - x^*| = o(|x^k - x^*|)$. Furthermore, if*

$$\alpha := \frac{|d^+ - d^-|}{\min\{|d^+|, |d^-|\}} < 1,$$

then $\{x^k\}$ converges to x^* Q-linearly with Q-factor α; If f is strongly semismooth at x^*, then the sequence $\{x^k\}$ converges to x^* 3-step Q-quadratically.

When the values d^- and d^+ have the same sign, Theorem 2.5 states that the sequence x^k exhibits 3-step Q-superlinear (Q-quadratic if $f(\cdot)$ is strongly semismooth) convergence. On the other hand, when $d^- \cdot d^+ < 0$, the sequence $\{x^k\}$ converges to x^* with 2-step Q-superlinear (Q-quadratic if $f(\cdot)$ is strongly semismooth) convergence (Potra et al., 1998, Theorem 3.3). We point out that when $|d^+ - d^-|$ is small and $f(\cdot)$ is strongly semimsooth, we know from Theorem 2.5 that the secant method converges with a fast linear rate and 3-step Q-quadratic rate. We provide a numerical example slightly modified from (Potra et al., 1998, Equation (3.15)) to illustrate the convergence rates shown in Theorem 2.5. We test Algorithm 1 with $x^{-1} = 0.01$ and $x^0 = 0.005$ for finding the zero $x^* = 0$ of

$$f(x) = \begin{cases} x(x+1) & \text{if } x < 0, \\ -\beta x(x-1) & \text{if } x \geq 0, \end{cases} \tag{2.27}$$

where β is chosen from $\{1.1, 1.5, 2.1\}$. The numerical results are shown in Table (2.1), which coincide with the theoretical results.

Table 2.1: The numerical performance of finding the zero of (2.27). Case I: $\beta = 1.1,\ d^+ = 1.1,\ d^- = 1,$ and $\alpha = 0.1$; Case II: $\beta = 1.5,\ d^+ = 1.5,\ d^- = 1,$ and $\alpha = 0.5$; Case III: $\beta = 2.1,\ d^+ = 2.1,\ d^- = 1,$ and $\alpha = 1.1$.

Case	Iter	1	2	3	4	5	6	7	8
I	x	-5.1e-5	-4.3e-6	2.2e-10	-2.2e-11	-1.8e-12	4.1e-23	-4.1e-24	-3.4e-25
II	x	-5.1e-5	-1.7e-5	8.4e-10	-4.2e-10	-1.1e-10	4.5e-20	-2.2e-20	-5.6e-21
III	x	-5.1e-5	-2.6e-5	1.3e-9	-1.5e-9	-5.1e-10	7.4e-19	-8.2e-19	-2.8e-19

The secant method for strongly semismooth equations has an R-order of at least $\sqrt[3]{2}$ when $d^- \cdot d^+ > 0$, and an R-order of at least $\sqrt{2}$ when $d^- \cdot d^+ < 0$, as follows from the next proposition.

Proposition 2.6 (Potra, Qi, and Sun (1998, Lemma 4.1))**.** *If $\{x^k\}$ is a convergent sequence with limit x^* and satisfies*

$$|x^{k+p} - x^*| = O(|x^k - x^*|^r),$$

then the R-order of convergence of $\{x^k\}$ is at least $\sqrt[p]{r}$.

2.4 The alternating direction method of multipliers

In this section, we will introduce an important and well known iteration method, the alternating direction method of multipliers (ADMM, Glowinski and Marroco (1975); Gabay and Mercier (1976)), for solving the problem (CP(ϱ)) studied in this thesis.

We begin this section with a more general form. Consider the convex optimization problem characterized by the following separable structure

$$\begin{aligned} \min_{x\in\mathcal{X},\, y\in\mathcal{Y}} \quad & f(x) + g(y) \\ \text{s.t.} \quad & Ax + By = c, \end{aligned} \tag{2.28}$$

where $f : \mathcal{X} \to (-\infty, +\infty]$ and $g : \mathcal{Y} \to (-\infty, +\infty]$ are closed proper convex functions, $A : \mathcal{X} \to \mathcal{Z}$ and $B : \mathcal{Y} \to \mathcal{Z}$ are linear operators. The ADMM solves problem (2.28) by the procedures given in Algorithm 2. When $\tau = 1$, Eckstein and Bertsekas

Algorithm 2 The ADMM for solving (2.28)

Input: $(x^0, y^0, z^0) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$.
Initialization: Set $k = 0$.
while A termination criterion is not met **do**
  Step 1. Set

$$\begin{cases} x^{k+1} = \arg\min_{x\in\mathcal{X}} f(x) - \langle z^k, Ax\rangle + \frac{\rho}{2}\|Ax + By^k - c\|^2, \\ y^{k+1} = \arg\min_{y\in\mathcal{Y}} g(y) - \langle z^k, By\rangle + \frac{\rho}{2}\|Ax^{k+1} + By - c\|^2, \end{cases} \tag{2.29}$$

  where $\rho > 0$ is the penalty parameter.
  Step 2. Set

$$z^{k+1} = z^k - \tau\rho(Ax^{k+1} + By^{k+1} - c), \tag{2.30}$$

  where $\tau \in (0, (1 + \sqrt{5})/2)$ is the step length.
  Step 3. Set $k = k + 1$.
end while
Output: (x^k, y^k, z^k).

(1992) has shown that the ADMM, being a special case of the Douglas–Rachford splitting (Gabay, 1983), is an application of the proximal point algorithm on the dual problem for a specially constructed operator. When $B = I$ and A is surjective, the global convergence of the ADMM with $\tau \in (0, (1+\sqrt{5})/2)$ has been demonstrated by Glowinski and Oden (1985) and Fortin and Glowinski (1983).

We consider the following constraint qualification (CQ): There exists $(x, y) \in \mathrm{ri}(\mathrm{dom} f \times \mathrm{dom} g) \cap Q$, where Q is the constraint set in (2.28). Under this CQ, we know from (Rockafellar, 1970, Corollary 28.3.1) that $(\bar{x}, \bar{y}) \in \mathcal{X} \times \mathcal{Y}$ is an optimal solution to (2.28) if and only if there exists a Lagrange multiplier $\bar{z} \in \mathcal{Z}$ such that

$$A^*\bar{z} \in \partial f(\bar{x}), \quad B^*\bar{z} \in \partial g(\bar{y}), \quad A\bar{x} + B\bar{y} - c = 0. \tag{2.31}$$

Besides, any $\bar{z} \in \mathcal{Z}$ is an optimal solution to the dual of (2.28). Furthermore, since $f(\cdot)$ and $g(\cdot)$ are both closed proper convex functions, it is known from (Rockafellar and Wets, 2009, Theorem 12.17) (the subdifferential mappings of the closed proper convex functions are maximal monotone) that there are two self adjoint and positive semidefinite operators, Σ_f and Σ_g, such that, for any $x, \hat{x} \in \mathrm{dom} f$, $v \in \partial f(x)$ and $\hat{v} \in \partial f(\hat{x})$,

$$f(x) \geq f(\hat{x}) + \langle \hat{v}, x - \hat{x} \rangle + \frac{1}{2}\|x - \hat{x}\|^2_{\Sigma_f} \text{ and } \langle v - \hat{v}, x - \hat{x} \rangle \geq \|x - \hat{x}\|^2_{\Sigma_f},$$

and for any $y, \hat{y} \in \mathrm{dom} g$, $w \in \partial f(y)$ and $\hat{w} \in \partial f(\hat{y})$,

$$f(y) \geq f(\hat{y}) + \langle \hat{w}, y - \hat{y} \rangle + \frac{1}{2}\|y - \hat{y}\|^2_{\Sigma_g} \text{ and } \langle w - \hat{w}, y - \hat{y} \rangle \geq \|y - \hat{y}\|^2_{\Sigma_g}.$$

We can then introduce the following convergence result of ADMM for (2.28) from (Fazel et al., 2013, Theorem B.1).

Theorem 2.6. *Assume that the solution set to* (2.28) *is nonempty and that the CQ holds. If both* $\Sigma_f + \rho A^*A$ *and* $\Sigma_g + \rho B^*B$ *are positive definite, then the sequences* $\{(x^k, y^k)\}$ *and* $\{z^k\}$ *generated by Algorithm 2 converge to an optimal solution of* (2.28) *and to an optimal solution of the dual of* (2.28), *respectively.*

Now, recall the main problem (CP(ϱ)) that we are interested

$$\min_{x \in \mathbb{R}^n} \{p(x) \mid \|Ax - b\| \leq \varrho\}.$$

It is clear that we can assume $Ax - b \neq 0$ for all $x \in \mathbb{R}^n$; otherwise, the problem would become trivial and lose its significance. Note that, we can rewrite (CP(ϱ))

into an equivalent form

$$\begin{aligned} \min_{x\in\mathbb{R}^n,\, w\in\mathbb{R}^m} \quad & p(x) + \delta_D(w) \\ \text{s.t.} \quad & Ax - b = w, \end{aligned} \tag{2.32}$$

where $D = \{w \in \mathbb{R}^m \mid \|w\| \leq \varrho\}$. We will not solve this equivalent problem using ADMM, as n is generally much larger than m (the number of rows of A). In such cases, using ADMM to address the primal problem is typically more computationally expensive than solving its dual.

We have $\delta_D^*(y) = \varrho\|y\|$, since $\delta_D^*(y) = \sup_{s\in D}\{\langle s, y\rangle\} = \sup_{s\in D}\{\|s\|\|y\|\} = \varrho\|y\|$. Then, we can readily obtain the dual of (2.32) as

$$\begin{aligned} \min_{y\in\mathbb{R}^m, z\in\mathbb{R}^n} \quad & \varrho\|y\| + \langle b, y\rangle + p^*(z) \\ \text{s.t.} \quad & A^T y + z = 0. \end{aligned} \tag{2.33}$$

However, when applying ADMM to (2.33), the variable y is updated in k-th iteration by solving the following problem:

$$\min_{y\in\mathbb{R}^m} \ \varrho\|y\| + \langle b - Ax^k, y\rangle + \frac{\rho}{2}\|A^T y + z^k\|^2,$$

which may not be straightforward to compute. Thus, we introduce another variable $s \in \mathbb{R}^m$ to rewrite (2.33) into:

$$\begin{aligned} \min_{y\in\mathbb{R}^m, s\in\mathbb{R}^m, z\in\mathbb{R}^n} \quad & \varrho\|s\| + \langle b, y\rangle + p^*(z) \\ \text{s.t.} \quad & \begin{pmatrix} A^T & 0 \\ -I & I \end{pmatrix} \begin{pmatrix} y \\ s \end{pmatrix} + \begin{pmatrix} I \\ 0 \end{pmatrix} z = 0. \end{aligned} \tag{2.34}$$

Then, through some straightforward derivation, we obtain Algorithm 3 from Algorithm 2 by treating $\begin{pmatrix} y \\ s \end{pmatrix}$ as the first variable and z as the second one. Moreover, from Theorem 2.6, we know that under the assumption that the solution set of (2.33) is nonempty and that the CQ holds for (2.33), the sequences $\{(y^k, z^k)\}$ and $\{x^k\}$ generated by Algorithm 3 converge to an optimal solution of (2.33) and to an optimal solution of (CP(ϱ)), respectively.

Algorithm 3 The ADMM for solving (2.34)

1: **Input**: $(y^0, s^0, z^0, x^0, w^0) \in \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m$; $\rho > 0$; and $\tau \in (0, (1+\sqrt{5})/2)$.
2: **Initialization**: Set $k = 0$.
3: **while** A termination criterion is not met **do**
4: **Step 1**. Update y^{k+1} by solving

$$\rho(AA^T + I)y = \rho s^k + A(x^k - \rho z^k) - w^k - b.$$

5: **Step 2**. Set

$$\begin{cases} s^{k+1} = \mathrm{Prox}_{\varrho\|\cdot\|/\rho}(y^{k+1} + w^k/\rho), \\ z^{k+1} = \mathrm{Prox}_{p^*/\rho}(x^k/\rho - A^T y^{k+1}), \end{cases} \tag{2.35}$$

6: **Step 3**. Set

$$\begin{cases} x^{k+1} = x^k - \tau\rho(A^T y^{k+1} + z^{k+1}); \\ w^{k+1} = w^k - \tau\rho(s^{k+1} - y^{k+1}). \end{cases} \tag{2.36}$$

7: **Step 4**. Set $k = k + 1$.
8: **end while**
9: **Output**: $(y^k, s^k, z^k, x^k, w^k)$.

Chapter 3

The adaptive sieving technique for sparse optimization problems

In this chapter, we will focus on the unconstrained sparse optimization problems with a dimension reduction technique introduced. This technique, called adaptive sieving, addresses the original sparse optimization problem by solving several reduced problems with much smaller dimensions compared to the original. Several dual based approaches exist for solving the subproblems within the adaptive sieving strategy. However, we may encounter situations where m (the number of rows of A) is much larger than the dimension of the reduced problem. In such scenarios, dual based algorithms may not be the most efficient option. To further enhance computational efficiency, we will introduce a smoothing Newton method for the primal problem. Additionally, we will present a warm-started path-following adaptive sieving technique specifically designed to tackle extremely large scale problems.

3.1 The adaptive sieving technique

In this section, we will first introduce the adaptive sieving technique developed in (Yuan et al., 2023, 2022) for solving unconstrained sparse optimization problems. This technique addresses the original problem by solving a sequence of reduced problems. The dimensionality of these reduced problems is significantly smaller than that of the original problem. We will conduct extensive numerical experiments on the precision matrix estimation problem to demonstrate that this technique can significantly improve the overall efficiency.

3.1.1 The adaptive sieving technique for sparse optimization problems

We introduce the adaptive sieving technique for solving sparse optimization problems of the following form:

$$\min_{x\in\mathbb{R}^n} \left\{\Phi(x) + P(x)\right\}, \tag{3.1}$$

where $\Phi:\mathbb{R}^n\to\mathbb{R}$ is a continuously differentiable convex function, and $P:\mathbb{R}^n\to(-\infty,+\infty]$ is a proper closed convex function. We assume that the convex composite optimization problem (3.1) has at least one solution. We define the proximal residual function $R:\mathbb{R}^n\to\mathbb{R}^n$ as

$$R(x) := x - \mathrm{Prox}_P(x-\nabla\Phi(x)), \quad x\in\mathbb{R}^n. \tag{3.2}$$

The norm of $R(x)$ is a standard measurement for the quality of an obtained solution, and x is a solution to (3.1) if and only if $R(x)=0$.

Let $I\subseteq[n]$ be an index set. We consider the following constrained optimization problem with the index set I:

$$\min_{x\in\mathbb{R}^n} \left\{\Phi(x) + P(x) \mid x_{I^c}=0\right\}, \tag{3.3}$$

where $I^c=[n]\backslash I$ is the complement set of I. A key fact is that, a solution to (3.3) is also a solution to (3.1) if there exists a solution $\bar{x}$ to (3.1) such that $\mathrm{supp}(\bar{x})\subseteq I$. The adaptive sieving technique is motivated by this fact. Specifically, starting with a reasonable guessing $I_0\subseteq[n]$, the adaptive sieving technique is an adaptive strategy to refine the current index set I_k based on a solution to (3.3) with $I=I_k$. We present the details of the adaptive sieving technique for solving (3.1) in Algorithm 4.

Algorithm 4 The adaptive sieving strategy for solving (3.1)

Input: an initial index set $I_0 \subseteq [n]$, a given tolerance $\epsilon \geq 0$ and a given positive integer $k_{\max}$ (e.g., $k_{\max} = 500$).
Output: a solution x^* to the problem (3.1) satisfying $\|R(x^*)\| \leq \epsilon$.
1. Find

$$x^0 \in \arg\min_{x \in \mathbb{R}^n} \Big\{ \Phi(x) + P(x) - \langle \delta^0, x \rangle \;\mid\; x_{I_0^c} = 0 \Big\}, \tag{3.4}$$

where $\delta^0 \in \mathbb{R}^n$ is an error vector such that $\|\delta^0\| \leq \epsilon$ and $(\delta^0)_{I_0^c} = 0$.
2. Compute $R(x^0)$ and set $s = 0$.
while $\|R(x^s)\| > \epsilon$ **do**
  3.1. Create J_{s+1} as

$$J_{s+1} = \Big\{ j \in I_s^c \;\mid\; (R(x^s))_j \neq 0 \Big\}. \tag{3.5}$$

  If $J_{s+1} = \emptyset$, let $I_{s+1} \leftarrow I_s$; otherwise, let k be a positive integer satisfying $k \leq \min\{|J_{s+1}|, k_{\max}\}$ and define

$$\widehat{J}_{s+1} = \Big\{ j \in J_{s+1} \;\Big|\; |(R(x^s))_j| \text{ is among the first } k \text{ largest values in } \{|(R(x^s))_i|\}_{i \in J_{s+1}} \Big\}.$$

  Update I_{s+1} as:

$$I_{s+1} \leftarrow I_s \cup \widehat{J}_{s+1}.$$

  3.2. Solve the constrained problem:

$$x^{s+1} \in \arg\min_{x \in \mathbb{R}^n} \Big\{ \Phi(x) + P(x) - \langle \delta^{s+1}, x \rangle \;\mid\; x_{I_{s+1}^c} = 0 \Big\}, \tag{3.6}$$

  where $\delta^{s+1} \in \mathbb{R}^n$ is an error vector such that $\|\delta^{s+1}\| \leq \epsilon$ and $(\delta^{s+1})_{I_{s+1}^c} = 0$.
  3.3: Compute $R(x^{s+1})$ and set $s \leftarrow s + 1$.
end while
return: Set $x^* = x^s$.

It is worthwhile mentioning that, in Algorithm 4, the error vectors δ^0, $\{\delta^{s+1}\}$ in (3.4) and (3.6) are not given but imply that the corresponding minimization problems can be solved inexactly. We can just take $\delta^s = 0$ (for $s \geq 0$) if we solve the reduced subproblems exactly. The following proposition shows that we can obtain an inexact solution by solving a reduced problem with a much smaller dimension.

Proposition 3.1 (Yuan, Lin, Sun, and Toh (2023, Proposition 1))**.** *For any given* *nonnegative integer s, the updating rule of x^s in Algorithm 4 can be interpreted in the procedure as follows. Let M_s be a linear map from $\mathbb{R}^{|I_s|}$ to $\mathbb{R}^n$ defined as*

$$(M_s z)_{I_s} = z, \quad (M_s z)_{I_s^c} = 0, \quad z \in \mathbb{R}^{|I_s|},$$

and Φ^s, P^s be functions from $\mathbb{R}^{|I_s|}$ to $\mathbb{R}$ defined as $\Phi^s(z) := \Phi(M_s z)$, $P^s(z) := P(M_s z)$ for all $z \in \mathbb{R}^{|I_s|}$. Then $x^s \in \mathbb{R}^n$ can be computed as

$$(x^s)_{I_s} := \mathrm{Prox}_{P^s}(\hat{z} - \nabla \Phi^s(\hat{z})),$$

and $(x^s)_{I_s^c} = 0$, where $\hat{z}$ is an approximate solution to the problem

$$\min_{z \in \mathbb{R}^{|I_s|}} \left\{ \Phi^s(z) + P^s(z) \right\}, \tag{3.7}$$

which satisfies

$$\|\hat{z} - \mathrm{Prox}_{P^s}(\hat{z} - \nabla \Phi^s(\hat{z})) + \nabla \Phi^s(\mathrm{Prox}_{P^s}(\hat{z} - \nabla \Phi^s(\hat{z}))) - \nabla \Phi^s(\hat{z})\| \leq \epsilon, \tag{3.8}$$

and ϵ is the parameter given in Algorithm 4.

The finite termination property of Algorithm 4 for solving (3.1) is shown in the following proposition.

Proposition 3.2 (Yuan, Lin, Sun, and Toh (2023, Theorem 1))**.** *For any given initial index set $I_0 \subseteq [n]$ and tolerance $\epsilon \geq 0$, the while loop in Algorithm 4 will terminate after a finite number of iterations.*

The high efficiency of the adaptive sieving technique for solving a wide class of sparse optimization problems in the form of (3.1) has been demonstrated in (Yuan et al., 2023, 2022; Li et al., 2023; Wu et al., 2023), such as the (group) Lasso penalized least square problem and the (exclusive) Lasso logistic regression problem.

3.1.2 Numerical experiments for the precision matrix estimation problem

In this section, we will present in detail the adaptive sieving technique for addressing the precision matrix estimation problem to demonstrate the performance of this technique (Li et al., 2023). An ℓ_1-penalized D-trace loss estimator was proposed in (Zhang and Zou, 2014; Liu and Luo, 2015) for estimating the precision matrix (or inverse covariance matrix). This estimator is derived from solving a convex composite optimization problem that incorporates a quadratic loss function along with an ℓ_1-regularized penalty:

$$\min_{\Omega \in \mathbb{S}^p} \left\{ \frac{1}{2} \mathrm{tr}(\Omega \widehat{\Sigma} \Omega^T) - \mathrm{tr}(\Omega) + \lambda \|\Omega\|_{1,\mathrm{off}} \right\}, \tag{3.9}$$

where $\mathbb{S}^p$ is the space of $p \times p$ real symmetric matrices and $\|\cdot\|_{1,\mathrm{off}}$ is the off-diagonal ℓ_1-norm, i.e., $\|\Omega\|_{1,\mathrm{off}} = \sum_{i \neq j} |\Omega_{i,j}|$.

We will develop a dual based approach to solve problem (3.9). To facilitate the designing of the dual approach, we write problem (3.9) equivalently as

$$\min_{\Omega \in \mathbb{S}^p} \left\{ \frac{1}{2} \|\Omega A\|_F^2 - \langle \Omega, I_p \rangle + \lambda \|\Omega\|_{1,\mathrm{off}} \right\}, \tag{3.10}$$

where A is a real matrix with rank n such that $AA^T = \widehat{\Sigma}$. Note that instead of applying the singular value decomposition (SVD) on $\widehat{\Sigma}$, the matrix A can be efficiently obtained by applying a thin SVD on the $p \times n$ dimensional centered data matrix. The thin SVD requires significantly less space and time than the full SVD, especially in the high-dimensional setting. Without loss of generality, we assume that A is a $p \times n$ matrix with rank n. For later use, we denote $\theta(\Omega) := \|\Omega\|_{1,\mathrm{off}}, \forall\, \Omega \in \mathbb{S}^p$. Moreover, we further denote the optimal solution set of (3.10) by Θ_λ, and the associated proximal residual mapping by

$$R_\lambda(\Omega) := \Omega - \mathrm{Prox}_{\lambda\theta}(\Omega - h(\Omega)), \quad \forall\, \Omega \in \mathbb{S}^p,$$

where $h(\Omega) := \frac{1}{2}(\Omega\widehat{\Sigma} + \widehat{\Sigma}\Omega) - I_p$ with I_p being the p dimensional identity matrix, and δ_{B_λ} is the indicator function with $B_\lambda = \{Z \in \mathbb{S}^p \mid Z_{ii} = 0, |Z_{ij}| \leq \lambda, i, j =$

$1, \cdots, p,\ i \neq j\}$, i.e., $\delta_{B_\lambda}(Z) = 0$ for any $Z \in B_\lambda$ and $\delta_{B_\lambda}(Z) = +\infty$ otherwise. We know that $\widetilde{\Omega} \in \Theta_\lambda$ if and only if $R_\lambda(\widetilde{\Omega}) = 0$.

We call our algorithm MARS, since it is designed for **M**atrix estimation via an **A**daptive sieving **R**eduction strategy and a **S**emismooth Newton augmented Lagrangian algorithm. Then, we will conduct several tests to illustrate the performance of our MARS. For comparison, we consider several popular solvers including scio (Liu and Luo, 2015), EQUAL (Wang and Jiang, 2020), glasso (Friedman et al., 2008), and QUIC (Hsieh et al., 2014). We point out that the main purpose of presenting the performances of "glasso" and "QUIC", which are designed for the graphical lasso model, is not for comparison. Note that, the graphical lasso model was developed much earlier and has been widely used in many research areas for years, and less attention was given to the D-trace loss estimator. The main reason that we compare MARS with "glasso" and "QUIC" is to provide users with a more intuitive demonstration that there is an alternative and more efficient choice for estimating the precision matrix with our developed package. Since the existing popular methods are mainly first order methods and the stopping criteria of those algorithms are different from each other and also ours, for better comparison, we will also test some other algorithms for solving (3.10) (all the details of those algorithms can be found in (Li et al., 2023, Section 6.1)). Specifically, we will conduct tests with a second-order algorithm, namely a semismooth Newton augmented Lagrangian method (SSNAL), and two kinds of alternating direction methods of multipliers (ADMM), where one is derived by solving the sub-problem inexactly (iADMM) and the other derived by solving it exactly (eADMM).

Before proceeding to the experiments, we provide some explanations about our MARS. In our MARS, we use the relative KKT residual

$$\eta = \frac{\|R(\Omega)\|_F}{1 + \|h\left(\Omega\right)\|_F + \|\Omega\|_F}$$

to measure the accuracy of the generated solution Ω. That is, we use η to decide whether our MARS should be stopped. Unless otherwise specified, we set the stopping tolerance to 10^{-4} for all the solvers/algorithms except EQUAL in the following experiments. Based on several tests, the stopping tolerance of EQUAL is set to 10^{-6}. The reason for such an adjustment is that their stopping criterion is determined by the distance between two solutions in two consecutive iterations, and a slightly larger stopping tolerance may cause the generated solution to be too far from the optimal solution set, in terms of the relative KKT residual.

All the numerical results in this section are obtained by running Microsoft R Open 4.0.2 on a Windows workstation (Intel(R) Core(TM) i7-10700 CPU @2.90GHz 2.00GHz RAM 32GB). For simplicity, we will use R to represent Microsoft R Open 4.0.2.

