ALGORITHMS FOR LARGE SCALE NUCLEAR NORM MINIMIZATION AND CONVEX QUADRATIC SEMIDEFINITE PROGRAMMING PROBLEMS

JIANG KAIFENG

(B.Sc., NJU, China)

A THESIS SUBMITTED

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF MATHEMATICS

NATIONAL UNIVERSITY OF SINGAPORE

2011

To my parents

Acknowledgements

I would like to express my sincerest thanks to my supervisor Professor Toh Kim Chuan for his invaluable guidance and perpetual encouragement and support. I have benefited intellectually from his fresh ideas and piercing insights in scientific research, as well as many enjoyable discussions we had during the past four years. He always encourages me to do research independently, even though sometimes I was lack of confidence in myself. I am very grateful to him for providing me extensive training in the field of numerical computation. I am greatly indebted to him.

I would like to thank Professor Sun Defeng for his great effort on conducting weekly optimization research seminars, which have significantly enriched my knowledge of the theory, algorithms and applications of optimization. His amazing depth of knowledge and tremendous expertise in optimization has greatly facilitated my research progress. I feel very honored to have an opportunity of doing research with him.

I would like to thank Professor Zhao Gongyun for his instruction on mathematical programming, which is the first module I took during my first year in NUS. His excellent teaching style helps me to gain broad knowledge of numerical optimization and software. I am very thankful to him for sharing with me his wonderful mathematical insights and research experience in the field of optimization.

I would like to thank Department of Mathematics and National University of Singapore for providing me excellent research conditions and scholarship to complete my PhD study. I also would like to thank Faculty of Science for providing me financial support for attending the 2011 SIAM conference on optimization in Darmstadt, Germany.

Finally, I would like to thank all my friends in Singapore for their long-time encouragement and support. Many thanks go to Dr. Liu Yongjin, Dr. Zhao Xinyuan, Dr. Li Lu, Dr. Gao Yan, Dr. Yang Junfeng, Ding Chao, Miao Weimin, Gong Zheng, Shi Dongjian, Wu Bin, Chen Caihua, Li Xudong, Du Mengyu for their helpful discussions in many interesting optimization topics related to my research.

Contents

Acknowledgements **v**

Summary **xi**

1 Introduction **1**

1.1 Nuclear norm regularized matrix least squares problems 2

1.1.1 Existing models and related algorithms 3

1.1.2 Motivating examples 6

1.2 Convex semidefinite programming problems 7

1.3 Contributions of the thesis 12

1.4 Organization of the thesis 15

2 Preliminaries **17**

2.1 Notations 17

2.2 Metric projectors 19

2.3 The soft thresholding operator 20

2.4 The smoothing counterpart 28

3 Nuclear norm regularized matrix least squares problems 37

3.1 The general proximal point algorithm 37

3.2 A partial proximal point algorithm 41

3.3 Convergence analysis of the partial PPA 50

3.4 An inexact smoothing Newton method for inner subproblems 54

3.4.1 Inner subproblems 54

3.4.2 An inexact smoothing Newton method 57

3.4.3 Constraint nondegeneracy and quadratic convergence 59

3.5 Efficient implementation of the partial PPA 66

4 A semismooth Newton-CG method for unconstrained inner subproblems 71

4.1 A semismooth Newton-CG method 72

4.2 Convergence analysis 77

4.3 Symmetric matrix problems 78

5 An inexact APG method for linearly constrained convex SDP 81

5.1 An inexact accelerated proximal gradient method 82

5.1.1 Specialization to the case where $g = \delta(\cdot \,|\, \Omega)$ 89

5.2 Analysis of an inexact APG method for (P) 91

5.2.1 Boundedness of $\{p_k\}$ 99

5.2.2 A semismooth Newton-CG method 101

6 Numerical Results 105

6.1 Numerical Results for nuclear norm minimization problems 105

6.2 Numerical Results for linearly constrained QSDP problems 125

7 Conclusions 133

Bibliography **135**

Summary

This thesis focuses on designing efficient algorithms for solving large scale structured matrix optimization problems, which have many applications in a wide range of fields, such as signal processing, system identification, image compression, molecular conformation, sensor network localization and so on. We introduce a partial proximal point algorithm, in which only some of the variables appear in the quadratic proximal term, for solving nuclear norm regularized matrix least squares problems with linear equality and inequality constraints. We establish the global and local convergence of our proposed algorithm based on the results for the general partial proximal point algorithm. The inner subproblems, reformulated as a system of semismooth equations, are solved by an inexact smoothing Newton method, which is proved to be quadratically convergent under the constraint nondegeneracy condition, together with the strong semismoothness property of the soft thresholding operator.

As a special case where the nuclear norm regularized matrix least squares problem has equality constraints only, we introduce a semismooth Newton-CG method to solve the unconstrained inner subproblem in each iteration. We show that the positive definiteness of the generalized Hessian of the objective function in the inner subproblem is equivalent to the constraint nondegeneracy of the corresponding primal problem, which is a key property for applying the semismooth Newton-CG method to solve the inner subproblems efficiently. The global and local superlinear (quadratic) convergence of the semismooth Newton-CG method is also established.

To solve large scale convex quadratic semidefinite programming (QSDP) problems, we extend the accelerated proximal gradient (APG) method to the inexact setting where the subproblem in each iteration is progressively solved with sufficient accuracy. We show that the inexact APG method enjoys the same superior convergent rate of $O(1/k^2)$ as the exact version.

Extensive numerical experiments on a variety of large scale nuclear norm regularized matrix least squares problems show that our proposed partial proximal point algorithm is very efficient and robust. We can successfully find a low rank approximation of the target matrix while maintaining the desired linear structure of the original system. Numerical experiments on some large scale convex QSDP problems demonstrate the high efficiency and robustness of the proposed inexact APG algorithm. In particular, our inexact APG algorithm can efficiently solve the H-weighted nearest correlation matrix problem, where the given weight matrix H is highly ill-conditioned.

Chapter 1

Introduction

In this thesis, we focus on designing algorithms for solving large scale structured matrix optimization problems. In particular, we are interested in nuclear norm regularized matrix least squares problems and linearly constrained convex semidefinite programming problems. Let $\Re^{p\times q}$ be the space of all $p \times q$ matrices equipped with the standard trace inner product and its induced Frobenius norm $\|\cdot\|$. The general structured matrix optimization problem we consider in this thesis can be stated as follows:

$$\min\Big\{ f(X) + g(X) \ : \ X \in \Re^{p\times q}\Big\}, \tag{1.1}$$

where $f : \Re^{p\times q} \to \Re$ and $g : \Re^{p\times q} \to \Re \cup \{+\infty\}$ are proper, lower semi-continuous convex functions (possibly nonsmooth). In many applications, such as statistical regression and machine learning, f is a loss function which measures the difference between the observed data and the value provided by the model. The quadratic loss function, e.g., the linear least squares loss function, is a common choice. The function g, which is generally nonsmooth, favors certain desired properties of the computed solution, and it can be chosen by the user based on the available prior information about the target matrix. In practice, the data matrix X, which describes the original system, has some or all of the following properties:

1. The computed solution X should be positive semidefinite;

2. In order to reduce the complexity of the whole system, X should be of low rank;

3. Some entries of X are in the confidence interval which indicates the reliability of the statistical estimation;

4. All entries of X should be nonnegative because they correspond to physically nonnegative quantities such as density or image intensity;

5. X belongs to some special classes of matrices, e.g., Hankel matrices arising from linear system realization, (doubly) stochastic matrices which describe the transition probability of a Markov chain, and so on.

1.1 Nuclear norm regularized matrix least squares problems

In the first part of this thesis, we consider the following nuclear norm regularized matrix least squares problem with linear equality and inequality constraints:

$$\begin{aligned} \min_{X \in \Re^{p\times q}} \quad & \frac{1}{2}\|\mathcal{A}(X) - b\|^2 + \langle C, X\rangle + \rho\|X\|_* \\ \text{s.t.} \quad & \mathcal{B}(X) \in d + \mathcal{Q}, \end{aligned} \tag{1.2}$$

where $\|X\|_*$ denotes the nuclear norm of X defined as the sum of all its singular values, $\mathcal{A} : \Re^{p\times q} \to \Re^m$ and $\mathcal{B} : \Re^{p\times q} \to \Re^s$ are linear maps, $C \in \Re^{p\times q}, b \in \Re^m, d \in \Re^s$, ρ is a given positive parameter, and $\mathcal{Q} = \{0\}^{s_1} \times \Re^{s_2}_+$ is a polyhedral convex cone with $s = s_1 + s_2$. In this case, the convex functions f and g in (1.1) are of the following forms:

$$f(X) = \frac{1}{2}\|\mathcal{A}(X) - b\|^2 + \langle C, X\rangle \quad \text{and} \quad g(X) = \rho\|X\|_* + \delta(X \mid \mathcal{D}_1),$$

where $\mathcal{D}_1 = \{X \in \Re^{p\times q} \mid \mathcal{B}(X) \in d + \mathcal{Q}\}$ is the feasible set of (1.2) and $\delta(\cdot \mid \mathcal{D}_1)$ is the indicator function on the set $\mathcal{D}_1$. In many applications, such as signal processing [68, 111, 112, 129], molecular structure modeling for protein folding [86, 87, 122] and computation of the greatest common divisor (GCD) of unvariate polynomials [27, 62] from computer algebra, we need to find a low rank approximation of a given target matrix while preserving certain structures. The nuclear norm function has been widely used as a regularizer which favors a low rank solution of (1.2). In [25], Chu, Funderlic and Plemmons addressed some theoretical and numerical issues concerning structured low rank approximation problems. In many data analysis problems, the collected empirical data, possibly contaminated by noise, usually do not have the specified structure or the desired low rank. So it is important to find the nearest low rank approximation of the given matrix while maintaining the underlying structure of the original system. In practice, the data to be analyzed is very often nonnegative such as those corresponding to concentrations or intensity values, and it would be preferable to take into account such structural constraints.

1.1.1 Existing models and related algorithms

In this subsection, we give a brief review of existing models involving the nuclear norm function and related variants. Recently there are intensive studies on the following affine rank minimization problem:

$$\min\Big\{\operatorname{rank}(X) \ : \ \mathcal{A}(X) \ = \ b, \ X \in \Re^{p\times q}\Big\}. \tag{1.3}$$

The problem (1.3) has many applications in diverse fields, see, e.g., [1, 2, 19, 37, 44, 82, 102]. (Note that there are some special rank approximation problems that have known solutions. For example, the low rank approximation of a given matrix in Frobenius norm can be derived via singular value decomposition by the classic Eckart-Young Theorem [35].) However, this affine rank minimization problem is generally an NP-hard nonconvex optimization problem. A tractable heuristic introduced in [36, 37] is to minimize the nuclear norm over the same constraints as in (1.3):

$$\min\Big\{\|X\|_* \ : \ \mathcal{A}(X) \ = \ b, \ X \in \Re^{p\times q}\Big\}. \tag{1.4}$$

The nuclear norm function is the greatest convex function majorized by the rank function over the unit ball of matrices with operator norm at most one. In [19, 21, 51, 63, 101, 102], the authors established remarkable results which state that under suitable incoherence assumptions, a $p \times q$ matrix of rank r can be recovered with high probability from uniformly random sampled entries of size slightly larger than $O((p+q)r)$ by solving (1.4). A frequently used alternative to (1.4) for accommodating problems with noisy data is to consider solving the following matrix least squares problem with nuclear norm regularization (see [77, 121]):

$$\min \left\{ \frac{1}{2}\|\mathcal{A}(X) - b\|^2 + \rho\|X\|_* \ : \ X \in \Re^{p\times q} \right\}, \tag{1.5}$$

where ρ is a given positive parameter. It is known that (1.4) or (1.5) can be equivalently reformulated as a semidefinite programming (SDP) problem (see [36, 102]), which has one $(p+q)\times(p+q)$ semidefinite constraint and m linear equality constraints. One can use standard interior-point method based semidefinite programming solvers such as SeDuMi [114] and SDPT3 [119] to solve this SDP problem. However, these solvers are not suitable for problems with large $p+q$ or m since in each iteration of these solvers, a large and dense Schur complement equation must be solved for computing the search direction even when the data is sparse.

To overcome the difficulties faced by interior-point methods, several algorithms have been proposed to solve (1.4) or (1.5) directly. In [102], Recht, Fazel and Parrilo considered the projected subgradient method to solve (1.4). However, the convergence of the projected subgradient method considered in [102] is still unknown since problem (1.4) is a nonsmooth problem, and the convergence is observed to be very slow for large scale matrix completion problems. Recht, Fazel and Parrilo [102] also considered the method of using the low-rank factorization technique introduced by Burer and Monteiro [15, 16] to solve (1.4). The advantage of this method is that it requires less computer memory for solving large scale problems. However, the potential difficulty of this method is that the low rank factorization formulation is nonconvex and the rank of the optimal matrix is generally unknown. In [17], Cai, Candès and Shen proposed a singular value thresholding (SVT) algorithm for solving the following Tikhonov regularized version of (1.4):

$$\min\Big\{\tau\|X\|_* + \frac{1}{2}\|X\|^2 \ : \ \mathcal{A}(X) \ = \ b, \ X \in \Re^{p\times q}\Big\}, \tag{1.6}$$

where τ is a given positive parameter. The SVT algorithm is a gradient method applied to the dual problem of (1.6). Ma, Goldfarb and Chen [77] proposed a fixed point algorithm with continuation (FPC) for solving (1.5) and a Bregman iterative algorithm for solving (1.4). Their numerical results on randomly generated matrix completion problems demonstrated that the FPC algorithm is much more efficient than the semidefinite programming solver SDPT3. In [121], Toh and Yun proposed an accelerated proximal gradient algorithm (APG), which terminates in $O(1/\sqrt{\varepsilon})$ iterations for achieving ε-optimality (in terms of the function value), to solve the unconstrained matrix least squares problem (1.5). Their numerical results show that the APG algorithm is highly efficient and robust in solving large-scale random matrix completion problems. In [71], Liu, Sun and Toh considered the following nuclear norm minimization problem with linear and second order cone constraints:

$$\min\Big\{\|X\|_* \ : \ \mathcal{A}(X) \in b + \mathcal{K}, \ X \in \Re^{p\times q}\Big\}, \tag{1.7}$$

where $\mathcal{K} = \{0\}^{m_1} \times \mathcal{K}^{m_2}$, and $\mathcal{K}^{m_2}$ stands for the m_2-dimensional second order cone (or ice-cream cone, or Lorentz cone) defined by

$$\mathcal{K}^{m_2} \ := \ \{x = (x_0; \overline{x}) \in \Re \times \Re^{m_2-1} : \|\overline{x}\| \leq x_0\}.$$

They developed three inexact proximal point algorithms (PPA) in the primal, dual and primal-dual forms with comprehensive convergence analysis built upon the classic results of the general PPA established by Rockafellar [108, 107]. Their numerical results demonstrated the efficiency and robustness of these three forms of PPA in solving randomly generated matrix completion problems and real matrix completion problems. Moreover, they showed that the SVT algorithm [17] is just one outer iteration of the exact primal PPA, and the Bregman iterative method [77] is a special case of the exact dual PPA.

However, all the above mentioned models and related algorithms cannot address the following goal: given the observed data matrix (possibly contaminated by noise), we want to find the nearest low rank approximation of the target matrix while maintaining the prescribed structure of the original system. In particular, the APG method considered in [121] cannot be applied directly to solve (1.2).

1.1.2 Motivating examples

A strong motivation for proposing the model (1.2) arises from finding the nearest low rank approximation of transition matrices. For a given data matrix $\widetilde{P}$ which describes the full distribution of a random walk through the entire data set, the problem of finding the low rank approximation of $\widetilde{P}$ can be stated as follows:

$$\min_{X \in \Re^{n\times n}} \left\{ \frac{1}{2}\|X - \widetilde{P}\|^2 + \rho\|X\|_* \, : \, Xe = e, \; X \geq 0 \right\}, \tag{1.8}$$

where $e \in \Re^n$ is the vector of all ones and $X \geq 0$ denotes the condition that all entries of X are nonnegative. In [70], Lin proposed the Latent Markov Analysis (LMA) approach for finding the reduced rank approximations of transition matrices. The LMA is applied to clustering such that the inferred cluster relationships can be described probabilistically by the reduced-rank transition matrix. In [24], Chennubhotla exploited the spectral properties of the Markov transition matrix to obtain low rank approximation of the original transition matrix in order to develop a fast eigen-solver for spectral clustering. Another application of finding the low rank approximation of the transition matrix comes from computing the personalized PageRank [6] which describes the backlink-based page quality around user-selected pages. In many applications, since only partial information of the original transition matrix is available, it is also important to estimate the missing entries of $\widetilde{P}$. For example, transition probabilities between different credit ratings play a crucial role in the credit portfolio management. If our primary interest is in a specific group, the number of observations of available rating transitions is very small. Due to lack of rating data, it is important to estimate the rating transition matrix in the presence of missing data [5, 59].

Another strong motivation for considering the model (1.2) comes from finding low rank approximations of doubly stochastic matrices with a prescribed entry. A matrix $M \in \Re^{n\times n}$ is called doubly stochastic if it is nonnegative and all its row and column sums are equal to one. Then the problem for matching the first moment of M with sparsity pattern $\mathcal{E}$ can be stated as follows:

$$\min_{X\in\Re^{n\times n}} \left\{\frac{1}{2}\|X_{\mathcal{E}} - \widetilde{M}_{\mathcal{E}}\|^2 + \rho\|X\|_* \,:\, Xe = e,\, X^T e = e,\, X_{11} = M_{11},\, X \geq 0\right\}, \quad (1.9)$$

where $\widetilde{M}_{\mathcal{E}}$ dentoes the partially observed data (possibly with noise). This problem arose from numerical simulation of large circuit networks. In order to reduce the complexity of the simulation of the whole system, the Padé approximation with Krylov subspace method, such as the Lanczos algorithm, is a useful tool for generating a lower order approximation to the linear system matrix which describes the large linear network [3]. The tridiagonal matrix M produced by the Lanczos algorithm generally is not doubly stochastic. If the original system matrix is doubly stochastic, then we need to find a low rank approximation of M such that it is doubly stochastic and matches the first moment of M.