Now, we will use some real data sets to demonstrate the promising performance of our MARS for generating a solution path. The publicly available data sets we are going to use include a prostate data set (`https://web.stanford.edu/~hastie/CASI_files/DATA/prostate.html`) and a breast cancer data set (Hess et al., 2006), which can be found on (`https://bioinformatics.mdanderson.org/public-datasets/`). The prostate data set contains two groups, the first one is 6033 genetic activity measurements of 50 control subjects and the other is that of 52 prostate cancer subjects. Thus, the number of variables contained in the precision matrix that needs to be estimated is more than 18 million. As for the breast cancer data set, it contains the measurements of 22283 genes with 133 subjects, where 99 of them are labeled as residual disease (RD) and the remaining 34 subjects are labeled as pathological complete response (pCR). For this data set, the estimated precision matrix contains about 250 million parameters.

After standardizing the two groups of the prostate data set, we use MARS, SSNAL, EQUAL, and scio to generate solution paths for the two groups separately. We should note that, when λ is too small, there may not exist optimal solutions for the precision matrix estimator. Therefore, before going further to the main comparison tests, we should conduct some pretests to find a suitable smallest λ. By comparing the objective value and η (details can be dound in (Li et al., 2023)), we conclude that MARS and SSNAL can outperform both EQUAL and scio since all the η are smaller than the set tolerance 10^{-4}. Although both MARS and SSNAL can generate satisfactory solutions, from Table 3.1, we find that MARS is much more efficient. In particular, the computation time of SSNAL to generate the solution path is more than 14 times that of MARS in the Control group and more than 18 times that of MARS in the Cancer group. This can also be seen in Figure 3.1, which illustrates that MARS has high efficiency in generating solutions for each λ. Besides, we obtain the final precision matrix estimations of the two different groups through 5-fold cross-validation, and the corresponding graphs are shown in Figure 3.2. From this figure, we can clearly see that the genes of the control group and the cancer group have different connections.

Table 3.1: The computation time (seconds) of different algorithms for generating a solution path with the prostate data set.

	MARS	SSNAL	EQUAL	scio
control group	113.6	1629.88	1325.67*	5690.73+
cancer group	112.43	2108.06	1273.22*	5949.71+

1. The symbol "$*$" indicates that none of the relative KKT residuals of EQUAL is less than 10^{-3}.
2. The symbol "+" indicates that, due to out of memory, the time here does not include the time for generating estimations by scio with the two smallest λ.

Next, we will test the performance of MARS and glasso on the breast cancer data set. We follow the same assumption stated in (Cai et al., 2011) that this gene measurements data are normally distributed with $N(\mu_k, \Sigma)$, $k = 1, 2$, where Σ is the same for RD group and pCR group, but the means are different. Some two-sample

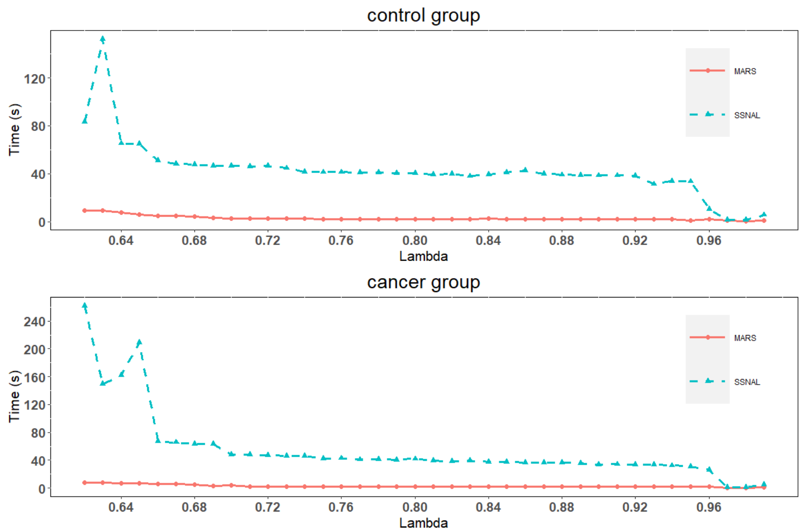


Figure 3.1: The computation time of MARS and SSNAL for each λ with the prostate data set.

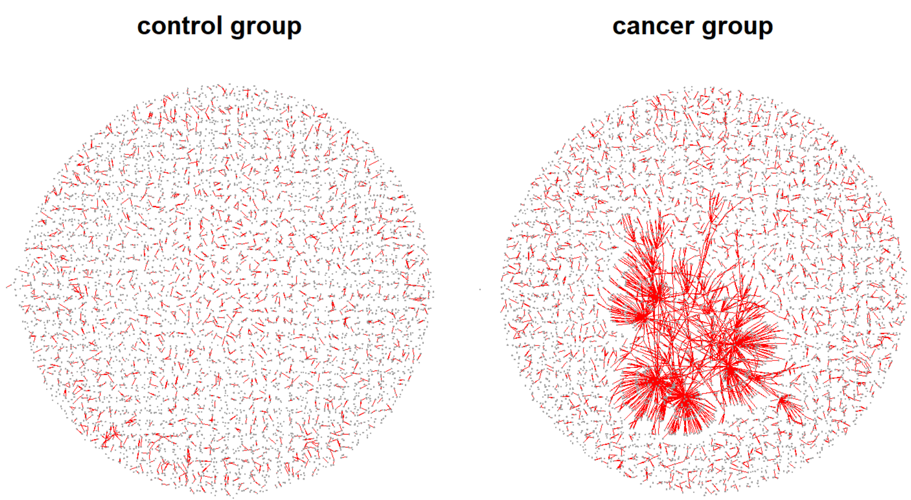


Figure 3.2: The estimated graphs chosen by five-fold cross-validation generated by MARS with the prostrate data set.

t-tests are performed with given p-value tolerances, which are set to 0.005, 0.01, 0.05, 0.1, and 1, to obtain the most significant genes (with a smaller p-value). Under those set p-values, the numbers of chosen genes are 1228, 1646, 3640, 5418, and 22283 respectively. Note that, the last one contains all the genes with nearly 250 million parameters. We point out that, the λ paths for all the tests, except the test with p-value tolerance 0.05, are set from $\lambda_{\min}$ to 1 by 0.01, where $\lambda_{\min}$ is decided by some pre-tests with the D-trace estimator. When the p-value tolerance is set to 0.05, if the gap between two subsequent regularization parameters in the path is 0.01, glasso will fail due to insufficient memory, so we set the λ gap for this test to 0.02. The regularization parameters for each test are chosen by five-fold cross-validation, and the total computation times are concluded in Table 3.2. The stopping tolerance for MARS and glasso is set to 10^{-4} to ensure that all the relative KKT residuals are less than 10^{-4}. The estimated graphs obtained by MARS and glasso with p-value tolerance 0.005, 0.01, and 0.01 can be found in Figure 3.3. From this figure, we notice that the graphs obtained by MARS and glasso are similar to each other, but the times taken by MARS are obviously less than those taken by glasso. Especially when the p-value tolerance is 0.05, the total computation time of glasso is more than 20 times that of MARS. Besides, Figure 3.3 also shows the estimated graphs obtained by MARS when the p-value tolerances are 0.1 and 1, but the figure for the latter one only plots the connections among the most significant 5418 genes.

3.2 Algorithms for addressing the subproblems in the adaptive sieving strategy

In this section, we explore various algorithms that address the subproblem within the adaptive sieving strategy, with a particular emphasis on problems of the form

Table 3.2: Test results of MARS and glasso on the breast cancer data sets with different p-value tolerances.

p-value tolerance	No. of genes	time (mins) including cross-validation		No. of λ
		MARS	glasso	
0.005	1228	20.26	71.51	63
0.01	1646	23.06	159.79	60
0.05	3640	58.32	1257.81	28
0.1	5418	150.54	–	54
1	22283	553.35	–	29

The symbol "–" indicates out of memory.

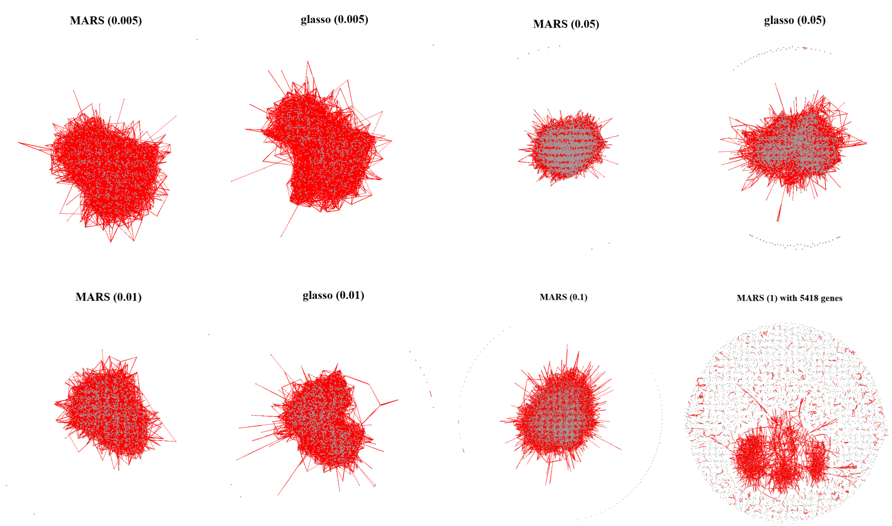


Figure 3.3: The estimated graphs for the breast cancer data set chosen by five-fold cross-validation with using MARS and glasso under different p-value tolerances.

presented in ($\mathrm{P_{LS}}(\lambda)$) with different penalty function $p(\cdot)$. Recall that

$$\min_{x\in\mathbb{R}^n} \left\{\frac{1}{2}\|Ax-b\|^2+\lambda p(x)\right\}. \tag{$\mathrm{P_{LS}}(\lambda)$}$$

There are numerous algorithms that tackle this problem with different $p(\cdot)$ by focusing on the dual problem. Then, we will first provide a summary of these methods. It is important to note that the dual problem has been addressed previously because, in practical applications, the dimension n is typically much larger than m (the number of rows in A). Consequently, solving the dual problem is often more advantageous in terms of computational complexity. However, with the introduction of the adaptive sieving strategy, we may encounter situations where m is greater than n. In such cases, addressing the dual problem no longer offers significant advantages. Therefore, we propose a smoothing Newton method to directly solve the primal problem, particularly when $p(\cdot) = \|\cdot\|_1$, while deferring the consideration of other cases for future research.

3.2.1 Dual based approaches

In this subsection, we will summarize several dual based approaches for addressing ($\mathrm{P_{LS}}(\lambda)$) with different penalty functions $p(\cdot)$.

Let $\lambda > 0$. The dual of ($\mathrm{P_{LS}}(\lambda)$) is

$$\max_{y\in\mathbb{R}^m, u\in\mathbb{R}^n} \left\{ -\frac{1}{2}\|y\|^2 + \langle b, y\rangle - \lambda p^*(u) \mid A^T y - \lambda u = 0 \right\}. \tag{$\mathrm{D_{LS}}(\lambda)$}$$

The semismooth Newton augmented Lagrangian method (SSNAL, Li et al. (2018b)) has demonstrated exceptional performance in solving the ℓ_1 penalized least squares problem (($\mathrm{P_{LS}}(\lambda)$) with $p(\cdot) = \|\cdot\|_1$) comparied to other algorithms, such as ADMM and the accelerated proximal gradient (APG) algorithm (Nesterov, 1983; Beck and Teboulle, 2009). It addresses the dual problem ($\mathrm{D_{LS}}(\lambda)$) using an inexact augmented Lagrangian method, with each subproblem being solved by a semismooth Newton method. Moreover, we know that SSNAL is of an asymptotic superlinear convergence rate for solving ($\mathrm{D_{LS}}(\lambda)$) when $p(\cdot) = \|\cdot\|_1$. Given $\rho > 0$, the augmented Lagrangian function associated to the dual problem is

$$L_\rho(y,u,x) := \frac{1}{2}\|y\|^2 - \langle b, y\rangle + \lambda p^*(u) + \frac{\rho}{2}\|A^T y - \lambda\|^2, \quad \forall\,(y,u,x) \in \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^n.$$

Then, in the k-th iteration of the inexact augmented Lagrangian method, (y^{k+1}, u^{k+1}) is updated by approximately solving $\min_{y,u} L_{\rho^k}(y,u,x^k)$ with the semismooth Newton method (Zhao et al., 2010), while also exploring second-order sparsity to enhance efficiency. Then, when $p(x) = \sum_{i=1}^{n} \gamma_i |x|_{(i)}$, $x \in \mathbb{R}^n$ with given parameters $\gamma_1 \geq \gamma_2 \geq \cdots \geq \gamma_n \geq 0$ with $\gamma_1 > 0$, where $|x|_{(1)} \geq |x|_{(2)} \geq \cdots \geq |x|_{(n)}$, Luo et al. (2019) introduced a semismooth Newton-based augmented Lagrangian method (Newt-ALM) to solve $(\mathrm{D_{LS}}(\lambda))$. When the penalty function is the non-overlapping group lasso regularization, i.e., $p(x) = \sum_{l=1}^{g} w_l \|x_{G_l}\|$, where for any $l = 1, 2, \cdots, g$, $w_l > 0$ and $G_l \subseteq \{1, 2, \cdots, n\}$ is the index set that includes all the features in the l-th group, Zhang et al. (2020) proposed a Hessian-based algorithm that implements a superlinearly convergent inexact semismooth Newton method.

The algorithms mentioned above can be employed to solve the subproblems in the adaptive sieving strategy. However, since the adaptive sieving strategy constructs reduced problems that may have a dimensionality smaller than m, the dual based algorithm may not be the optimal choice in this context. To enhance efficiency further, we will introduce a smoothing Newton method for directly solving the primal problem in the following section. This method will be integrated with a dual based algorithm to address the subproblems within the adaptive sieving strategy.

3.2.2 A smoothing Newton method for the primal problem

In this subsection, we will introduce a smoothing Newton method to solve the unconstrained optimization problem $(\mathrm{P_{LS}}(\lambda))$. Additionally, we will provide a detailed description of this algorithm specifically for the case when $p(\cdot) = \|\cdot\|_1$.

Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be a locally Lipschitz continuous function. Consider the equation

$$F(x) = 0, \quad x \in \mathbb{R}^n. \tag{3.11}$$

Let $G : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ be a locally Lipschitz continuous function such that

$$G(\epsilon, \tilde{x}) \to F(x) \quad \text{as} \quad (\epsilon, \tilde{x}) \to (0, x), \tag{3.12}$$

and the function $G(\cdot,\cdot)$ is continuously differentiable around $(\epsilon, x) \in \mathbb{R} \times \mathbb{R}^n$ except when $\epsilon = 0$. Denote $E : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R} \times \mathbb{R}^n$ as

$$E(\epsilon, x) := \begin{pmatrix} \epsilon \\ G(\epsilon, x) \end{pmatrix}, \quad \forall\, (\epsilon, x) \in \mathbb{R} \times \mathbb{R}^n. \tag{3.13}$$

Then, we solve (3.11) by solving

$$E(\epsilon, x) = 0. \tag{3.14}$$

Define the merit function $\psi : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}_+$ by

$$\phi(\epsilon, x) := \|E(\epsilon, x)\|^2, \quad (\epsilon, x) \in \mathbb{R} \times \mathbb{R}^n. \tag{3.15}$$

Given a scalar $r \in (0, 1)$. Denote

$$\zeta(\epsilon, x) := r \min\{1, \phi(\epsilon, x)\}, \quad (\epsilon, x) \in \mathbb{R} \times \mathbb{R}^n. \tag{3.16}$$

Then, the detailed iterations of the inexact smoothing Newton method are given in Algorithm 5.

The convergence properties of the inexact smoothing Newton method are presented below, and more information can be found in (Gao and Sun, 2009).

Theorem 3.1. *Assume that, for any $(\epsilon, x) \in \mathbb{R}_{++} \times \mathbb{R}^n$, $E'(\epsilon, x)$ is nonsingular. Then Algorithm 5 is well defined and generates an infinite sequence $\{(\epsilon^k, x^k)\}$ with $(\epsilon^k, x^k) \in \mathcal{N} := \{(\epsilon, x) \mid \epsilon \geq \zeta(\epsilon, x)\hat{\epsilon}\}$ such that any accumulation point $(\bar{\epsilon}, \bar{x})$ of $\{(\epsilon^k, x^k)\}$ is a solution of $E(\epsilon, x) = 0$.*

Theorem 3.2. *Assume that, for any $(\epsilon, x) \in \mathbb{R}_{++} \times \mathbb{R}^n$, $E'(\epsilon, x)$ is nonsingular, and $(\bar{\epsilon}, \bar{x})$ is an accumulation point $(\bar{\epsilon}, \bar{x})$ generated by Algorithm 5. Further assume that $E(\cdot,\cdot)$ is semismooth at $(\bar{\epsilon}, \bar{x})$ and all $V \in \partial_B E(\bar{\epsilon}, \bar{x})$ are nonsingular. Then the sequence $\{(\epsilon^k, x^k)\}$ converges to $(\bar{\epsilon}, \bar{x})$ superlinearly, that is,*

$$\|(\epsilon^{k+1} - \bar{\epsilon}, x^{k+1} - \bar{x})\| = o(\|(\epsilon^k - \bar{\epsilon}, x^k - \bar{x})\|).$$

Algorithm 5 The inexact smoothing Newton method for solving (3.14)

Input: Let $\hat{\epsilon} \in (0, \infty)$ and $\eta \in (0, 1)$ be such that

$$\delta := \sqrt{2} \max\{r\hat{\epsilon}, \eta\} < 1.$$

Choose constants $\rho \in (0, 1)$, $\sigma \in (0, 1/2)$, $\tau \in (0, 1)$, and $\hat{\tau} \in [1, \infty)$. Set $\epsilon^0 := \hat{\epsilon}$. Let $x^0 \in \mathbb{R}^m$ be an arbitrary point.

Initialization: Set $k = 0$. If $E(\epsilon^k, x^k) = 0$, then stop. Otherwise, calculate

$$\zeta_k := r \min\{1, \phi(\epsilon^k, y^k)\}, \quad \eta_k := \min\{\tau, \hat{\tau}\|E(\epsilon^k, y^k)\|\}.$$

while A termination criterion is not met **do**

Step 1. Solve

$$E(\epsilon^k, x^k) + E'(\epsilon^k, x^k) \begin{bmatrix} \Delta\epsilon^k \\ \Delta x^k \end{bmatrix} = \begin{bmatrix} \zeta_k \hat{\epsilon} \\ 0 \end{bmatrix} \tag{3.17}$$

approximately such that

$$\|R_k\| \le \min\{\eta_k \|G(\epsilon^k, x^k) + G'_\epsilon(\epsilon^k, x^k)\Delta\epsilon^k\|_2, \eta\|E(\epsilon^k, x^k)\|\}, \tag{3.18}$$

where $\Delta\epsilon^k := -\epsilon^k + \zeta_k \hat{\epsilon}$ and

$$R_k := G(\epsilon^k, x^k) + G'(\epsilon^k, x^k) \begin{bmatrix} \Delta\epsilon^k \\ \Delta x^k \end{bmatrix}.$$

Step 2. Let l_k be the smallest nonnegative integer l such that

$$\phi(\epsilon^k + \rho^l \Delta\epsilon^k, x^k + \rho^l \Delta y^k) \le [1 - 2\sigma(1 - \delta)\rho^l]\phi(\epsilon^k, x^k). \tag{3.19}$$

Define $(\epsilon^{k+1}, x^{k+1}) := (\epsilon^k + \rho^{l_k}\Delta\epsilon^k, x^k + \rho^{l_k}\Delta x x^k)$.

Step 3. k = k+1;

end while

Output: (ϵ^k, x^k).

Furthermore, if $E(\cdot, \cdot)$ is strongly semismooth at $(\bar{\epsilon}, \bar{x})$, then the sequence (ϵ^k, x^k) converges to $(\bar{\epsilon}, \bar{x})$ quadratically, that is,

$$\|(\epsilon^{k+1} - \bar{\epsilon}, x^{k+1} - \bar{x})\| = O(\|(\epsilon^k - \bar{\epsilon}, x^k - \bar{x})\|^2).$$

In the remainder of this section, we will specifically consider the ℓ_1-penalized least squares problem

$$\min_{x\in\mathbb{R}^n} \left\{\frac{1}{2}\|Ax-b\|^2+\lambda\|x\|_1\right\}, \tag{3.20}$$

and provide further details on the smoothing Newton method used to solve it. We can see that the solution set to (3.20) is nonempty and bounded, since it is level bounded (Rockafellar and Wets, 2009, Theorem 1.9). Besides, $x\in\mathbb{R}^n$ is an optimal solution to (3.20) if and only if

$$F(x)=x-\mathrm{Prox}_{\lambda\|\cdot\|_1}(x-A^TAx+A^Tb)=0. \tag{3.21}$$

Since $\mathrm{Prox}_{\lambda\|\cdot\|_1}(\cdot)$ is globally Lipschitz continuous, it follows that $F(\cdot)$ is also globally Lipschitz continuous. Consequently, we can solve (3.21) using the inexact smoothing Newton method previously introduced.

Note that,

$$\mathrm{Prox}_{\lambda\|\cdot\|_1}(x)=\max\{x-\lambda,0\}-\max\{-x-\lambda,0\},\quad \forall\, x\in\mathbb{R}^n.$$

Additionally, there is a well known smoothing function called the Huber function, which approximates $\max\{(w,0)\}$, $w\in\mathbb{R}$ by

$$\theta(\epsilon,w):=\begin{cases}0, & \text{if } w\le -|\epsilon|/2,\\ w, & \text{if } w\ge |\epsilon|/2,\\ \frac{(w+|\epsilon|/2)^2}{2|\epsilon|}, & \text{otherwise,}\end{cases}\qquad \forall\,(\epsilon,w)\in\mathbb{R}\times\mathbb{R}.$$

Therefore, we can use the function $\Theta(\epsilon,u):=\begin{pmatrix}\theta(\epsilon,u_1)\\ \vdots\\ \theta(\epsilon,u_n)\end{pmatrix}$, $\forall\, u\in\mathbb{R}^n$ to approximate $F(\cdot)$ in the following manner: for any $(\epsilon,x)\in\mathbb{R}\times\mathbb{R}^n$,

$$\widetilde{F}(\epsilon,x)=x-\Theta(\epsilon,f(x)-\lambda)+\Theta(\epsilon,-f(x)-\lambda), \tag{3.22}$$

where $f(x):=x-A^TAx+A^Tb$. Note that, when $\epsilon\neq 0$, the partial derivatives of $G(\cdot,\cdot)$ respect to ϵ and u are

$$\Theta'_\epsilon(\epsilon,u)=\mathrm{diag}\left(\theta'_\epsilon(\epsilon,u_1),\cdots,\theta'_\epsilon(\epsilon,u_n)\right)\ \text{ and }\ \Theta'_u(\epsilon,u)=\mathrm{diag}\left(\theta'_w(\epsilon,u_1),\cdots,\theta'_w(\epsilon,u_n)\right),$$

where

$$\theta'_\epsilon(\epsilon, w) = \begin{cases} 0, & \text{if } |w| \geq \frac{|\epsilon|}{2}, \\ \frac{1}{8} - \frac{w^2}{2\epsilon^2}, & \text{if } |w| < \frac{|\epsilon|}{2} \,\&\, \epsilon > 0, \\ \frac{w^2}{2\epsilon^2} - \frac{1}{8}, & \text{if } |w| < \frac{|\epsilon|}{2} \,\&\, \epsilon < 0, \end{cases} \text{ and } \theta'_w(\epsilon, w) = \begin{cases} 0, & \text{if } w \leq -\frac{|\epsilon|}{2}, \\ 1, & \text{if } w \geq \frac{|\epsilon|}{2}, \\ \frac{w}{|\epsilon|} + \frac{1}{2}, & \text{otherwise.} \end{cases}$$

Before discussing the properties of $\widetilde{F}(\cdot, \cdot)$, we present the following definition for a P_0 (P) matrix.

Definition 3.1 (P_0 (P) matrix)**.** *A matrix $M \in \mathbb{R}^{n\times n}$ is referred to as a P_0 (P) matrix if and only if all its principle minors are nonnegative (or positive).*

Proposition 3.3. *Let $\widetilde{F}(\cdot, \cdot)$ be defined by* (3.22)*. Then the following properties hold.*

(a) $\widetilde{F}(\cdot, \cdot)$ is globally Lipschitz continuous on $\mathbb{R} \times \mathbb{R}^n$.

(b) $\widetilde{F}(\cdot, \cdot)$ continuously differentiable around any $(\epsilon, x) \in \mathbb{R} \times \mathbb{R}^n$, except when $\epsilon = 0$. Moreover, for any fixed $\epsilon \neq 0$, $\widetilde{F}'_x(\epsilon, x)$, $x \in \mathbb{R}$ is a P_0 matrix.

(c) $\widetilde{F}(\cdot, \cdot)$ is strongly semismooth at $(0, x)$, $x \in \mathbb{R}^n$.

Proof. (a) Given that the Huber function is globally Lipschitz continuous, it follows that $\widetilde{F}(\cdot, \cdot)$ is also globally Lipschitz continuous.

(b) It is clear from the definition of the Huber function that $\widetilde{F}(\cdot, \cdot)$ continuously differentiable around any $(\epsilon, x) \in \mathbb{R} \times \mathbb{R}^n$, except when $\epsilon = 0$.