1.2 Convex semidefinite programming problems

In the second part of this thesis, we consider the following linearly constrained convex semidefinite programming problem:

$$\begin{array}{rl} \min\limits_{X\in\mathcal{S}^n} & f(X) \\ \text{s.t.} & \mathcal{A}(X) \;=\; b, \\ & X \succeq \; 0, \end{array} \quad (1.10)$$

where f is a smooth convex function on $\mathcal{S}^n$, $\mathcal{A} : \mathcal{S}^n \to \mathcal{R}^m$ is a linear map, $b \in \mathcal{R}^m$, and $\mathcal{S}^n$ is the space of $n \times n$ symmetric matrices equipped with the standard trace inner product. The notation $X \succeq 0$ means that X is positive semidefinite. In this case, the function g in (1.1) takes the form: $g(X) = \delta(X \,|\, \mathcal{D}_2)$, where $\mathcal{D}_2 = \{X \in \mathcal{S}^n \,|\, \mathcal{A}(X) = b,\, X \succeq 0\}$ is the feasible set of (1.10). Let $\mathcal{A}^*$ be the adjoint of $\mathcal{A}$. The dual problem associated with (1.10) is given by

$$\begin{array}{ll} \max & f(X) - \langle \nabla f(X),\, X\rangle + \langle b,\, p\rangle \\ \text{s.t.} & \nabla f(X) - \mathcal{A}^* p - Z = 0, \\ & p \in \Re^m,\ Z \succeq 0,\ X \succeq 0. \end{array} \tag{1.11}$$

The problem (1.10) contains the following important special case of convex quadratic semidefinite programming (QSDP):

$$\min\Big\{\frac{1}{2}\langle X,\ \mathcal{Q}(X)\rangle + \langle C,\, X\rangle \ :\ \mathcal{A}(X) = b, X \succeq 0\Big\}, \tag{1.12}$$

where $\mathcal{Q} : \mathcal{S}^n \to \mathcal{S}^n$ is a given self-adjoint positive semidefinite linear operator and $C \in \mathcal{S}^n$. The Lagrangian dual problem of (1.12) is given by

$$\max\Big\{ -\frac{1}{2}\langle X,\ \mathcal{Q}(X)\rangle + \langle b, p\rangle \ :\ \mathcal{A}^*(p) - \mathcal{Q}(X) + Z = C, Z \succeq 0\Big\}. \tag{1.13}$$

A typical example of QSDP is the nearest correlation matrix problem [55], where given a symmetric matrix $U \in \mathcal{S}^n$ and a linear map $\mathcal{L} : \mathcal{S}^n \to \mathcal{R}^{n\times n}$, we want to solve

$$\min\Big\{\frac{1}{2}\|\mathcal{L}(X-U)\|^2 \ :\ \text{Diag}(X) = e,\ X \succeq 0\Big\}, \tag{1.14}$$

where $e \in \Re^n$ is the vector of all ones. If we let $\mathcal{Q} = \mathcal{L}^*\mathcal{L}$ and $C = -\mathcal{L}^*\mathcal{L}(U)$ in (1.14), then we get the QSDP problem (1.12). A well studied special case of (1.14) is the W-weighted nearest correlation matrix problem, where $\mathcal{L} = W^{1/2} \circledast W^{1/2}$ for a given $W \in \mathcal{S}^n_{++}$ and $\mathcal{Q} = W \circledast W$. Note that for $U \in \Re^{n\times r}$, $V \in \Re^{n\times s}$, $U \circledast V : \Re^{r\times s} \to \mathcal{S}^n$ is the symmetrized Kronecker product linear map defined by $U \circledast V(M) = (UMV^T + VM^TU^T)/2$.

There are several methods available for solving (1.14), which include the alternating projection method [55], the quasi-Newton method [78], the inexact semismooth Newton-CG method [97] and the inexact interior-point method [120]. All these methods, excluding the inexact interior-point method, rely critically on the fact that the projection of a given matrix $X \in \mathcal{S}^n$ onto $\mathcal{S}^n_+$ has an analytical formula with respect to the norm $\|W^{1/2}(\cdot)W^{1/2}\|$. However, all above mentioned techniques cannot be extended to efficiently solve the H-weighted case [55] of (1.14), where $\mathcal{L}(X) = H \circ X$ for some $H \in \mathcal{S}^n$ with nonnegative entries and $\mathcal{Q}(X) = (H \circ H) \circ X$, with "$\circ$" denoting the Hardamard product of two matrices defined by $(A \circ B)_{ij} = A_{ij}B_{ij}$. In [50], a H-weighted kernel matrix completion problem of the form

$$\min\Big\{\|H \circ (X - U)\| \mid \mathcal{A}(X) = b,\ X \succeq 0\Big\} \tag{1.15}$$

is considered, where $U \in \mathcal{S}^n$ is a given kernel matrix with missing entries. The aforementioned methods are not well suited for the H-weighted case of (1.14) because there is no explicitly computable formula for the following problem

$$\min\Big\{\frac{1}{2}\|H \circ (X - U)\|^2 \ :\ X \succeq 0\Big\}, \tag{1.16}$$

where $U \in \mathcal{S}^n$ is a given matrix. To tackle the H-weighted case of (1.14), Toh [118] proposed an inexact interior-point method for a general convex QSDP including the H-weighted nearest correlation matrix problem. Recently, Qi and Sun [98] introduced an augmented Lagrangian dual method for solving the H-weighted version of (1.14), where the inner subproblem was solved by a semismooth Newton-CG (SSNCG) method. In her PhD thesis, Zhao [137] designed a semismooth Newton-CG augmented Lagrangian method and analyzed its convergence for solving convex quadratic programming over symmetric cones. The augmented Lagrangian dual method avoids solving (1.16) directly and it can be much faster than the inexact interior-point method [118]. However, if the weight matrix H is very sparse or ill-conditioned, the conjugate gradient (CG) method would have great difficulty in solving the linear system of equations in the semismooth Newton method, and the augmented Lagrangian method would not be efficient or even fail. Another drawback of the augmented Lagrangian dual method in [98] is that the computed solution X usually is not positive semidefinite. A post processing step is generally needed to make the computed solution positive semidefinite.

Another example of QSDP comes from the civil engineering problem of estimating a positive semidefinite stiffness matrix for a stable elastic structure from r measurements of its displacements $\{u_1, \ldots, u_r\} \subset \Re^n$ in response to a set of static loads $\{f_1, \ldots, f_r\} \subset \Re^n$ [130]. In this application, one is interested in the QSDP problem:

$$\min\Big\{ \|f - \mathcal{L}(X)\|^2 \ : \ X \succeq 0 \Big\}, \tag{1.17}$$

where $\mathcal{L} : \mathcal{S}^n \to \Re^{n\times r}$ is defined by $\mathcal{L}(X) = XU$, and $f = [f_1, \ldots, f_r]$, $U = [u_1, \ldots, u_r]$. In this case, the corresponding map $\mathcal{Q} = \mathcal{L}^*\mathcal{L}$ is given by $\mathcal{Q}(X) = (XB + BX)/2$ with $B = UU^T$.

The main purpose of the second part of this thesis is to design an efficient algorithm to solve the problem (1.10). The algorithm we propose here is based on the APG method of Beck and Teboulle [4] (the method is called FISTA in [4]), where in the kth iteration with iterate $\overline{X}_k$, a subproblem of the following form must be solved:

$$\min\Big\{ \langle \nabla f(\overline{X}_k),\, X - \overline{X}_k\rangle + \frac{1}{2}\langle X - \overline{X}_k,\, \mathcal{H}_k(X - \overline{X}_k)\rangle : \mathcal{A}(X) = b,\ X \succeq 0 \Big\}, \tag{1.18}$$

where $\mathcal{H}_k : \mathcal{S}^n \to \mathcal{S}^n$ is a given self-adjoint positive definite linear operator. In FISTA [4], $\mathcal{H}_k$ is restricted to $L\mathcal{I}$, where $\mathcal{I} : \mathcal{S}^n \to \mathcal{S}^n$ denotes the identity map and L is a Lipschitz constant of ∇f. More significantly, for FISTA in [4], the subproblem (1.18) must be solved exactly to generate the next iterate X_{k+1}. In this thesis, we design an inexact APG method which overcomes the two limitations just mentioned. Specifically, in our inexact algorithm, the subproblem (1.18) is only solved approximately and $\mathcal{H}_k$ is not restricted to be a scalar multiple of $\mathcal{I}$. In addition, we are able to show that if the subproblem (1.18) is progressively solved with sufficient accuracy, then the number of iterations needed to achieve ε-optimality (in terms of the function value) is also proportional to $1/\sqrt{\varepsilon}$, just as in the exact version of the APG method.

Another strong motivation for designing an inexact APG algorithm comes from the recent paper [22], which considered the following regularized inverse problem:

$$\min_{x\in\Re^p} \left\{ \frac{1}{2}\|\Phi x - y\|^2 + \lambda \|x\|_{\mathcal{B}} \right\}, \tag{1.19}$$

where $\Phi : \mathcal{R}^p \to \mathcal{R}^n$ is a given linear map and $\|x\|_{\mathcal{B}}$ is the atomic norm induced by a given compact set of atoms $\mathcal{B}$ in $\mathcal{R}^p$. It appears that the APG algorithm is highly suitable for solving (1.19). Note that in each iteration of the APG algorithm, a subproblem of the form

$$\min_{z\in\Re^p} \left\{ \mu\|z\|_{\mathcal{B}} + \frac{1}{2}\|z - x\|^2 \right\} \equiv \min_{y\in\Re^p} \left\{ \frac{1}{2}\|y - x\|^2 \mid \|y\|_{\mathcal{B}}^* \le \mu \right\}$$

must be solved, where $\|\cdot\|_{\mathcal{B}}^*$ is the dual norm of $\|\cdot\|_{\mathcal{B}}$. However, for most choices of $\mathcal{B}$, the subproblem does not admit an analytical solution and has to be solved numerically. As a result, the subproblem is never solved exactly. In fact, it may be computationally very expensive to solve the subproblem to high accuracy. Our inexact APG algorithm thus has the attractive computational advantage that the subproblems need only be solved with progressively better accuracy while still maintaining the global iteration complexity.

Finally we should mention that the fast gradient method of Nesterov [90] has also been extended in [30] to the problem

$$\min\{f(x) \mid x \in Q\}, \tag{1.20}$$

where the function f is convex (not necessarily smooth) on the closed convex set Q, and is equipped with the so-called first-order (δ, L)-oracle where for any $y \in Q$, we can compute a pair $(f_{\delta,L}(y), g_{\delta,L}(y))$ such that

$$0 \le f(x) - f_{\delta,L}(y) - \langle g_{\delta,L}(y),\, x - y\rangle \le \frac{L}{2}\|x - y\|^2 + \delta \quad \forall\, x \in Q.$$

In the inexact-oracle fast gradient method in [30], the subproblem of the form

$$\min\left\{ \langle g_{\delta,L}(y),\, x - y\rangle + \frac{L}{2}\|x - y\|^2 \mid x \in Q \right\}$$

in each iteration must be solved exactly. Thus the kind of the inexactness considered in [30] is very different from what we consider in this thesis.

1.3 Contributions of the thesis

In the first part of this thesis, we study a partial proximal point algorithm (PPA) for solving (1.2), in which only some of the variables appear in the quadratic proximal term. Based on the results of the general partial PPA studied by Ha [52], we analyze the global and local convergence of our proposed partial PPA for solving (1.2). In [52], Ha presented a modification of the general PPA studied by Rockafellar [108], in which only some variables appear in the proposed iterative procedure. The partial PPA was further analyzed by Bertsekas and Tseng [11], in which the close relation between the partial PPA and some parallel algorithms in convex programming was revealed. In [60], Ibaraki and Fukushima proposed two variants of the partial proximal method of multipliers for solving convex programming problems with linear constraints only, in which the objective function is separable. The convergence analysis of their proposed two variants of algorithms is built upon the results of the partial PPA by Ha [52]. We note that the proposed partial PPA requires solving an inner subproblem with linear inequality constraints at each iteration. To handle the inequality constraints, Gao and Sun [42] recently designed a quadratically convergent inexact smoothing Newton method, which was used to solve the least squares semidefinite programming with equality and inequality constraints. Their numerical results demonstrated the high efficiency of the inexact smoothing Newton method. This strongly motivated us to use the inexact smoothing Newton method to solve inner subproblems for achieving fast convergence. For the inner subproblem, due to the presence of inequality constraints, we reformulate the problem as a system of semismooth equations. By defining a smoothing function for the soft thresholding operator, we then introduce an inexact smoothing Newton method to solve the semismooth system, where at each iteration the BiCGStab iterative solver is used to approximately solve the generated linear system. Based on the classic results of nonsmooth analysis by Clarke [26], we study the properties of the epigraph of the nuclear norm function, and develop a constraint nondegeneracy condition, which provides a theoretical foundation for the analysis of the quadratic convergence of the inexact smoothing Newton method.

When the nuclear norm regularized matrix least squares problem (1.2) has equality constraints only, we introduce a semismooth Newton-CG method, which is preferable to the inexact smoothing Newton method for solving unconstrained inner subproblems. We are able to show that the positive definiteness of the generalized Hessian of the objective function of inner subproblems is equivalent to the constraint nondegeneracy of the corresponding primal problems, which is an important property for successfully applying the semismooth Newton-CG method to solve inner subproblems. The quadratic convergence of the semismooth Newton-CG method is established under the constraint nondegeneracy condition, together with the strong semismoothness property of the soft thresholding operator.

In the second part of this thesis, we focus on designing an efficient algorithm for solving the linearly constrained convex semidefinite programming problem (1.10). In recent years there are intensive studies on the theories, algorithms and applications of large scale structured matrix optimization problems. The accelerated proximal gradient (APG) method, first proposed by Nesterov [90], later refined by Beck and Teboulle [4], and studied in a unifying manner by Tseng [123], has proven to be highly efficient in solving some classes of large scale structured convex optimization problems. The method has superior convergent rate of $O(1/k^2)$ over the classical projected gradient method [47, 67]. Our proposed algorithm is based on the APG method introduced by Beck and Teboulle [4] (named FISTA in [4]), where the subproblem of the form in (1.18) must be solved in each iteration. A limitation of the FISTA method in [4] is that the positive definite linear operator $\mathcal{H}_k$ is restricted to $L\mathcal{I}$, where $\mathcal{I} : \mathcal{S}^n \to \mathcal{S}^n$ denotes the identity map and L is a Lipschitz constant of ∇f. Note that the number of iterations needed by FISTA to achieve ε-optimality (in terms of the function value) is proportional to $\sqrt{L/\varepsilon}$. In many applications, the Lipschitz constant L of ∇f is very large, which will cause the FISTA method to converge very slowly for obtaining a good approximate solution. A more significant limitation of the FISTA method in [4] is that the subproblem (1.18) must be solved exactly to generate the next iterate. However, the subproblem (1.18) generally does not admit an analytical solution, and it could be computationally expensive to solve the subproblem to high accuracy. In this thesis, we design an inexact APG method which is able to overcome the two limitations just mentioned. Specifically, our inexact APG algorithm has the attractive computational advantages that the subproblem (1.18) needs only be solved approximately and $\mathcal{H}_k$ is not restricted to be a scalar multiple of $\mathcal{I}$. In the kth iteration, we are able to choose the positive definite linear operator of the form $\mathcal{H}_k = W_k \circledast W_k$, where $W_k \in \mathcal{S}^n_{++}$. Then the subproblem (1.18) can be solved very efficiently by the semismooth Newton-CG method introduced by Qi and Sun in [97] with warm start using the iterate from the previous iteration, and our inexact APG algorithm can be much more efficient than the state-of-the-art algorithm (the augmented Lagrangian method in [98]) for solving some large scale convex QSDP problems arising from the H-weighted case of the nearest correlation matrix problem (1.14). For the augmented Lagrangian method in [98], when the map $\mathcal{Q}$ associated with the weight matrix H is highly ill-conditioned, then the CG method has great difficulty in solving the ill-conditioned linear system of equations obtained by the semismooth Newton method. In addition, we are able to show that if the subproblem (1.18) is progressively solved with sufficient accuracy, then our inexact APG method enjoys the same superior convergent rate of $O(1/k^2)$ as the exact version.

It seems that the APG algorithm is very suited for solving the nuclear norm regularized matrix least squares problem (1.2). In the kth iteration of the APG method with iterate $\overline{X}_k$, a subproblem of the following form must be solved:

$$\min_{X\in\Re^{p\times q}} \Big\{ \langle \nabla f(\overline{X}_k),\, X - \overline{X}_k\rangle + \frac{L}{2}\|X - \overline{X}_k\|^2 + \rho\|X\|_* \,:\, \mathcal{B}(X) \,\in\, d + \mathcal{Q} \Big\}, \quad (1.21)$$

where $f(X) = \frac{1}{2}\|\mathcal{A}(X) - b\|^2 + \langle C,\, X\rangle$ and L is the Lipschitz constant of ∇f. One significant limitation of the APG algorithm is that the subproblem (1.21) must be solved exactly to generate the next iterate X_{k+1}. The convergence of the APG algorithm with inexact solution of the subproblem (1.21) is still unknown, and we leave it as an interesting topic for future research.