Fix $\epsilon \neq 0$. Then the partial derivative of $\widetilde{F}(\cdot, \cdot)$ respect to x can be computed by

$$\widetilde{F}'_x(\epsilon, x) = I - U\left(I - A^T A\right) = I - U + UA^T A, \tag{3.23}$$

where $U = \Theta'_u(\epsilon, f(x) - \lambda) + \Theta'_u(\epsilon, -f(x) - \lambda)$. The positive semidefiniteness of $A^T A$ implies that it is also a P_0 matrix. Therefore, by (Cottle et al., 1992, Theorem 3.4.2), we know that, for any $h \neq 0$, there is an index i, such that $h_i \neq 0$ and

$$\langle h_i, (A^T Ah)_i \rangle \geq 0.$$

Besides, it can be observed that for all $i \in \{1, \cdots, n\}$, the i-th diagonal components u_i of U lies within the interval $[0, 1]$. Consequently, we obtain that

$$\langle h_i, (\widetilde{F}'_x(\epsilon, x)h)_i \rangle \geq 0,$$

which implies that $\widetilde{F}'_x(\epsilon, x)$ is a P_0 matrix.

(c) Since piecewise affine functions and twice continuously differentiable functions are strongly semismooth, along with Proposition 2.5, we can conclude that $\widetilde{F}(\cdot, \cdot)$ is strongly semismooth at $(0, x)$.

□

Let $\kappa \in (0, +\infty)$ be a given scalar. Define

$$G(\epsilon, x) := \widetilde{F}(\epsilon, x) + \kappa|\epsilon|x, \quad \forall\, (\epsilon, x) \in \mathbb{R} \times \mathbb{R}^n. \tag{3.24}$$

The reason why we define $G(\cdot, \cdot)$ by adding a term to $\widetilde{F}(\cdot, \cdot)$ is to ensure that, for any $\epsilon \neq 0$, $G'_x(\epsilon, \cdot)$ is a P matrix (by Proposition 3.3 and (Cottle et al., 1992, Theorem 3.4.2)). Define

$$E(\epsilon, x) := \begin{pmatrix} \epsilon \\ G(\epsilon, x) \end{pmatrix} = \begin{pmatrix} \epsilon \\ \widetilde{F}(\epsilon, x) + \kappa|\epsilon|x \end{pmatrix}, \quad \forall (\epsilon, x) \in \mathbb{R} \times \mathbb{R}^n. \tag{3.25}$$

We can then apply Algorithm 5 to solve the equation $E(\epsilon, x) = 0$, and the convergence properties are detailed in the following theorem.

Theorem 3.3. *Algorithm 5 is well defined and generates an infinite sequence* $\{(\epsilon^k, x^k)\}$ *with* $(\epsilon^k, x^k) \in \mathcal{N} := \{(\epsilon, x) \mid \epsilon \geq \zeta(\epsilon, x)\hat{\epsilon}\}$ *such that any accumulation point* $(\bar{\epsilon}, \bar{x})$ *of* $\{(\epsilon^k, x^k)\}$ *is a solution of* $E(\epsilon, x) = 0$ *and* $\lim_{k\to\infty} \phi(\epsilon^k, x^k) = 0$. *Moreover,* $\{(\epsilon^k, x^k)\}$ *is bounded.*

Proof. From Proposition 3.3 (b) and the definition of $G(\cdot,\cdot)$, we know that, for any $(\epsilon, x) \in \mathbb{R}_{++} \times \mathbb{R}^n$, $G'_x(\epsilon, x)$ is a P matrix, which implies that $E'(\epsilon, x)$ is also a P matrix. Then from Theorem 3.1, we have that Algorithm 5 is well defined and generates an infinite sequence $\{(\epsilon^k, x^k)\}$ with $(\epsilon^k, x^k) \in \mathcal{N}$ such that any accumulation point $(\bar{\epsilon}, \bar{x})$ of $\{(\epsilon^k, x^k)\}$ is a solution of $E(\epsilon, x) = 0$.

From the design of Algorithm 5, we know that $\{\phi(\epsilon^k, x^k)\}$ is a decreasing sequence, thus $\lim_{k\to\infty} \phi(\epsilon^k, x^k)$ exists. Denote $\bar{\phi} := \lim_{k\to\infty} \phi(\epsilon^k, x^k) \geq 0$. Suppose that $\bar{\phi} > 0$. Then there exists $\tilde{\epsilon} > 0$ such that, for any $k \geq 0$, we have $\epsilon^k \geq \tilde{\epsilon}$. Using a similar argument as in the proof of (Gao and Sun, 2009, Theorem 4.1), we have that for any $x \geq 0$,

$$\{x \in \mathbb{R}^n \mid \|G(\epsilon, x)\| \leq v,\, \epsilon \in [\tilde{\epsilon}, \hat{\epsilon}]\}$$

is bounded. This implies that the set $\{(\epsilon^k, x^k)\}$ is bounded, and therefore it has at least one accumulation point, which is a solution to the equation $E(\epsilon, x) = 0$. However, this contradicts the assumption that $\bar{\phi} > 0$. Therefore, we conclude that $\bar{\phi} = 0$.

Since the objective function in (3.20) is level bounded, the solution set is nonempty and compact (Rockafellar and Wets, 2009, Theorem 1.9). It follows from (Rockafellar, 1970, Corollary 31.2.1) that the solution set to the dual of (3.20) is nonempty, and that the optimal solutions to the primal and dual are equivalent and finite. Moreover, the strong convexity of the dual implies that the solution set of the dual is bounded. Therefore, the solution set of $E(\epsilon, x) = 0$ is also nonempty and compact. It then follows from (Ravindran and Gowda, 2001, Theorem 2.5) that $\{(\epsilon^k, x^k)\}$ is bounded. □

Let $(\bar{\epsilon}, \bar{x})$ be an accumulation point of $\{(\epsilon^k, x^k)\}$ generated by Algorithm 5. From the above theorem we know that $\bar{\epsilon} = 0$ and $F(\bar{x}) = 0$, which indicates that $\bar{x}$ is an optimal solution to (3.20). Let $\bar{y} = b - A\bar{x}$ and $\bar{u} = A^T\bar{y}/\lambda$. It follows from

(Rockafellar, 1970, Corollary 31.2.1) and the strong convexity of the dual problem ($\mathrm{D_{LS}}(\lambda)$) that $(\bar{y}, \bar{u})$ is the unique optimal solution to ($\mathrm{D_{LS}}(\lambda)$).

For deriving the quadratic convergence of Algorithm 5, we need to consider the linear independence constraint qualification (LICQ) to the dual problem ($\mathrm{D_{LS}}(\lambda)$). In the nonlinear programming, the LICQ requires that the gradients of the active inequality constraints and the gradients of the equality constraints are linearly independent at the relevant point. After some straightforward calculations, we find that the constraint qualification holds at $\bar{u} \in \mathbb{R}^n$ if

$$A_{J(\bar{u})}^T A_{J(\bar{u})} \succ 0, \qquad \text{where } J(\bar{u}) := \{j \in \{1, \cdots, n\} \mid |\bar{u}_j| = 1\}. \tag{3.26}$$

Proposition 3.4. *Let $\widetilde{F}(\cdot,\cdot)$ be defined by* (3.22)*. Assume that the constraint nondegeneracy* (3.26) *holds at $\bar{u}$. Then for any $W \in \partial \widetilde{F}(0, \bar{x})$, there exists $i \in \{1, \cdots, n\}$ such that for any nonzero $h \in \mathbb{R}^n$,*

$$h_i \left(W(0, h)\right)_i > 0. \tag{3.27}$$

Proof. Suppose that there exists $h \neq 0$ such that

$$\max_i h_i \left(W(0, h)\right)_i \leq 0. \tag{3.28}$$

We have that

$$W(0, h) = h - D^1(0, h - A^T A h) - D^2(0, h - A^T A h) = h - D(0, h - A^T A h),$$

where $D^1 \in \partial_B \Theta(0, f(\bar{x}) - \lambda)$, $D^2 \in \partial_B \Theta(0, -f(\bar{x}) - \lambda)$, and $D = D^1 + D^2$. By some simple calculations, we know that there is a nonnegative vector $d \in \mathbb{R}^n$ with

$$d_i = \begin{cases} 1, & \text{if } i \in J(\bar{u}), \\ 0, & \text{otherwise,} \end{cases}$$

such that

$$\big(D(0, h - A^T A h)\big)_i = d_i h_i - d_i (A^T A h)_i, \quad i = 1, \cdots, n.$$

Then, from (3.28), we have that

$$\begin{cases} h_i(A^T Ah)_i \leq 0, & \text{if } i \in J(\bar{u}), \\ h_i = 0, & \text{otherwise}, \end{cases} \tag{3.29}$$

which implies that $\langle h_{J(\bar{u})}, A^T_{J(\bar{u})} A_{J(\bar{u})} \langle h_{J(\bar{u})} \rangle = 0$. This contradicts the assumption in (3.26), thereby concluding the proof. □

Theorem 3.4. *Let* $(\bar{\epsilon}, \bar{x})$ *be an accumulation point of* $\{(\epsilon^k, y^k)\}$ *generated by Algorithm 5. Assume that the constraint nondegeneracy* (3.26) *holds at* $\bar{u}$. *Then the sequence* $\{(\epsilon^k, x^k)\}$ *converges to* $(\bar{\epsilon}, \bar{x})$ *quadratically, i.e.,*

$$\|(\epsilon^{k+1} - \bar{\epsilon}, x^{k+1} - \bar{x})\| = O(\|(\epsilon^k - \bar{\epsilon}, x^k - \bar{x})\|^2).$$

Proof. It follows from Proposition 3.3 (c) and the fact that the modulus function $|\cdot|$ is strongly semismooth on $\mathbb{R}$ that $E(\cdot, \cdot)$ is strongly semismooth at $(\bar{\epsilon}, \bar{x})$.

Let $V \in \partial_B E(\bar{\epsilon}, \bar{x})$ be arbitrarily chosen. We can then derive from Proposition 3.4 and (3.25) that, for any $d \in \mathbb{R}^{n+1}$, there exists $i \in \{1, \cdots, n, n+1\}$ such that

$$d_i(Vd)_i > 0.$$

Therefore, according to (Cottle et al., 1992, Theorem 3.3.4), we conclude that V is a P matrix, which implies that it is nonsingular. By applying the result from Theorem 3.2, we can directly obtain the conclusion of this theorem. □

3.3 A warm-started path-following adaptive sieving technique

Modern data driven applications have posed great challenges for solving the corresponding large scale optimization problems under restrictive efficiency and memory constraints. Inspired by the recently developed level set method (Li et al., 2018b), we design an efficient warm-started path-following algorithm in this section to address the computational challenges for solving large scale sparse optimization problems (3.1). In particular, we obtain a solution to the original problem by sequentially solving a set of regularized problems, where the adaptive sieving and warm-start strategies are naturally incorporated to efficiently solve these problems. Consequently, we can take advantage of the solution sparsity and only solve a sequence of reduced problems with much smaller problem dimensions, which are computationally and memory efficient.

3.3.1 A warm-started path-following adaptive sieving technique for large scale sparse optimization problems

For any given $\lambda > 0$, instead of directly solving the original problem, we adopt an iterative approach by solving a sequence of problems with penalty parameters $\lambda_1 > \lambda_2 > \cdots > \lambda$. Each problem is warm-started and significantly smaller in dimensionality compared to the original problem, achieved by implementing the adaptive sieving strategy. Although there could be several possible ways to determine the penalty parameter sequence, we present a straightforward implementation based on numerical observations. While this approach may not be the optimal solution, it offers a practical alternative. Our experiments on (3.20) revealed that the relationship between solution sparsity and the penalty parameter is not linear. Rather, it generally follows a functional form that resembles inverse proportionality. We then drew inspiration from the bisection method to construct the sequence of penalty parameters within the interval $[\lambda, \lambda_0]$, where $\lambda_0 > \lambda$. At the i-th iteration, we set the current point as $(\lambda_{i-1} + \lambda)/2$. Then the penalty parameter will gradually approach λ until the absolute difference between the current point and λ is less than a specified tolerance $\tau > 0$. At that point, we set the final penalty parameter to be λ itself. Now, we provide the detailed steps of our algorithm in Algorithm 6.

We will demonstrate the numerical performance of Algorithm 6 on two widely

Algorithm 6 A warm-started path-following algorithm for solving ($\mathrm{P_{LS}}(\lambda)$)

Input: a given penalty parameter $\lambda > 0$, an initial maximum penalty parameter $\lambda_0 > \lambda$, an initial point $x^0 \in \mathbb{R}^n$, a given tolerance $\tau > 0$ and a given positive integer $k_{\max}$. Let $i = 1$.
Output: an approximate solution x^* to the problem ($\mathrm{P_{LS}}(\lambda)$).
while $\lambda_i > \lambda$ **do**
  if $|\lambda_{i-1} - \lambda| > \tau$ **then**

$$\lambda_i = \frac{\lambda + \lambda_{i-1}}{2};$$

  else
    Set $\lambda_i = \lambda$.
  end if
  3.1: Create I^{i-1} as

$$I^{i-1} = \Big\{ j \in \{1, \cdots, n\} \ \mid \ ((x^{i-1}))_j \neq 0 \Big\}. \tag{3.30}$$

  If $|I^{i-1}| \leq k_{\max}$, go to step 3.2; otherwise, define I^{i-1} as the union of all indices of the $k_{\max}$ largest absolute values in x^{i-1}.
  3.2: Call Algorithm 4 with the penalty parameter λ_i, initializing the starting point as x^{i-1} and the initial set as I^{i-1}. Then set x^i as the corresponding output.
  3.3: Set $i \leftarrow i + 1$.
end while
return: Set $x^* = x^i$.

used models. The first model is the ℓ_1 penalized least squares problem (3.20). The other model is the ℓ_1 penalized logistic regression problem, which is a combination of the logistic regression (Cox, 1958) and the ℓ_1 penalty to perform variable selection. It can be expressed as follows:

$$\min_{x \in \mathbb{R}^n} \left\{ \sum_{i=1}^{m} \log\Big(1 + \exp\big(- b_i(A_i^T x)\big)\Big) + \lambda \|x\|_1 \right\}, \tag{3.31}$$

where $A_i \in \mathbb{R}^n$, $i \in 1, \cdots, m$ (which is the i-th column of the feature matrix A) and the label vector $b \in \{-1, 1\}^m$ are given. The ℓ_1 penalized logistic regression model is extensively employed for feature selection (Ng, 2004; Wainwright et al., 2006; Ravikumar et al., 2010), predictive classification (Ryali et al., 2010; Liang et al., 2013), and others.

We will use (3.20) as an example to discuss a practical way to choose the initial penalty parameter λ_1. We know that for all $\lambda \geq \lambda_{\max} > 0$, the origin is an optimal solution to (3.20) (Li et al., 2024, Proposition 3.1 (i)). Here, $\lambda_{\max}$ is defined as the value of the gauge function over $\partial p(0)$ on $A^T b$, given by

$$\lambda_{\max} = \gamma(A^T b \mid \partial p(0)). \tag{3.32}$$

Therefore, for any given $\lambda \in (0, \lambda_{\max})$, we can set $\lambda_0 = \lambda_{\max}$ and then obtain

$$\lambda_1 = \frac{\lambda + \lambda_{\max}}{2}.$$

It is important to note that the value of λ_1 provided above may not be the optimal choice, as there are scenarios where a smaller initial penalty parameter could be more beneficial. For example, if the solution obtained with the specified λ_1 is overly sparse, selecting a smaller initial penalty parameter could enhance computational efficiency by minimizing unnecessary iterations. For model (3.31), the only difference compared to the above case is that $\lambda_{\max} = \gamma(-A^T b/2 \mid \partial p(0))$.

Then, we will perform extensive numerical experiments to demonstrate the high efficiency of our algorithm in solving two popular models (3.20) and (3.31) on real applications compared to some state-of-the-art algorithms. Since our algorithm is a warm-started path-following algorithm implementing the adaptive sieving strategy, we refer to it as WarmPAS. To evaluate the efficiency of WarmPAS, we compare it to two other algorithms:

1. SSNAL, a standalone algorithm;

2. AS, which employs the adaptive sieving strategy with each subproblem solved by SSNAL (and the smoothing Newton algorithm (Section 3.2.2) for (3.20) only).

For any approximate solution $\hat{x}$, we measure the accuracy of the solution using the following relative KKT residual:

$$\eta := \frac{\|\hat{x} - \text{Prox}_{\lambda p}(\hat{x} - \nabla f(\hat{x}))\|}{1 + \|\hat{x}\| + \|\nabla f(\hat{x})\|}.$$

Let nnz(x) denotes the number of nonzeros in the solution vector $x \in \mathbb{R}^n$. It is defined as $\text{nnz}(x) := \min\left\{k \mid \sum_{i=1}^{k} |x|_{(i)} \geq 0.999\|x\|_1\right\}$. The numerical results presented in this section are generated using MATLAB R2023a on a Windows workstation equipped with the following specifications: a 12-core Intel(R) Core(TM) i7-12700 (2.10GHz) processor and 64 GB of RAM.

3.3.2 Numerical experiments on the ℓ_1 penalized least squares problem

In this subsection, we will evaluate the superior performance of our algorithm in solving the large scale ℓ_1 penalized least squares problem (3.20) on several UCI datasets, comparing it with AS and SSNAL. The UCI datasets used in this subsection are sourced from the UCI Machine Learning Repository, as mentioned in (Li et al., 2018a,b).

Table 3.3: The computation time of WarmPAS, AS, and SSNAL in solving the large scale ℓ_1 penalized linear regression problem on some UCI datasets, where the penalty parameter λ is set to $c\|A^T b\|$ and the stopping tolerance is set to 10^{-6}.

Name	m	n	sparsity(A)	c	nnz(x)	Time (s)		
						WarmPAS	AS	SSNAL
E2006.train	16087	150360	0.0083	1e-6	25	0.515	1.844	6.392
				2e-7	520	1.312	3.563	20.487
log1p.E2006.train	16087	4272227	0.0014	5e-4	39	3.111	3.784	57.620
				1e-4	599	5.189	6.769	135.249
E2006.test	3308	150358	0.0092	1e-6	51	0.156	0.187	1.766
				2e-7	692	4.767	8.454	11.565
log1p.E2006.test	3308	4272226	0.0016	5e-4	49	1.406	6.178	24.194
				1e-4	1081	2.656	29.559	62.396
pyrim5	74	201376	0.5405	5e-5	78	0.500	0.719	1.516
				5e-6	96	0.610	0.907	1.797
triazines4	186	635376	0.6569	5e-3	475	1.813	1.563	16.886
				1e-4	261	5.160	6.373	108.964

The experiments were conducted with a stopping tolerance of 10^{-6} for all algorithms. Additionally, the penalty parameter λ was determined as $c\|A^T b\|$, where $c > 0$ was chosen to ensure a reasonable number of nonzero solutions. Table 3.3 displays the computational performance of WarmPAS, AS, and SSNAL when applied to solve the large-scale ℓ_1 penalized linear regression problem (3.20) on the UCI datasets, where all the algorithms achieved a solution within the specified stopping tolerance. The results clearly demonstrate that WarmPAS significantly outperforms SSNAL across all tests, achieving a remarkable speed improvement of up to 26 times. In comparison to AS, WarmPAS outperforms it in all but one test, where it is slightly slower by approximately 1.16 times. In all other tests, however, WarmPAS shows speed enhancements of up to 11 times faster than AS.

3.3.3 Numerical experiments on the ℓ_1 penalized logistic regression problems

In this subsection, we will demonstrate the superior performance of our algorithm in solving the ℓ_1 penalized logistic regression problem (3.31) on several large scale real datasets. All the data sets used in here are sourced from the following website: https://www.csie.ntu.edu.tw/ cjlin/libsvmtools/datasets/.

Similar to the previous subsection, we will compare our algorithm, WarmPAS, with AS and SSNAL, and set the penalty parameter λ as $c\|A^T b\|$, where $c > 0$ is chosen to maintain a reasonable number of nonzero solutions. The stopping tolerance for all tests is set to 10^{-6}. The numerical results of the experiments are provided in Table 3.4 and all the algorithms achieved a solution within the specified stopping tolerance. It is important to highlight that our algorithm demonstrates superior performance, particularly in terms of efficiency compared to SSNAL, achieving a time speed-up of over 50 times in the best-case scenario. Even with the adaptive sieving strategy employed by AS to exploit the sparsity of the solution, our algorithm still outperforms it, achieving a speed-up of up to 5 times.

Table 3.4: The computation time of WarmPAS, AS and SSNAL for solving the large scale ℓ_1 penalized logistic regression problem on some real data sets, where the penalty parameter λ is set to $c\|A^T b\|$ and the stopping tolerance is set to 10^{-6}.

Name	m	n	sparsity(A)	c	nnz(x)	Time (s) WarmPAS	AS	SSNAL
news20.binary	19996	1355191	3.36e-4	8e-3	491	2.453	4.767	32.474
				5e-3	691	4.336	6.470	41.805
rcv1_train.binary	20242	47236	1.57e-3	1e-2	275	0.750	1.969	2.814
				5e-3	450	1.594	3.736	4.876
real-sim	72309	20958	2.45e-3	1e-2	239	2.203	5.923	13.597
				5e-3	424	5.221	8.752	23.892
kddb-raw-libsvm.t	748401	1163024	7.74e-6	1e-3	37	5.048	10.473	85.082
				7e-4	50	8.797	22.590	120.696
kdda.t	510302	2014669	1.87e-5	5e-2	46	3.422	8.689	116.998
				2e-2	80	6.392	35.539	326.797
kddb.t	748401	2990384	9.81e-6	1e-2	53	6.126	19.100	178.954
				7e-3	80	10.657	28.885	299.886

Chapter 4

SMOP: A root finding based secant method for solving the sparse optimization problem with least squares constraints

In this chapter, we will develop an efficient sieving based **S**ecant **M**ethod for solving the sparse **O**ptimization **P**roblem (CP(ϱ)), called SMOP, by finding the root of the equation (E_φ). We begin this chapter by discussing the properties of the value function $\varphi(\cdot)$, which are essential for designing the secant method to find the root of (E_φ). In the subsequent section, we will study the HS-Jacobian (Han and Sun, 1997) of the value function $\varphi(\cdot)$. It is important to note that the HS-Jacobian introduced in (Han and Sun, 1997) provides a solid foundation for our research. With these preparations in place, we will then be ready to introduce the secant method for solving the main problem (CP(ϱ)) in the next two sections. In particular, under the assumption that $p(\cdot)$ is a polyhedral gauge function, we show that the secant method converges at least 3-step Q-quadratically for solving (E_φ), and if $\partial_{\mathrm{B}}\varphi(\lambda^*)$ is a singleton, the secant method converges superlinearly with Q-order at least $(1+\sqrt{5})/2$. Furthermore, for a general strongly semismooth function $\varphi(\cdot)$, if $\partial\varphi(\lambda^*)$ is a singleton and nondegenerate, the secant method converges superlinearly with R-order of at least $(1+\sqrt{5})/2$.

4.1 The properties of the value function $\varphi(\cdot)$

In this section, we explore several useful properties of the value function $\varphi(\cdot)$, particularly its (strong) semismoothness property when $p(\cdot)$ is a (polyhedral) gauge function. This preliminary analysis lays the groundwork for the subsequent design of the secant method to solve (CP(ϱ)).

Since $p(\cdot)$ is assumed to be a nonnegative positively homogeneous convex function such that $p(0) = 0$, i.e. a gauge function (Rockafellar, 1970, Section 15), we have that $0 \in \partial p(0)$.

Proposition 4.1. *Assume that $\lambda_\infty \in (0 + \infty)$. It holds that*

(a) for all $\lambda \geq \lambda_\infty$, $y(\lambda) = b$ and $0 \in \Omega(\lambda)$;

(b) the value function $\varphi(\cdot)$ is nondecreasing on $(0, +\infty)$ and for any $0 < \lambda_1 < \lambda_2 < +\infty$, $\varphi(\lambda_1) = \varphi(\lambda_2)$ implies $p(x(\lambda_1)) = p(x(\lambda_2))$, where for any $\lambda > 0$, $x(\lambda)$ is an optimal solution to $\mathrm{P}_{\mathrm{LS}}(\lambda)$.

Proof. (a) Since $0 \in \partial p(0)$, for all $\lambda > \lambda_\infty$, it holds that

$$A^T b/\lambda \in \partial p(0),$$

which implies that $0 \in \Omega(\lambda)$. Since $\lambda_\infty > 0$ and $\partial p(0)$ is closed, we know

$$A^T b/\lambda_\infty \in \partial p(0),$$

which implies that $0 \in \Omega(\lambda_\infty)$. Therefore, for all $\lambda \geq \lambda_\infty$, $0 \in \Omega(\lambda)$ and $y(\lambda) = b$.

(b) Let $0 < \lambda_1 < \lambda_2 < \infty$ be arbitrarily chosen. Let $x(\lambda_1) \in \Omega(\lambda_1)$ and $x(\lambda_2) \in \Omega(\lambda_2)$. Then, we have

$$\frac{1}{2}\|Ax(\lambda_1) - b\|^2 + \lambda_1 p(x(\lambda_1)) \leq \frac{1}{2}\|Ax(\lambda_2) - b\|^2 + \lambda_1 p(x(\lambda_2)), \quad (4.1)$$

$$\frac{1}{2}\|Ax(\lambda_2) - b\|^2 + \lambda_2 p(x(\lambda_2)) \leq \frac{1}{2}\|Ax(\lambda_1) - b\|^2 + \lambda_2 p(x(\lambda_1)), \quad (4.2)$$

which implies that

$$(\lambda_1 - \lambda_2)(p(x(\lambda_1)) - p(x(\lambda_2))) \leq 0. \tag{4.3}$$

Since $\lambda_1 - \lambda_2 < 0$, we know that $p(x(\lambda_1)) \geq p(x(\lambda_2))$. It follows from (4.1) that

$$\frac{1}{2}\|Ax(\lambda_1) - b\|^2 \leq \frac{1}{2}\|Ax(\lambda_2) - b\|^2 + \lambda_1(p(x(\lambda_2)) - p(x(\lambda_1))) \leq \frac{1}{2}\|Ax(\lambda_2) - b\|^2,$$

which implies that $p(x(\lambda_1)) = p(x(\lambda_2))$ if $\varphi(\lambda_1) = \varphi(\lambda_2)$. This completes the proof of the proposition. □

Due to Proposition 4.1, we can apply the bisection method to solve (E_φ) and for any $\epsilon > 0$ we can obtain a solution λ_ε satisfying $|\lambda_\varepsilon - \lambda^*| \leq \varepsilon$ in $O(\log(1/\varepsilon))$ iterations, where λ^* is a solution to (E_φ). We will design a more efficient secant method for solving (E_φ) later. To achieve this goal, we first study the (strong) semismoothness property of $\varphi(\cdot)$.