1.4 Organization of the thesis

The thesis is organized as follows: in Chapter 2, we present some preliminaries that are critical for subsequent discussions. We show that the soft thresholding operator is strongly semismooth everywhere, and define a smoothing function of the soft thresholding operator. In Chapter 3, we introduce a partial proximal point algorithm for solving nuclear norm regularized matrix least squares problems with equality and inequality constraints. The inner subproblems, reformulated as a system of semismooth equations, are solved by a quadratically convergent inexact smoothing Newton method. In Chapter 4, we introduce a quadratically convergent semismooth Newton-CG method to solve unconstrained inner subproblems. In Chapter 5, we design an inexact APG algorithm for solving convex QSDP problems, and show that it enjoys the same superior worst-case iteration complexity as the exact counterpart. In Chapter 6, numerical experiments conducted on a variety of large scale nuclear norm minimization and convex QSDP problems show that our proposed algorithms are very efficient and robust. We give the final conclusion of the thesis and discuss a few future research directions in Chapter 7.

Chapter 2

Preliminaries

In this chapter, we give a brief introduction on some basic concepts such as semismooth functions, the B-subdifferential and Clarke's generalized Jacobian of Lipschitz functions. These concepts and properties will be critical for our subsequent discussions.

2.1 Notations

Let $\Re^{p\times q}$ be the space of all $p \times q$ matrices equipped with the standard trace inner product $\langle X, Y\rangle = \mathrm{Tr}(X^T Y)$ and its induced Frobenius norm $\|\cdot\|$. Without loss of generality, we assume $p \leq q$ throughout this thesis. For a given $X \in \Re^{p\times q}$, its nuclear norm $\|X\|_*$ is defined as the sum of all its singular values and its operator norm $\|X\|_2$ is defined as the largest singular value of X. We use the notation $X \geq 0$ to denote that X is a nonnegative matrix, i.e., all entries of X are nonnegative. We let $\mathcal{S}^n$ be the space of all $n\times n$ symmetric matrices, $\mathcal{S}^n_+$ be the cone of symmetric positive semidefinite matrices and $\mathcal{S}^n_{++}$ be the set of symmetric positive definite matrices. We use the notation $X \succeq 0$ to denote that X is a symmetric positive semidefinite matrix. For $U \in \Re^{n\times r}$, $V \in \Re^{n\times s}$, $U \circledast V : \Re^{r\times s} \to \mathcal{S}^n$ is the symmetrized Kronecker product linear map defined by $U \circledast V(M) = (UMV^T + VM^TU^T)/2$. Let $\alpha \subseteq \{1, \ldots, p\}$ and $\beta \subseteq \{1, \ldots, q\}$ be index sets, and X be an $p \times q$ matrix. The cardinality of α

is denoted by $|\alpha|$. We use the notation $X_{\alpha\beta}$ to denote the $|\alpha| \times |\beta|$ submatrix of X formed by selecting the corresponding rows and columns of X indexed by α and β, respectively. For any $X \in \Re^{p\times q}$, $\mathrm{Diag}(X)$ denotes the vector that is the main diagonal of X. For any $x \in \Re^p$, $\mathrm{Diag}(x)$ denotes the diagonal matrix whose ith diagonal element is given by x_i.

Definition 2.1. We say $F : \Re^m \longrightarrow \Re^l$ is directionally differentiable at $x \in \Re^m$ if

$$F'(x;h) := \lim_{t\to 0^+} \frac{F(x+th) - F(x)}{t} \quad \text{exists}$$

for all $h \in \Re^m$ and F is directionally differentiable if F is directionally differentiable at every $x \in \Re^m$.

Let $F : \Re^m \longrightarrow \Re^l$ be a locally Lipschitz function. By Redemacher's theorem [109, Section 9.J], F is Fréchet differentiable almost everywhere. Let D_F denote the set of points in $\Re^m$ where F is differentiable. The Bouligand subdifferential of F at $x \in \Re^m$ is defined by

$$\partial_B F(x) := \{V : \; V = \lim_{k\to\infty} F'(x^k), \; x^k \longrightarrow x, \; x^k \in D_F\},$$

where $F'(x)$ denotes the Jacobian of F at $x \in D_F$. Then the Clarke's [26] generalized Jacobian of F at $x \in \Re^m$ is defined as the convex hull of $\partial_B F(x)$, i.e.,

$$\partial F(x) = \mathrm{conv}\{\partial_B F(x)\}.$$

From [100, Lemma 2.2], we know that if F is directionally differentiable in a neighborhood of $x \in \Re^m$, then for any $h \in \Re^m$, there exists $\mathcal{V} \in \partial F(x)$ such that $F'(x;h) = \mathcal{V}h$. The following concept of semismoothness was first introduced by Mifflin [83] for functionals and was extended by Qi and Sun [100] to vector-valued functions.

Definition 2.2. We say that F is semismooth at x if

1. F is directionally differentiable at x; and

2. for any $h \in \Re^m$ and $V \in \partial F(x+h)$ with $h \to 0$,

$$F(x+h) - F(x) - Vh = o(\|h\|).$$

Furthermore, F is said to be strongly semismooth at x if F is semismooth at x and for any $h \in \Re^m$ and $V \in \partial F(x+h)$ with $h \to 0$,

$$F(x+h) - F(x) - Vh = O(\|h\|^2).$$

2.2 Metric projectors

Let K be a closed convex set in a finite dimensional real Hilbert space $\mathcal{X}$ equipped with a scalar inner product $\langle \cdot, \cdot \rangle$ and its induced norm $\| \cdot \|$. Let $\Pi_K : \mathcal{X} \to \mathcal{X}$ denote the metric projector over K, i.e., for any $y \in \mathcal{X}$, $\Pi_K(y)$ is the unique optimal solution to the following convex optimization problem:

$$\begin{array}{ll} \min & \frac{1}{2}\langle x - y, x - y \rangle \\ \text{s.t.} & x \in K. \end{array} \tag{2.1}$$

It is well known [134] that the metric projector $\Pi_K(\cdot)$ is Lipschitz continuous with modulus 1 and $\|\Pi_K(\cdot)\|^2$ is continuously differentiable. Hence, $\Pi_K(\cdot)$ is almost everywhere Fréchet differentiable in $\mathcal{X}$ and for every $y \in \mathcal{X}$, $\partial \Pi_K(y)$ is well defined. The following lemma [81, Proposition 1] provides the general properties of $\partial \Pi_K(\cdot)$.

Lemma 2.1. *Let $K \subseteq \mathcal{X}$ be a closed convex set. Then, for any $y \in \mathcal{X}$ and $\mathcal{V} \in \partial \Pi_K(y)$, it holds that*

(i) *$\mathcal{V}$ is self-adjoint.*

(ii) $\langle h, \mathcal{V}h \rangle \geq 0 \quad \forall h \in \mathcal{X}$.

(iii) $\langle \mathcal{V}h, h - \mathcal{V}h \rangle \geq 0 \quad \forall h \in \mathcal{X}$.

For $X \in \mathcal{S}^n$, let $X_+ = \Pi_{\mathcal{S}^n}(X)$ be the metric projection of X onto $\mathcal{S}^n_+$ under the standard trace inner product. Assume that X has the following spectral decomposition

$$X = Q\Lambda Q^T, \tag{2.2}$$

where Λ is the diagonal matrix with diagonal entries consisting of the eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k > 0 \geq \lambda_{k+1} \geq \cdots \geq \lambda_n$ of X and Q is a corresponding orthogonal matrix of eigenvectors. Then

$$X_+ = Q\Lambda_+ Q^T,$$

where Λ_+ is a diagonal matrix whose diagonal entries are the nonnegative parts of the respective diagonal entries of Λ. Furthermore, Sun and Sun [115] show that $\Pi_{\mathcal{S}^n_+}(\cdot)$ is strongly semismooth everywhere in $\mathcal{S}^n$. Define the operator $\mathcal{U} : \mathcal{S}^n \longrightarrow \mathcal{S}^n$ by

$$\mathcal{U}(X)[M] = Q(\Omega \circ (Q^T M Q))Q^T, \quad M \in S^n,$$

where " $\circ$ " denotes the Hadamard product of two matrices and

$$\Omega = \begin{bmatrix} E_k & \overline{\Omega} \\ \overline{\Omega}^T & 0 \end{bmatrix}, \quad \overline{\Omega}_{ij} = \frac{\lambda_i}{\lambda_i - \lambda_j}, \; i \in \{1, \ldots, k\}, \; j \in \{k+1, \ldots, n\},$$

where E_k is the square matrix of ones with dimension k (the number of positive eigenvalues), and the matrix $\overline{\Omega}$ has all its entries lying in the interval $[0, 1]$. By Pang, Sun and Sun [94, Lemma 11], $\mathcal{U}$ is an element of the set $\partial\Pi_{\mathcal{S}^n_+}(X)$.

2.3 The soft thresholding operator

In this section, we shall show that the soft thresholding operator [17, 71] is strongly semismooth everywhere. Let $Y \in \Re^{p\times q}$ admit the following singular value decomposition (SVD):

$$Y = U[\Sigma \; 0]V^T, \tag{2.3}$$

where $U \in \Re^{p\times p}$ and $V \in \Re^{q\times q}$ are orthogonal matrices, $\Sigma = \mathrm{Diag}(\sigma_1, \cdots, \sigma_p)$, and $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_p \geq 0$ are singular values of Y being arranged in non-increasing order. For each threshold $\rho > 0$, the soft thresholding operator D_ρ is defined as follows:

$$D_\rho(Y) = U[\Sigma_\rho \ 0]V^T, \tag{2.4}$$

where $\Sigma_\rho = \mathrm{Diag}((\sigma_1 - \rho)_+, \ldots, (\sigma_p - \rho)_+)$.

Lemma 2.2. *Let $G : \mathcal{S}^n \to \mathcal{S}^n$ be defined by*

$$G(X) = (X - \rho I)_+ - (-X - \rho I)_+\,, \ X \in \mathcal{S}^n\,.$$

Then G is strongly semismooth everywhere on $\mathcal{S}^n$.

Proof. This follows directly from the strong semismoothness of $(\cdot)_+ : \mathcal{S}^n \to \mathcal{S}^n$ [115]. □

Decompose $V \in \Re^{q\times q}$ into the form $V = [V_1 \ \ V_2]$, where $V_1 \in R^{q\times p}$ and $V_2 \in \Re^{q\times(q-p)}$. Let the orthogonal matrix $Q \in \Re^{(p+q)\times(p+q)}$ be defined by

$$Q := \frac{1}{\sqrt{2}}\begin{bmatrix} U & U & 0 \\ V_1 & -V_1 & \sqrt{2}V_2 \end{bmatrix}, \tag{2.5}$$

and $\Xi : \Re^{p\times q} \to \mathcal{S}^{p+q}$ be defined by

$$\Xi(Y) := \begin{bmatrix} 0 & Y \\ Y^T & 0 \end{bmatrix}, \quad Y \in \Re^{p\times q}\,. \tag{2.6}$$

Then, by [49, Section 8.6], we know that the symmetric matrix $\Xi(Y)$ has the following spectral decomposition:

$$\Xi(Y) = Q\begin{bmatrix} \Sigma & 0 & 0 \\ 0 & -\Sigma & 0 \\ 0 & 0 & 0 \end{bmatrix} Q^T\,, \tag{2.7}$$

i.e., the eigenvalues of $\Xi(Y)$ are $\pm\sigma_i, i = 1, \ldots, p$, and 0 of multiplicity $q-p$. Define $g_\rho : \Re \to \Re$ by

$$g_\rho(t) := (t-\rho)_+ - (-t-\rho)_+ = \begin{cases} t-\rho & \text{if } t > \rho \\ 0 & \text{if } -\rho \le t \le \rho \\ t+\rho & \text{if } t < -\rho \end{cases}, \; t \in \Re. \tag{2.8}$$

For any $W = P\text{Diag}(\lambda_1, \cdots, \lambda_{(p+q)})P^T \in \mathcal{S}^{p+q}$, define $G_\rho : \mathcal{S}^{p+q} \to \mathcal{S}^{p+q}$ by

$$G_\rho(W) := P\text{Diag}(g_\rho(\lambda_1), \cdots, g_\rho(\lambda_{(p+q)}))P^T = (W-\rho I)_+ - (-W-\rho I)_+ \,.$$

Then, from Lemma 2.2, we have that $G_\rho(\cdot)$ is strongly semismooth everywhere in $\mathcal{S}^{p+q}$. By direct calculations, we have

$$\Psi(Y) \;:=\; G_\rho(\Xi(Y)) \;=\; Q \begin{bmatrix} \Sigma_\rho & 0 & 0 \\ 0 & -\Sigma_\rho & 0 \\ 0 & 0 & 0 \end{bmatrix} Q^T = \begin{bmatrix} 0 & D_\rho(Y) \\ D_\rho(Y)^T & 0 \end{bmatrix}. \tag{2.9}$$

Theorem 2.3. *The function $D_\rho(\cdot)$ is strongly semismooth everywhere in $\Re^{p\times q}$.*

Proof. Let $Y \in \Re^{p\times q}$ admit the SVD as in (2.3). We have known that $G_\rho(\cdot)$ is strongly semismooth in $\mathcal{S}^{p+q}$. This, together with (2.9), proves that $D_\rho(\cdot)$ is strongly semismooth at Y. Since Y is arbitrarily chosen, we have that $D_\rho(\cdot)$ is strongly semismooth everywhere in $\Re^{p\times q}$. □

Note that (2.9) provides an easy way to calculate the derivative, if exists, of D_ρ at Y. We define the following three index sets:

$$\alpha := \{1, \ldots, p\}, \quad \gamma := \{p+1, \ldots, 2p\}, \quad \beta := \{2p+1, \ldots, p+q\}. \tag{2.10}$$

For any $\lambda = (\lambda_1, \ldots, \lambda_{(p+q)})^T \in \mathcal{R}^{p+q}$ and $\lambda_i \ne \pm\rho, i = 1, \ldots, p+q$, we denote by Ω the $(p+q)\times(p+q)$ first divided difference symmetric matrix of $g_\rho(\cdot)$ at λ [12] whose (i,j)th entry is:

$$\Omega_{ij} = \begin{cases} \dfrac{g_\rho(\lambda_i) - g_\rho(\lambda_j)}{\lambda_i - \lambda_j} & \text{if } \lambda_i \ne \lambda_j, \\ g'_\rho(\lambda_i) & \text{if } \lambda_i = \lambda_j. \end{cases}$$

Proposition 2.4. *Let $Y \in \Re^{p\times q}$ admit the SVD as in (2.3). If $\sigma_i \neq \rho, i = 1, \ldots, p$, then, for any $H \in \mathcal{R}^{p\times q}$, it holds that:*

$$D'_\rho(Y)H = \frac{1}{2}U\Big[\left(\Omega_{\alpha\alpha} \circ (H_1 + H_1^T) + \Omega_{\alpha\gamma} \circ (H_1 - H_1^T)\right) V_1^T + 2\left(\Omega_{\alpha\beta} \circ H_2\right) V_2^T\Big], \tag{2.11}$$

where $H_1 = U^T H V_1$ and $H_2 = U^T H V_2$.

Proof. Since $\sigma_i \neq \rho, i = 1, \ldots, p$, from (2.7) and (2.9) we obtain the first divided difference matrix for $g_\rho(\cdot)$:

$$\Omega = \begin{pmatrix} \Omega_{\alpha\alpha} & \Omega_{\alpha\gamma} & \Omega_{\alpha\beta} \\ \Omega_{\alpha\gamma}^T & \Omega_{\gamma\gamma} & \Omega_{\gamma\beta} \\ \Omega_{\alpha\beta}^T & \Omega_{\gamma\beta}^T & \Omega_{\beta\beta} \end{pmatrix}, \tag{2.12}$$

where

$$(\Omega_{\alpha\alpha})_{ij} = \Omega_{ij} = \begin{cases} \dfrac{(\sigma_i - \rho)_+ - (\sigma_j - \rho)_+}{\sigma_i - \sigma_j} & \text{if } \sigma_i \neq \sigma_j \\ g'_\rho(\sigma_i) & \text{if } \sigma_i = \sigma_j \end{cases}, \text{ for } i, j = 1, \ldots, p,$$

$$(\Omega_{\alpha\gamma})_{ij} = \Omega_{i(j+p)} = \begin{cases} \dfrac{(\sigma_i - \rho)_+ + (\sigma_j - \rho)_+}{\sigma_i + \sigma_j} & \text{if } \sigma_i \neq 0 \text{ or } \sigma_j \neq 0 \\ 0 & \text{if } \sigma_i = \sigma_j = 0 \end{cases}, \text{ for } i, j = 1, \ldots, p,$$

$$(\Omega_{\alpha\beta})_{ij} = \Omega_{i(j+2p)} = \begin{cases} \dfrac{(\sigma_i - \rho)_+}{\sigma_i} & \text{if } \sigma_i \neq 0 \\ 0 & \text{if } \sigma_i = 0 \end{cases}, \text{ for } i = 1, \ldots, p,\ j = 1, \ldots, q - p,$$

$$\Omega_{\gamma\gamma} = \Omega_{\alpha\alpha}, \quad \Omega_{\gamma\beta} = \Omega_{\alpha\beta}, \quad \text{and } \Omega_{\beta\beta} = 0.$$

Note that $\Omega_{\alpha\alpha} = \Omega_{\alpha\alpha}^T$ and $\Omega_{\alpha\gamma} = \Omega_{\alpha\gamma}^T$. Then based on the famous result of Löwner [73], , we have from (2.9) that for any $H \in \Re^{p\times q}$

$$\Psi'(Y)H = G'_\rho(\Xi(Y))\Xi(H) = Q\big[\Omega \circ (Q^T \Xi(H) Q)\big]Q^T.$$