We focus on the case where $p(\cdot)$ is a gauge function, as this is a very common case. In most of the applications, $p(\cdot)$ is a norm function, which is automatically a gauge function. We will leave the study of the (strong) semismoothness of $\varphi(\cdot)$ for a general $p(\cdot)$ as future work.

Since $p(\cdot)$ is a gauge function, then $p^*(\cdot) = \delta(\cdot \,|\, \partial p(0))$ and the optimization problem $\mathrm{D_{LS}}(\lambda)$ is equivalent to

$$\max_{y \in \mathbb{R}^m} \left\{ -\frac{1}{2}\|y\|^2 + \langle b, y \rangle \mid \lambda^{-1} y \in Q \right\}, \tag{4.4}$$

where

$$Q := \{z \in \mathbb{R}^m \mid A^T z \in \partial p(0)\}. \tag{4.5}$$

Then by performing a variable substitution, we have the following useful observation about the solution mapping to (4.4).

Proposition 4.2. *Let $p(\cdot)$ be a gauge function, for any $0 < \lambda < +\infty$, the unique solution to* (4.4) *can be written as*

$$y(\lambda) = \lambda\,\Pi_Q(\lambda^{-1}b) = \Pi_{\lambda Q}(b). \tag{4.6}$$

The following proposition is useful in understanding the semismoothness of $y(\cdot)$ and $\varphi(\cdot)$ even if $p(\cdot)$ is non-polyhedral. Part (b) of the proposition is generalized from (Li et al., 2018b, Proposition 1 (iv)) ($p(\cdot)$ is assumed to be a polyhedral gauge function in (Li et al., 2018b)) and we provide a more explicit proof that does not rely on the piecewise linearity of the solution mapping $y(\cdot)$ in (4.6).

Proposition 4.3. *Let $p(\cdot)$ be a gauge function. It holds that*

(a) the functions $y(\cdot)$ and $\varphi(\cdot)$ are locally Lipschitz continuous on $(0, +\infty)$;

(b) if $0 < \lambda_\infty < +\infty$, $\varphi(\cdot)$ is strictly increasing on $(0, \lambda_\infty]$;

(c) if the set Q is tame, $\varphi(\cdot)$ is semismooth on $(0, +\infty)$;

(d) if Q is globally subanalytic, $\varphi(\cdot)$ is γ-order semismooth on $(0, +\infty)$ for some $\gamma > 0$.

Proof. For convenience, we denote $\tilde{y}(\lambda) := \Pi_Q(\lambda^{-1}b)$ for any $\lambda > 0$.

(a) Since $\Pi_Q(\cdot)$ is Lipschitz continuous with modulus 1, both $\tilde{y}(\cdot)$ and $y(\cdot)$ are locally Lipschitz continuous on $(0, +\infty)$. Therefore, $\varphi(\cdot) = \|y(\cdot)\|$ is locally Lipschitz continuous on $(0, +\infty)$.

(b) It follows from Proposition 4.1 that $\varphi(\cdot)$ is nondecreasing. We will now prove that $\varphi(\cdot)$ is strictly increasing on $(0, \lambda_\infty]$. We prove it by contradiction. Assume that there exist $0 < \lambda_1 < \lambda_2 \leq \lambda_\infty$ such that $\varphi(\lambda_1) = \varphi(\lambda_2)$. Let $x(\lambda_1) \in \Omega(\lambda_1)$ and $x(\lambda_2) \in \Omega(\lambda_2)$ be arbitrarily chosen. From Proposition 4.1 (ii), we know that

$$p(x(\lambda_1)) = p(x(\lambda_2)),$$

which implies that $x(\lambda_1) \in \Omega(\lambda_2)$ and $x(\lambda_2) \in \Omega(\lambda_1)$. Therefore, we get

$$y(\lambda_1) = b - Ax(\lambda_1) = b - Ax(\lambda_2) = y(\lambda_2).$$

Thus, by using the facts $0 \in \lambda_2 Q$, $y(\lambda_2) = \Pi_{\lambda_2 Q}(b)$, and $\lambda_1^{-1}\lambda_2 y(\lambda_2) = \lambda_1^{-1}\lambda_2 y(\lambda_1) \in \lambda_2 Q$, we obtain from the properties of the metric projector $\Pi_{\lambda_2 Q}(\cdot)$ that

$$\langle b - y(\lambda_2), 0 - y(\lambda_2)\rangle \leq 0, \quad \langle b - y(\lambda_2), \lambda_1^{-1}\lambda_2 y(\lambda_2) - y(\lambda_2)\rangle \leq 0.$$

Therefore,

$$\langle b - y(\lambda_1), y(\lambda_1)\rangle = \langle b - y(\lambda_2), y(\lambda_2)\rangle = 0.$$

Since $\lambda_1 < \lambda_\infty$ and $y(\lambda_1) = \Pi_{\lambda_1 Q}(b)$, we know that $y(\lambda_1) \neq b$. Hence,

$$\langle b - y(\lambda_1), \lambda_\infty^{-1}\lambda_1 b - y(\lambda_1)\rangle = \lambda_\infty^{-1}\lambda_1 \|b - y(\lambda_1)\|^2 > 0.$$

However, $\lambda_\infty^{-1}\lambda_1 b \in \lambda_1 Q$ and $y(\lambda_1) = \Pi_{\lambda_1 Q}(b)$ imply that

$$\langle b - y(\lambda_1), \lambda_\infty^{-1}\lambda_1 b - y(\lambda_1)\rangle \leq 0,$$

which is a contradiction. This contradiction shows that $\varphi(\cdot)$ is strictly increasing on $(0, \lambda_\infty]$.

(c) Since $\tilde{y}(\cdot)$ is locally Lipschitz continuous on $(0, \lambda_\infty)$, it follows from (Bolte et al., 2009) that $\tilde{y}(\cdot)$ is semismooth on $(0, +\infty)$ if Q is a tame set. Since $\|\cdot\|$ is strongly semismooth, we know that $g_1(\cdot) = \|\tilde{y}(\cdot)\|$ is semismooth on $(0, \lambda_\infty)$. Therefore, $\varphi(\cdot)$ is semismooth on $(0, \lambda_\infty)$ (Facchinei and Pang, 2003b, Proposition 7.4.4, Proposition 7.4.8).

(d) The γ-order semismoothness of $\varphi(\cdot)$ can be proved similarly as for (c). □

The above proposition can be used to prove the semismoothness of $\varphi(\cdot)$ for a wide class of functions $p(\cdot)$. For example, the next corollary shows the semismoothness of $\varphi(\cdot)$ when $p(\cdot)$ is the nuclear norm function defined on $\mathbb{R}^{d\times n}$, using the fact that $\partial p(0)$ is linear matrix inequality representable (Recht et al., 2010).

Corollary 4.1. *Denote the adjoint of the linear operator* $\mathcal{A} : \mathbb{R}^{d\times n} \to \mathbb{R}^m$ *as* $\mathcal{A}^*$*. Let* $p(\cdot) = \|\cdot\|_*$ *be the nuclear norm function defined on* $\mathbb{R}^{d\times n}$*. Then* $Q = \{z \in \mathbb{R}^m \mid \mathcal{A}^* z \in \partial p(0)\}$ *is a tame set and* $\Pi_Q(\cdot)$ *is semismooth.*

Next, we will show that $\varphi(\cdot)$ can be strongly semismooth for a class of important instances of $p(\cdot)$.

Proposition 4.4. *Let* $p(\cdot)$ *be a gauge function. Define* $\Phi(x) := \frac{1}{2}\|Ax - b\|^2$, $x \in \mathbb{R}^n$ *and*

$$H(x, \lambda) := x - \mathrm{Prox}_p(x - \lambda^{-1}\nabla\Phi(x)), \quad (x, \lambda) \in \mathbb{R}^n \times \mathbb{R}_{++}.$$

For any $(x, \lambda) \in \mathbb{R}^n \times \mathbb{R}_{++}$*, denote* $\partial_x H(x, \lambda)$ *as the Canonical projection of* $\partial H(x, \lambda)$ *onto* $\mathbb{R}^n$*. It holds that*

(a) if $\Pi_{\partial p(0)}(\cdot)$ *is strongly semismooth and* $\partial_x H(\bar{x}, \bar{\lambda})$ *is nondegenerate at some* $(\bar{x}, \bar{\lambda})$ *satisfying* $H(\bar{x}, \bar{\lambda}) = 0$*, then* $y(\cdot)$ *and* $\varphi(\cdot)$ *are strongly semismooth at* $\bar{\lambda}$*;*

(b) if $p(\cdot)$ *is further assumed to be polyhedral, the function* $y(\cdot)$ *is piecewise affine and* $\varphi(\cdot)$ *is strongly semismooth on* $\mathbb{R}_{++}$*.*

Proof. (a) It follows from the Moreau identity (Rockafellar, 1970, Theorem 31.5) that for any $(x, \lambda) \in \mathbb{R}^n \times \mathbb{R}_{++}$,

$$\begin{aligned} H(x, \lambda) &= x - ((x - \lambda^{-1}\nabla\Phi(x)) - \mathrm{Prox}_{p^*}(x - \lambda^{-1}\nabla\Phi(x))) \\ &= \lambda^{-1}\nabla\Phi(x) + \Pi_{\partial p(0)}(x - \lambda^{-1}\nabla\Phi(x)). \end{aligned}$$

The rest of the proof can be obtained from the fact that $\nabla\Phi(\cdot)$ is linear and the Implicit Function Theorem for semismooth functions (Theorem 2.4).

(b) When $p(\cdot)$ is a polyhedral gauge function, we know that the set Q defined in (4.5) is a convex polyhedral set (Rockafellar, 1970, Theorem 19.3) and the projector $\Pi_Q(\cdot)$ is piecewise affine (Facchinei and Pang, 2003b, Proposition 4.1.4). Therefore, $y(\cdot)$ is a piecewise affine function on $(0, +\infty)$. Then both $y(\cdot)$ and $\varphi(\cdot)$ are strongly semismooth on $(0, +\infty)$ (Facchinei and Pang, 2003b, Proposition 7.4.7, Proposition 7.4.4, Proposition 7.4.8). □

Remark 4.1. *We make some remarks on the assumptions in Part (a) of Proposition 4.4. On one hand, the strong semismoothness of the projector* $\Pi_K(\cdot)$ *has been proved for some important non-polyhedral closed convex sets* K*, such as the positive semidefinite cone (Sun and Sun, 2002), the second-order cone (Chen et al., 2003), and the* ℓ_2 *norm ball (Zhang et al., 2020, Lemma 2.1). On the other hand, the assumption of the nondegeneracy of* $\partial_x H(\cdot, \cdot)$ *at the concerned point is closely related to the important concept of strong regularity of the KKT system of* $(\mathrm{P}_{\mathrm{LS}}(\lambda))$*. One can refer to the Monograph (Bonnans and Shapiro, 2000) and the references therein for a general discussion, and to (Sun, 2006; Chan and Sun, 2008) for the semidefinite programming problems.*

4.2 The HS-Jacobian of $\varphi(\cdot)$ for some special $p(\cdot)$

In this section, we assume by default that $0 < \lambda_\infty < +\infty$. Inspired by the generalized Jacobian for the projector over a polyhedral set derived by Han and Sun (1997), which we call the HS-Jacobian, we will derive the HS-Jacobian of the value function $\varphi(\cdot)$. As an important implication, we will prove that the Clarke Jacobian of $\varphi(\cdot)$ at any $\lambda \in (0, \lambda_\infty)$ is positive. Note that the open interval $(0, \lambda_\infty)$ contains the solution λ^* to (E_φ). Additionally, we also study a special case of the polyhedral gauge function, the k-norm function. This is because the derivation for the general polyhedral gauge function does not easily yield the HS-Jacobian of $\varphi(\cdot)$ at $\lambda \in (0, \lambda_\infty)$ and its nondegeneracy.

4.2.1 When $p(\cdot)$ is a polyhedral gauge function

Let $p(\cdot)$ be a polyhedral gauge function. Then the set $\partial p(0)$ is polyhedral (Rockafellar, 1970, Theorem 19.2), which can be assumed without loss of generality, to take the form of

$$\partial p(0) := \{u \in \mathbb{R}^n \mid Bu \leq d\} \tag{4.7}$$

for some $B \in \mathbb{R}^{q \times n}$ and $d \in \mathbb{R}^q$.

Let $\lambda \in (0, \lambda_\infty)$ be arbitrarily chosen. Let $(y(\lambda), u(\lambda))$ be the unique solution to $(\mathrm{D_{LS}}(\lambda))$ with the parameter λ. Here, we denote $(y, u) = (y(\lambda), u(\lambda))$ to simplify our notation and hide the dependency on λ. Then there exists $x \in \Omega(\lambda)$ such that (y, u, x) satisfies the following KKT system:

$$u = \Pi_{\partial p(0)}(u + x), \quad y - b + Ax = 0, \quad A^T y - \lambda u = 0. \tag{4.8}$$

Therefore, u is the unique solution to the following optimization problem

$$\min_{z \in \mathbb{R}^n} \left\{ \frac{1}{2}\|z - (u + x)\|^2 \mid Bz \leq d \right\} \tag{4.9}$$

and there exists $\xi \in \mathbb{R}^q$ such that (u, ξ) satisfies the following KKT system for (4.9):

$$B^T \xi - x = 0, \quad Bu - d \leq 0, \quad \xi \geq 0, \quad \xi^T (Bu - d) = 0. \tag{4.10}$$

As a result, there exists $(x, \xi) \in \mathbb{R}^n \times \mathbb{R}^q$ such that (y, u, x, ξ) satisfies the following augmented KKT system

$$\begin{cases} B^T \xi - x = 0, \quad Bu - d \leq 0, \quad \xi \geq 0, \quad \xi^T (Bu - d) = 0, \\ y - b + Ax = 0, \quad A^T y - \lambda u = 0. \end{cases} \tag{4.11}$$

Let $M(\lambda)$ be the set of Lagrange multipliers associated with (y, u) defined as

$$M(\lambda) := \{(x, \xi) \in \mathbb{R}^n \times \mathbb{R}^q \mid (y, u, x, \xi) \text{ satisfies } (4.11)\}.$$

Since $x = B^T\xi$, we obtain the following system by eliminating the variable x in (4.11):

$$\begin{cases} Bu - d \le 0, \quad \xi \ge 0, \quad \xi^T(Bu-d) = 0, \\ y - b + \widehat{A}\xi = 0, \quad A^T y - \lambda u = 0, \end{cases} \tag{4.12}$$

where $\widehat{A} = AB^T \in \mathbb{R}^{m\times q}$. Denote

$$\widehat{M}(\lambda) := \{\xi \in \mathbb{R}^q \,|\, (y,u,\xi) \text{ satisfies (4.12)}\}\,. \tag{4.13}$$

Then, the set $M(\lambda)$ is equivalent to

$$M(\lambda) = \left\{(x,\xi) \in \mathbb{R}^n \times \mathbb{R}^q \,|\, x = B^T\xi,\ \xi \in \widehat{M}(\lambda)\right\}. \tag{4.14}$$

Denote the active set of u as

$$I(u) := \{i \in [q] \,|\, B_{i:}u - d_i = 0\}. \tag{4.15}$$

For any $\lambda \in (0, \lambda_\infty)$, we define

$$\mathcal{B}(\lambda) := \left\{K \subseteq [q] \,|\, \exists\, \xi \in \widehat{M}(\lambda) \text{ s.t. } \mathrm{supp}(\xi) \subseteq K \subseteq I(u) \text{ and } \mathrm{rank}(\widehat{A}_{:K}) = |K|\right\}. \tag{4.16}$$

Since the polyhedral set $\widehat{M}(\lambda)$ does not contain a line, this implies that $\widehat{M}(\lambda)$ has at least one extreme point $\bar{\xi}$ (Rockafellar, 1970, Corollary 18.5.3). Note that $0 < \lambda < \lambda_\infty$ and $x \neq 0$, which implies that $\bar{\xi} \neq 0$ and $\mathcal{B}(\lambda)$ is nonempty.

Define the HS-Jacobian of $y(\cdot)$ as

$$\mathcal{H}(\lambda) := \left\{h^K \in \mathbb{R}^m \,|\, h^K = \widehat{A}_{:K}(\widehat{A}_{:K}^T\widehat{A}_{:K})^{-1}d_K,\ K \in \mathcal{B}(\lambda)\right\}, \quad \lambda \in (0,\lambda_\infty), \tag{4.17}$$

where d_K is the subvector of d indexed by K. For notational convenience, for any $\lambda \in (0, \lambda_\infty)$ and $K \in \mathcal{B}(\lambda)$, denote

$$P^K = I - \widehat{A}_{:K}(\widehat{A}_{:K}^T\widehat{A}_{:K})^{-1}\widehat{A}_{:K}^T. \tag{4.18}$$

Define

$$\mathcal{V}(\lambda) := \left\{t \in \mathbb{R} \,|\, t = \lambda \|h\|^2/\varphi(\lambda),\ h \in \mathcal{H}(\lambda)\right\}, \quad \lambda \in \mathcal{D}, \tag{4.19}$$

where $\mathcal{D} = \{\lambda \in (0, \lambda_\infty) \,|\, \varphi(\lambda) > 0\}$. The following lemma is proved by following the same line as in (Han and Sun, 1997, Lemma 2.1).

Lemma 4.1. *Let $\bar{\lambda} \in (0, \lambda_\infty)$ be arbitrarily chosen. It holds that*

$$y(\bar{\lambda}) = P^K b + \bar{\lambda} h^K, \quad \forall\, h^K \in \mathcal{H}(\bar{\lambda}). \tag{4.20}$$

Moreover, there exists a positive scalar ς such that $\mathcal{N}(\bar{\lambda}) := (\bar{\lambda} - \varsigma, \bar{\lambda} + \varsigma) \subseteq (0, \lambda_\infty)$ and for all $\lambda \in \mathcal{N}(\bar{\lambda})$,

(a) $\mathcal{B}(\lambda) \subseteq \mathcal{B}(\bar{\lambda})$ and $\mathcal{H}(\lambda) \subseteq \mathcal{H}(\bar{\lambda})$;

(b) $y(\lambda) = y(\bar{\lambda}) + (\lambda - \bar{\lambda})h, \quad \forall\, h \in \mathcal{H}(\lambda)$.

Proof. Choose a sufficiently small $\varsigma > 0$ such that $\mathcal{N}(\bar{\lambda}) := (\bar{\lambda} - \varsigma, \bar{\lambda} + \varsigma) \subseteq (0, \lambda_\infty)$ and let $\lambda \in \mathcal{N}(\bar{\lambda}) \backslash \bar{\lambda}$ be arbitrarily chosen. Denote $(\bar{y}, \bar{u}) = (y(\bar{\lambda}), u(\bar{\lambda}))$ and $(y, u) = (y(\lambda), u(\lambda))$ for notational simplicity.

Let $K \in \mathcal{B}(\bar{\lambda})$ be arbitarily chosen. From (4.12), there exists $\bar{\xi} \in \widehat{M}(\bar{\lambda})$ with $\mathrm{supp}(\bar{\xi}) \subseteq K \subseteq I(\bar{u})$ such that $(\bar{y}, \bar{u}, \bar{\xi})$ satisfies

$$\bar{y} = b - \widehat{A}_{:K} \bar{\xi}_K, \quad BA^T \bar{y} = \widehat{A}^T \bar{y} = \bar{\lambda} B \bar{u}, \quad B_{K:} \bar{u} = d_K, \quad \xi_{K^c} = 0, \tag{4.21}$$

where K^c is the complement of K, which implies

$$\bar{\lambda} d_K = \widehat{A}_{:K}^T (b - \widehat{A}_{:K} \bar{\xi}_K).$$

Since $\widehat{A}_{:K}$ is of full column rank, we have

$$\bar{\xi}_K = (\widehat{A}_{:K}^T \widehat{A}_{:K})^{-1} (\widehat{A}_{:K}^T b - \bar{\lambda} d_K).$$

Consequently, we have

$$\bar{y} = P^K b + \bar{\lambda} h^K,$$

where $h^K = \widehat{A}_{:K}(\widehat{A}_{:K}^T\widehat{A}_{:K})^{-1}d_K \in \mathcal{H}(\bar{\lambda})$.

(a) It follows from Proposition 4.3 that $u(\cdot)$ is locally Lipschitz continuous, which implies that $I(u) \subseteq I(\bar{u})$. Next, we prove that $\mathcal{B}(\lambda) \subseteq \mathcal{B}(\bar{\lambda})$. If not, then there exists a sequence $\{\lambda_k\}_{k\geq 1} \subseteq \mathcal{N}(\bar{\lambda})$ converges to $\bar{\lambda}$, such that for all k, there is an index set $K^k \in \mathcal{B}(\lambda_k)\backslash\mathcal{B}(\bar{\lambda})$. Denote the solution to ($\mathrm{D_{LS}}(\lambda)$) with the parameter λ_k as (y^k, u^k). Since there exist only finitely many choices for the index sets in $\mathcal{B}(\cdot)$, if necessary by taking a subsequence we assume that the index sets K^k are identical for all $k \geq 1$. Denote the common index set as $\tilde{K}$. Then, the matrix $\widehat{A}_{:\tilde{K}}$ has full column rank and there exists $\xi^k \in \widehat{M}(\lambda_k)$ (and $(B^T\xi^k, \xi^k) \in M(\lambda_k)$) such that $\mathrm{supp}(\xi^k) \subseteq \tilde{K} \subseteq I(u^k)$ but $\tilde{K} \notin \mathcal{B}(\bar{\lambda})$. Since $I(u^k) \subseteq I(\bar{u})$, then there is no $\xi \in \mathcal{B}(\bar{\lambda})$ such that $\mathrm{supp}(\xi) \subseteq \tilde{K}$. However, since $\xi^k \in \widehat{M}(\lambda^k)$, it satisfies

$$y^k - b + \widehat{A}_{:\tilde{K}}\xi^k_{\tilde{K}} = 0. \tag{4.22}$$

As $y(\cdot)$ is locally Lipschitz continuous and $\widehat{A}_{:\tilde{K}}$ is of full column rank, the sequence $\{\xi^k\}_{k\geq 1}$ is bounded. Let $\tilde{\xi}$ be an accumulation point of $\{\xi^k\}_{k\geq 1}$, then $\tilde{\xi} \in \mathcal{B}(\bar{\lambda})$ and $\mathrm{supp}(\tilde{\xi}) \subseteq \tilde{K}$. This is a contradiction. Therefore, $\mathcal{B}(\lambda) \subseteq \mathcal{B}(\bar{\lambda})$. From the definition of $\mathcal{H}(\cdot)$ in (4.17), we also have $\mathcal{H}(\lambda) \subseteq \mathcal{H}(\bar{\lambda})$.

(b) Let $K \in \mathcal{B}(\lambda)$ be arbitarily chosen. It follows from (4.20) that

$$y = P^K b + \lambda h^K,$$

where $h^K = \widehat{A}_{:K}(\widehat{A}_{:K}^T\widehat{A}_{:K})^{-1}d_K \in \mathcal{H}(\lambda)$. Since $K \in \mathcal{B}(\lambda) \subseteq \mathcal{B}(\bar{\lambda})$ and $h^K \in \mathcal{H}(\lambda) \subseteq \mathcal{H}(\bar{\lambda})$, in a same vein, we have

$$\bar{y} = P^K b + \bar{\lambda} h^K.$$

As a result, for all $h \in \mathcal{H}(\lambda)$,

$$y = \bar{y} + (\lambda - \bar{\lambda})h.$$

We complete the proof of the lemma. □

Next, we prove the nondegeneracy of $\partial\varphi(\bar{\lambda})$ for any $\bar{\lambda} \in (0, \lambda_\infty)$, which is important for analyzing the convergence rates of the secant method for solving (E_φ).

Theorem 4.1. *Let $p(\cdot)$ be a polyhedral gauge function. Assume that $0 < \lambda_\infty < +\infty$. For any $\bar{\lambda} \in (0, \lambda_\infty)$, it holds that*

(a) for any integer $k \geq 1$, the function $\varphi(\cdot)$ is piecewise C^k in an open interval containing $\bar{\lambda}$;

(b) all $v \in \partial\varphi(\bar{\lambda})$ are positive.

Proof. Choose a sufficiently small $\varsigma > 0$ such that $\mathcal{N}(\bar{\lambda}) = (\bar{\lambda} - \varsigma, \bar{\lambda} + \varsigma) \subseteq (0, \lambda_\infty)$ and $\mathcal{B}(\lambda) \subseteq \mathcal{B}(\bar{\lambda})$ for any $\lambda \in \mathcal{N}(\bar{\lambda})$. Let $\lambda \in \mathcal{N}(\bar{\lambda})\backslash\bar{\lambda}$ and $K \in \mathcal{B}(\lambda)$ be arbitrarily chosen. Denote

$$h^K = \widehat{A}_{:K}(\widehat{A}_{:K}^T \widehat{A}_{:K})^{-1} d_K.$$

Then, we have $h^K \in \mathcal{H}(\lambda) \subseteq \mathcal{H}(\bar{\lambda})$.