Note that

$$
\begin{aligned}
Q^T \Xi(H) Q &= \frac{1}{2}\begin{bmatrix} U^T & V_1^T \\ U^T & -V_1^T \\ 0 & \sqrt{2}V_2^T \end{bmatrix} \begin{bmatrix} 0 & H \\ H^T & 0 \end{bmatrix} \begin{bmatrix} U & U & 0 \\ V_1 & -V_1 & \sqrt{2}V_2 \end{bmatrix} \\
&= \frac{1}{2}\begin{bmatrix} H_1 + H_1^T & H_1^T - H_1 & \sqrt{2}H_2 \\ H_1 - H_1^T & -(H_1 + H_1^T) & \sqrt{2}H_2 \\ \sqrt{2}H_2^T & \sqrt{2}H_2^T & 0 \end{bmatrix}, \qquad (2.13)
\end{aligned}
$$

where $H_1 = U^T H V_1$ and $H_2 = U^T H V_2$. By simple algebraic calculations, we have that

$$
\Psi'(Y)H = Q\big[\Omega \circ (Q^T \Xi(H) Q)\big] Q^T = \begin{bmatrix} 0 & M_{12} \\ M_{12}^T & 0 \end{bmatrix}, \qquad (2.14)
$$

where

$$
M_{12} = \frac{1}{2} U \Big[\big(\Omega_{\alpha\alpha} \circ (H_1 + H_1^T) + \Omega_{\alpha\gamma} \circ (H_1 - H_1^T)\big) V_1^T + 2\,(\Omega_{\alpha\beta} \circ H_2)\, V_2^T \Big].
$$

Since

$$
\Psi'(Y)H = \begin{bmatrix} 0 & D'_\rho(Y)H \\ (D'_\rho(Y)H)^T & 0 \end{bmatrix},
$$

we have from (2.14) that

$$
D'_\rho(Y)H = \frac{1}{2} U \Big(\big[\Omega_{\alpha\alpha} \circ (H_1 + H_1^T) + \Omega_{\alpha\gamma} \circ (H_1 - H_1^T)\big] V_1^T + 2(\Omega_{\alpha\beta} \circ H_2) V_2^T \Big).
$$

□

Next, we give a characterization of the generalized Jacobian of $D_\rho(\cdot)$, which was presented in [131, Lemma 2.3.6 and Proposition 2.3.7]. For any $\lambda = (\lambda_1, \ldots, \lambda_{(p+q)})^T \in \Re^{p+q}$, let $\lambda_i = \sigma_i$ for $i \in \alpha$, $\lambda_i = -\sigma_{i-p}$ for $i \in \gamma$, and $\lambda_i = 0$ for $i \in \beta$. For each threshold $\rho > 0$, we decompose the index set α into the following three subindex sets:

$$
\alpha_1 := \{i \,|\, \sigma_i > \rho, i \in \alpha\}, \ \alpha_2 := \{i \,|\, \sigma_i = \rho, i \in \alpha\}, \ \alpha_3 := \{i \,|\, \sigma_i < \rho, i \in \alpha\}. \qquad (2.15)
$$

Let Γ denote the following $(p+q)\times(p+q)$ symmetric matrix

$$\Gamma = \begin{pmatrix} \Gamma_{\alpha\alpha} & \Gamma_{\alpha\gamma} & \Gamma_{\alpha\beta} \\ \Gamma_{\alpha\gamma}^T & \Gamma_{\gamma\gamma} & \Gamma_{\gamma\beta} \\ \Gamma_{\alpha\beta}^T & \Gamma_{\gamma\beta}^T & \Gamma_{\beta\beta} \end{pmatrix}, \tag{2.16}$$

whose (i,j)th entry is given by

$$\Gamma_{ij} = \begin{cases} \dfrac{g_\rho(\lambda_i) - g_\rho(\lambda_j)}{\lambda_i - \lambda_j} & \text{if } \lambda_i \neq \lambda_j, \\ 1 & \text{if } \lambda_i = \lambda_j \text{ and } |\lambda_i| > \rho, \\ \in \partial g_\rho(\lambda_i) = [0,1] & \text{if } \lambda_i = \lambda_j \text{ and } |\lambda_i| = \rho, \\ 0 & \text{if } \lambda_i = \lambda_j \text{ and } |\lambda_i| < \rho. \end{cases} \tag{2.17}$$

Theorem 2.5. *Let $Y \in \Re^{p\times q}$ admit the SVD as in (2.3). Then, for any $\mathcal{V} \in \partial_B \Psi(Y)$, one has*

$$\mathcal{V}(H) = Q(\Gamma \circ (Q^T \Xi(H) Q))Q^T \quad \forall\, H \in \Re^{p\times q}. \tag{2.18}$$

Moreover, for any $\mathcal{W} \in \partial_B D_\rho(Y)$ and any $H \in \Re^{p\times q}$, we have

$$\mathcal{W}(H) = \frac{1}{2} U\Big[\big(\Gamma_{\alpha\alpha} \circ (H_1 + H_1^T) + \Gamma_{\alpha\gamma} \circ (H_1 - H_1^T)\big) V_1^T + 2\,(\Gamma_{\alpha\beta} \circ H_2)\, V_2^T \Big], \tag{2.19}$$

where $H_1 = U^T H V_1$, $H_2 = U^T H V_2$, and

$$\Gamma_{\alpha\alpha} = \begin{pmatrix} E_{\alpha_1\alpha_1} & E_{\alpha_1\alpha_2} & \tau_{\alpha_1\alpha_3} \\ E_{\alpha_2\alpha_1} & \nu_{\alpha_2\alpha_2} & 0 \\ \tau_{\alpha_1\alpha_3}^T & 0 & 0 \end{pmatrix}, \qquad \begin{array}{c} \nu_{ij} = \nu_{ji} \in [0,1] \text{ for } i,j \in \alpha_2, \\ \tau_{ij} = \dfrac{\sigma_i - \rho}{\sigma_i - \sigma_j}, \text{ for } i \in \alpha_1,\ j \in \alpha_3, \end{array}$$

$$\Gamma_{\alpha\gamma} = \begin{pmatrix} \omega_{\alpha_1\alpha_1} & \omega_{\alpha_1\alpha_2} & \omega_{\alpha_1\alpha_3} \\ \omega_{\alpha_1\alpha_2}^T & 0 & 0 \\ \omega_{\alpha_1\alpha_3}^T & 0 & 0 \end{pmatrix}, \ \omega_{ij} := \frac{(\sigma_i - \rho)_+ + (\sigma_j - \rho)_+}{\sigma_i + \sigma_j}, \text{ for } i \in \alpha_1, j \in \alpha,$$

$$\Gamma_{\alpha\beta} = \begin{pmatrix} \mu_{\alpha_1\bar\beta} \\ 0 \end{pmatrix}, \ \bar\beta = \beta - 2p = \{1,\ldots,q-p\},\ \mu_{ij} = \frac{\sigma_i - \rho}{\sigma_i}, \text{ for } i \in \alpha_1,\ j \in \bar\beta.$$

Define the operator $\mathcal{W}^0 : \Re^{p\times q} \to \Re^{p\times q}$ by

$$\mathcal{W}^0(H) = \frac{1}{2}U\Big[\left(\Gamma^0_{\alpha\alpha}\circ(H_1+H_1^T)+\Gamma_{\alpha\gamma}\circ(H_1-H_1^T)\right)V_1^T + 2\left(\Gamma_{\alpha\beta}\circ H_2\right)V_2^T\Big], \tag{2.20}$$

where

$$\Gamma^0_{\alpha\alpha} = \begin{pmatrix} E_{\alpha_1\alpha_1} & E_{\alpha_1\alpha_2} & \tau_{\alpha_1\alpha_3} \\ E_{\alpha_2\alpha_1} & 0 & 0 \\ \tau^T_{\alpha_1\alpha_3} & 0 & 0 \end{pmatrix},$$

we can easily have that $\mathcal{W}^0$ is an element in $\partial_B D_\rho(Y)$.

In the following, we show that all elements of the generalized Jacobian $\partial D_\rho(\cdot)$ are self-adjoint and positive semidefinite. First we prove the following useful lemma.

Lemma 2.6. *Let $Y \in \Re^{p\times q}$ admit the SVD as in (2.3). Then the unique minimizer of the following problem*

$$\min\Big\{\|X-Y\|^2 \ : \ X \in B_\rho := \{Z\in\Re^{p\times q} \ : \ \|Z\|_2 \le \rho\}\Big\} \tag{2.21}$$

is $X^* = \Pi_{B_\rho}(Y) = U[\min(\Sigma,\rho)\ 0]V^T$, where $\min(\Sigma,\rho) = \mathrm{Diag}(\min(\sigma_1,\rho),\ldots,\min(\sigma_p,\rho))$.

Proof. Obviously problem (2.21) has an unique optimal solution which is equal to $\Pi_{B_\rho}(Y)$. For any $Z \in B_\rho$ with the SVD as in (2.3), we have that $\sigma_i(Z) \le \rho,\ i = 1,\ldots,p$. Since $\|\cdot\|$ is unitarily invariant, by [12, Exercise IV.3.5], we have that

$$\|Y-Z\|^2 \ge \sum_{i\in\alpha_1}(\sigma_i(Y)-\sigma_i(Z))^2 + \sum_{i\in\alpha_2\cup\alpha_3}(\sigma_i(Y)-\sigma_i(Z))^2 \ge \sum_{i\in\alpha_1}(\sigma_i(Y)-\rho)^2.$$

Since

$$\|Y-X^*\|^2 = \sum_{i\in\alpha_1}(\sigma_i(Y)-\rho)^2,$$

we have that

$$\|Y-Z\|^2 \ge \|Y-X^*\|^2 \quad \text{for any } Z\in B_\rho.$$

Thus $X^* = U[\min(\Sigma,\rho)\ 0]V^T$ is the unique optimal solution. □

Note that the above lemma has also been proved in [96] with a different proof. From the above lemma, we have that $\boxed{D_\rho(Y) = Y - \Pi_{B_\rho}(Y)}$, which implies that

$\Pi_{B_\rho}(\cdot)$ is also strongly semismooth everywhere in $\Re^{p\times q}$. Then we have the following proposition.

Proposition 2.7. *For any $Y \in \Re^{p\times q}$ and $\mathcal{V} \in \partial D_\rho(Y)$, it holds that*

(a) $\mathcal{V}$ is self-adjoint.

(b) $\langle H, \mathcal{V}H\rangle \geq 0 \quad \forall H \in \Re^{p\times q}$.

(c) $\langle \mathcal{V}H, H - \mathcal{V}H\rangle \geq 0 \quad \forall H \in \Re^{p\times q}$.

Proof. (a) Since $D_\rho(Y) = Y - \Pi_{B_\rho}(Y)$, for any $\mathcal{V} \in \partial D_\rho(Y)$, there exists $\mathcal{W} \in \partial\Pi_{B_\rho}(Y)$ such that for any $H \in \Re^{p\times q}$,

$$\mathcal{V}H = H - \mathcal{W}H.$$

By (i) of Lemma 2.1, we have that $\mathcal{W}$ is self-adjoint, which implies that $\mathcal{V}$ is self-adjoint.

(b) It is a simple conclusion of (c).

(c) Since for any $H \in \Re^{p\times q}$

$$\langle \mathcal{V}H, H - \mathcal{V}H\rangle = \langle H - \mathcal{W}H, \mathcal{W}H\rangle \geq 0,$$

where the above inequality follows from (iii) of Lemma 2.1, the third inequality holds. □

Next, we shall show that even though the soft thresholding operator $D_\rho(\cdot)$ is not differentiable everywhere, $\|D_\rho(\cdot)\|^2$ is continuously differentiable. First we summarize some well-known properties of Moreau-Yosida [88, 132] regularization. Assume that $\mathcal{Y}$ is a finite-dimensional real Hilbert space. Let $f : \mathcal{Y} \to (-\infty, +\infty]$ be a proper lower semicontinuous convex function. For a given $\sigma > 0$, the Moreau-Yosida regularization of f is defined by

$$F_\sigma(y) = \min\left\{ f(x) + \frac{1}{2\sigma}\|x - y\|^2 \;:\; x \in \mathcal{Y}\right\}. \tag{2.22}$$

It is well known that F_σ is a continuously differentiable convex function on $\mathcal{Y}$ and for any $y \in \mathcal{Y}$

$$\nabla F_\sigma(y) = \frac{1}{\sigma}(y - x(y)),$$

where $x(y)$ denotes the unique optimal solution of (2.22). It is well known that $x(\cdot)$ is globally Lipschitz continuous with modulus 1 and ∇F_σ is globally Lipschitz continuous with modulus $1/\sigma$.

Proposition 2.8. *Let $\Theta(Y) = \frac{1}{2}\|D_\rho(Y)\|^2$, where $Y \in \Re^{p\times q}$. Then $\Theta(Y)$ is continuously differentiable and*

$$\nabla\Theta(Y) = D_\rho(Y). \tag{2.23}$$

Proof. It is already known that the following minimization problem

$$F(Y) \ = \ \min\Big\{\rho\|X\|_* + \frac{1}{2}\|X - Y\|^2 \ : \ X \in \Re^{p\times q}\Big\},$$

has an unique optimal solution $X = D_\rho(Y)$ (see, [17, 77]). From the properties of Moreau-Yosida regularization, we know that $D_\rho(\cdot)$ is globally Lipschitz continuous with modulus 1 and $F(Y)$ is continuously differentiable with

$$\nabla F(Y) = Y - D_\rho(Y). \tag{2.24}$$

Since $D_\rho(Y)$ is the unique optimal solution, we have that

$$F(Y) = \rho\|D_\rho(Y)\|_* + \frac{1}{2}\|D_\rho(Y) - Y\|^2 = \frac{1}{2}\|Y\|^2 - \frac{1}{2}\|D_\rho(Y)\|^2. \tag{2.25}$$

This, together with (2.24), implies that $\Theta(Y)$ is continuously differentiable with

$$\nabla\Theta(Y) = D_\rho(Y).$$

□

2.4 The smoothing counterpart

Next, we shall discuss the smoothing counterpart of the soft thresholding operator $D_\rho(\cdot)$. Let $\phi_H(\varepsilon, t) : \Re \times \Re \to \Re$ be defined by the following Huber smoothing function for $(t)_+ = \max\{t, 0\}$

$$\phi_H(\varepsilon, t) = \begin{cases} t & \text{if } t \geq \frac{|\varepsilon|}{2}, \\ \frac{1}{2|\varepsilon|}(t + \frac{|\varepsilon|}{2})^2 & \text{if } -\frac{|\varepsilon|}{2} < t < \frac{|\varepsilon|}{2}, \\ 0 & \text{if } t \leq -\frac{|\varepsilon|}{2}, \end{cases} \quad (\varepsilon, t) \in \Re \times \Re. \tag{2.26}$$

Then the smoothing function for $g_\rho(\cdot)$ in (2.8) is defined as follows:

$$\phi_\rho(\varepsilon, t) = \phi_H(\varepsilon, t - \rho) - \phi_H(\varepsilon, -t - \rho), \quad (\varepsilon, t) \in \Re \times \Re. \tag{2.27}$$

From the above definition, we know that $\phi_\rho(\varepsilon, \cdot)$ is an odd function about $t \in \Re$. Let $Y \in \Re^{p \times q}$ admit the SVD as in (2.3). For any $\varepsilon \in \Re$, the smoothing function for $G_\rho(\Xi(Y))$ in (2.9) is defined as follows:

$$\Psi_\rho(\varepsilon, Y) := \mathcal{G}_\rho(\varepsilon, \Xi(Y)) = Q \begin{bmatrix} \Sigma_{\phi_\rho} & 0 & 0 \\ 0 & -\Sigma_{\phi_\rho} & 0 \\ 0 & 0 & 0 \end{bmatrix} Q^T, \tag{2.28}$$

where $\Sigma_{\phi_\rho} = \text{Diag}(\phi_\rho(\varepsilon, \sigma_1), \ldots, \phi_\rho(\varepsilon, \sigma_p))$. By direct calculations, we have

$$\mathcal{G}_\rho(\varepsilon, \Xi(Y)) = \begin{bmatrix} 0 & \Phi_\rho(\varepsilon, Y) \\ (\Phi_\rho(\varepsilon, Y))^T & 0 \end{bmatrix},$$

where

$$\Phi_\rho(\varepsilon, Y) = U\big[\text{Diag}(\phi_\rho(\varepsilon, \sigma_1), \ldots, \phi_\rho(\varepsilon, \sigma_p)) \;\; 0\big] V^T, \tag{2.29}$$

which is a smoothing function for the soft thresholding operator $D_\rho(Y)$. Note that when $\varepsilon = 0, \mathcal{G}_\rho(0, \Xi(Y)) = G_\rho(\Xi(Y))$ and $\Phi_\rho(0, Y) = D_\rho(Y)$. For any $\lambda = (\lambda_1, \ldots, \lambda_{(p+q)})^T \in \Re^{p+q}$, let $\lambda_i = \sigma_i$ for $i \in \alpha$, $\lambda_i = -\sigma_{i-p}$ for $i \in \gamma$, and $\lambda_i = 0$ for $i \in \beta$. When $\varepsilon \neq 0$ or $\sigma_i \neq \rho, i = 1, \ldots, p$, we use $\Lambda(\varepsilon, \lambda) \in \mathcal{S}^{p+q}$ to denote the following first divided difference symmetric matrix for $\phi_\rho(\varepsilon, \cdot)$ at λ

$$\Lambda(\varepsilon, \lambda) = \begin{bmatrix} \Lambda_{\alpha\alpha} & \Lambda_{\alpha\gamma} & \Lambda_{\alpha\beta} \\ \Lambda_{\alpha\gamma}^T & \Lambda_{\gamma\gamma} & \Lambda_{\gamma\beta} \\ \Lambda_{\alpha\beta}^T & \Lambda_{\gamma\beta}^T & \Lambda_{\beta\beta} \end{bmatrix}, \tag{2.30}$$

where

$$
\begin{aligned}
(\Lambda_{\alpha\alpha})_{ij} &= \begin{cases} \dfrac{\phi_\rho(\varepsilon,\sigma_i)-\phi_\rho(\varepsilon,\sigma_j)}{\sigma_i-\sigma_j} & \text{if } \sigma_i \neq \sigma_j \\ (\phi_\rho)'_{\sigma_i}(\varepsilon,\sigma_i) & \text{if } \sigma_i = \sigma_j \end{cases}, \ \text{for } i,j=1,\ldots,p, \\
(\Lambda_{\alpha\gamma})_{ij} &= \begin{cases} \dfrac{\phi_\rho(\varepsilon,\sigma_i)+\phi_\rho(\varepsilon,\sigma_j)}{\sigma_i+\sigma_j} & \text{if } \sigma_i \neq 0 \text{ or } \sigma_j \neq 0 \\ (\phi_\rho)'_{\sigma_i}(\varepsilon,\sigma_i) & \text{if } \sigma_i = \sigma_j = 0 \end{cases}, \ \text{for } i,j=1,\ldots,p, \\
(\Lambda_{\alpha\beta})_{ij} &= \begin{cases} \dfrac{\phi_\rho(\varepsilon,\sigma_i)}{\sigma_i} & \text{if } \sigma_i \neq 0 \\ (\phi_\rho)'_{\sigma_i}(\varepsilon,\sigma_i) & \text{if } \sigma_i = 0 \end{cases}, \ \text{for } i=1,\ldots,p,\ j=1,\ldots,q-p, \\
(\Lambda_{\beta\beta})_{ij} &= (\phi_\rho)'_t(\varepsilon,0), \ \text{for } i,j=1,\ldots,q-p.
\end{aligned}
$$