Now, we prove Part (a) of the theorem. From the fact

$$\langle P^K b, h^K \rangle = 0$$

and Lemma 4.1, we know that

$$\varphi(\lambda) = \sqrt{\|P^K b\|^2 + \lambda^2 \|h^K\|^2} \quad \text{and} \quad \varphi(\bar{\lambda}) = \sqrt{\|P^K b\|^2 + (\bar{\lambda})^2 \|h^K\|^2}.$$

Define $\varphi^K : \mathbb{R} \to \mathbb{R}_+$ by

$$\varphi^K(s) := \sqrt{\|P^K b\|^2 + s^2 \|h^K\|^2}, \quad s \in \mathbb{R}. \tag{4.23}$$

From Proposition 4.3, we know that $\varphi(\cdot)$ is strictly increasing on $(0, \lambda_\infty]$. Therefore, it holds that

$$\varphi^K(\lambda) = \varphi(\lambda) \neq \varphi(\bar{\lambda}) = \varphi^K(\bar{\lambda}), \tag{4.24}$$

which implies that $h^K \neq 0$. Since $K \in \mathcal{B}(\lambda)$ is arbitrarily chosen, we obtain that

$$h^K \neq 0, \forall K \in \mathcal{B}(\lambda). \tag{4.25}$$

Thus, for any integer $k \geq 1$, $\varphi^K(\cdot)$ is C^k on $\mathcal{N}(\bar{\lambda})$.

Denote $\bar{\mathcal{B}} = \bigcup_{\lambda \in \mathcal{N}(\bar{\lambda}) \setminus \bar{\lambda}} \mathcal{B}(\lambda)$. We know that $\bar{\mathcal{B}} \subseteq \mathcal{B}(\bar{\lambda})$ and $|\bar{\mathcal{B}}|$ is finite. Moreover,

$$\varphi(s) \in \{\varphi^K(s)\}_{K \in \bar{\mathcal{B}}}, \quad \forall\, s \in \mathcal{N}(\bar{\lambda}), \tag{4.26}$$

which implies that for any $k \geq 1$, $\varphi(\cdot)$ is piecewise C^k on $\mathcal{N}(\bar{\lambda})$.

Next, we prove Part (b) of the theorem. It follows from (4.25) that for any $K \in \bar{\mathcal{B}}$ and $\lambda > 0$,

$$h^K \neq 0 \quad \text{and} \quad (\varphi^K)'(\lambda) = \lambda \|h^K\|^2 / \varphi^K(\lambda) > 0. \tag{4.27}$$

By (Clarke, 1983, Theorem 2.5.1) and the upper semicontinuity of $\partial_{\mathrm{B}}\varphi(\cdot)$, we have

$$\partial\varphi(\bar{\lambda}) \subseteq \mathrm{conv}(\{\bar{\lambda}\|h^K\|^2/\varphi(\bar{\lambda}) \mid K \in \bar{\mathcal{B}}\}), \tag{4.28}$$

which implies

$$v > 0, \quad \forall\, v \in \partial\varphi(\bar{\lambda}).$$

We complete the proof of the theorem. □

The following proposition shows that for the least squares constrained Lasso problem, $\partial_{\mathrm{HS}}\varphi(\bar{\lambda})$ is positive for any $\bar{\lambda} \in (0, \lambda_\infty)$.

Proposition 4.5. *Suppose that $p(\cdot)$ is a polyhedral gauge function and $\partial p(0)$ has the expression as in (4.7). Assume that $0 < \lambda_\infty < +\infty$ and let $\bar{\lambda} \in (0, \lambda_\infty)$ be arbitrarily chosen. Let $\mathcal{B}(\bar{\lambda})$ and $\mathcal{V}(\bar{\lambda})$ be the sets defined as in (4.16) and (4.19) for $\lambda = \bar{\lambda}$. If $d_K \neq 0$ for all $K \in \mathcal{B}(\bar{\lambda})$, then $v > 0$ for all $v \in \mathcal{V}(\bar{\lambda})$. Moreover, $d_K \neq 0$ for all $K \in \mathcal{B}(\bar{\lambda})$ when $p(\cdot) = \|\cdot\|_1$.*

Proof. Recall that $\mathcal{B}(\bar{\lambda})$ is nonempty. Let $K \in \mathcal{B}(\bar{\lambda})$ be arbitrarily chosen. We know that $\widehat{A}_{:K}$ is of full column rank. Denote $h^K = \widehat{A}_{:K}(\widehat{A}_{:K}^T \widehat{A}_{:K})^{-1} d_K \in \mathcal{H}(\bar{\lambda})$. Since $d_K \neq 0$, it holds that

$$\|h_K\|^2 = \langle \widehat{A}_{:K}(\widehat{A}_{:K}^T \widehat{A}_{:K})^{-1} d_K, \widehat{A}_{:K}(\widehat{A}_{:K}^T \widehat{A}_{:K})^{-1} d_K \rangle = \langle d_K, (\widehat{A}_{:K}^T \widehat{A}_{:K})^{-1} d_K \rangle > 0.$$

Therefore, it follows from Lemma 4.1 and the facts $\bar{\lambda} > 0$ and $\langle P^K b, h^K \rangle = 0$ that

$$\varphi(\bar{\lambda}) = \sqrt{\|P^K b\|^2 + \bar{\lambda}^2 \|h^K\|^2} > 0,$$

which implies that

$$v^K = \bar{\lambda}\|h^K\|^2/\varphi(\bar{\lambda}) \in \mathcal{V}(\bar{\lambda}) \quad \text{and} \quad v^K > 0.$$

Since $K \in \mathcal{B}(\bar{\lambda})$ is arbitrarily chosen, we know that $v > 0$ for all $v \in \mathcal{V}(\bar{\lambda})$.

When $p(\cdot) = \|\cdot\|_1$, the set $\partial p(0)$ has the representation of

$$\partial p(0) = \{u \in \mathbb{R}^n \mid -1 \leq u_i \leq 1,\ i \in [n]\}.$$

In other words, $B = [I_n \ \ -I_n]^T \in \mathbb{R}^{2n \times n}$ and $d = e_{2n}$. Therefore, $d_K \neq 0$ for any $\bar{\lambda} \in (0, \lambda_\infty)$ and $K \in \mathcal{B}(\bar{\lambda})$. □

4.2.2 When $p(\cdot)$ is a k-norm function

In this section, we focus on the discussion of the HS-Jacobian of $\varphi(\cdot)$ when $p(\cdot)$ is the k-norm function. While the k-norm function is indeed a particular example of a polyhedral gauge function, we adopt a direct approach to studying results similar to those previously obtained, rather than transforming the set $\partial p(0)$ into the form presented in (4.7). This is motivated by the challenges of constructing the set $\partial p(0)$ in that form. Additionally, this method is essential for deriving the HS-Jacobian of $\varphi(\cdot)$, which is necessary for employing the semismooth Newton method to solve (CP(ϱ)) and to compare its performance with our secant method in Chapter 5.

Let $n > 2$. For any given $k \in \{1, \cdots, n\}$, the k-norm function $p(\cdot)$ is defined by

$$p(x) := \|x\|_{(k)} = \sum_{i=1}^{k} |x_{(i)}|, \quad \forall\, x \in \mathbb{R}^n,$$

where $|x_{(1)}| \geq \cdots \geq |x_{(k)}|$ are the k largest absolute values of the elements in x. Specifically, when $k = 1$, the k-norm function reduces to the infinity norm function, i.e., $p(x) = \|x\|_\infty$, and when $k = n$, it becomes the ℓ_1-norm function, i.e., $p(x) = \|x\|_1$. We now turn our attention to the k-norm penalized least squares problem ($\mathrm{P}_{\mathrm{LS}}(\lambda)$).

Recall that, for any $\lambda > 0$, the dual problem of ($\mathrm{P}_{\mathrm{LS}}(\lambda)$) is given by:

$$\max_{y\in\mathbb{R}^n, u\in\mathbb{R}^n} \left\{ -\frac{1}{2}\|y\|^2 + \langle b, y\rangle + p^*(u) \mid \; A^T y - \lambda u = 0 \right\}, \tag{$\mathrm{D}_{\mathrm{LS}}(\lambda)$}$$

where we know from (Rockafellar, 1970, Corollary 13.2.1) that

$$p^*(\cdot) = \delta(\cdot \mid \partial p(0)). \tag{4.29}$$

Moreover, it follows from (Watson, 1992, Theorem 1) that

$$\partial p(0) = \{u \in \mathbb{R}^n \mid \|u\|_{(k)}^* \leq 1\}, \tag{4.30}$$

where the dual norm $\|\cdot\|_{(k)}^*$ of $\|\cdot\|_{(k)}$ is $\|u\|_{(k)}^* = \max\{\|u\|_1/k, \|u\|_\infty\}$, $u \in \mathbb{R}^n$. Therefore, the KKT system associated with ($\mathrm{P}_{\mathrm{LS}}(\lambda)$) and ($\mathrm{D}_{\mathrm{LS}}(\lambda)$) is

$$\begin{cases} y - b + Ax = 0, \\ u = \Pi_{\partial p(0)}(u + x), \qquad\qquad (y, u, x) \in \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^m. \\ A^T y - \lambda u = 0, \end{cases} \tag{4.31}$$

For any u and x with $u = \Pi_{\partial p(0)}(u + x)$, it follows from (4.30) that u is the unique optimal solution to

$$\min_{w\in\mathbb{R}^n} \left\{ \frac{1}{2}\|w - (u + x)\|^2 \mid \|u\|_1 \leq k,\ \|u\|_\infty \leq 1 \right\}. \tag{4.32}$$

It then follows from (Rockafellar, 1970, Theorem 28.2) that there exist some $\alpha, \beta \in \mathbb{R}$ such that

$$\begin{cases} x \in \alpha\partial\|u\|_1 + \beta\partial\|u\|_\infty, \\ \|u\|_1 \le k, \ \ \alpha \ge 0, \\ \alpha(\|u\|_1 - k) = 0, \\ \|u\|_\infty \le 1, \ \ \beta \ge 0, \\ \beta(\|u\|_\infty - 1) = 0. \end{cases} \tag{4.33}$$

Let $\lambda \in (0, \lambda_\infty)$. Since $\lambda < \lambda_\infty = \max\{\|A^T b\|_1/k, \ \|A^T b\|_\infty\}$, it is not possible for both α and β to be equal to 0 at the same time. This is because when $\alpha = 0$ and $\beta = 0$, from (4.31) and (4.33), we have that $\|A^T b\|_1 \le k\lambda$ and $\|A^T b\|_\infty \le \lambda$, which contradicts to the assumption that $\lambda < \lambda_\infty$. Also, we know from (4.33) that $u \ne 0$.

Let $(y, u) \in \mathbb{R}^m \times \mathbb{R}^n$ denotes the unique optimal solution to $(\mathrm{D_{LS}}(\lambda))$. It follows from an example in (Rockafellar, 1970, Section 23, Page 215) that

$$\partial\|u\|_\infty = \mathrm{conv}\{\mathrm{sign}(u_j)e_j \mid j \in J_u\}, \quad J_u = \{j \mid |u_j| = 1\}, \tag{4.34}$$

where e_j represents the vector that forms the j-th column of the $n \times n$ identity matrix. For any integer $i \in \{1, \cdots, n\}$, we know from the inclusion $x \in \alpha\partial\|u\|_1 + \beta\partial\|u\|_\infty$ and (4.34), that

$$x_i \in \begin{cases} \{\mathrm{sign}(u_i) \cdot (\alpha + \beta \cdot \omega_i \cdot e_i)\}, & \text{if } |u_i| = 1, \\ [-\alpha, \alpha], & \text{if } u_i = 0, \\ \{\mathrm{sign}(u_i) \cdot \alpha\}, & \text{if } 0 < |u_i| < 1, \end{cases} \tag{4.35}$$

where $\omega_i \ge 0$, and $\sum_i \omega_i = 1$. Denote

$$I(u) = I^0(u) \cup I^1(u), \quad \bar{I}(u) := \{1, \cdots, n\} \backslash I(u),$$

where

$$I^0(u) := \{i \mid u_i = 0\}, \quad I^1(u) := \{i \mid |u_i| = 1\}.$$

Additionally, from (4.35), we have

$$\big\{\mathrm{sign}(u_i) \cdot [x_i - \mathrm{sign}(u_i) \cdot \alpha], \ i \in I^1(u)\big\} \in \beta\Delta_{|I^1(u)|},$$

where $\Delta_{|I^1(u)|}$ represents the $|I^1(u)|$ dimensional simplex, with $|I^1(u)|$ representing the cardinality of the set $I^1(u)$. Therefore,

$$\begin{cases} \beta = \sum_{i \in I^1(u)} \text{sign}(u_i) \cdot x_i - \alpha \cdot |I^1(u)|, \\ \text{sign}(u_i) \cdot x_i - \alpha \geq 0, \quad i \in I^1(u). \end{cases}$$

Consequently, (4.31) becomes

$$\begin{cases} y - b + Ax = 0, \\ \beta = \sum_{i \in I^1(u)} \text{sign}(u_i) \cdot x_i - \alpha \cdot |I^1(u)|, \\ \text{sign}(u_i) \cdot x_i - \alpha \geq 0, \quad i \in I^1(u), \\ x_i \leq \alpha, \quad i \in I^0(u), \\ x_i \geq -\alpha, \quad i \in I^0(u), \\ x_i = \text{sign}(u_i) \cdot \alpha, \quad i \in \bar{I}(u), \\ \alpha \geq 0, \ \beta \geq 0, \\ \alpha(\|u\|_1 - k) = 0, \\ \|u\|_\infty \leq 1, \ \|u\|_1 \leq k, \\ \beta(\|u\|_\infty - 1) = 0, \\ A^T y - \lambda u = 0. \end{cases} \tag{4.36}$$

Let $\mathcal{M}(\lambda)$ be the nonempty set of multipliers $(x, \alpha, \beta) \in \mathbb{R}^n \times \mathbb{R}_+ \times \mathbb{R}_+$ that satisfies the KKT condition (4.36). Denote

$$S(x, \alpha) := \{i \mid |x_i| \neq \alpha\}.$$

We then define the family $\mathcal{B}(\lambda)$ of indices $\{1, \cdots, n\}$ as follows: $K \in \mathcal{B}(\lambda)$ if and only if $S(x, \alpha) \subseteq K \subseteq I(u)$, for some $(x, \alpha, \beta, y, u, \lambda)$ satisfying the KKT system (4.36), and the vectors $\{A_i, i \in K\}$ are linearly independent. Let A_K denote the matrix consisting of the columns of A indexed by K. The family $\mathcal{B}(\lambda)$ is non-empty due to the existence of an non-zero extreme point of $\mathcal{M}(\lambda)$. Denote $\overline{K} = \{1, \cdots, n\} \backslash K$.

Define

$$\mathcal{H}(\lambda) = \left\{ h \in \mathbb{R}^n \mid h = r(u, x, K) \cdot P_K A_{\overline{K}} \text{sign}(x_{\overline{K}}) + A_K (A_K^T A_K)^{-1} u_K, K \in \mathcal{B}(\lambda) \right\},$$

where

$$r(u,x,K) = \begin{cases} \frac{\|u_{\overline{K}}\|_1 - \langle A_{\overline{K}}\mathrm{sign}(x_{\overline{K}}), A_K(A_K^T A_K)^{-1} u_K\rangle}{\|A_{\overline{K}}\mathrm{sign}(x_{\overline{K}})\|_{P_K}^2}, & \text{if } \|A_{\overline{K}}\mathrm{sign}(x_{\overline{K}})\|_{P_K}^2 \neq 0, \\ 0, & \text{otherwise,} \end{cases}$$

and

$$P_K = I - A_K(A_K^T A_K)^{-1} A_K^T.$$

Proposition 4.6. *Let* $\bar{\lambda} \in (0, \lambda_\infty)$.

(a) There exists a neighborhood $\mathcal{N}(\bar{\lambda})$ *of* $\bar{\lambda}$ *such that*

$$\mathcal{B}(\lambda) \subseteq \mathcal{B}(\bar{\lambda}), \quad \mathcal{H}(\lambda) \subseteq \mathcal{H}(\bar{\lambda}), \quad \forall\, \lambda \in \mathcal{N}(\bar{\lambda}). \tag{4.37}$$

(b) When $\mathcal{B}(\lambda) \subseteq \mathcal{B}(\bar{\lambda})$*, then*

$$y - \bar{y} = (\lambda - \bar{\lambda}) \cdot h, \quad \forall\, h \in \mathcal{H}(\lambda), \tag{4.38}$$

where y *and* $\bar{y}$ *denote the unique optimal solutions to* $(\mathrm{D_{LS}}(\lambda))$ *and* $(\mathrm{D_{LS}}(\bar{\lambda}))$*, respectively.*

Proof. Since $\bar{\lambda} \in (0, \lambda_\infty)$, we may choose $\mathcal{N}(\bar{\lambda})$ to be sufficiently small such that $\lambda \in (0, \lambda_\infty)$ holds for all $\lambda \in \mathcal{N}(\bar{\lambda})$. Hence, $\mathcal{B}(\lambda)$, $\mathcal{H}(\lambda)$ and $\mathcal{T}(\lambda)$ are well-defined. Let $(\bar{y}, \bar{u})$ denote the unique optimal solution to the corresponding $(\mathrm{D_{LS}}(\bar{\lambda}))$.

(a) According to the definition of $\mathcal{H}$, it is only necessary to establish that $\mathcal{B}(\lambda) \subseteq \mathcal{B}(\bar{\lambda})$ in this proof. We prove that $\mathcal{B}(\lambda) \subseteq \mathcal{B}(\bar{\lambda})$ using a proof by contradiction. Suppose that there is a sequence $\{\lambda^k\}$ converging to $\bar{\lambda}$ such that $K^k \in \mathcal{B}(\lambda^k)$ but $K^k \notin \mathcal{B}(\bar{\lambda})$ for all k. Since the number of such index sets is finite, we can assume, by taking a subsequence if necessary, that K^k is the same for all k. Let us denote this common index set by K. From the definition of $\mathcal{B}$, we know that $\{A_i, i \in K\}$ are linearly independent and there is $(x^k, \alpha^k, \beta^k) \in \mathcal{M}(\lambda^k)$ such that

$$S(x^k, \alpha^k) \subseteq K \subseteq I(u^k).$$

By refining the subsequence of $\{\lambda^k\}$ appropriately, we can chose $(x^k, \alpha^k) \in \mathcal{M}(\lambda^k)$ such that for all k, the signs of the entries of $x^k_{\overline{K}}$ coincide with each other.

We will now establish that the sequence $\{x^k, \alpha^k, \beta^k\}$ is bounded. For any λ, since the objective value of $(\mathrm{P_{LS}}(\lambda))$ tends to ∞ as $|x| \to \infty$, the solution set of $(\mathrm{P_{LS}}(\lambda))$ is nonempty and compact (Rockafellar and Wets, 2009, Theorem 1.9). Let Ω_λ be the optimal solution set to $(\mathrm{P_{LS}}(\lambda))$. For any $0 < \lambda_1 < \lambda_2 < \lambda_\infty$ and any $x_1 \in \Omega_{\lambda_1}$, $x_2 \in \Omega_{\lambda_2}$, we have

$$\begin{aligned} \frac{1}{2}\|Ax_1 - b\|^2 + \lambda_1 p(x_1) \leq \frac{1}{2}\|Ax_2 - b\|^2 + \lambda_1 p(x_2), \\ \frac{1}{2}\|Ax_2 - b\|^2 + \lambda_2 p(x_2) \leq \frac{1}{2}\|Ax_1 - b\|^2 + \lambda_2 p(x_1). \end{aligned} \tag{4.39}$$

By adding the two inequalities in (4.39) and considering that $\lambda_1 < \lambda_2$, we can conclude that

$$(\lambda_2 - \lambda_1)(p(x_2) - p(x_1)) \leq 0,$$

which in turn implies that $p(x_1) \leq p(x_2)$. Therefore, for any $\kappa > p(x_{\{1\}})$ with $x_{\{1\}} \in \Omega_{\max\{\lambda^k\}}$, all elements of the sequence $\{x^k\}$ are contained within the closed and bounded convex level set $\{x \mid p(x) \leq \kappa\}$. The closeness property follows from the fact that $p(\cdot)$ is closed (Rockafellar, 1970, Theorem 7.1), and the property of being bounded corresponds to the behavior of $p(x)$ approaching infinity as the absolute value of x approaches infinity. From the observation of (4.36), it follows that the sequence $\{x^k, \alpha^k, \beta^k\}$ is bounded. Since $\lambda^k \to \bar{\lambda}$, there exists an accumulation point of the set $\{x^k, \alpha^k, \beta^k\}$ that belongs to $\mathcal{M}(\bar{\lambda})$ and satisfies $K \in \mathcal{B}(\bar{\lambda})$. This is a contradiction.

(b) Let $K \in \mathcal{B}(\lambda)$. Let (y, u) be the unique solution to $\mathrm{D}_{\mathrm{LS}}(\lambda)$ and $(x, \alpha, \beta) \in \mathcal{M}(\lambda)$ with $S(x, \alpha) \subseteq K \subseteq I(u)$. From (4.36), we have

$$\begin{cases} y = b - A_K x_K - \alpha \cdot A_{\overline{K}} \mathrm{sign}(x_{\overline{K}}), \\ A_K^T y = \lambda u_K, \quad A_{\overline{K}}^T y = \lambda u_{\overline{K}}. \end{cases}$$

Then, substituting y into the last two equations, we obtain

$$\begin{cases} x_K = (A_K^T A_K)^{-1} \left[A_K^T b - \alpha \cdot A_K^T A_{\overline{K}} \mathrm{sign}(x_{\overline{K}})\right] - \lambda (A_K^T A_K)^{-1} u_K, \\ \lambda u_{\overline{K}} = A_{\overline{K}}^T b - (A_{\overline{K}}^T A_K) x_K - \alpha \cdot A_{\overline{K}}^T A_{\overline{K}} \mathrm{sign}(x_{\overline{K}}). \end{cases}$$

This implies that

$$\lambda(u_{\overline{K}} - A_{\overline{K}}^T A_K (A_K^T A_K)^{-1} u_K) = A_{\overline{K}}^T P_K b - \alpha \cdot A_{\overline{K}}^T P_K A_{\overline{K}} \mathrm{sign}(x_{\overline{K}}).$$

Moreover, since $\lambda \|u_{\overline{K}}\|_1 = \langle \mathrm{sign}(x_{\overline{K}}), \lambda u_{\overline{K}} \rangle$, we then have

$$\lambda \left(\|u_{\overline{K}}\|_1 - \langle A_{\overline{K}} \mathrm{sign}(x_{\overline{K}}), A_K (A_K^T A_K)^{-1} u_K \rangle\right) = \langle A_{\overline{K}} \mathrm{sign}(x_{\overline{K}}), P_K b \rangle - \alpha \cdot \|A_{\overline{K}} \mathrm{sign}(x_{\overline{K}})\|_{P_K}^2.$$

Let $(\bar{y}, \bar{u})$ be the unique solution to $\mathrm{D}_{\mathrm{LS}}(\bar{\lambda})$ and $(\bar{x}, \bar{\alpha}, \bar{\beta}) \in \mathcal{M}(\bar{\lambda})$.

If $\|A_{\overline{K}} \mathrm{sign}(x_{\overline{K}})\|_{P_K}^2 \neq 0$, we have that

$$\alpha = \frac{\langle A_{\overline{K}} \mathrm{sign}(x_{\overline{K}}), P_K b \rangle - \lambda \left(\|u_{\overline{K}}\|_1 - \langle A_{\overline{K}} \mathrm{sign}(x_{\overline{K}}), A_K (A_K^T A_K)^{-1} u_K \rangle\right)}{\|A_{\overline{K}} \mathrm{sign}(x_{\overline{K}})\|_{P_K}^2}.$$

It is important to note that α cannot be zero. If $\alpha = 0$, then, based on the definition of $\mathcal{B}(\lambda)$, it follows that $x_{\overline{K}}$ must equal 0. However, this contradicts the assumption that $\|A_{\overline{K}} \mathrm{sign}(x_{\overline{K}})\|_{P_K}^2 \neq 0$. Moreover, from (4.36), we have that $\|u\|_1 = k$. It then follows from the above discussion that

$$y = P_K[b - \alpha \cdot A_{\overline{K}} \mathrm{sign}(x_{\overline{K}})] + \lambda A_K (A_K^T A_K)^{-1} u_K.$$

Note that $K \in \mathcal{B}(\bar{\lambda})$ and $\mathrm{sign}(x_{\overline{K}}) = \mathrm{sign}(\bar{x}_{\overline{K}})$. Thus, $P_K A_{\overline{K}} \mathrm{sign}(x_{\overline{K}}) = P_K A_{\overline{K}} \mathrm{sign}(\bar{x}_{\overline{K}})$. Therefore, we have

$$\bar{y} = P_K[b - \bar{\alpha} \cdot A_{\overline{K}} \mathrm{sign}(x_{\overline{K}})] + \bar{\lambda} A_K (A_K^T A_K)^{-1} \bar{u}_K,$$

where

$$\bar{\alpha} = \frac{\langle A_{\overline{K}}\text{sign}(x_{\overline{K}}), P_K b\rangle - \bar{\lambda}\left(\|\bar{u}_{\overline{K}}\|_1 - \langle A_{\overline{K}}\text{sign}(x_{\overline{K}}), A_K(A_K^T A_K)^{-1}\bar{u}_K\rangle\right)}{\|A_{\overline{K}}\text{sign}(x_{\overline{K}})\|_{P_K}^2}.$$

A similar argument can be made in the case $\|A_{\overline{K}}\text{sign}(\bar{x}_{\overline{K}})\|_{P_K}^2 \neq 0$ to show that $\bar{\alpha} \neq 0$ and hence that $\|\bar{u}\|_1 = k$. Moreover, we know that $u_K = \bar{u}_K$. Thereby $\|u_{\overline{K}}\|_1 = \|\bar{u}_{\overline{K}}\|_1$. As a result, we have

$$y - \bar{y} = (\lambda - \bar{\lambda}) \cdot \left\{ r(u, x, K) \cdot P_K A_{\overline{K}}\text{sign}(x_{\overline{K}}) + A_K(A_K^T A_K)^{-1} u_K \right\}. \quad (4.40)$$

If $\|A_{\overline{K}}\text{sign}(x_{\overline{K}})\|_{P_K}^2 = 0$, we know that $P_K A_{\overline{K}}\text{sign}(x_{\overline{K}}) = 0$ (note that $P_K^2 = P_K$), which implies that

$$y = P_K b + \lambda A_K(A_K^T A_K)^{-1} u_K \text{ and } y = P_K b + \bar{\lambda} A_K(A_K^T A_K)^{-1} \bar{u}_K$$

with $\bar{u}_K = u_K$. This completes the proof.