Since $\phi_\rho(\varepsilon,\cdot)$ is an odd function, we can easily obtain the following results:

$$\Lambda_{\alpha\alpha} = \Lambda_{\gamma\gamma}, \quad \Lambda_{\alpha\gamma} = (\Lambda_{\alpha\gamma})^T, \quad \Lambda_{\gamma\beta} = \Lambda_{\alpha\beta},$$

and $(\Lambda(\varepsilon,\lambda))_{ij} \in [0,1]$ for all $i,j = 1,\ldots,p+q$. Then based on the famous result of Löwner [73], we know that for any $H \in \Re^{p\times q}$,

$$(\Psi_\rho)'_Y(\varepsilon,Y)H = (\mathcal{G}_\rho)'_{\Xi(Y)}(\varepsilon,\Xi(Y))\Xi(H) = Q\big[\Lambda(\varepsilon,\lambda)\circ(Q^T\Xi(H)Q)\big]Q^T, \tag{2.31}$$

where " $\circ$ " denotes the Hadamard product and $Q^T\Xi(H)Q$ takes the form as in (2.13). By simple algebraic calculations, we have that

$$(\Psi_\rho)'_Y(\varepsilon,Y)H = Q(\Lambda(\varepsilon,\lambda)\circ(Q^T\Xi(H)Q))Q^T = \begin{bmatrix} 0 & A_{12} \\ A_{12}^T & 0 \end{bmatrix},$$

where

$$A_{12} = U\Big(\Lambda_{\alpha\alpha}\circ\frac{H_1+H_1^T}{2} + \Lambda_{\alpha\gamma}\circ\frac{H_1-H_1^T}{2}\Big)V_1^T + U(\Lambda_{\alpha\beta}\ \circ H_2)V_2^T,$$

$H_1 = U^THV_1$ and $H_2 = U^THV_2$. When $\varepsilon \neq 0$ or $\sigma_i \neq \rho, i = 1,\ldots,p$, the partial derivative of $\Psi_\rho(\cdot,\cdot)$ with respect to ε can be computed by

$$(\Psi_\rho)'_\varepsilon(\varepsilon,Y) = Q\begin{bmatrix} D(\varepsilon,\Sigma) & 0 & 0 \\ 0 & -D(\varepsilon,\Sigma) & 0 \\ 0 & 0 & 0 \end{bmatrix} Q^T = \begin{bmatrix} 0 & UD(\varepsilon,\Sigma)V_1^T \\ V_1D(\varepsilon,\Sigma)U^T & 0 \end{bmatrix}, \tag{2.32}$$

where

$$D(\varepsilon,\Sigma) = \mathrm{Diag}((\phi_\rho)'_\varepsilon(\varepsilon,\sigma_1),\ldots,(\phi_\rho)'_\varepsilon(\varepsilon,\sigma_p)). \tag{2.33}$$

Since

$$(\Psi_\rho)'(\varepsilon,Y)(\tau,H) = \begin{bmatrix} 0 & (\Phi_\rho)'(\varepsilon,Y)(\tau,H) \\ ((\Phi_\rho)'(\varepsilon,Y)(\tau,H))^T & 0 \end{bmatrix},$$

for any $(\tau,H)\in\Re\times\Re^{p\times q}$, we have

$$\boxed{\begin{aligned}(\Phi_\rho)'(\varepsilon,Y)(\tau,H) = U\Big(\Lambda_{\alpha\alpha}\circ\frac{H_1+H_1^T}{2}+\Lambda_{\alpha\gamma}\circ\frac{H_1-H_1^T}{2}+\tau D(\varepsilon,\Sigma)\Big)V_1^T \\ +\, U(\Lambda_{\alpha\beta}\circ H_2)V_2^T.\end{aligned}} \tag{2.34}$$

Thus, $\Phi_\rho(\cdot,\cdot)$ is continuously differentiable around $(\varepsilon,Y)\in\Re\times\Re^{p\times q}$ if $\varepsilon\neq 0$ or $\sigma_i\neq\rho, i=1,\ldots,p$. Furthermore, $\Phi_\rho(\cdot,\cdot)$ is globally Lipschitz continuous and strongly semismooth at any $(0,Y)\in\Re\times\Re^{p\times q}$ [76].

Define $(\Phi_\rho)_{|\alpha_2|}:\Re\times\Re^{|\alpha_2|\times|\alpha_2|}\to\Re^{|\alpha_2|\times|\alpha_2|}$ by replacing the dimension p and q in the definition of $\Phi_\rho:\Re\times\Re^{p\times q}\to\Re^{p\times q}$ with $|\alpha_2|$, respectively, where the index set α_2 is defined as in (2.15). As in the case for $\Phi_\rho(\cdot,\cdot)$, the mapping $(\Phi_\rho)_{|\alpha_2|}(\cdot,\cdot)$ is also Lipschitz continuous. Then the Clarke's generalized Jacobian $\partial\Phi_\rho(0,Y)$ of Φ_ρ at $(0,Y)$ and $\partial(\Phi_\rho)_{|\alpha_2|}(0,Z)$ of $(\Phi_\rho)_{|\alpha_2|}$ at $(0,Z)\in\Re\times\Re^{|\alpha_2|\times|\alpha_2|}$ are both well defined.

Next, we will give a characterization of the generalized Jacobian $\partial\Phi_\rho(0,Y)$ of Φ_ρ at $(0,Y)\in\Re\times\Re^{p\times q}$. Let $\mathcal{D}_{\Phi_\rho}$ be the set of points in $\Re\times\Re^{p\times q}$ at which Φ_ρ is differentiable. Suppose that $\mathcal{N}$ is any set of of Lebesgue measure zero in $\Re\times\Re^{p\times q}$. Then

$$\partial\Phi_\rho(0,Y) = \mathrm{conv}\Big\{\lim_{(\varepsilon^k,Y^k)\to(0,Y)}(\Phi_\rho)'(\varepsilon^k,Y^k)\ :\ (\varepsilon^k,Y^k)\in\mathcal{D}_{\Phi_\rho},\ (\varepsilon^k,Y^k)\notin\mathcal{N}\Big\}. \tag{2.35}$$

Note that $\partial\Phi_\rho(0,Y)$ does not depend on the choice of the null set $\mathcal{N}$ [126, Theorem 4].

Proposition 2.9. *Let $Y\in\Re^{p\times q}$ admit the SVD as in (2.3). Then, for any $\mathcal{V}\in$*

$\partial\Phi_\rho(0,Y)$, there exists $\mathcal{V}_{|\alpha_2|} \in \partial(\Phi_\rho)_{|\alpha_2|}(0,\rho I_{|\alpha_2|})$ such that

$$\begin{aligned}\mathcal{V}(\tau,H) &= U\begin{bmatrix} (\widetilde{H}_1)_{\alpha_1\alpha_1} & (\widetilde{H}_1)_{\alpha_1\alpha_2} & \Omega_{\alpha_1\alpha_3}\circ(\widetilde{H}_1)_{\alpha_1\alpha_3} \\ (\widetilde{H}_1)^T_{\alpha_1\alpha_2} & \mathcal{V}_{|\alpha_2|}(\tau,(\widetilde{H}_1)_{\alpha_2\alpha_2}) & 0 \\ \Omega^T_{\alpha_1\alpha_3}\circ(\widetilde{H}_1)^T_{\alpha_1\alpha_3} & 0 & 0 \end{bmatrix} V_1^T \\ &\quad + U\Big[(\Gamma_{\alpha\gamma}\circ\frac{H_1-H_1^T}{2})V_1^T + (\Gamma_{\alpha\beta}\circ H_2)V_2^T\Big] \end{aligned} \tag{2.36}$$

for all $(\tau,H)\in\Re\times\Re^{p\times q}$, where $\Omega_{\alpha_1\alpha_3},\Gamma_{\alpha\gamma}$ and $\Gamma_{\alpha\beta}$ are defined as follows, respectively,

$$(\Omega_{\alpha_1\alpha_3})_{ij} := \frac{\sigma_i-\rho}{\sigma_i-\sigma_j}, \quad \text{for } i\in\alpha_1,\ j\in\alpha_3, \tag{2.37}$$

$$\Gamma_{\alpha\gamma} := \begin{bmatrix} \omega_{\alpha_1\alpha_1} & \omega_{\alpha_1\alpha_2} & \omega_{\alpha_1\alpha_3} \\ \omega^T_{\alpha_1\alpha_2} & 0 & 0 \\ \omega^T_{\alpha_1\alpha_3} & 0 & 0 \end{bmatrix},\ \omega_{ij} := \frac{(\sigma_i-\rho)_+ + (\sigma_j-\rho)_+}{\sigma_i+\sigma_j},\ \text{for } i\in\alpha_1, j\in\alpha, \tag{2.38}$$

$$\Gamma_{\alpha\beta} := \begin{bmatrix} \mu_{\alpha_1\bar\beta} \\ 0 \end{bmatrix},\ \bar\beta := \beta - 2p = \{1,\ldots,q-p\},\ \mu_{ij} := \frac{\sigma_i-\rho}{\sigma_i},\ \text{for } i\in\alpha_1,\ j\in\bar\beta, \tag{2.39}$$

$I_{|\alpha_2|}$ is an identity matrix of size $|\alpha_2|$, $H_1 = U^THV_1, H_2 = U^THV_2$, and $\widetilde{H}_1 = \frac{1}{2}(H_1+H_1^T)$.

Proof. Let $\mathcal{N} := \{0\}\times\Re^{p\times q}$ which has Lebesgue measure zero in $\Re\times\Re^{p\times q}$ and

$$\partial_{\mathcal{N}}\Phi_\rho(0,Y) := \Big\{\lim_{k\to\infty}(\Phi_\rho)'(\varepsilon^k,Y^k)\ :\ (\varepsilon^k,Y^k)\to(0,Y),\ \varepsilon^k\neq 0\Big\}. \tag{2.40}$$

Then, from (2.35), we have

$$\partial\Phi_\rho(0,Y) = \text{conv}(\partial_{\mathcal{N}}\Phi_\rho(0,Y)).$$

First, we give a characterization of all elements in the set $\partial_{\mathcal{N}}\Phi_\rho(0,Y)$. For any $\mathcal{V}\in\partial_{\mathcal{N}}\Phi_\rho(0,Y)$, there exists a sequence $\{(\varepsilon^k,Y^k)\}\to(0,Y)$ with $\varepsilon^k\neq 0$ such that Φ_ρ is differential at (ε^k,Y^k) and for any $(\tau,H)\in\Re\times\Re^{p\times q}$,

$$\mathcal{V}(\tau,H) = \lim_{k\to\infty}(\Phi_\rho)'(\varepsilon^k,Y^k)(\tau,H).$$

Since $\varepsilon^k \neq 0$, we have that Ψ_ρ defined by (2.28) is differential at (ε^k, Y^k) and

$$\begin{aligned}\lim_{k\to\infty}(\Psi_\rho)'(\varepsilon^k, Y^k)(\tau, H) &= \begin{bmatrix} 0 & \lim_{k\to\infty}(\Phi_\rho)'(\varepsilon^k, Y^k)(\tau, H) \\ (\lim_{k\to\infty}(\Phi_\rho)'(\varepsilon^k, Y^k)(\tau, H))^T & 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 & \mathcal{V}(\tau, H) \\ (\mathcal{V}(\tau, H))^T & 0 \end{bmatrix}.\end{aligned}$$

Let $Y^k = U^k[\Sigma^k\ 0](V^k)^T$ be the SVD of Y^k, where $U^k \in \Re^{p\times p}$ and $V^k \in \Re^{q\times q}$ are orthogonal matrices, $\Sigma^k = \mathrm{Diag}(\sigma_1^k, \cdots, \sigma_p^k)$, and $\sigma_1^k \geq \sigma_2^k \geq \cdots \geq \sigma_p^k \geq 0$ are singular values of Y^k being arranged in non-increasing order. Writing each Σ^k in the same format as Σ:

$$\Sigma^k = \begin{bmatrix} \Sigma^k_{\alpha_1} & 0 & 0 \\ 0 & \Sigma^k_{\alpha_2} & 0 \\ 0 & 0 & \Sigma^k_{\alpha_3} \end{bmatrix},$$

we have $\Sigma = \lim_{k\to\infty} \Sigma^k$, which implies that $\Sigma^k_{\alpha_1} - \rho I_{|\alpha_1|}$ and $\Sigma^k_{\alpha_3} - \rho I_{|\alpha_3|}$ are nonsingular matrices for all k sufficiently large and $\lim_{k\to\infty} \Sigma^k_{\alpha_2} = \Sigma_{\alpha_2} = \rho I_{|\alpha_2|}$. For each k, let $\lambda^k = (\lambda_1^k, \ldots, \lambda_{(p+q)}^k)^T \in \Re^{p+q}$, where $\lambda_i^k = \sigma_i^k$ for $i \in \alpha$, $\lambda_i^k = -\sigma_{i-p}^k$ for $i \in \gamma$, and $\lambda_i^k = 0$ for $i \in \beta$. Let $\Lambda^k \equiv \Lambda^k(\varepsilon^k, \lambda^k)$ be defined by (2.30) and $D^k \equiv D(\varepsilon^k, \Sigma^k)$ be defined by (2.33), respectively. Then, for any $(\tau, H) \in \Re \times \Re^{p\times q}$, we obtain from (2.31) and (2.32) that

$$(\Psi_\rho)'(\varepsilon^k, Y^k)(\tau, H) = Q^k \left[\Lambda^k \circ \left((Q^k)^T \Xi(H) Q^k\right) + \tau \begin{pmatrix} D^k & 0 & 0 \\ 0 & -D^k & 0 \\ 0 & 0 & 0 \end{pmatrix} \right] (Q^k)^T, \tag{2.41}$$

where Q^k has the form as in (2.5). By taking subsequences if necessary, we may assume that $\{U^k\}$ and $\{V^k\}$ are both convergent sequences with limits $U = \lim_{k\to\infty} U^k$ and $V = \lim_{k\to\infty} V^k$ (clearly $Y = U[\Sigma\ 0]V^T$). Since both $\{\Lambda^k\}$ and $\{D^k\}$ are uniformly bounded, by taking subsequences further if necessary, we may assume that both $\{\Lambda^k\}$ and $\{D^k\}$ converge. Let $M = \lim_{k\to\infty} \Lambda^k \circ ((Q^k)^T \Xi(H) Q^k) = \lim_{k\to\infty} \Lambda^k \circ (Q^T \Xi(H) Q)$.

Taking limits on both sizes of (2.41), we obtain that

$$Q^T \begin{bmatrix} 0 & \mathcal{V}(\tau, H) \\ (\mathcal{V}(\tau, H))^T & 0 \end{bmatrix} Q = \begin{bmatrix} M_{\alpha\alpha} & M_{\alpha\gamma} & M_{\alpha\beta} \\ M_{\alpha\gamma}^T & M_{\gamma\gamma} & M_{\gamma\beta} \\ M_{\alpha\beta}^T & M_{\gamma\beta}^T & M_{\beta\beta} \end{bmatrix} + \tau \begin{pmatrix} \lim_{k\to\infty} D^k & 0 & 0 \\ 0 & -\lim_{k\to\infty} D^k & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{2.42}$$

By simple calculations, we have $\lim_{k\to\infty} \Lambda_{\alpha\gamma}^k = \Gamma_{\alpha\gamma}$, $\lim_{k\to\infty} \Lambda_{\alpha\beta}^k = \Gamma_{\alpha\beta}$ and $\lim_{k\to\infty} \Lambda_{\beta\beta}^k = 0$, where $\Gamma_{\alpha\gamma}$ and $\Gamma_{\alpha\beta}$ are of the forms as in (2.38) and (2.39), respectively. Then we obtain that

$$M_{\alpha\alpha} = \begin{bmatrix} (\widetilde{H}_1)_{\alpha_1\alpha_1} & (\widetilde{H}_1)_{\alpha_1\alpha_2} & \Omega_{\alpha_1\alpha_3} \circ (\widetilde{H}_1)_{\alpha_1\alpha_3} \\ (\widetilde{H}_1)_{\alpha_1\alpha_2}^T & \lim_{k\to\infty} (\Lambda_{\alpha\alpha}^k)_{\alpha_2\alpha_2} \circ (\widetilde{H}_1)_{\alpha_2\alpha_2} & 0 \\ \Omega_{\alpha_1\alpha_3}^T \circ (\widetilde{H}_1)_{\alpha_1\alpha_3}^T & 0 & 0 \end{bmatrix},$$

$$M_{\alpha\gamma} = \Gamma_{\alpha\gamma} \circ \frac{1}{2}(H_1^T - H_1),\ M_{\alpha\beta} = \Gamma_{\alpha\beta} \circ (\frac{1}{\sqrt{2}} H_2),\ M_{\gamma\gamma} = -M_{\alpha\alpha},\ M_{\gamma\beta} = M_{\alpha\beta},\ M_{\beta\beta} = 0,$$

and

$$\lim_{k\to\infty} D^k = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \lim_{k\to\infty} D_{\alpha_2\alpha_2}^k & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

where $\Omega_{\alpha_1\alpha_3}$ is of the form as in (2.37), $H_1 = U^T H V_1, H_2 = U^T H V_2, \widetilde{H}_1 = \frac{1}{2}(H_1 + H_1^T)$, and

$$D_{\alpha_2\alpha_2}^k = \mathrm{Diag}\big((\phi_\rho)'_\varepsilon(\varepsilon_k, \sigma_{|\alpha_1|+1}^k), \ldots, (\phi_\rho)'_\varepsilon(\varepsilon_k, \sigma_{|\alpha_1|+|\alpha_2|}^k)\big).$$

By applying (2.34) to $(\Phi_\rho)_{|\alpha_2|}$ at $(\varepsilon^k, \Sigma_{\alpha_2}^k)$, for any $(\tau, \Delta H) \in \Re \times \Re^{|\alpha_2|\times|\alpha_2|}$, we have

$$(\Phi_\rho)'_{|\alpha_2|}(\varepsilon^k, \Sigma_{\alpha_2}^k)(\tau, \Delta H) = (\Lambda_{\alpha\alpha}^k)_{\alpha_2\alpha_2} \circ \frac{\Delta H + (\Delta H)^T}{2} + (\Lambda_{\alpha\gamma}^k)_{\alpha_2\alpha_2} \circ \frac{\Delta H - (\Delta H)^T}{2} + \tau D_{\alpha_2\alpha_2}^k.$$

Since both $\{\Lambda^k\}$ and $\{D^k\}$ converge, we obtain that $\lim_{k\to\infty} (\Phi_\rho)'_{|\alpha_2|}(\varepsilon^k, \Sigma_{\alpha_2}^k)(\tau, \Delta H)$ exists for any $(\tau, \Delta H) \in \Re \times \Re^{|\alpha_2|\times|\alpha_2|}$, which implies that $\lim_{k\to\infty} (\Phi_\rho)'_{|\alpha_2|}(\varepsilon^k, \Sigma_{\alpha_2}^k)$ exists. By the definition of $\partial_{\mathcal{N}}(\Phi_\rho)_{|\alpha_2|}(0, \rho I_{|\alpha_2|})$, which is analogous to the one defined in (2.40), we have that there exists $\mathcal{V}_{|\alpha_2|} \in \partial_{\mathcal{N}}(\Phi_\rho)_{|\alpha_2|}(0, \rho I_{|\alpha_2|})$ such that

$$\mathcal{V}_{|\alpha_2|}(\tau, \Delta H) = \lim_{k\to\infty} (\Phi_\rho)'_{|\alpha_2|}(\varepsilon^k, \Sigma_{\alpha_2}^k)(\tau, \Delta H) \quad \forall\ (\tau, \Delta H) \in \Re \times \Re^{|\alpha_2|\times|\alpha_2|}.$$

In particular, let $\Delta H = (\widetilde{H}_1)_{\alpha_2\alpha_2}$ which is symmetric, we have

$$\mathcal{V}_{|\alpha_2|}(\tau, (\widetilde{H}_1)_{\alpha_2\alpha_2}) = \lim_{k\to\infty} (\Phi_\rho)'_{|\alpha_2|}(\varepsilon^k, \Sigma^k_{\alpha_2})(\tau, (\widetilde{H}_1)_{\alpha_2\alpha_2}) = \lim_{k\to\infty} (\Lambda^k_{\alpha\alpha})_{\alpha_2\alpha_2} \circ (\widetilde{H}_1)_{\alpha_2\alpha_2} + \tau D^k_{\alpha_2\alpha_2}.$$

Then we have

$$M_{\alpha\alpha} + \tau \lim_{k\to\infty} D^k = \begin{bmatrix} (\widetilde{H}_1)_{\alpha_1\alpha_1} & (\widetilde{H}_1)_{\alpha_1\alpha_2} & \Omega_{\alpha_1\alpha_3} \circ (\widetilde{H}_1)_{\alpha_1\alpha_3} \\ (\widetilde{H}_1)^T_{\alpha_1\alpha_2} & \mathcal{V}_{|\alpha_2|}(\tau, (\widetilde{H}_1)_{\alpha_2\alpha_2}) & 0 \\ \Omega^T_{\alpha_1\alpha_3} \circ (\widetilde{H}_1)^T_{\alpha_1\alpha_3} & 0 & 0 \end{bmatrix}. \tag{2.43}$$

By simple algebraic calculations, we obtain from (2.42) that

$$\begin{aligned} \mathcal{V}(\tau, H) &= U \begin{bmatrix} (\widetilde{H}_1)_{\alpha_1\alpha_1} & (\widetilde{H}_1)_{\alpha_1\alpha_2} & \Omega_{\alpha_1\alpha_3} \circ (\widetilde{H}_1)_{\alpha_1\alpha_3} \\ (\widetilde{H}_1)^T_{\alpha_1\alpha_2} & \mathcal{V}_{|\alpha_2|}(\tau, (\widetilde{H}_1)_{\alpha_2\alpha_2}) & 0 \\ \Omega^T_{\alpha_1\alpha_3} \circ (\widetilde{H}_1)^T_{\alpha_1\alpha_3} & 0 & 0 \end{bmatrix} V_1^T \\ &\quad + U\Big[(\Gamma_{\alpha\gamma} \circ \frac{H_1 - H_1^T}{2})V_1^T + (\Gamma_{\alpha\beta} \circ H_2)V_2^T\Big]. \end{aligned}$$

Since $\partial\Phi_\rho(0, Y) = \mathrm{conv}(\partial_{\mathcal{N}}\Phi_\rho(0, Y))$ and $\partial(\Phi_\rho)_{|\alpha_2|}(0, \rho I_{|\alpha_2|}) = \mathrm{conv}(\partial_{\mathcal{N}}(\Phi_\rho)_{|\alpha_2|}(0, \rho I_{|\alpha_2|}))$, from the above equality, we conclude that (2.36) holds. □

Next, we present a useful inequality for elements in $\partial\Phi_\rho(0, Y)$, which is analogous to Proposition 2.7 (c) for the soft thresholding operator $D_\rho(\cdot)$.

Proposition 2.10. *For any $\mathcal{V} \in \partial\Phi_\rho(0, Y)$, it holds that*

$$\langle H - \mathcal{V}(0, H), \mathcal{V}(0, H)\rangle \geq 0 \quad \forall H \in \Re^{p\times q}. \tag{2.44}$$

Proof. First, we show that for any $\mathcal{V} \in \partial_{\mathcal{N}}\Phi_\rho(0, Y)$ defined by (2.40), inequality (2.44) holds. For any $\mathcal{V} \in \partial_{\mathcal{N}}\Phi_\rho(0, Y)$, there exists a sequence $\{(\varepsilon^k, Y^k)\} \to (0, Y), \varepsilon^k \neq 0$ such that Φ_ρ is differential at (ε^k, Y^k) and for any $H \in \Re^{p\times q}$, $\mathcal{V}(0, H) = \lim_{k\to\infty} (\Phi_\rho)'(\varepsilon^k, Y^k)(0, H)$. Since $\varepsilon^k \neq 0$, we have that Ψ_ρ defined by (2.28) is differentiable at (ε^k, Y^k),

$$\begin{aligned}
&\langle H, (\Phi_\rho)'(\varepsilon^k, Y^k)(0,H)\rangle = \frac{1}{2}\langle \Xi(H),\ \Xi((\Phi_\rho)'(\varepsilon^k, Y^k)(0,H))\rangle \\
&= \frac{1}{2}\langle \Xi(H), (\Psi_\rho)'(\varepsilon^k, Y^k)(0,H)\rangle = \frac{1}{2}\langle \Xi(H),\ Q^k\Big(\Lambda^k \circ ((Q^k)^T\Xi(H)Q^k)\Big)(Q^k)^T\rangle \\
&= \frac{1}{2}\langle \widetilde{H}_k, \Lambda^k \circ \widetilde{H}_k\rangle = \frac{1}{2}\sum_{i=1}^{p+q}\sum_{j=1}^{p+q} \Lambda^k_{ij}\ (\widetilde{H}_k)^2_{ij},
\end{aligned}$$

where $\widetilde{H}_k = (Q^k)^T\Xi(H)Q^k$ and the linear map $\Xi(\cdot)$ is defined by (2.6), and

$$\begin{aligned}
&\langle (\Phi_\rho)'(\varepsilon^k, Y^k)(0,H), (\Phi_\rho)'(\varepsilon^k, Y^k)(0,H)\rangle \\
&= \frac{1}{2}\langle (\Psi_\rho)'(\varepsilon^k, Y^k)(0,H), (\Psi_\rho)'(\varepsilon^k, Y^k)(0,H)\rangle \\
&= \frac{1}{2}\langle Q^k\Big(\Lambda^k \circ ((Q^k)^T\Xi(H)Q^k)\Big)(Q^k)^T, Q^k\Big(\Lambda^k \circ ((Q^k)^T\Xi(H)Q^k)\Big)(Q^k)^T\rangle \\
&= \frac{1}{2}\langle \Lambda^k \circ \widetilde{H}_k, \Lambda^k \circ \widetilde{H}_k\rangle = \frac{1}{2}\sum_{i=1}^{p+q}\sum_{j=1}^{p+q}(\Lambda^k_{ij})^2\ (\widetilde{H}_k)^2_{ij}.
\end{aligned}$$

Since $\Lambda^k_{ij} \in [0,1]$ for all $i,j = 1,\ldots,p+q$, we have

$$\langle H - (\Phi_\rho)'(\varepsilon^k, Y^k)(0,H), (\Phi_\rho)'(\varepsilon^k, Y^k)(0,H)\rangle = \frac{1}{2}\sum_{i=1}^{p+q}\sum_{j=1}^{p+q}(\Lambda^k_{ij} - (\Lambda^k_{ij})^2)(\widetilde{H}_k)^2_{ij} \geq 0.$$

Hence

$$\langle H - \mathcal{V}(0,H), \mathcal{V}(0,H)\rangle \geq 0 \quad \forall H \in \Re^{p\times q}.$$

Let $\mathcal{V} \in \partial\Phi_\rho(0,Y)$. Then, by Carathéodory's theorem, there exists a positive κ and $\mathcal{V}^i \in \partial_{\mathcal{N}}\Phi_\rho(0,Y), i = 1,\ldots,\kappa$ such that $\mathcal{V} = \sum_{i=1}^{\kappa} t_i\mathcal{V}^i$, where $t_i \geq 0, i = 1,\ldots,\kappa$ and $\sum_{i=1}^{\kappa} t_i = 1$. Define $\theta(Y) := \langle Y, Y\rangle, Y \in \Re^{p\times q}$. By convexity, we have that for any $H \in \Re^{p\times q}$

$$\theta(\mathcal{V}(0,H)) = \theta(\sum_{i=1}^{\kappa} t_i\mathcal{V}^i(0,H)) \leq \sum_{i=1}^{\kappa} t_i\theta(\mathcal{V}^i(0,H)) = \sum_{i=1}^{\kappa} t_i\langle \mathcal{V}^i(0,H), \mathcal{V}^i(0,H)\rangle,$$

which implies

$$\langle \mathcal{V}(0,H), \mathcal{V}(0,H)\rangle \leq \sum_{i=1}^{\kappa} t_i\langle H, \mathcal{V}^i(0,H)\rangle = \langle H, \sum_{i=1}^{\kappa} t_i\mathcal{V}^i(0,H)\rangle = \langle H, \mathcal{V}(0,H)\rangle.$$

Hence, (2.44) holds. □

Chapter 3

Nuclear norm regularized matrix least squares problems

In this chapter, we introduce a partial proximal point algorithm, in which only some of the variables appear in the quadratic proximal term, for solving nuclear norm regularized matrix least squares problems with equality and inequality constraints. Due to the presence of inequality constraints, the inner subproblem is reformulated as a system of semismooth equations which are then solved by an inexact smoothing Newton method. We prove that the inexact smoothing Newton method is quadratically convergent under a constraint nondegeneracy condition, together with the strong semi-smoothness property of the soft thresholding operator.

3.1 The general proximal point algorithm

Let $\mathcal{Z}$ be a finite dimensional real Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and its induced norm $\| \cdot \|$. Let $\mathcal{T} : \mathcal{Z} \to \mathcal{Z}$ be a set-valued map. We define its domain, image and graph, respectively, as follows:

$$\mathrm{Dom}\,(\mathcal{T}) := \{z \in \mathcal{Z} \mid \mathcal{T}(z) \neq \emptyset\},$$
$$\mathrm{Im}\,(\mathcal{T}) := \bigcup_{z\in\mathcal{Z}} \mathcal{T}(z),$$

and

$$\mathrm{Graph}\,(T) := \{(z, w) \in \mathcal{Z} \times \mathcal{Z} \mid w \in \mathcal{T}(z)\}.$$

The multifunction $\mathcal{T} : \mathcal{Z} \to \mathcal{Z}$ is said to be a monotone operator if

$$\langle z - z',\, w - w'\rangle \geq 0 \quad \text{whenever} \quad w \in \mathcal{T}(z), w' \in \mathcal{T}(z'). \tag{3.1}$$

It is said to be strongly monotone with modulus $\alpha > 0$ if

$$\langle z - z',\, w - w'\rangle \geq \alpha\|z - z'\|^2 \quad \text{whenever} \quad w \in \mathcal{T}(z), w' \in \mathcal{T}(z'). \tag{3.2}$$

The multifunction $\mathcal{T}$ is said to be maximal monotone if it is monotone and its graph Graph (T) is not properly contained in the graph of any other monotone operator. For any maximal monotone operator $\mathcal{T} : \mathcal{Z} \to \mathcal{Z}$, we define the mapping $\mathcal{T}^{-1}$ by

$$\mathcal{T}^{-1}(w) = \big\{z \in \mathcal{Z} \,|\, w \in \mathcal{T}(z)\big\}. \tag{3.3}$$

It is obvious that $\mathcal{T}^{-1}$ is also maximal monotone. We shall say that $\mathcal{T}^{-1}$ is Lipschitz continuous at the origin (with modulus $a \geq 0$) [108] if there is a unique solution $\bar{z}$ to $0 \in \mathcal{T}(z)$ and for some $\tau > 0$ we have

$$\|z - \bar{z}\| \leq a\|w\| \quad \text{whenever} \quad z \in \mathcal{T}^{-1}(w) \quad \text{and} \quad \|w\| \leq \tau. \tag{3.4}$$

Many problems can be formulated as finding an element z such that $0 \in \mathcal{T}(z)$, where $\mathcal{T} : \mathcal{Z} \to \mathcal{Z}$ is a maximal monotone operator. An important example is the following convex programming problem

$$\min_{z\in\mathcal{Z}} f(z), \tag{3.5}$$

where $f : \mathcal{Z} \to (-\infty, +\infty]$ is a proper lower semicontinuous convex function. Let $\mathcal{T} = \partial f$ be the subgradient of f. It is well known that ∂f is maximal monotone.

Moreover, a point $\bar{z} \in \mathcal{Z}$ solves the minimization problem (3.5) if and only if $0 \in \partial f(\bar{z})$. To solve inclusion problems with maximal monotone operators, Rockafellar [108, 107] proposed the general inexact proximal point algorithm (PPA). Given a sequence of parameters σ_k such that

$$0 < \sigma_k \uparrow \sigma_\infty \leq +\infty, \tag{3.6}$$

and an initial point $z^0 \in \mathcal{Z}$, the general PPA generates a sequence $\{z^k\}$ in $\mathcal{Z}$ by the following scheme:

$$z^{k+1} \approx \overline{P}_{\sigma_k}(z^k) \;:=\; (I + \sigma_k \mathcal{T})^{-1}(z^k). \tag{3.7}$$

This algorithm is based upon the fact that the proximal mapping $\overline{P}_{\sigma_k}$ is single-valued and nonexpansive [84]. Rockafellar [108] shows that under certain mild assumptions the sequence $\{z^k\}$ converges to a particular solution z^* for the problem $0 \in \mathcal{T}(z)$. When applied to the minimization problem (3.5), the above approximate rule reduces to

$$z^{k+1} \approx \arg\min_{z \in \mathcal{Z}} \Big\{ f(z) + \frac{1}{2\sigma_k}\|z - z^k\|^2 \Big\}. \tag{3.8}$$

The attractive feature of this approach is that the objective function in (3.8) is strongly convex, which suggests that we may apply an indirect method for solving (3.8) based on the duality theory for convex programming.

Let $\mathcal{Z} = \mathcal{X} \times \mathcal{Y}$, where $\mathcal{X}$ and $\mathcal{Y}$ are two finite dimensional real Hilbert spaces each equipped with a scalar product $\langle \cdot, \cdot \rangle$ and its induced norm $\|\cdot\|$. Suppose now that $z \in \mathcal{Z}$ is partitioned into two components $z = (x, y)$, where $x \in \mathcal{X}$ and $y \in \mathcal{Y}$. Then the approximate rule of the general PPA for solving (3.5) is given by

$$(x^{k+1}, y^{k+1}) \approx \arg\min_{(x,y) \in \mathcal{X} \times \mathcal{Y}} \Big\{ f(x, y) + \frac{1}{2\sigma_k}\|(x, y) - (x^k, y^k)\|^2 \Big\}. \tag{3.9}$$

However, in many applications we may only want to add a quadratic proximal term for only one variable, say y. Then (x^{k+1}, y^{k+1}) is generated by approximately solving the following minimization problem

$$\min_{(x,y) \in \mathcal{X} \times \mathcal{Y}} \Big\{ f(x, y) + \frac{1}{2\sigma_k}\|y - y^k\|^2 \Big\}. \tag{3.10}$$

Note that the objective function in (3.10) may not be strongly convex in (x, y). But in the case that f is already strongly convex in x for all $y \in \mathcal{Y}$, then the problem (3.10) could be easier to solver than (3.9). In [52], Ha proposed a partial PPA to solve the inclusion problem in two variables

$$0 \in \mathcal{T}(x, y), \tag{3.11}$$

in which only one of the variables is involved in the quadratic proximal term. Below we give a brief review of the idea of the partial PPA proposed by Ha [52].