□

Then, we define the HS-Jacobian of $\varphi(\cdot)$ at $\lambda \in (0, \lambda_\infty)$ as follows

$$\mathcal{V}(\lambda) = \left\{ v \in \mathbb{R} \mid v = \lambda \|h\|^2 / \varphi(\lambda), h \in \mathcal{H}(\lambda) \right\}.$$

From (4.38) and the strictly increasing property of $\varphi(\cdot)$ on $(0, \lambda_\infty]$, we know that for any $\lambda \in (0, \lambda_\infty)$ and $h \in \mathcal{H}(\lambda)$, $\|h\| > 0$ holds. Therefore, for any $\lambda \in (0, \lambda_\infty)$, the value $v \in \mathcal{V}(\lambda)$ is also strictly greater than 0. Finally, we remark that when the function $p(\cdot)$ is the ℓ_1 norm, which is a special case of the k-norm function, then for any $\lambda \in (0, \lambda_\infty)$, we have

$$\mathcal{H}(\lambda) = \left\{ h \in \mathbb{R}^n \mid h = A_K(A_K^T A_K)^{-1} u_K, K \in \mathcal{B}(\lambda) \right\}. \quad (4.41)$$

This is because, when $p(\cdot)$ is the ℓ_1 norm function, we have $\text{sign}(x_{\overline{K}}) = 0$.

4.3 A fast convergent secant method for semismooth equations

With all preparations finalized, this section will analyze the convergence properties of the secant method. Specifically, we will demonstrate that the sequence $\{\lambda^k\}$ generated by the secant method converges 3-step Q-superlinearly (quadratically) to a solution λ^* of $\varphi(\lambda) = \varrho$, $\lambda \in (0, +\infty)$ when $\varphi(\cdot)$ is (strongly) semismooth. Moreover, if $p(\cdot)$ is polyhedral and $\partial\varphi(\lambda^*)$ is a singleton, then $\{\lambda^k\}$ converges to λ^* Q-superlinearly with a Q-order of at least $\frac{1+\sqrt{5}}{2}$.

We start our discussion of the secant method for a general semismooth equation. Let $f : \mathbb{R} \to \mathbb{R}$ be a locally Lipschitz continuous function which is semismooth at a solution x^* to the following equation

$$f(x) = 0. \tag{4.42}$$

Then, we will analyze the convergence of the secant method described in Algorithm 1 with two generic starting points x^{-1} and x^0.

The convergence results of Algorithm 1 are given in the following proposition. The proof can be obtained by following the procedure in the proof of (Potra et al., 1998, Theorem 3.2).

Proposition 4.7. *Suppose that $f : \mathbb{R} \to \mathbb{R}$ is semismooth at a solution x^* to (4.42). Let d^- and d^+ be the lateral derivatives of f at x^* as defined in (2.21). If d^- and d^+ are both positive (or negative), then there are two neighborhoods $\mathcal{U}$ and $\mathcal{N}$ of x^*, $\mathcal{U} \subseteq \mathcal{N}$, such that for $x^{-1}, x^0 \in \mathcal{U}$, Algorithm 1 is well defined and produces a sequence of iterates $\{x^k\}$ such that $\{x^k\} \subseteq \mathcal{N}$. The sequence $\{x^k\}$ converges to x^* 3-step Q-superlinearly, i.e., $|x^{k+3} - x^*| = o(|x^k - x^*|)$. Moreover, it holds that*

(a) $|x^{k+1} - x^*| \leq \frac{|d^+ - d^- + o(1)|}{\min\{|d^+|, |d^-|\} + o(1)} |x^k - x^*|$ *for* $k \geq 0$*;*

(b) *if* $\alpha := \frac{|d^+ - d^-|}{\min\{|d^+|,|d^-|\}} < 1$*, then* $\{x^k\}$ *converges to* x^* *Q-linearly with Q-factor* α*;*

(c) *if* f *is* γ*-order semismooth at* x^* *for some* $\gamma > 0$*, then* $|x^{k+3} - x^*| = O(|x^k - x^*|^{1+\gamma})$ *for sufficiently large* k*; the sequence* $\{x^k\}$ *converges to* x^* *3-step quadratically if* f *is strongly semismooth at* x^**.*

Here, we only consider the case for $d^+ \cdot d^- > 0$ since the function $\varphi(\cdot)$ we are interested in is nondecreasing. For the case $d^+ \cdot d^- < 0$, one can refer to (Potra et al., 1998, Theorem 3.3).

Note that (Potra et al., 1998, Lemma 4.1) implies that the sequence $\{x^k\}$ generated by Algorithm 1 converges suplinearly with R-order at least $\sqrt[3]{2}$. Next, we will prove that the sequence $\{x^k\}$ generated by Algorithm 1 converges superlinearly to a solution x^* to (4.42) with R-order at least $(1+\sqrt{5})/2$ when f is strongly semismooth at x^* and $\partial f(x^*)$ is a singleton and nondegenerate.

Proposition 4.8. *Let* x^* *be a solution to* (4.42)*. Let* $\{x^k\}$ *be the sequence generated by* (1) *for solving* (4.42)*. For* $k \geq -1$*, denote* $e_k := x^k - x^*$ *and assume that* $|e_k| > 0$*. For* $k \geq -1$*, denote* $c_k := |e_k|/(|e_{k-1}||e_{k-2}|)$*. Assume that* $\partial f(x^*)$ *is a singleton and nondegenerate. It holds that*

(a) *if* f *is semismooth at* x^**, the sequence* $\{x^k\}$ *converges to* x^* *Q-superlinearly;*

(b) *if* f *is strongly semismooth at* x^**, then either one of the following two properties is satisfied: (1)* $\{x^k\}$ *converges to* x^* *superlinearly with Q-order at least* $(1+\sqrt{5})/2$*; (2)* $\{x^k\}$ *converges to* x^* *superlinearly with R-order at least* $(1+\sqrt{5})/2$ *and for any constant* $\underline{C} > 0$*, there exists a subsequence* $\{c_{i_k}\}$ *satisfying* $c_{i_k} < \underline{C} i_k^{-i_k}$*.*

Proof. Let $\mathcal{N}$ and $\mathcal{U}$ be the neighborhoods of x^* specified in Theorem 4.7. Assume that $x^{-1}, x^0 \in \mathcal{U}$. Then Algorithm 1 is well defined and it generates a sequence

$\{x^k\} \subseteq \mathcal{N}$ which converges to x^*. Denote $\partial f(x^*) = \{v^*\}$ for some $v^* \neq 0$. Let d^- and d^+ be the lateral derivatives of f at x^* as defined in (2.21). Then,

$$d^+ = d^- = v^*.$$

Let K_1 be a sufficiently large integer. For all $k \geq K_1$, we have

$$e_{k+1} = \delta_f(x^k, x^{k-1})^{-1}[\delta_f(x^k, x^{k-1}) - \delta_f(x^k, x^*)]e_k. \tag{4.43}$$

(a) Assume that f is semismooth at x^*. We estimate

$$\delta_f(x^k, x^{k-1})^{-1}[\delta_f(x^k, x^{k-1}) - \delta_f(x^k, x^*)]$$

by considering the following two cases.

(a-a) $x^k, x^{k-1} > x^*$ or $x^k, x^{k-1} < x^*$. From Lemma 2.1, we obtain that

$$\begin{aligned} |e_{k+1}| &= |\delta_f(x^k, x^{k-1})^{-1}[\delta_f(x^k, x^{k-1}) - \delta_f(x^k, x^*)]e_k| \\ &= |(v^* + o(1))^{-1}[(v^* + o(1)) - (v^* + o(1))]e_k| \\ &= o(|e_k|). \end{aligned}$$

(a-b) $x^{k-1} < x^* < x^k$ or $x^k < x^* < x^{k-1}$. We will consider the first case. The second case can be treated similarly. By Lemma 2.1, it holds that

$$\begin{aligned} \delta_f(x^k, x^{k-1}) &= \frac{f(x^k) - f(x^*) + f(x^*) - f(x^{k-1})}{x^k - x^{k-1}} \\ &= \frac{(v^* e_k + o(|e_k|)) - (v^* e_{k-1} + o(|e_{k-1}|))}{x^k - x^{k-1}} \\ &= v^* + o(1). \end{aligned}$$

Therefore,

$$\begin{aligned} |e_{k+1}| &= |\delta_f(x^k, x^{k-1})^{-1}[\delta_f(x^k, x^{k-1}) - \delta_f(x^k, x^*)]e_k| \\ &= |(v^* + o(1))^{-1}||(v^* + o(1)) - (v^* + o(1))||e_k| \\ &= o(|e_k|). \end{aligned}$$

Thus, we prove that the sequence $\{x^k\}$ converges to x^* Q-superlinearly.

(b) Now, assume that f is strongly semismooth at x^*. We build the recursion for e_k for sufficiently large integers k by considering the following two cases.

(b-a) $x^k, x^{k-1} > x^*$ or $x^k, x^{k-1} < x^*$. From Lemma 2.1, we obtain

$$\begin{aligned}
|e_{k+1}| &= |\delta_f(x^k, x^{k-1})^{-1}[\delta_f(x^k, x^{k-1}) - \delta_f(x^k, x^*)]e_k| \\
&= |v^* + O(|e_k| + |e_{k-1}|)|^{-1}|(O(|e_k| + |e_{k-1}|))||e_k| \\
&= O(|e_k|(|e_k| + |e_{k-1}|)).
\end{aligned}$$

(b-b) $x^{k-1} < x^* < x^k$ or $x^k < x^* < x^{k-1}$. We will consider the first case. The second case can be treated similarly. By Lemma 2.1, it holds that

$$\begin{aligned}
\delta_f(x^k, x^{k-1}) &= \frac{f(x^k) - f(x^*) + f(x^*) - f(x^{k-1})}{x^k - x^{k-1}} \\
&= \frac{(v^* e_k + O(|e_k|^2)) - (v^* e_{k-1} + O(|e_{k-1}|^2))}{x^k - x^{k-1}} \\
&= v^* + O(|e_k| + |e_{k-1}|)
\end{aligned}$$

and

$$\begin{aligned}
|e_{k+1}| &= |\delta_f(x^k, x^{k-1})^{-1}[\delta_f(x^k, x^{k-1}) - \delta_f(x^k, x^*)]e_k| \\
&= |v^* + O(|e_k| + |e_{k-1}|)|^{-1}|O(|e_k| + |e_{k-1}|)||e_k| \\
&= O(|e_k|(|e_k| + |e_{k-1}|)).
\end{aligned}$$

Therefore, for sufficiently large integers k, we have

$$|e_{k+1}| = O(|e_k|^2 + |e_k||e_{k-1}|) \tag{4.44}$$

and

$$\limsup_{k\to\infty} c_k = O(1/|v^*|) < +\infty. \tag{4.45}$$

Then, there exists a constant $\bar{C} > 0$ and a positive integer K_2 such that

$$|e_{k+1}| \leq \bar{C}|e_k||e_{k-1}|, \quad \forall\, k \geq K_2.$$

Therefore, it follows from (Ortega and Rheinboldt, 1970, Theorem 9.2.9) that $\{x^k\}$ converges to x^* with R-order at least $(1+\sqrt{5})/2$.

If there exists a constant $\underline{C} > 0$ such that $c_k \geq \underline{C}k^{-k}$ for all k sufficiently large, it follows from (Potra, 1989, Corollary 3.1) that $\{x^k\}$ converges to x^* Q-superlinearly with Q-order at least $(1+\sqrt{5})/2$. We complete the proof. □

Proposition 4.9. *Let $p(\cdot)$ be a polyhedral gauge function and λ^* be the solution to (E_φ). Assume that $0 < \lambda_\infty < +\infty$. If $\partial\varphi(\lambda^*)$ is a singleton, the sequence $\{\lambda_k\}$ generated by Algorithm 1 for solving (E_φ) converges to λ^* Q-superlinearly with Q-order at least $(1+\sqrt{5})/2$.*

Proof. The assumption $\partial\varphi(\lambda^*)$ is a singleton implies that $\varphi(\cdot)$ is strictly differentiable at λ^* (Clarke, 1983, Proposition 2.2.4). It follows from Proposition 4.1 that $\varphi'(\lambda^*) > 0$ and $\varphi(\cdot)$ is piecewise C^k for any positive integer $k \geq 1$ in a neighborhood of λ^*.

Choose a sufficiently small $\varsigma > 0$ such that $\mathcal{N}(\lambda^*) := (\lambda^* - \varsigma, \lambda^* + \varsigma) \subseteq (0, \lambda_\infty)$ and $\mathcal{B}(\lambda) \subseteq \mathcal{B}(\lambda^*)$ for any $\lambda \in \mathcal{N}(\lambda^*)$. Denote $\bar{\mathcal{B}} = \bigcup_{\lambda\in\mathcal{N}(\lambda^*)\setminus\lambda^*} \mathcal{B}(\lambda)$. Let $K \in \bar{\mathcal{B}}$ be arbitrarily chosen. Define $\varphi^K : \mathbb{R} \to \mathbb{R}_+$ by

$$\varphi^K(s) := \sqrt{\|P^K b\|^2 + s^2\|h^K\|^2}, \quad s \in \mathbb{R}.$$

By choosing a smaller ς if necessary, we assume that $\{\varphi^K\}_{K\in\bar{\mathcal{B}}}$ is a minimal local representation for $\varphi(\cdot)$ at λ^*. Therefore, it follows from (Qi and Tseng, 2007, Theorem 2) that

$$\varphi'(\lambda^*) = (\varphi^K)'(\lambda^*) = \lambda^*\|h^K\|^2/\varphi(\lambda^*), \quad \forall K \in \bar{\mathcal{B}}.$$

Since $\varphi'(\lambda^*) > 0$, we have

$$\|h^K\| = \|h^{K'}\|,\ \|P^K b\| = \|P^{K'} b\|, \quad \forall K, K' \in \bar{\mathcal{B}}.$$

Therefore, $\varphi(\cdot)$ is C^k on $\mathcal{N}(\lambda^*)$ for any integer $k \geq 1$. It follows from (Traub, 1964, Example 6.1) that $\{\lambda_k\}$ converges to λ^* Q-superlinearly with Q-order at least $(1+\sqrt{5})/2$. □

We give the following example to show that a function satisfying the assumptions in (b) of Theorem 4.8 is not necessarily piecewise smooth:

$$f(x) = \begin{cases} \kappa x, & \text{if } x < 0, \\ -\frac{1}{3}\left(\frac{1}{4^k}\right) + (1 + \frac{1}{2^k})x, & \text{if } x \in \left[\frac{1}{2^{k+1}}, \frac{1}{2^k}\right] \quad \forall k \geq 0, \\ 2x - \frac{1}{3} & \text{if } x > 1, \end{cases} \tag{4.46}$$

where κ is a given constant.

Proposition 4.10. *The function $f(\cdot)$ defined in (4.46) is strongly semismooth at $x = 0$ but not piecewise smooth in the neighborhood of $x = 0$.*

Proof. By the construction of $f(\cdot)$, we know that it is not piecewise smooth in the neighborhood of $x = 0$ since there are infinitely many non-differentiable points. Next, we show that $f(\cdot)$ is strongly semismooth at $x = 0$.

Firstly, it is not difficult to verify that $f(\cdot)$ is Lipschitz continuous with modulus $L = \max\{|\kappa|,\ 2\}$. Secondly, we know that $f'(0; -1) = \kappa$, and for any integer $k \geq 0$, it holds that

$$1 + \frac{1}{3 \times 2^k} \leq f(x)/x \leq \left(1 + \frac{1}{3 \times 2^{k-1}}\right) \quad \forall\, x \in \left[\frac{1}{2^{k+1}}, \frac{1}{2^k}\right],$$

which implies that

$$f'(x; 1) = \lim_{x \downarrow 0} f(x)/x = 1.$$

Therefore, $f(\cdot)$ is directionally differentiable at $x = 0$. Note that both $\partial f(0)$ and $\partial_{\mathrm{B}} f(0)$ are singleton when $\kappa = 1$.

Next, we show that $f(\cdot)$ is strongly G-semismooth at $x = 0$. On the one hand, for any $x < 0$, we have

$$|f(x) - f(0) - \kappa x| \; = |\kappa x - \kappa x| \; = 0.$$

On the other hand, for any integer $k \geq 1$ and $x \in \left[\frac{1}{2^{k+1}}, \frac{1}{2^k}\right]$, we know

$$1 + 2^{-k} \leq |v| \leq 1 + 2^{1-k}, \quad \forall\, v \in \partial f(x),$$

which implies that

$$|f(x) - f(0) - v(x)x| = \left| -\frac{1}{3}\left(\frac{1}{4^k}\right) + \left(1 + \frac{1}{2^k}\right) x - v(x)x \right| \leq \frac{1}{2^{k-1}} x \leq 4x^2.$$

Therefore,

$$|f(x) - f(0) - v(x)x| = O(|x|^2), \quad \forall\, x \to 0.$$

The proof of the proposition is completed. □

4.4 The level set methods for the sparse optimization problem with least squares constraints

In this section, we propose a globally convergent secant method for solving (CP(ϱ)) via finding the root of (E_φ), called SMOP. Alternatively, we can replace the secant method with a semismooth Newton method based on the HS-Jacobian, when the HS-Jacobian is available. We refer to this variant as NMOP. To better illustrate the efficiency of SMOP, we will compare its performance to NMOP for solving (E_φ) on the least squares constrained Lasso problem, where the HS-Jacobian is available. The numerical results will show that the secant method and the semismooth Newton method are comparable for solving the least squares constrained Lasso problem, which also demonstrates the high efficiency of the secant method even for the case that the HS-Jacobian can be computed.

The detailed steps of SMOP are described in Algorithm 7 and the convergence properties are given in Theorem 4.2.

Algorithm 7 A globally convergent secant method for (CP(ϱ))

1: **Input**: $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^n$, $\mu \in (0,1)$, $\lambda_{-1}, \lambda_0, \lambda_1$ in $(0, \lambda_\infty)$ satisfying $\varphi(\lambda_0) > \varrho$, and $\varphi(\lambda_{-1}) < \varrho$.
2: **Initialization**: Set $i = 0$, $\underline{\lambda} = \lambda_{-1}$, and $\overline{\lambda} = \lambda_0$.
3: **for** $k = 1, 2, \dots$ **do**
4: Compute

$$\hat{\lambda}_{k+1} = \lambda_k - \frac{\lambda_k - \lambda_{k-1}}{\varphi(\lambda_k) - \varphi(\lambda_{k-1})}(\varphi(\lambda_k) - \varrho). \tag{4.47}$$

5: **if** $\hat{\lambda}_{k+1} \in [\lambda_{-1}, \lambda_0]$ **then**
6: Compute $x(\hat{\lambda}_{k+1})$ and $\varphi(\hat{\lambda}_{k+1})$. Set $i = i + 1$.
7: **if** either (i) or (ii) holds: (i) $i \geq 3$ and $|\varphi(\hat{\lambda}_{k+1}) - \varrho| \leq \mu|\varphi(\lambda_{k-2}) - \varrho|$ (ii) $i < 3$, **then** set $\lambda_{k+1} = \hat{\lambda}_{k+1}$, $x(\lambda_{k+1}) = x(\hat{\lambda}_{k+1})$; **else** go to line 9.
8: **else**
9: **if** $\varphi(\hat{\lambda}_{k+1}) > \varrho$, **then** set $\overline{\lambda} = \min\{\overline{\lambda}, \hat{\lambda}_{k+1}\}$; **else** set $\underline{\lambda} = \max\{\underline{\lambda}, \hat{\lambda}_{k+1}\}$.
10: Set $\lambda_{k+1} = 1/2(\overline{\lambda} + \underline{\lambda})$. Compute $x(\lambda_{k+1})$ and $\varphi(\lambda_{k+1})$. Set $i = 0$.
11: **end if**
12: **if** $\varphi(\lambda_{k+1}) > \varrho$, **then** set $\overline{\lambda} = \min\{\overline{\lambda}, \lambda_{k+1}\}$; **else** set $\underline{\lambda} = \max\{\underline{\lambda}, \lambda_{k+1}\}$.
13: **end for**
14: **Output**: $x(\lambda_k)$ and λ_k.

Theorem 4.2. *Let $p(\cdot)$ be a gauge function and assume that $0 < \lambda_\infty < +\infty$. Denote λ^* as the solution to (E_φ). Then Algorithm 7 is well defined and the sequences $\{\lambda_k\}$ and $\{x(\lambda_k)\}$ converge to λ^* and a solution $x(\lambda^*)$ to* (CP(ϱ)), *respectively. Denote $e_k = \lambda_k - \lambda^*$ for all $k \geq 1$. Suppose that both d^+ and d^- of $\varphi(\cdot)$ at λ^* as defined in* (2.21) *are positive, the following properties hold for all sufficiently large integer k:*

(a) If $\varphi(\cdot)$ is semismooth at λ^*, then $|e_{k+3}| = o(|e_k|)$;

(b) if $\varphi(\cdot)$ is γ-order semismooth at λ^* for some $\gamma > 0$, then $|e_{k+3}| = O(|e_k|^{1+\gamma})$;

(c) if $\partial\varphi(\lambda^*)$ is a singleton and $\varphi(\cdot)$ is semismooth at λ^*, then $\{e_k\}$ converges to zero Q-superlinearly; if $p(\cdot)$ is further assumed to be polyhedral and $\partial\varphi(\lambda^*)$ is a singleton, then $\{e_k\}$ converges to zero superlinearly with Q-order $(1+\sqrt{5})/2$.

Proof. When $p(\cdot)$ is a gauge function, we know from Proposition 4.3 that $\varphi(\cdot)$ is strictly increasing on $(0, \lambda_\infty]$, which implies that the sequences $\{\hat{\lambda}_k\}$ and $\{\lambda_k\}$ generated in Algorithm 7 are well defined. For any $k \geq 1$, if we run the algorithm for three more iterations, then it holds that either (1) $(\overline{\lambda} - \underline{\lambda})$ will reduce at least half; or (2) $|\varphi(\lambda_{k+3}) - \varrho| \leq \mu|\varphi(\lambda_k) - \varrho|$. Therefore, the sequence $\{\lambda_k\}$ will converge to λ^*. Suppose that $\varphi(\cdot)$ is semismooth at λ^* and both d^+ and d^- are positive. We know from Theorem 4.7 that there exists a positive integer $k_{\max}$ such that for all $k \geq k_{\max}$, $\hat{\lambda}_k \in [\lambda_{-1}, \lambda_0]$ and

$$|\hat{\lambda}_{k+3} - \lambda^*| = o(|\lambda_k - x^*|). \tag{4.48}$$

Therefore, it follows from Lemma 2.1 that

$$\frac{|\varphi(\hat{\lambda}_{k+3}) - \varrho|}{|\varphi(\lambda_k) - \varrho|} = \frac{\delta_\varphi(\hat{\lambda}_{k+3}, \lambda^*)}{\delta_\varphi(\lambda_k, \lambda^*)} \times \frac{|\hat{\lambda}_{k+3} - \lambda^*|}{|\lambda_k - \lambda^*|} \leq \frac{\max\{d^+, d^-\} + o(1)}{\min\{d^+, d^-\} + o(1)} \times o(1).$$

Thus, for all $k \geq k_{\max}$,

$$|\varphi(\hat{\lambda}_{k+3}) - \varrho| \leq \mu|\varphi(\lambda_k) - \varrho|.$$

The rest of the proof of this theorem follows from Theorem 4.7 and Proposition 4.9. $\square$

Chapter 5

Numerical experiments

In this chapter, we will present numerical results to demonstrate the high efficiency of our proposed SMOP. We will focus on solving (CP(ϱ)) with two objective functions: (1) the ℓ_1 penalty: $p(x) = \|x\|_1$, $x \in \mathbb{R}^n$; (2) the sorted ℓ_1 penalty: $p(x) = \sum_{i=1}^n \gamma_i |x|_{(i)}$, $x \in \mathbb{R}^n$ with given parameters $\gamma_1 \geq \gamma_2 \geq \cdots \geq \gamma_n \geq 0$ and $\gamma_1 > 0$, where $|x|_{(1)} \geq |x|_{(2)} \geq \cdots \geq |x|_{(n)}$, which serve as illustrative examples to highlight the efficiency of our algorithm. It is worthwhile mentioning that the sorted ℓ_1 penalty is not separable. For demonstration purposes only, we will test the performance of SMOP when $p(\cdot)$ is a non-polyhedral function at the last section.

Table 5.1: Statistics of the UCI test instances.

Problem idx	Name	m	n	Sparsity(A)	norm(b)
1	E2006.train	16087	150360	0.0083	452.8605
2	log1p.E2006.train	16087	4272227	0.0014	452.8605
3	E2006.test	3308	150358	0.0092	221.8758
4	log1p.E2006.test	3308	4272226	0.0016	221.8758
5	pyrim5	74	201376	0.5405	5.7768
6	triazines4	186	635376	0.6569	9.1455
7	bodyfat7	252	116280	1.0000	16.7594
8	housing7	506	77520	1.0000	547.3813

In our numerical experiments, we measure the accuracy of the obtained solution

$\tilde{x}$ for (CP(ϱ)) by the following relative residual:

$$\eta := \frac{|\tilde{\varphi} - \varrho|}{\max\{1,\ \varrho\}},$$

where $\tilde{\varphi} := \|A\tilde{x} - b\|$. We test all algorithms on datasets from UCI Machine Learning Repository as in (Li et al., 2018a,b), which are originally obtained from the LIBSVM datasets (Chih-Chung, 2011). Table 5.1 presents the statistics of the tested UCI instances. All our computational results are obtained using MATLAB R2023a on a Windows workstation with the following specifications: 12-core Intel(R) Core(TM) i7-12700 (2.10GHz) processor, and 64 GB of RAM. In all the tables presented in this section, nnz(x) represents the number of elements in the solution x obtained by SMOP (with a stopping tolerance of 10^{-6}) for solving (CP(ϱ)) that have an absolute value greater than 10^{-8}. Besides, We denote BMOP (NMOP) as the root finding based bisection method (hybrid of the bisection method and the semismooth Newton method) for solving the optimization problem (CP(ϱ)).

5.1 The ℓ_1 penalized problems with least squares constraints

In this section, we focus on the problem (CP(ϱ)) with $p(\cdot) = \|\cdot\|_1$. We will compare the efficiency of SMOP to the state-of-the-art SSNAL-LSM algorithm (Li et al., 2018b), SPGL1 solver (Van den Berg and Friedlander, 2008, 2019) and ADMM (Section 2.4). Moreover, we perform experiments to demonstrate that our secant method is considerably more efficient than the bisection method for root finding while performing on par with the semismooth Newton method, where the HS-Jacobian is computable.

In practice, we have multiple choices for solving the subproblems in SMOP. In our experiments, we use the squared smoothing Newton method (Section 3.2.2) and

Table 5.2: The values of c to obtain $\varrho = c\|b\|$ for the ℓ_1 penalized problems with least squares constraints. In this table, $c_{LS} = \frac{\lambda^*}{\|A^T b\|_\infty}$ represents the regularization parameter for the corresponding $\mathrm{P_{LS}}(\lambda^*)$, where the optimal solution λ^* to $\varphi(\lambda) = \varrho$ is obtained by SMOP.

Test	idx	1	2	3	4	5	6	7	8
I	c	0.1	0.1	0.08	0.08	0.05	0.1	0.001	0.1
	nnz(x)	339	110	246	405	79	655	107	148
	c_{LS}	2.6-7	2.8-4	4.2-7	2.1-4	5.7-3	2.8-3	1.1-6	1.3-3
II	c	0.09	0.09	0.06	0.06	0.015	0.03	0.0001	0.04
	nnz(x)	1387	1475	884	1196	92	497	231	377
	c_{LS}	1.1-7	6.2-5	1.7-7	9.6-5	3.0-4	5.6-5	3.8-8	3.0-5

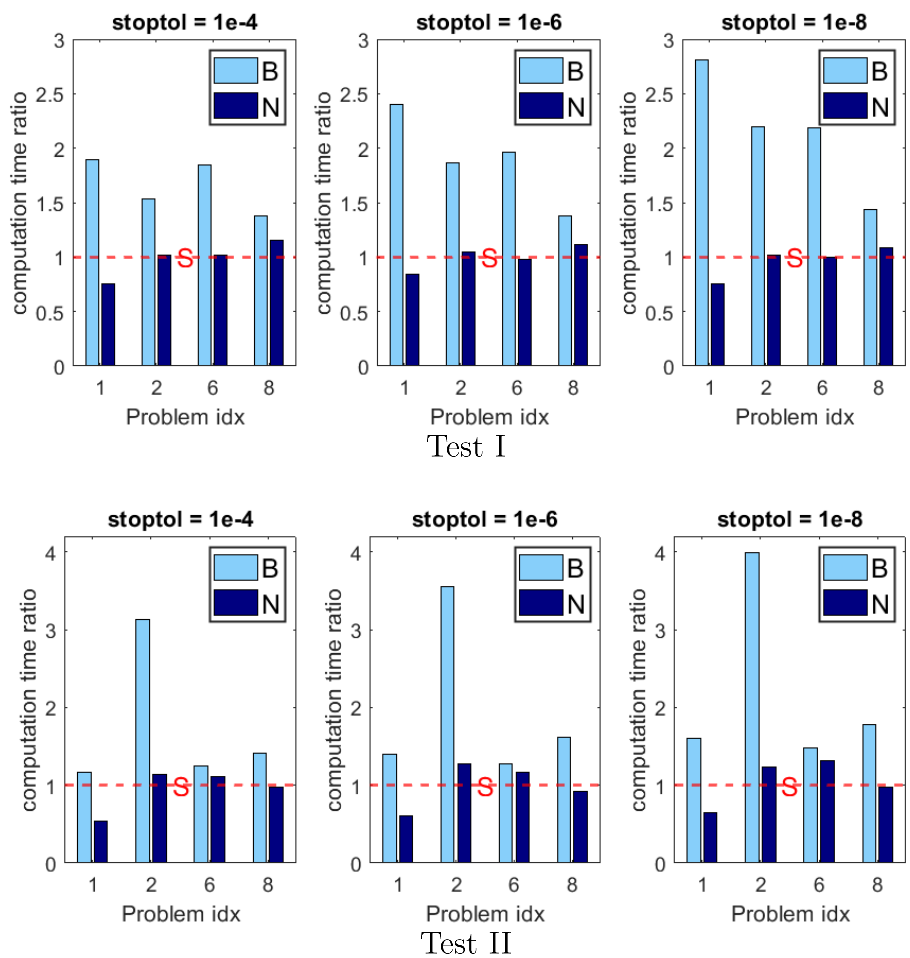


Figure 5.1: The ratio of the computation time between BMOP (B) and NMOP (N) to the computation time of SMOP in solving (CP(ϱ)) with the ℓ_1 regularization.

SSNAL (Li et al., 2018a) to solve the subproblems in SMOP. The maximum number of iterations for SPGL1, SSNAL, and ADMM is set to 100,000, while for SMOP and SSNAL-LSM, the maximum number of iterations of the outermost loop is set to 200. Additionally, we have set the maximum running time to 1 hour. To select ϱ, we use the values of c in Table 5.2 for each instance listed in Table 5.1, and let $\varrho = c\|b\|$. At last, we point out that the adaptive sieving technique is not employed in the SSNAL-LSM.

Table 5.3: The performance of SMOP (A1), SSNAL-LSM (A2), SPGL1 (A3) and ADMM (A4), in solving the ℓ_1 penalized problems with least squares constraints (CP(ϱ)) with $\varrho = c\|b\|$, where the specific value of c for each problem is listed in Table 5.2. In the table, the underline is used to mark cases where the algorithm fails to reach the given tolerance. For simplicity, we omit the "e" in the scientific notation.

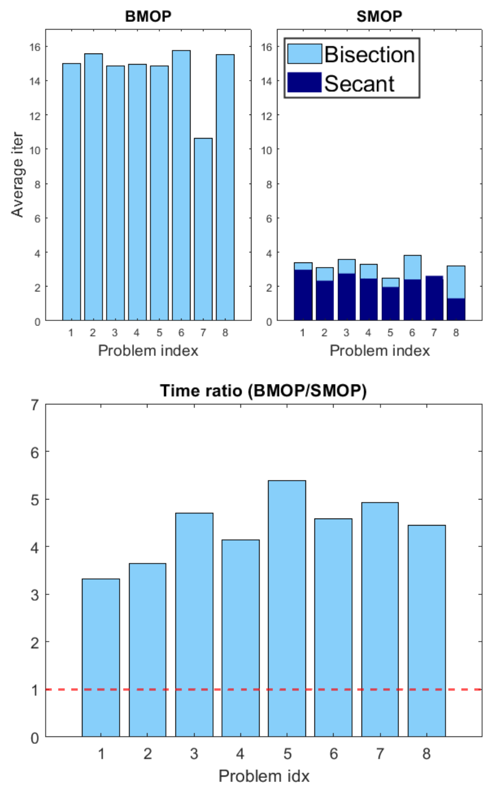
idx	time (s)				η				outermost iter			
	A1	A2	A3	A4	A1	A2	A3	A4	A1	A2	A3	A4
	Test I with stoptol = 10^{-4}											
1	1.39+0	2.18+2	3.51+2	4.22+2	2.3-5	4.9-5	1.0-4	1.0-4	24	29	7342	2049
2	2.29+0	5.12+2	1.45+3	6.84+2	3.1-6	7.8-5	9.0-5	8.7-5	12	16	3445	1470
3	4.02-1	5.83+1	3.21+2	8.87+1	9.4-6	2.6-5	1.0-4	1.0-4	24	30	21094	4918
4	1.59+0	2.06+2	7.19+2	9.90+1	1.2-5	7.3-5	9.5-5	1.3-5	13	15	3174	854
5	2.73-1	1.20+1	9.81+0	5.63+0	6.9-6	5.4-6	7.4-5	2.2-5	6	14	498	273
6	2.32+0	1.74+2	3.35+2	1.01+2	5.8-6	4.4-5	9.1-5	7.5-5	9	17	1987	571
7	4.35-1	9.12+0	8.98+0	8.59+0	2.8-5	5.9-5	9.8-5	9.9-5	15	18	539	583
8	2.99-1	9.07+0	1.29+1	7.94+0	2.6-5	8.6-5	1.0-4	9.0-5	10	14	515	424
	Test I with stoptol = 10^{-6}											
1	1.45+0	3.22+2	1.51+3	7.06+2	2.5-7	6.1-8	9.9-7	1.0-6	25	36	28172	3539
2	2.52+0	6.68+2	1.75+3	3.42+3	9.9-8	3.5-8	9.2-7	9.9-7	13	24	4155	8725
3	4.12-1	7.40+1	2.11+3	1.81+2	1.1-8	2.3-7	6.2-6	1.0-6	25	35	100000	10100
4	1.72+0	3.40+2	1.04+3	4.03+2	1.3-9	5.7-7	7.2-7	7.9-7	14	26	4584	3820
5	2.93-1	1.61+1	4.58+1	3.95+2	1.0-7	6.0-8	9.1-7	9.8-7	7	19	2468	20155
6	2.47+0	2.13+2	8.24+2	2.31+3	3.0-7	4.0-7	8.2-7	3.4-7	10	23	5578	13672
7	4.68-1	1.18+1	9.11+0	1.85+1	1.9-9	9.6-7	2.7-7	9.9-7	17	22	544	1250
8	3.28-1	1.45+1	3.84+1	4.40+1	2.4-7	8.4-8	4.0-7	8.7-7	11	24	1539	2427
	Test II with stoptol = 10^{-4}											
1	7.26+0	4.51+2	1.38+3	6.12+2	3.0-6	4.6-5	1.0-4	1.0-4	26	30	27775	3014
2	6.79+0	1.54+3	1.32+3	4.01+2	1.8-5	3.6-5	9.7-5	6.8-5	14	21	3000	733
3	3.51+0	1.84+2	1.50+3	1.34+2	1.3-5	2.3-5	8.7-2	1.0-4	25	29	100000	7333
4	2.91+0	6.91+2	6.23+2	4.94+1	7.5-6	3.6-6	9.6-5	5.8-5	14	22	2694	385
5	6.23-1	1.53+1	8.65+0	2.01+1	2.8-5	7.9-6	6.6-5	9.5-5	9	13	395	1000
6	9.02+0	3.46+2	3.60+3	3.82+2	6.8-6	3.7-5	7.6-2	9.9-5	12	17	24924	2232
7	1.50+0	1.59+1	3.06+2	3.39+1	1.6-5	8.7-6	9.9-5	9.8-5	12	18	19820	2340
8	2.37+0	1.90+1	1.69+2	1.19+1	1.4-6	8.9-5	9.1-5	9.8-5	13	18	5914	644
	Test II with stoptol = 10^{-6}											
1	7.23+0	5.96+2	3.60+3	8.82+2	3.7-9	2.9-7	3.6-2	1.0-6	27	35	62384	4453
2	7.37+0	1.85+3	2.04+3	1.46+3	1.4-7	3.9-7	9.7-7	1.0-7	15	27	4688	3464
3	3.59+0	2.36+2	1.49+3	1.99+2	8.1-10	8.3-7	8.7-2	1.0-6	26	36	100000	11051
4	3.02+0	8.44+2	1.37+3	2.18+2	3.1-9	4.3-7	9.9-7	6.0-7	15	28	5912	1980
5	6.37-1	2.49+1	4.14+2	1.48+2	2.4-7	3.0-8	8.7-7	9.7-7	10	22	22091	7592
6	9.37+0	4.25+2	3.60+3	3.60+3	5.4-11	6.7-7	7.5-2	1.2-7	14	22	25158	21556
7	1.59+0	2.09+1	3.37+2	8.54+1	3.2-7	1.9-8	8.8-7	9.7-7	13	23	21523	5817
8	2.39+0	2.68+1	1.65+3	3.34+1	4.5-7	6.9-7	8.8-7	9.8-7	14	26	59147	1834

We compare SMOP to SPGL1, SSNAL-LSM and ADMM to solve (CP(ϱ)) with the tolerances of 10^{-4} and 10^{-6}, respectively. The test results are presented in Table

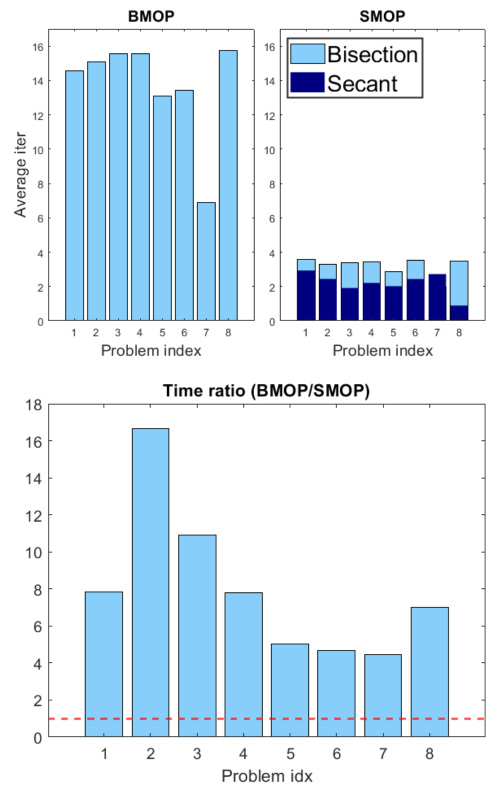
5.3. These results indicate that SMOP successfully solves all the tested instances and outperforms SSNAL-LSM, SPGL1 and ADMM. It can be seen from Table 5.3 that SMOP can achieve a speed-up of up to 1,000 times compared to SPGL1 for the problems that can be solved by SPGL1 (a significant number of instances cannot be solved by SPGL1 to the required accuracy). Regarding ADMM, SMOP remains significantly superior in terms of efficiency for all cases, with a speed-up of over 1300 times. Furthermore, compared to SSNAL-LSM, SMOP also shows vast superiority, with efficiency improvements up to more than 260 times. Note that SPGL1 has two modes: the primal mode (denoted by SPGL1) and the hybrid mode (denoted by SPGL1_H). We do not print the results of SPGL1_H since SPGL1 outperforms SPGL1_H in most of the cases in our tests.

Subsequently, we perform numerical experiments to compare the performance of the secant method to the bisection method and the HS-Jacobian based semismooth Newton method for finding the root of (E_φ) to further illustrate the efficiency of SMOP. Figure 5.1 presents the ratio of the computation time between BMOP and NMOP to the computation time of SMOP on solving (CP(ϱ)) for some instances. The numerical results show that SMOP easily beats BMOP, with a large margin when a higher precision solution is required. The results also reveal that SMOP performs comparably to NMOP for solving the ℓ_1 penalized least squares constrained problem, in which the HS-Jacobian are computable. This is another strong evidence to illustrate the efficiency of SMOP.

Next we perform tests on BMOP and SMOP to generate a solution path for (CP(ϱ)) involving multiple choices of tuning parameters $\varrho > 0$. In this test, we solve (CP(ϱ)) with $\varrho_i = c_i \cdot c\|b\|$, $i = 1, \ldots, 100$, where $c_i = 1.5 - 0.5 \times (i-1)/99$ and c is the same constant as in Table 5.2. In this test, we apply the warm-start strategy to both algorithms. The average iteration numbers of BMOP and SMOP and the ratio of the computation time of BMOP to the computation time of SMOP are shown in Figure



(a) Test I



(b) Test II

Figure 5.2: The performance of BMOP and SMOP in generating a solution path for (CP(ϱ)) with the ℓ_1 regularization and stopping tolerance 10^{-6}.

5.2 with a tolerance of 10^{-6}. From this figure, it is evident that utilizing the secant method for root-finding significantly reduces the number of iterations by around 4 times. The reduction in iterations results in a substantial decrease in computation time for SMOP, which is typically less than one-third of the time required by BMOP.

5.2 The sorted ℓ_1 penalized problems with least squares constraints

In this section, we will present the numerical results of SMOP in solving the sorted ℓ_1 penalized problems with least squares constraints (CP(ϱ)). For comparison purposes, we also conducted tests on Newt-ALM-LSM (similar to SSNAL-LSM, but with the subproblems solved by Newt-ALM (Luo et al., 2019)) and ADMM (Section 2.4) for (CP(ϱ)).

Table 5.4: The performance of SMOP (A1), Newt-ALM-LSM (A2) and ADMM (A4), in solving the sorted ℓ_1 penalized problems with least squares constraints (CP(ϱ)) with $\varrho = c\|b\|$. In the table, $c_{LS} = \frac{\lambda^*}{\|A^T b\|_\infty}$ represents the regularization parameter for the corresponding $\mathrm{P_{LS}}(\lambda^*)$, where the optimal solution λ^* to $\varphi(\lambda) = \varrho$ is obtained by SMOP. The stopping tolerance is set to 10^{-6} and the underline is used to mark cases where the algorithm fails to reach the given tolerance. For simplicity, we omit the "e" in the scientific notation.

idx	c | nnz(x) | c_{LS}	time (s)	η	outermost iter
		A1 | A2 | A4	A1 | A2 | A4	A1 | A2 | A4
Test I				
2	0.15 | 3 | 2.4-2	3.84+0 | 1.34+2 | 3.60+3	1.1-7 | 5.3-7 | 2.8-1	8 | 21 | 8637
4	0.1 | 3 | 4.8-3	4.79+0 | 1.35+2 | 3.60+3	6.0-7 | 8.9-7 | 2.9-4	10 | 17 | 28891
5	0.1 | 113 | 1.9-2	6.29−1 | 4.98+1 | 4.23+2	1.0-7 | 4.5-7 | 1.5-7	7 | 22 | 17974
6	0.15 | 413 | 1.0-2	3.10+0 | 2.43+2 | 3.60+3	2.7-7 | 1.6-7 | 1.9-4	9 | 21 | 19071
7	0.002 | 22 | 1.9-5	3.56−1 | 1.67+1 | 2.44+1	3.6-9 | 6.0-7 | 9.9-7	14 | 22 | 1616
8	0.15 | 95 | 6.9-3	6.06−1 | 2.55+1 | 1.57+2	1.3-7 | 7.7-7 | 9.0-7	10 | 23 | 8329
Test II				
1	0.1 | 339 | 2.6-7	2.53+1 | 1.40+2 | 5.13+2	2.9-7 | 5.6-7 | 1.0-6	25 | 34 | 2490
2	0.095 | 629 | 1.0-4	5.39+1 | 4.82+2 | 2.87+3	1.7-7 | 2.9-7 | 9.4-7	17 | 27 | 6770
3	0.08 | 246 | 4.2-7	4.98+0 | 6.54+1 | 1.60+2	2.0-8 | 7.1-7 | 1.0-6	25 | 36 | 8491
4	0.07 | 758 | 1.4-4	2.26+1 | 4.26+2 | 5.86+2	4.0-8 | 9.0-7 | 9.8-7	16 | 27 | 4550
5	0.02 | 95 | 5.7-4	2.05+0 | 9.87+1 | 3.58+2	3.2-8 | 5.6-7 | 7.6-7	11 | 20 | 15582
6	0.05 | 997 | 5.5-4	2.32+1 | 1.04+3 | 3.60+3	8.4-7 | 2.1-7 | 3.5-6	10 | 23 | 19159
7	0.001 | 107 | 1.1-6	1.02+0 | 2.85+1 | 1.30+1	5.9-8 | 6.9-9 | 9.5-7	17 | 22 | 826
8	0.08 | 206 | 4.3-4	3.38+0 | 1.03+2 | 5.58+1	5.7-9 | 7.4-7 | 3.8-7	13 | 25 | 2842

In our numerical experiments, we choose the parameters $\gamma_i = 1 - (i-1)/(n-1)$, $i = 1, \ldots, n$, in the sorted ℓ_1 penalty function $p(x) = \sum_{i=1}^n \gamma_i |x|_{(i)}$, $x \in \mathbb{R}^n$. The maximum iteration number is set to 200 for SMOP and Newt-ALM-LSM, and 100,000 for ADMM. The subproblems in SMOP are solved by Newt-ALM in this test. In all experiments within this section, the stopping tolerance is set to 10^{-6}. The results obtained with a tolerance of 10^{-4} are similar to those obtained with a tolerance of 10^{-6}, therefore, we will not present them in order to save space. We set $\varrho = c\|b\|$, where the values of c are specified in Table 5.4. The numerical results are presented in Table 5.4. From the table, it is evident that SMOP outperforms Newt-ALM-LSM and ADMM for all the cases. More specifically, SMOP can be up to around 80 times faster than Newt-ALM-LSM and up to more than 600 times

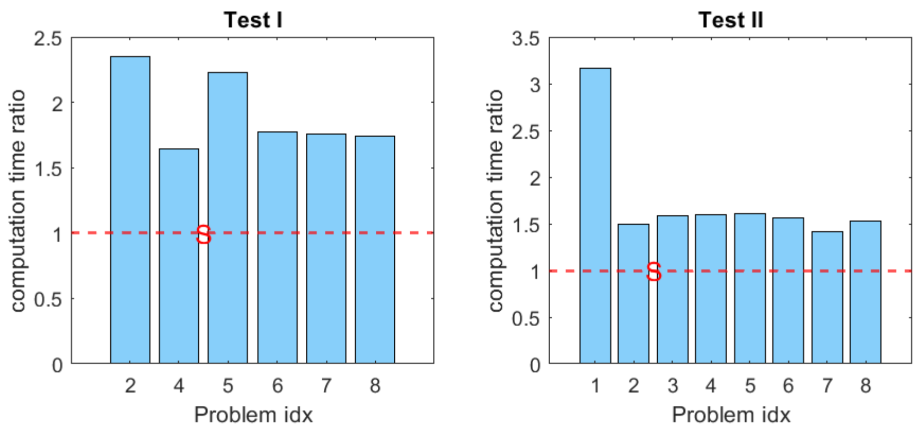


Figure 5.3: The ratio of the computation time between BMOP to the computation time of SMOP in solving (CP(ϱ)) with the sorted ℓ_1 regularization.

faster than ADMM for the problems that can be solved by ADMM. Additionally, Figure 5.3 presents the computation time ratio between BMOP and SMOP for both Test I and Test II. This also demonstrates the significance of the secant method in root-finding for achieving higher efficiency.

5.3 A group lasso penalized problems with least squares constraints

In this ection, we will present the numerical experiments conducted to solve a group lasso penalized problems with least squares constraints. The purpose of this demonstration is to illustrate the potential and high efficiency of our proposed secant method in solving the equation $\varphi(\lambda) = \varrho$ for the non-polyhedral function penalized problems with least squares constraints. We will compare our algorithm, SMOP, with other state-of-the-art algorithms to demonstrate its high efficiency and robustness.

We consider the following penalty function $p(\cdot)$ in this section:

$$p(x) = \sum_{t=1}^{l} \sqrt{x_{2t-1}^2 + x_{2t}^2}, \quad x \in \mathbb{R}^{2l}. \tag{5.1}$$

For the purpose of demonstration, we will keep using the UCI dataset that was utilized in the previous two subsections. However, it is necessary to ensure that the

Table 5.5: The values of c to obtain $\varrho = c\|b\|$ for the group lasso penalized problems with least squares constraints. In the table, $c_{LS} = \frac{\lambda^*}{\|A^T b\|_\infty}$ represents the regularization parameter for the corresponding $\mathrm{P_{LS}}(\lambda^*)$, where the optimal solution λ^* to $\varphi(\lambda) = \varrho$ is obtained by SMOP.

	idx	c	nnz(x)	c_{LS}
Test I	4	0.1	6	4.4-3
	5	0.1	50	2.4-2
	6	0.15	138	1.3-2
	7	0.002	28	2.4-5
	8	0.15	66	8.4-3
Test II	1	0.105	95	7.5-7
	3	0.08	403	4.3-7
	4	0.08	731	2.2-4
	5	0.02	120	9.1-4
	6	0.05	372	6.3-4
	7	0.001	186	1.3-6
	8	0.08	260	4.9-4

value of n is even. Next, we group the i-th and $(i+1)$-th elements together for all $i = 1, 3, \cdots, n-1$. The values of c utilized to obtain $\varrho = c\|b\|$ are presented in Table 5.5. In SMOP, the subproblems are solved by SSNAL (Zhang et al., 2020). The maximum iteration number for both SMOP and SSNAL-LSM is set to 200, while for SPGL1 and ADMM, their maximum iteration number is set to 100,000. As for the maximum running time, it remains set at 1 hour. Next, we will compare SMOP with the state-of-the-art algorithms SSNAL-LSM, SPGL1, and ADMM. The results of the tests are presented in Table 5.6. From the table, it is evident that SMOP outperforms SSNAL-LSM, SPGL1, and ADMM with speed-ups of up to 300, 900, and 1,100, respectively. In addition, Figure 5.4 illustrates the ratio of computation time between BMOP and SMOP. This figure clearly shows that using the secant method can greatly enhance overall efficiency, resulting in a speed improvement of approximately 1.5-4 times, even when dealing with the non-polyhedral penalty function (5.1).