Let $\Pi : \mathcal{X} \times \mathcal{Y} \to \mathcal{X} \times \mathcal{Y}$ be the orthogonal projection of $\mathcal{X} \times \mathcal{Y}$ onto $\{0\} \times \mathcal{Y}$, i.e.,

$$\Pi(x, y) = (0, y), \quad \forall\, (x, y) \in \mathcal{X} \times \mathcal{Y}.$$

Let $T : \mathcal{X} \times \mathcal{Y} \to \mathcal{X} \times \mathcal{Y}$ be a maximal monotone operator. To solve the inclusion problem $0 \in \mathcal{T}(x, y)$, from a given initial point $(x^0, y^0) \in \mathcal{X} \times \mathcal{Y}$, the exact partial PPA generates a sequence $\{(x^k, y^k)\}$ by the following scheme:

$$(x^{k+1}, y^{k+1}) \in (\Pi + \sigma_k \mathcal{T})^{-1}(0, y^k), \tag{3.12}$$

where the sequence $\{\sigma_k\}$ satisfies (3.6). Let $P_{\sigma_k} := (\Pi + \sigma_k \mathcal{T})^{-1}\Pi$. Then (3.12) can be written as

$$(x^{k+1}, y^{k+1}) \in P_{\sigma_k}(x^k, y^k). \tag{3.13}$$

Note that if Π is replaced by the identity map I, then P_{σ_k} would be the standard proximal map $\overline{P}_{\sigma_k}$ of $\mathcal{T}$ in (3.7). In general, the mapping P_{σ_k} is neither single-valued nor nonexpansive. However, by [52, Proposition 2], we know that the second component of $P_{\sigma_k}(x^k, y^k)$ is uniquely determined and nonexpansive. For practical purpose, the following general approximation criteria were introduced in [52]:

$$\|(x^{k+1}, y^{k+1}) - (u^{k+1}, v^{k+1})\| \le \varepsilon_k, \quad \varepsilon_k > 0, \ \sum_{k=0}^{\infty} \varepsilon_k < \infty, \tag{3.14a}$$

$$\|(x^{k+1}, y^{k+1}) - (u^{k+1}, v^{k+1})\| \le \delta_k \|(x^{k+1}, y^{k+1}) - (x^k, y^k)\|, \tag{3.14b}$$

$$\|y^{k+1} - v^{k+1}\| \le \delta_k \|y^{k+1} - y^k\|, \quad \delta_k > 0, \ \sum_{k=0}^{\infty} \delta_k < \infty, \tag{3.14c}$$

where

$$(u^{k+1}, v^{k+1}) \in P_{\sigma_k}(x^k, y^k).$$

In [52], Ha showed that under a mild assumption, namely, $0 \in \operatorname{int} \operatorname{Im}(T)$, the sequence $\{(x^k, y^k)\}$ generated by the partial PPA under criterion (3.14a) is bounded and any of its cluster point is a solution to (3.11). Moreover, the sequence $\{y^k\}$ converges weakly to $\bar{y}$, which is the second component of a solution to (3.11). If, in addition, (3.14b) and (3.14c) are also satisfied and $\mathcal{T}^{-1}$ is Lipschitz continuous at the origin, then the sequence $\{(x^k, y^k)\}$ converges locally at least at a linear rate whose ratio tends to zero as $\sigma_k \to +\infty$. For more discussion of the convergence analysis of the partial PPA, see [52, Theorem 1 & 2].

3.2 A partial proximal point algorithm

In this section, we consider the following nuclear norm regularized matrix least squares problem with linear equality and inequality constraints:

$$\begin{array}{cl} \min\limits_{X \in \Re^{p\times q}} & \dfrac{1}{2}\|\mathcal{A}(X) - b\|^2 + \rho\|X\|_* + \langle C, X\rangle \\ \text{s.t.} & \mathcal{B}(X) \in d + \mathcal{Q}. \end{array} \tag{3.15}$$

where $\mathcal{A} : \Re^{p\times q} \to \Re^m$ and $\mathcal{B} : \Re^{p\times q} \to \Re^s$ are linear maps, $C \in \Re^{p\times q}, b \in \Re^m, d \in \Re^s$, ρ is a given positive parameter, and $\mathcal{Q} = \{0\}^{s_1} \times \Re^{s_2}_+$ is a polyhedral convex cone. Here, $s = s_1 + s_2$. It is easy to see that (3.15) can be rewritten as follows:

$$\begin{array}{cl} \min\limits_{u\in\Re^m, X \in \Re^{p\times q}} & f_\rho(u, X) := \dfrac{1}{2}\|u\|^2 + \rho\|X\|_* + \langle C, X\rangle \\ \text{s.t.} & \mathcal{A}(X) + u = b, \\ & \mathcal{B}(X) \in d + \mathcal{Q}. \end{array} \tag{3.16}$$

Note that the objective function $f_\rho(u, X)$ is strongly convex in u for all $X \in \Re^{p\times q}$. For any $X \in \Re^{p\times q}$ such that $\mathcal{B}(X) \in d + \mathcal{Q}$, let $u = b - \mathcal{A}(X)$, then $(u, X) \in \Re^m \times \Re^{p\times q}$ is a feasible solution of (3.16). Note that the map $(\mathcal{A}, \mathcal{I})$ in (3.16) is surjective, where $\mathcal{I}$ is an identity mapping from $\Re^m$ to $\Re^m$. For the convergence analysis, we assume the following Slater condition to hold throughout this chapter:

$$\begin{cases} \{\mathcal{B}_i\}_{i=1}^{s_1} \text{ are linearly independent and } \exists\, X^0 \in \Re^{p\times q} \\ \text{such that } \mathcal{B}_i(X^0) = d_i, i = 1,\dots,s_1 \text{ and } \mathcal{B}_i(X^0) > d_i, i = s_1+1,\dots,s. \end{cases} \tag{3.17}$$

Let $l(u,X;\zeta,\xi) : \Re^m \times \Re^{p\times q} \times \Re^m \times \Re^s \to \Re$ be the ordinary Lagrangian function for (3.16) in the extended form:

$$l(u,X;\zeta,\xi) := \begin{cases} f_\rho(u,X) + \langle \zeta, b - \mathcal{A}(X) - u\rangle + \langle \xi, d - \mathcal{B}(X)\rangle & \text{if } \xi \in \mathcal{Q}^*, \\ -\infty & \text{if } \xi \notin \mathcal{Q}^*, \end{cases} \tag{3.18}$$

where $\mathcal{Q}^* = \Re^{s_1} \times \Re^{s_2}_+$ is the dual cone of $\mathcal{Q}$. The essential objective function in (3.16) is

$$f(u,X) \;:=\; \sup_{\zeta\in\Re^m,\ \xi\in\Re^s} l(u,X;\zeta,\xi) = \begin{cases} f_\rho(u,X) & \text{if } (u,X) \in \mathcal{F}_P, \\ +\infty & \text{if } (u,X) \notin \mathcal{F}_P, \end{cases} \tag{3.19}$$

where $\mathcal{F}_P = \{(u,X) \in \Re^m \times \Re^{p\times q} \mid \mathcal{A}(X) + u = b, \mathcal{B}(X) \in d + \mathcal{Q}\}$ is the feasible set of (3.16). The dual problem of (3.16) is given by:

$$\begin{array}{ll} \max & g_\rho(\zeta,\xi) := -\frac{1}{2}\|\zeta\|^2 + \langle b,\zeta\rangle + \langle d,\xi\rangle \\ \text{s.t.} & \mathcal{A}^*(\zeta) + \mathcal{B}^*(\xi) + Z = C \\ & \|Z\|_2 \le \rho, \\ & \zeta \in \Re^m,\ \xi \in \mathcal{Q}^*, Z \in \Re^{p\times q}. \end{array} \tag{3.20}$$

As in Rockafellar [107], we define the following two maximal monotone operators

$$\mathcal{T}_f(u,X) = \{(v,Y) \in \Re^m \times \Re^{p\times q} :\ (v,Y) \in \partial f(u,X)\},$$
$$\mathcal{T}_l(u,X;\zeta,\xi) = \{(v,Y,y,z) \in \Re^m \times \Re^{p\times q} \times \Re^m \times \Re^s : (v,Y,-y,-z) \in \partial l(u,X;\zeta,\xi)\},$$

where $u \in \Re^m, X \in \Re^{p\times q}, \zeta \in \Re^m$, and $\xi \in \Re^s$. Note that since $f(u,X)$ is strongly convex in u with modulus 1 for all $X \in \Re^{p\times q}$, $\mathcal{T}_f$ is strongly monotone with modulus 1 with respect to the variable u [108, Proposition 6], i.e.,

$$\langle (u,X) - (u',X'),\, (v,Y) - (v',Y')\rangle \ge \|u - u'\|^2, \tag{3.21}$$

for all $(v, Y) \in \mathcal{T}_f(u, X)$ and $(v', Y') \in \mathcal{T}_f(u', X')$. From the definition of $\mathcal{T}_f$, we know that for any $(v, Y) \in \Re^m \times \Re^{p\times q}$,

$$\mathcal{T}_f^{-1}(v, Y) = \arg \min_{u\in\Re^m, X\in\Re^{p\times q}} \Big\{ f(u, X) - \langle v, u\rangle - \langle Y, X\rangle \Big\}.$$

Similarly, we have that for any $(v, Y, y, z) \in \Re^m \times \Re^{p\times q} \times \Re^m \times \Re^s$,

$$\mathcal{T}_l^{-1}(v, Y, y, z) = \arg \min_{\substack{u\in\Re^m \\ X\in\Re^{p\times q}}} \max_{\substack{\zeta\in\Re^m \\ \xi\in\Re^s}} \Big\{ l(u, X; \zeta, \xi) - \langle v, u\rangle - \langle Y, X\rangle + \langle y, \zeta\rangle + \langle z, \xi\rangle \Big\}.$$

Since $f(u, X)$ is strongly convex in u with modulus 1 for all $X \in \Re^{p\times q}$, we apply the partial PPA proposed by Ha [52] to the maximal monotone operator $\mathcal{T}_f$, in which only the variable X appears in the quadratic proximal term. Given a starting point $(u^0, X^0) \in \Re^m \times \Re^{p\times q}$, the inexact partial PPA generates a sequence $\{(u^k, X^k)\}$ by approximately solving the following problem

$$\min_{u\in\Re^m, X\in\Re^{p\times q}} \Big\{ f(u, X) + \frac{1}{2\sigma_k}\|X - X^k\| \Big\}. \tag{3.22}$$

We can easily have that any minimizer (u, X) of problem (3.22) satisfies

$$(0, 0) \in \partial f(u, X) + (0, \frac{1}{\sigma_k}(X - X^k)).$$

It follows that

$$(0, X^k) \in (0, X) + \sigma_k \mathcal{T}_f(u, X). \tag{3.23}$$

Let $\Pi : \Re^m \times \Re^{p\times q} \to \Re^m \times \Re^{p\times q}$ be the orthogonal projector of $\Re^m \times \Re^{p\times q}$ onto $\{0\} \times \Re^{p\times q}$, i.e.,

$$\Pi(u, X) = (0, X), \quad \forall\, (u, X) \in \Re^m \times \Re^{p\times q}.$$

Then (3.23) can be written as

$$\Pi(u^k, X^k) \in (\Pi + \sigma_k \mathcal{T}_f)(u, X),$$

which can also be equivalently written as

$$(u, X) \in (\Pi + \sigma_k \mathcal{T}_f)^{-1} \Pi(u^k, X^k).$$

Then we have that the set of minimizers of problem (3.22) can be expressed as $(\Pi + \sigma_k \mathcal{T}_f)^{-1}\Pi(u^k, X^k)$.

Next, for any parameter $\sigma > 0$, we show some properties of the mapping $Q_\sigma = (\Pi + \sigma \mathcal{T}_f)^{-1}$ in Proposition 3.1 and some properties of $P_\sigma = Q_\sigma \Pi = (\Pi + \sigma \mathcal{T}_f)^{-1}\Pi$ in Proposition 3.2. The proofs essentially follow the ideas in [60, Proposition 2 & 3].

Proposition 3.1. *For any given parameter $\sigma > 0$, let $Q_\sigma = (\Pi + \sigma \mathcal{T}_f)^{-1}$. Suppose that* $\mathrm{Dom}\,(\mathcal{T}_f) \neq \emptyset$. *Then we have the following properties:*

(i) The mapping Q_σ is single-valued in $\Re^m \times \Re^{p\times q}$.

(ii) For any $(u, X), (u', X') \in \Re^m \times \Re^{p\times q}$,

$$\|Q_\sigma(u, X) - Q_\sigma(u', X')\| \leq \frac{1}{\beta}\|(u, X) - (u', X')\|, \tag{3.24}$$

where $\beta = \min\{1, \sigma\}$.

Proof. (i) By [32, Theorem 2.7], it is enough to show that the map $\Pi + \sigma \mathcal{T}_f$ is maximal monotone and coercive. First we show that $\Pi + \sigma \mathcal{T}_f$ is maximal monotone. Since $\mathcal{T}_f$ is maximal monotone, $\sigma \mathcal{T}_f$ is also maximal monotone for any $\sigma > 0$. Since

$$\mathrm{Dom}\,(\mathcal{T}_f) \cap \mathrm{int}\,\mathrm{Dom}\,(\Pi) = \mathrm{Dom}\,(\mathcal{T}_f) \cap (\Re^m \times \Re^{p\times q}) \neq \emptyset,$$

we have from [105, Theorem 1] that $\Pi + \sigma \mathcal{T}_f$ is maximal monotone.

Next we show that $\Pi + \sigma \mathcal{T}_f$ is strongly monotone. For any $(v, Y) \in \mathcal{T}_f(u, X)$ and $(v', Y') \in \mathcal{T}_f(u', X')$, we have

$$(0, X) + \sigma(v, Y) \in (\Pi + \sigma \mathcal{T}_f)(u, X),$$
$$(0, X') + \sigma(v', Y') \in (\Pi + \sigma \mathcal{T}_f)(u', X').$$

Then, we have

$$\begin{aligned}
&\langle (u,X)-(u',X'),\, [(0,X)+\sigma(v,Y)]-[(0,X')+\sigma(v',Y')]\rangle \\
&= \langle X-X',\, X-X'\rangle + \sigma\langle (u,X)-(u',X'),\, (v,Y)-(v',Y')\rangle \\
&\geq \|X-X'\|^2 + \sigma\|u-u'\|^2 \\
&\geq \beta\|(u,X)-(u',X')\|^2,
\end{aligned}$$

where the first inequality follows from (3.21) and $\beta = \min\{1,\sigma\}$. This implies that $\Pi + \sigma\mathcal{T}_f$ is strongly monotone. Since the strong monotonicity of $\Pi + \sigma\mathcal{T}_f$ implies the coerciveness of $\Pi + \sigma\mathcal{T}_f$, we have that the mapping Q_σ is single-valued.

(ii) For any $(u,X) \in \Re^m \times \Re^{p\times q}$, let $(u_+,X_+) = Q_\sigma(u,X)$. Then from the definition of Q_σ we have

$$(u,X) \in (\Pi + \sigma\mathcal{T}_f)(u_+,X_+).$$

Then there exist some element $(v,Y) \in \mathcal{T}_f(u_+,X_+)$ such that

$$(u,X) = (0,X_+) + \sigma(v,Y). \tag{3.25}$$

Similarly, for any $(u',X') \in \Re^m \times \Re^{p\times q}$, we have

$$(u',X') = (0,X'_+) + \sigma(v',Y'), \tag{3.26}$$

where $(u'_+,X'_+) = Q_\sigma(u',X')$ and $(v',Y') \in \mathcal{T}_f(u'_+,X'_+)$. Since $\mathcal{T}_f$ is strong monotone with respect to the first component with modulus one, we have from (3.21) that

$$\langle (u_+,X_+)-(u'_+,X'_+),\, (v,Y)-(v',Y')\rangle \geq \|u_+-u'_+\|^2. \tag{3.27}$$

It follows from (3.25), (3.26) and (3.27) that

$$\begin{aligned}
&\langle (u,X)-(u',X'),\, (u_+,X_+)-(u'_+,X'_+)\rangle \\
&= \langle [(0,X_+)+\sigma(v,Y)]-[(0,X'_+)+\sigma(v',Y')],\, (u_+,X_+)-(u'_+,X'_+)\rangle \\
&= \|X_+-X'_+\|^2 + \sigma\langle (v,Y)-(v',Y'),\, (u_+,X_+)-(u'_+,X'_+)\rangle \\
&\geq \|X_+-X'_+\|^2 + \sigma\|u_+-u'_+\|^2 \geq \beta\|(u_+,X_+)-(u'_+,X'_+)\|^2,
\end{aligned}$$

where $\beta = \min\{1, \sigma\}$. Then we have

$$\|(u, X) - (u', X')\| \geq \beta\|(u_+, X_+) - (u'_+, X'_+)\| = \beta\|Q_\sigma(u, X) - Q_\sigma(u', X')\|,$$

which completes our proof. □

From the properties of the mapping Q_σ in Proposition 3.1, we can easily obtain the following properties of P_σ.