Table 5.6: The performance of SMOP (A1), SSNAL-LSM (A2), SPGL1 (A3) and ADMM (A4), in solving the group lasso penalized problems with least squares constraints (CP(ϱ)) with $\varrho = c\|b\|$. The stopping tolerance is set to 10^{-6} and the underline is used to mark cases where the algorithm fails to reach the given tolerance. For simplicity, we omit the "e" in the scientific notation.

idx	time (s)				η				outermost iter			
	A1	A2	A3	A4	A1	A2	A3	A4	A1	A2	A3	A4
Test I												
4	3.75+0	1.16+2	8.49+2	3.60+3	1.3−7	3.1-7	6.37-7	7.2-5	11	21	3024	22125
5	8.14−1	2.74+2	2.96+1	9.16+2	1.4−9	3.5-7	6.05-7	1.0-6	11	21	1319	38530
6	5.19+0	1.46+3	1.70+2	3.02+3	3.2-10	4.5-7	5.98-7	9.8-7	10	22	1086	15768
7	5.98−1	8.80+0	3.02+1	2.59+1	3.7−8	5.0-7	2.07-7	1.0-6	14	19	2102	1627
8	6.88−1	1.41+2	8.30+0	1.19+2	1.8−8	2.6-7	2.46-7	9.6-7	9	22	334	6211
Test II												
1	3.29+0	4.33+1	3.18+3	1.12+3	2.7-7	2.7-7	9.8-7	1.0-6	24	29	55596	5826
3	3.83+0	3.00+1	2.06+3	2.57+2	1.3-7	3.4-7	3.8-6	1.0-6	22	36	100000	13031
4	2.97+1	2.42+3	1.19+3	5.86+2	5.2-7	9.6-9	8.6-7	7.4-7	13	27	4241	3401
5	1.70+0	1.29+2	3.30+2	9.27+1	8.0-7	1.7-8	8.9-7	6.5-7	9	20	18001	3959
6	2.51+1	1.39+3	3.60+3	3.60+3	1.3-8	1.4-7	5.8-5	2.6-7	11	22	20646	19075
7	1.22+0	1.88+1	5.99+2	2.69+1	5.3-8	2.1-8	6.9-7	9.9-7	15	23	41578	1685
8	5.75+0	1.47+2	1.14+2	1.94+2	2.5-7	3.7-7	4.4-7	9.8-7	15	25	4373	9974

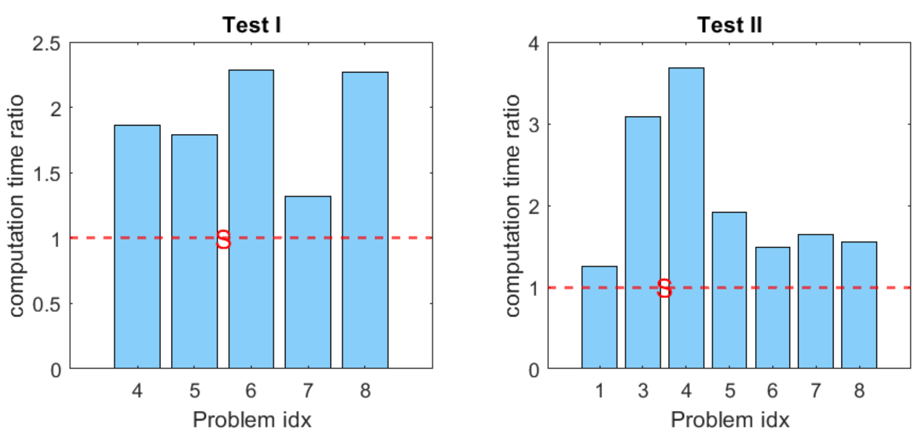


Figure 5.4: The ratio of the computation time between BMOP to the computation time of SMOP in solving (CP(ϱ)) with the group lasso regularization.

Chapter 6

Conclusions and future work

In this thesis, we have developed an efficient sieving based secant method for solving the sparse optimization problems with least squares constraints (CP(ϱ)).

In each iteration of the proposed algorithm, computing the value function $\varphi(\cdot)$ involves solving an unconstrained sparse optimization problem ($\mathrm{P}_{\mathrm{LS}}(\lambda)$). Consequently, we need to solve ($\mathrm{P}_{\mathrm{LS}}(\lambda)$) with a sequence of penalty parameters. This allows for the natural implementation of the adaptive sieving strategy (Yuan et al., 2023, 2022), where each reduced subproblem (except the first) is warm-started using the solution from the previous iteration. Note that, the dimension of the reduced subproblems could be much smaller than the number of rows in A, solving the primal problem may provide greater benefits than addressing the dual problem. We have introduced a smoothing Newton method to directly solve the primal problem, which will be combined with the existing dual-based algorithms to serve as solvers for the subproblems in the adaptive sieving strategy. Moreover, we have shown that, when $p(\cdot)$ is the ℓ_1 norm function, the smoothing Newton method converges quadratically to a solution of ($\mathrm{P}_{\mathrm{LS}}(\lambda)$) under the assumption that the LICQ condition for the dual problem holds.

When $p(\cdot)$ is a polyhedral gauge function, we have proven that for any $\bar{\lambda} \in (0, \lambda_\infty)$, all $v \in \partial\varphi(\bar{\lambda})$ are positive. Consequently, when $p(\cdot)$ is a polyhedral gauge function, the secant method can solve (E_φ) with at least a 3-step Q-quadratic convergence rate. We have demonstrated the high efficiency of our method for solving (CP(ϱ)) by two representative instances, specifically, the ℓ_1 and the sorted ℓ_1 penalized constrained problems. Moreover, our numerical results on the ℓ_1 penalized constrained problems, in which the $\partial_{\text{HS}}\varphi(\cdot)$ is computable as shown in Proposition 4.5, have verified that the efficiency of SMOP is not compromised compared to the performance of the HS-Jacobian based semismooth Newton method. This motivates us to use the secant method instead of the semismooth Newton method for solving (E_φ) regardless of the availability of the generalized Jacobians.

We point out that the work done in this thesis is far from comprehensive in addressing sparse optimization problems with least squares constraints. Below, we briefly list several research directions that deserve more explorations.

- Is it theoretically possible to extend the secant method for finding the root of (E_φ) when $p(\cdot)$ is a non-polyhedral function?
- When A is given in operator form rather than matrix form, can one find a dimension reduction technique that efficiently solves ($\text{P}_{\text{LS}}(\lambda)$)?
- Can one design a simpler and better algorithm than the smoothing Newton method to directly address the primal problem ($\text{P}_{\text{LS}}(\lambda)$)?

Bibliography

Aravkin, A. Y., Burke, J. V., Drusvyatskiy, D., Friedlander, M. P., and Roy, S. (2019), "Level-set methods for convex optimization," *Mathematical Programming*, 174, 359–390.

Baraldi, R., Kumar, R., and Aravkin, A. (2019), "Basis pursuit denoise with nonsmooth constraints," *IEEE Transactions on Signal Processing*, 67, 5811–5823.

Beck, A. and Teboulle, M. (2009), "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM Journal on Imaging Sciences*, 2, 183–202.

Bolte, J., Daniilidis, A., and Lewis, A. (2009), "Tame functions are semismooth," *Mathematical Programming*, 117, 5–19.

Bonnans, J. F. and Shapiro, A. (2000), *Perturbation Analysis of Optimization Problems*, Springer Series in Operations Research, Springer-Verlag, New York.

Cai, J.-F., Qu, X. B., Xu, W. Y., and Ye, G.-B. (2016), "Robust recovery of complex exponential signals from random Gaussian projections via low rank Hankel matrix reconstruction," *Applied and Computational Harmonic Analysis*, 41, 470–490.

Cai, T., Wang, L., and Xu, G. W. (2009), "Shifting inequality and recovery of sparse signals," *IEEE Transactions on Signal processing*, 58, 1300–1308.

Cai, T., Liu, W. D., and Luo, X. (2011), "A constrained ℓ_1 minimization approach to sparse precision matrix estimation," *Journal of the American Statistical Association*, 106, 594–607.

Candés, E. and Romberg, J. (2006), "Robust signal recovery from incomplete observations," in *2006 International Conference on Image Processing*, pp. 1281–1284, IEEE.

Candés, E. J. and Wakin, M. B. (2008), "An introduction to compressive sampling," *IEEE Signal Processing Magazine*, 25, 21–30.

Candés, E. J., Romberg, J. K., and Tao, T. (2006), "Stable signal recovery from incomplete and inaccurate measurements," *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences*, 59, 1207–1223.

Chan, Z. X. and Sun, D. F. (2008), "Constraint nondegeneracy, strong regularity, and nonsingularity in semidefinite programming," *SIAM Journal on Optimization*, 19, 370–396.

Chen, S. B. and Donoho, D. (1994), "Basis pursuit," in *Proceedings of 1994 28th Asilomar Conference on Signals, Systems and Computers*, vol. 1, pp. 41–44, IEEE.

Chen, S. S., Donoho, D. L., and Saunders, M. A. (2001), "Atomic decomposition by basis pursuit," *SIAM review*, 43, 129–159.

Chen, X. D., Sun, D. F., and Sun, J. (2003), "Complementarity functions and numerical experiments on some smoothing Newton methods for second-order-cone complementarity problems," *Computational Optimization and Applications*, 25, 39–56.

Chen, X. J., Qi, L. Q., and Sun, D. F. (1998), "Global and superlinear convergence of the smoothing Newton method and its application to general box constrained variational inequalities," *Mathematics of Computation*, 67, 519–540.

Chih-Chung, C. (2011), "LIBSVM: A library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, 2, 1–27.

Clarke, F. H. (1976), "On the inverse function theorem," *Pacific Journal of Mathematics*, 64, 97–102.

Clarke, F. H. (1983), *Optimization and Nonsmooth Analysis*, Wiley, New York.

Coste, M. (2000), *An Introduction to o-minimal Geometry*, Istituti Editoriali E Poligrafici Internazionali Pisa.

Cottle, R. W., Pang, J.-S., and Stone, R. E. (1992), *The linear complementarity problem*, SIAM.

Cox, D. R. (1958), "The regression analysis of binary sequences," *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 20, 215–232.

Ding, C., Sun, D. F., and Toh, K.-C. (2014), "An introduction to a class of matrix cone programming," *Mathematical Programming*, 144, 141–179.

Donoho, D. L. and Elad, M. (2003), "Optimally sparse representation in general (nonorthogonal) dictionaries via ℓ_1 minimization," *Proceedings of the National Academy of Sciences*, 100, 2197–2202.

Donoho, D. L. and Elad, M. (2006), "On the stability of the basis pursuit in the presence of noise," *Signal Processing*, 86, 511–532.

Donoho, D. L., Elad, M., and Temlyakov, V. N. (2005), "Stable recovery of sparse overcomplete representations in the presence of noise," *IEEE Transactions on information theory*, 52, 6–18.

Eckstein, J. and Bertsekas, D. P. (1992), "On the Douglas—Rachford splitting method and the proximal point algorithm for maximal monotone operators," *Mathematical Programming*, 55, 293–318.

El Ghaoui, L. (2012), "Safe feature elimination for the lasso and sparse supervised learning problems," *Pacific Journal of Optimization*, 8, 667.

Facchinei, F. and Pang, J.-S. (2003a), *Finite-dimensional Variational Inequalities and Complementarity Problems*, Springer.

Facchinei, F. and Pang, J.-S. (2003b), *Finite-dimensional Variational Inequalities and Complementarity Problems*, Springer Series in Operations Research, Springer-Verlag, New York.

Fazel, M., Pong, T. K., Sun, D. F., and Tseng, P. (2013), "Hankel matrix rank minimization with applications to system identification and realization," *SIAM Journal on Matrix Analysis and Applications*, 34, 946–977.

Fischer, A. (1997), "Solution of monotone complementarity problems with locally Lipschitzian functions," *Mathematical Programming*, 76, 513–532.

Fortin, M. and Glowinski, R. (1983), "Chapter III on decomposition-coordination methods using an augmented lagrangian," in *Studies in Mathematics and Its Applications*, vol. 15, pp. 97–146, Elsevier.

Friedlander, M. P. and Tseng, P. (2007), "Exact regularization of convex programs," *SIAM Journal on Optimization*, 18, 1326–1350.

Friedman, J., Hastie, T., and Tibshirani, R. (2008), "Sparse inverse covariance estimation with the graphical lasso," *Biostatistics*, 9, 432–441.

Gabay, D. (1983), "Chapter ix applications of the method of multipliers to variational inequalities," in *Studies in Mathematics and Its Applications*, vol. 15, pp. 299–331, Elsevier.

Gabay, D. and Mercier, B. (1976), "A dual algorithm for the solution of nonlinear variational problems via finite element approximation," *Computers and Mathematics with Applications*, 2, 17–40.

Gao, Y. and Sun, D. F. (2009), "Calibrating least squares covariance matrix problems with equality and inequality constraints," *SIAM Journal on Matrix Analysis and Applications*, 31, 1432–1457.

Glowinski, R. and Marroco, A. (1975), "Sur l'approximation, par éléments finis d'ordre un, et la résolution, par pénalisation-dualité d'une classe de problèmes de Dirichlet non linéaires," *Revue française d'automatique, informatique, recherche opérationnelle. Analyse numérique*, 9, 41–76.

Glowinski, R. and Oden, J. (1985), *Numerical Methods for Nonlinear Variational Problems*, Springer.

Han, J. Y. and Sun, D. F. (1997), "Newton and quasi-Newton methods for normal maps with polyhedral sets," *Journal of Optimization Theory and Applications*, 94, 659–676.

Hess, K. R., Anderson, K., Symmans, W. F., Valero, V., Ibrahim, N., Mejia, J. A., Booser, D., Theriault, R. L., Buzdar, A. U., Dempsey, P. J., et al. (2006), "Pharmacogenomic predictor of sensitivity to preoperative chemotherapy with paclitaxel and fluorouracil, doxorubicin, and cyclophosphamide in breast cancer," *Journal of clinical oncology*, 24, 4236–4244.

Hiriart-Urruty, J.-B. and Lemaréchal, C. (1993), *Convex Analysis and Minimization Algorithms II: Advanced Theory and Bundle Methods*, vol. 306, Springer-Verlag.

Hsieh, C.-J., Sustik, M. A., Dhillon, I. S., Ravikumar, P., et al. (2014), "QUIC: quadratic approximation for sparse inverse covariance estimation." *Journal of Machine Learning Research*, 15, 2911–2947.

Jin, Y. and Rao, B. D. (2010), "Algorithms for robust linear regression by exploiting the connection to sparse signal recovery," in *2010 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 3830–3833, IEEE.

Khanh, P. D., Mordukhovich, B. S., Phat, V. T., and Tran, D. B. (2023), "Generalized damped Newton algorithms in nonsmooth optimization via second-order subdifferentials," *Journal of Global Optimization*, 86, 93–122.

Kummer, B. (1991a), "An implicit-function theorem for $C^{0,1}$ -equations and parametric $C^{1,1}$-optimization," *Journal of Mathematical Analysis and Applications*, 158, 35–46.

Kummer, B. (1991b), "Lipschitzian inverse functions, directional derivatives, and applications in $C^{1,1}$-optimization," *Journal of Optimization Theory and Applications*, 70, 561–582.

Li, Q., Jiang, B. Y., and Sun, D. F. (2023), "MARS: A second-order reduction algorithm for high-dimensional sparse precision matrices estimation," *Journal of Machine Learning Research*, 24, 1–44.

Li, Q., Sun, D. F., and Yuan, Y. C. (2024), “An Efficient Sieving-Based Secant Method for Sparse Optimization Problems with Least-Squares Constraints,” *SIAM Journal on Optimization*, 34, 2038–2066.

Li, X. D., Sun, D. F., and Toh, K.-C. (2018a), “A highly efficient semismooth Newton augmented Lagrangian method for solving Lasso problems,” *SIAM Journal on Optimization*, 28, 433–458.

Li, X. D., Sun, D. F., and Toh, K.-C. (2018b), “On efficiently solving the subproblems of a level-set method for fused lasso problems,” *SIAM Journal on Optimization*, 28, 1842–1866.

Liang, Y., Liu, C., Luan, X.-Z., Leung, K.-S., Chan, T.-M., Xu, Z.-B., and Zhang, H. (2013), “Sparse logistic regression with a $L_{1/2}$ penalty for gene selection in cancer classification,” *BMC Bioinformatics*, 14, 1–12.

Liu, J., Zhao, Z., Wang, J., and Ye, J. P. (2014), “Safe screening with variational inequalities and its application to lasso,” in *International Conference on Machine Learning*, pp. 289–297, PMLR.

Liu, W. D. and Luo, X. (2015), “Fast and adaptive sparse precision matrix estimation in high dimensions,” *Journal of Multivariate Analysis*, 135, 153–162.

Luo, Z. Y., Sun, D. F., Toh, K.-C., and Xiu, N. H. (2019), “Solving the OSCAR and SLOPE Models Using a Semismooth Newton-Based Augmented Lagrangian Method,” *Journal of Machine Learning Research*, 20 (106), 1–25.

Lustig, M., Donoho, D., and Pauly, J. M. (2007), “Sparse MRI: The application of compressed sensing for rapid MR imaging,” *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 58, 1182–1195.

Lustig, M., Donoho, D. L., Santos, J. M., and Pauly, J. M. (2008), “Compressed sensing MRI,” *IEEE Signal Processing Magazine*, 25, 72–82.

Meng, F. W., Sun, D. F., and Zhao, G. (2005), “Semismoothness of solutions to generalized equations and the Moreau-Yosida regularization,” *Mathematical Programming*, 104, 561–581.

Mifflin, R. (1977), “Semismooth and semiconvex functions in constrained optimization,” *SIAM Journal on Control and Optimization*, 15, 959–972.

Moreau, J.-J. (1965), “Proximité et dualité dans un espace hilbertien,” *Bulletin de la Société mathématique de France*, 93, 273–299.

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

Ng, A. Y. (2004), "Feature selection, L_1 vs. L_2 regularization, and rotational invariance," in *Proceedings of the Twenty-First International Conference on Machine learning*, p. 78.

Ortega, J. M. and Rheinboldt, W. C. (1970), *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York-London.

Pang, J.-S. (1990), "Newton's method for B-differentiable equations," *Mathematics of Operations Research*, 15, 311–341.

Papageorgiou, G., Bouboulis, P., and Theodoridis, S. (2017), "Robust nonlinear regression: A greedy approach employing kernels with application to image denoising," *IEEE Transactions on Signal Processing*, 65, 4309–4323.

Potra, F. A. (1989), "On Q-order and R-order of convergence," *Journal of Optimization Theory and Applications*, 63, 415–431.

Potra, F. A., Qi, L. Q., and Sun, D. F. (1998), "Secant methods for semismooth equations," *Numerische Mathematik*, 80, 305–324.

Qi, L. Q. and Sun, D. F. (2002), "Smoothing functions and smoothing Newton method for complementarity and variational inequality problems," *Journal of Optimization Theory and Applications*, 113, 121–147.

Qi, L. Q. and Sun, J. (1993), "A nonsmooth version of Newton's method," *Mathematical Programming*, 58, 353–367.

Qi, L. Q. and Tseng, P. (2007), "On almost smooth functions and piecewise smooth functions," *Nonlinear Analysis: Theory, Methods & Applications*, 67, 773–794.

Qi, L. Q., Sun, D. F., and Zhou, G. L. (2000), "A new look at smoothing Newton methods for nonlinear complementarity problems and box constrained variational inequalities," *Math. Program.*, 87, 1–35.

Ravikumar, P., Wainwright, M. J., and Lafferty, J. D. (2010), "High-dimensional Ising model selection using ℓ_1-regularized logistic regression." *The Annals of Statistics*, pp. 1287–1319.

Ravindran, G. and Gowda, M. S. (2001), "Regularization of P0-functions in box variational inequality problems," *SIAM Journal on Optimization*, 11, 748–760.

Recht, B., Fazel, M., and Parrilo, P. A. (2010), "Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization," *SIAM Review*, 52, 471–501.

Richard, H. E. and George, C. M. (2001), *Computer Algebra Recipes: A Gourmet's Guide to the Mathematical Models of Science*, Springer, Berlin.

Rockafellar, R. T. (1970), *Convex Analysis*, Princeton University Press.

Rockafellar, R. T. and Wets, R. J.-B. (2009), *Variational Analysis*, vol. 317, Springer Science & Business Media.

Ryali, S., Supekar, K., Abrams, D. A., and Menon, V. (2010), "Sparse logistic regression for whole-brain classification of fMRI data," *NeuroImage*, 51, 752–764.

Scholtes, S. (2012), *Introduction to Piecewise Differentiable Equations*, Springer.

Shapiro, A. (1990), "On concepts of directional differentiability," *Journal of Optimization Theory and Applications*, 66, 477–487.

Sun, D. F. (2001), "A further result on an implicit function theorem for locally Lipschitz functions," *Operations Research Letters*, 28, 193–198.

Sun, D. F. (2006), "The strong second-order sufficient condition and constraint nondegeneracy in nonlinear semidefinite programming and their implications," *Mathematics of Operations Research*, 31, 761–776.

Sun, D. F. and Sun, J. (2002), "Semismooth matrix-valued functions," *Mathematics of Operations Research*, 27, 150–169.

Tibshirani, R. (1996), "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 58, 267–288.

Tibshirani, R., Saunders, M., Rosset, S., Zhu, J., and Knight, K. (2005), "Sparsity and smoothness via the fused lasso," *Journal of the Royal Statistical Society. Series B. Statistical Methodology*, 67, 91–108.

Tibshirani, R., Bien, J., Friedman, J., Hastie, T., Simon, N., Taylor, J., and Tibshirani, R. J. (2012), "Strong rules for discarding predictors in lasso-type problems," *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 74, 245–266.

Traub, J. F. (1964), *Iterative Methods for the Solution of Equations*, Prentice-Hall, Inc., Englewood Cliffs, N.J.

Tropp, J. A. (2006), "Just relax: Convex programming methods for identifying sparse signals in noise," *IEEE transactions on information theory*, 52, 1030–1051.

Ulbrich, M. (2011), *Semismooth Newton Methods for Variational Inequalities and Constrained Optimization Problems in Function Spaces*, SIAM.

Van den Berg, E. and Friedlander, M. P. (2008), "Probing the Pareto frontier for basis pursuit solutions," *SIAM Journal on Scientific Computing*, 31, 890–912.

Van den Berg, E. and Friedlander, M. P. (2011), "Sparse optimization with least-squares constraints," *SIAM Journal on Optimization*, 21, 1201–1229.

Van den Berg, E. and Friedlander, M. P. (2019), "SPGL1: A solver for large-scale sparse reconstruction," https://friedlander.io/spgl1.

Wainwright, M. J., Lafferty, J., and Ravikumar, P. (2006), "High-dimensional graphical model selection using ℓ_1-regularized logistic regression," *Advances in Neural Information Processing Systems*, 19.

Wang, C. and Jiang, B. Y. (2020), "An efficient ADMM algorithm for high dimensional precision matrix estimation via penalized quadratic loss," *Computational Statistics & Data Analysis*, 142, 106812.

Wang, J., Zhou, J., Wonka, P., and Ye, J. (2013), "Lasso screening rules via dual polytope projection," *Advances in Neural Information Processing Systems*, 26.

Watson, G. A. (1992), "Linear best approximation using a class of polyhedral norms," *Numerical Algorithms*, 2, 321–335.

Wu, C., Cui, Y., Li, D. H., and Sun, D. F. (2023), "Convex and Nonconvex Risk-Based Linear Regression at Scale," *INFORMS Journal on Computing*, 35, 797–816.

Xiang, Z. J., Wang, Y., and Ramadge, P. J. (2016), "Screening tests for lasso problems," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39, 1008–1027.

Ye, J. J. and Zhou, J. C. (2017), "Exact formulas for the proximal/regular/limiting normal cone of the second-order cone complementarity set," *Mathematical Programming*, 162, 33–50.

Yosida, K. (1964), *Functional Analysis*, vol. 123, Springer Science & Business Media.

Yuan, Y. C., Chang, T.-H., Sun, D. F., and Toh, K.-C. (2022), "A dimension reduction technique for large-scale structured sparse optimization problems with application to convex clustering," *SIAM Journal on Optimization*, 32, 2294–2318.

Yuan, Y. C., Lin, M. X., Sun, D. F., and Toh, K.-C. (2023), "Adaptive sieving: A dimension reduction technique for sparse optimization problems," *arXiv: 2306.17369*.

Zeng, Y. H., Yang, T. B., and Breheny, P. (2021), "Hybrid safe–strong rules for efficient optimization in lasso-type problems," *Computational Statistics & Data Analysis*, 153, 107063.

Zhang, T. and Zou, H. (2014), "Sparse precision matrix estimation via lasso penalized D-trace loss," *Biometrika*, 101, 103–120.

Zhang, Y. J., Zhang, N., Sun, D. F., and Toh, K.-C. (2020), "An efficient Hessian based algorithm for solving large-scale sparse group Lasso problems," *Mathematical Programming*, 179, 223–263.

Zhao, X.-Y., Sun, D. F., and Toh, K.-C. (2010), "A Newton-CG augmented Lagrangian method for semidefinite programming," *SIAM Journal on Optimization*, 20, 1737–1765.