Proposition 3.2. *For any given parameter $\sigma > 0$, let $P_\sigma = Q_\sigma\Pi = (\Pi + \sigma_k\mathcal{T}_f)^{-1}\Pi$. Suppose that* $\mathrm{Dom}\,(\mathcal{T}_f) \neq \emptyset$. *Then we have the following properties:*

(i) The mapping P_σ is single-valued in $\Re^m \times \Re^{p\times q}$.

(ii) For any $(u, X), (u', X') \in \Re^m \times \Re^{p\times q}$,

$$\|P_\sigma(u, X) - P_\sigma(u', X')\| \leq \frac{1}{\beta}\|X - X'\|, \tag{3.28}$$

where $\beta = \min\{1, \sigma\}$.

Proof. (i) It is obvious from Proposition 3.1 that the mapping P_σ is single-valued.

(ii) From (3.24), we have

$$\begin{aligned}\|P_\sigma(u, X) - P_\sigma(u', X')\| &= \|Q_\sigma\Pi(u, X) - Q_\sigma\Pi(u', X')\| \\ &\leq \frac{1}{\beta}\|\Pi(u, X) - \Pi(u', X')\| = \frac{1}{\beta}\|X - X'\|,\end{aligned}$$

which completes the proof. □

Since the operator P_{σ_k} is single-valued, the approximate rule of the partial PPA for solving problem (3.16) can be expressed as

$$(u^{k+1}, X^{k+1}) \approx P_{\sigma_k}(u^k, X^k) := (\Pi + \sigma_k\mathcal{T}_f)^{-1}\Pi(u^k, X^k), \tag{3.29}$$

where $P_{\sigma_k}(u^k, X^k)$ is defined by

$$P_{\sigma_k}(u^k, X^k) = \arg\min_{u\in\Re^m, X\in\Re^{p\times q}} \left\{f(u, X) + \frac{1}{2\sigma_k}\|X - X^k\|^2\right\}, \tag{3.30}$$

and $\{\sigma_k\}$ is a sequence satisfying (3.6).

Now we calculate the partial quadratic regularization of f in (3.30), which plays a key role in the study of the partial PPA for solving problem (3.16). For a given parameter $\sigma > 0$, the partial quadratic regularization of f in (3.19) associated with σ is given by

$$F_\sigma(X) \;=\; \min_{u\in\Re^m, Y\in\Re^{p\times q}} \Big\{ f(u,Y) + \frac{1}{2\sigma}\|Y - X\|^2 \Big\}. \tag{3.31}$$

From (3.19), we have

$$\begin{aligned}
F_\sigma(X) &= \min_{\substack{u\in\Re^m\\ Y\in\Re^{p\times q}}} \sup_{\substack{\zeta\in\Re^m\\ \xi\in\Re^s}} \Big\{ l(u,Y;\zeta,\xi) + \frac{1}{2\sigma}\|Y-X\|^2 \Big\} \\
&= \sup_{\substack{\zeta\in\Re^m\\ \xi\in\Re^s}} \min_{\substack{u\in\Re^m\\ Y\in\Re^{p\times q}}} \Big\{ l(u,Y;\zeta,\xi) + \frac{1}{2\sigma}\|Y-X\|^2 \Big\} \qquad (3.32)\\
&= \sup_{\substack{\zeta\in\Re^m\\ \xi\in\mathcal{Q}^*}} \min_{\substack{u\in\Re^m\\ Y\in\Re^{p\times q}}} \Big\{ f_\rho(u,Y) + \langle \zeta, b - \mathcal{A}(Y) - u\rangle + \langle \xi, d - \mathcal{B}(Y)\rangle + \frac{1}{2\sigma}\|Y-X\|^2 \Big\},
\end{aligned}$$

where the interchange of $\min_{u,Y}$ and $\sup_{\zeta,\xi}$ follows from the growth properties in (u,Y) [104, Theorem 37.3] and the third equality follows from (3.18). Then, we have

$$F_\sigma(X) \;=\; \sup_{\zeta\in\Re^m,\ \xi\in\mathcal{Q}^*} \Theta_\sigma^\rho(\zeta,\xi;X),$$

where

$$\begin{aligned}
\Theta_\sigma^\rho(\zeta,\xi;X) &:= \min_{\substack{u\in\Re^m\\ Y\in\Re^{p\times q}}} \Big\{ f_\rho(u,Y) + \langle \zeta, b - \mathcal{A}(Y) - u\rangle + \langle \xi, d - \mathcal{B}(Y)\rangle + \frac{1}{2\sigma}\|Y-X\|^2 \Big\} \\
&= \min_{Y\in\Re^{p\times q}} \Big\{ \rho\|Y\|_* + \langle C - \mathcal{A}^*\zeta - \mathcal{B}^*\xi, Y\rangle + \frac{1}{2\sigma}\|Y-X\|^2 \Big\} \\
&\quad + \min_{u\in\Re^m} \Big\{ \frac{1}{2}\|u\|^2 - \langle \zeta, u\rangle \Big\} + \langle b,\zeta\rangle + \langle d,\xi\rangle \\
&= -\frac{1}{2}\|\zeta\|^2 + \langle b,\zeta\rangle + \langle d,\xi\rangle + \min_{Y\in\Re^{p\times q}} \Big\{ \rho\|Y\|_* + \frac{1}{2\sigma}\|Y - W(\zeta,\xi;X)\| \Big\} \\
&\quad + \frac{1}{2\sigma}\|X\|^2 - \frac{1}{2\sigma}\|W(\zeta,\xi;X)\|^2 \\
&= -\frac{1}{2}\|\zeta\|^2 + \langle b,\zeta\rangle + \langle d,\xi\rangle + \frac{1}{2\sigma}\|X\|^2 - \frac{1}{2\sigma}\|D_{\rho\sigma}(W(\zeta,\xi;X))\|^2, \qquad (3.33)
\end{aligned}$$

where $W(\zeta,\xi;X) = X - \sigma(C - \mathcal{A}^*\zeta - \mathcal{B}^*\xi)$ and the last equality follows from (2.25). By the saddle point theorem [104, Theorem 28.3] and (3.32), we know that

$(\zeta(X), D_{\rho\sigma}(W(\zeta(X), \xi(X); X)))$ is the unique solution to (3.31) for any $(\zeta(X), \xi(X))$ such that

$$(\zeta(X), \xi(X)) \in \arg \sup_{\zeta \in \Re^m,\ \xi \in \mathcal{Q}^*} \Theta_\sigma^\rho(\zeta, \xi; X).$$

Then we have $F_\sigma(X) = \Theta_\sigma^\rho(\zeta(X), \xi(X); X)$.

Now we formally present the partial PPA for solving problem (3.16).

Algorithm 1: PPA. Given a positive parameter ρ and a tolerance $\varepsilon > 0$. Input $(u^0, X^0) \in \Re^m \times \Re^{p\times q}$ and $\sigma_0 > 0$. Set $k := 0$. Iterate:

Step 1. Compute an approximate maximizer

$$\Re^m \times \mathcal{Q}^* \ni (\zeta^{k+1}, \xi^{k+1}) \approx \arg \sup_{\zeta \in \Re^m, \xi \in \Re^s} \theta_{\sigma_k}^\rho(\zeta, \xi), \tag{3.34}$$

where

$$\theta_{\sigma_k}^\rho(\zeta, \xi) := \Theta_{\sigma_k}^\rho(\zeta, \xi; X^k) - \delta(\xi \mid \mathcal{Q}^*), \tag{3.35}$$

$\Theta_{\sigma_k}^\rho(\zeta, \xi; X^k)$ is defined as in (3.33) and $\delta(\cdot \mid \mathcal{Q}^*)$ is the indicator function over $\mathcal{Q}^*$.

Step 2. Compute $W^{k+1} := W(\zeta^{k+1}, \xi^{k+1}; X^k)$. Set

$$u^{k+1} = \zeta^{k+1},\ X^{k+1} = D_{\rho\sigma_k}(W^{k+1}),\ \text{and } Z^{k+1} = \frac{1}{\sigma_k}(D_{\rho\sigma_k}(W^{k+1}) - W^{k+1}).$$

Step 3. If $\|(X^k - X^{k+1})/\sigma_k\| \le \varepsilon$; stop; else; update σ_k; end.

Suppose that $(\bar\zeta(X^k), \bar\xi(X^k))$ is an optimal solution of the inner subproblem (3.34) for each X^k and $\sigma_k > 0$. Let P_{σ_k} be defined as in (3.30). Since P_{σ_k} is single-valued, we have $P_{\sigma_k}(u^k, X^k) = \left\{(\bar\zeta(X^k), D_{\rho\sigma_k}(W(\bar\zeta(X^k), \bar\xi(X^k); X^k)))\right\}$. In order to terminate (3.34) in the above PPA , we introduce the following stopping criteria:

$$\sup \theta^{\rho}_{\sigma_k}(\zeta,\xi) - \theta^{\rho}_{\sigma_k}(\zeta^{k+1},\xi^{k+1}) \leq \frac{\varepsilon_k^2}{4\sigma_k}, \tag{3.36a}$$

$$\|\zeta^{k+1} - \bar{\zeta}(X^k)\|^2 \leq \frac{1}{2}\varepsilon_k^2, \quad \varepsilon_k > 0, \ \sum_{k=0}^{\infty} \varepsilon_k < \infty, \tag{3.36b}$$

$$\sup \theta^{\rho}_{\sigma_k}(\zeta,\xi) - \theta^{\rho}_{\sigma_k}(\zeta^{k+1},\xi^{k+1}) \leq \frac{\delta_k^2}{2\sigma_k}\|X^{k+1} - X^k\|^2, \tag{3.36c}$$

$$\|\zeta^{k+1} - \bar{\zeta}(X^k)\|^2 \leq \delta_k^2\|\zeta^{k+1} - \zeta^k\|^2, \quad \delta_k > 0, \ \sum_{k=0}^{\infty} \delta_k < \infty, \tag{3.36d}$$

$$\mathrm{dist}(0, \partial\theta^{\rho}_{\sigma_k}(\zeta^{k+1},\xi^{k+1})) \leq \frac{\delta_k'}{\sigma_k}\|X^{k+1} - X^k\|, \quad 0 \leq \delta_k' \to 0. \tag{3.36e}$$

Note that $F_{\sigma_k}(X^k) = \sup \theta^{\rho}_{\sigma_k}(\zeta,\xi)$ and $\theta^{\rho}_{\sigma_k}(\zeta^{k+1},\xi^{k+1}) = \Theta^{\rho}_{\sigma_k}(\zeta^{k+1},\xi^{k+1};X^k)$. The following result reveals the relation between the estimation (3.36) and (3.14), which enables us to apply the convergence results of the partial PPA in [52, Theorem 1 & 2] to our partial PPA. The proof essentially follows the idea in [107, Proposition 6].

Proposition 3.3. *Suppose that $(\bar{\zeta}(X^k), \bar{\xi}(X^k))$ is an optimal solution of the inner subproblem (3.34). Let $(\bar{u}^{k+1}, \overline{X}^{k+1}) = (\bar{\zeta}(X^k), D_{\rho\sigma_k}(W(\bar{\zeta}(X^k), \bar{\xi}(X^k); X^k)))$ and $X^{k+1} = D_{\rho\sigma_k}(W(\zeta^{k+1}, \xi^{k+1}; X^k))$. Then one has*

$$\frac{1}{2\sigma_k}\|X^{k+1} - \overline{X}^{k+1}\|^2 \leq \sup \theta^{\rho}_{\sigma_k}(\zeta,\xi) - \theta^{\rho}_{\sigma_k}(\zeta^{k+1},\xi^{k+1}). \tag{3.37}$$

Proof. Since $\Theta^{\rho}_{\sigma}(\zeta,\xi;X)$ is convex in X and

$$\nabla_X \Theta^{\rho}_{\sigma_k}(\zeta^{k+1},\xi^{k+1};X^k) = \frac{1}{\sigma_k}(X^k - X^{k+1}),$$

the following inequality holds for any $Y \in \Re^{p\times q}$:

$$\begin{aligned}
&\Theta^{\rho}_{\sigma_k}(\zeta^{k+1},\xi^{k+1};X^k) + \langle \sigma_k^{-1}(X^k - X^{k+1}), Y - X^k\rangle \\
&\leq \Theta^{\rho}_{\sigma_k}(\zeta^{k+1},\xi^{k+1};Y) \ \leq \sup_{\zeta\in\Re^m,\xi\in\mathcal{Q}^*} \Theta^{\rho}_{\sigma_k}(\zeta,\xi;Y) = F_{\sigma_k}(Y) \\
&= \min_{\substack{u\in\Re^m \\ X\in\Re^{p\times q}}} \{f(u,X) + \frac{1}{2\sigma_k}\|X - Y\|^2\} \leq f(\bar{u}^{k+1}, \overline{X}^{k+1}) + \frac{1}{2\sigma_k}\|\overline{X}^{k+1} - Y\|^2.
\end{aligned} \tag{3.38}$$

We also know that

$$\begin{aligned}
\sup \theta^{\rho}_{\sigma_k}(\zeta,\xi) &= F_{\sigma_k}(X^k) = \min_{u\in\Re^m, X\in\Re^{p\times q}} \{f(u,X) + \frac{1}{2\sigma_k}\|X - X^k\|^2\} \\
&= f(\bar{u}^{k+1}, \overline{X}^{k+1})) + \frac{1}{2\sigma_k}\|\overline{X}^{k+1} - X^k\|^2,
\end{aligned} \tag{3.39}$$

which together with (3.38) and the fact that $\theta^{\rho}_{\sigma_k}(\zeta^{k+1},\xi^{k+1}) = \Theta^{\rho}_{\sigma_k}(\zeta^{k+1},\xi^{k+1};X^k)$, implies that

$$\begin{aligned}
& \sup \theta^{\rho}_{\sigma_k}(\zeta,\xi) - \theta^{\rho}_{\sigma_k}(\zeta^{k+1},\xi^{k+1}) \\
\geq\ & \frac{1}{2\sigma_k}\Big[\|\overline{X}^{k+1} - X^k\|^2 - \|\overline{X}^{k+1} - Y\|^2 - 2\langle X^{k+1} - X^k, Y - X^k\rangle\Big] \\
=\ & \frac{1}{2\sigma_k}\Big[2\langle \overline{X}^{k+1} - X^{k+1}, Y - X^k\rangle - \|Y - X^k\|^2\Big] \\
=\ & \frac{1}{2\sigma_k}\Big[-\|(\overline{X}^{k+1} + X^k - X^{k+1}) - Y\|^2 + \|\overline{X}^{k+1} - X^{k+1}\|^2\Big]. \qquad (3.40)
\end{aligned}$$

Since this inequality holds for all $Y \in \Re^{p\times q}$, by taking the maximum of (3.40) in Y, we have

$$\sup \theta^{\rho}_{\sigma_k}(\zeta,\xi) - \theta^{\rho}_{\sigma_k}(\zeta^{k+1},\xi^{k+1}) \geq \frac{1}{2\sigma_k}\|\overline{X}^{k+1} - X^{k+1}\|^2,$$

which proves our assertion. □

3.3 Convergence analysis of the partial PPA

In this section, we show the global convergence and local convergence of the partial PPA for solving (3.16), mainly based upon the convergence results of Ha [52, Theorem 1 & 2].

Proposition 3.4. *Consider the function $f(u,X)$ defined in (3.19). Suppose that for some $\lambda > 0$, the following parameterized problem perturbed by $(v,Y) \in \Re^m \times \Re^{p\times q}$*

$$\min_{u\in\Re^m, X\in\Re^{p\times q}} \Big\{ f(u,X) - \langle u,\, v\rangle - \langle X,\, Y\rangle \Big\} \qquad (3.41)$$

has an optimal solution whenever $\max\{\|v\|, \|Y\|\} \leq \lambda$. *Then we have*

$$0 \in \operatorname{int} \operatorname{Im}(\mathcal{T}_f). \qquad (3.42)$$

Proof. Since for each $(v,Y) \in \Re^m \times \Re^{p\times q}$ such that $\max\{\|v\|, \|Y\|\} \leq \lambda$ the parameterized problem (3.41) has an optimal solution $(\bar{u}, \overline{X})$, we have that

$$0 \in \partial f(\bar{u}, \overline{X}) - (v, Y),$$

which implies that $(v, Y) \in \partial f(\bar{u}, \overline{X}) \subseteq \operatorname{Im}(\mathcal{T}_f)$. Therefore, we have $0 \in \operatorname{int}\operatorname{Im}(\mathcal{T}_f)$.

□

Remark 3.5. *In many applications, we have that the linear term $\langle C, X\rangle$ is absent in the objective function $f_\rho(u, X) = \frac{1}{2}\|u\|^2 + \rho\|X\|_* + \langle C, X\rangle$, i.e., $C = 0$ (see examples in Section 6.1). Since*

$$\rho\|X\|_* - \langle X, Y\rangle \geq \rho\|X\|_* - \|Y\|_2\|X\|_* = (\rho - \|Y\|_2)\|X\|_*,$$

we have that the function $f_\rho(u, X)$ perturbed by $(v, Y) \in \Re^m \times \Re^{p\times q}$ with $\|Y\|_2 < \rho$

$$f_\rho(u, X) - \langle u, v\rangle - \langle X, Y\rangle = \frac{1}{2}\|u\|^2 + \rho\|X\|_* - \langle u, v\rangle - \langle X, Y\rangle$$

is coercive. Therefore, if $C = 0$ and $\lambda > 0$ is small enough, the parameterized problem (3.41) has an optimal solution for any $(v, Y) \in \Re^m \times \Re^{p\times q}$ such that $\max\{\|v\|, \|Y\|\} \leq \lambda$, *which implies that* $0 \in \operatorname{int}\operatorname{Im}(\mathcal{T}_f)$.

Theorem 3.6. (**Global convergence**) *Suppose that the hypotheses in Proposition 3.4 are satisfied. Let the partial PPA be executed with the stopping criterion (3.36a) and (3.36b). Then the generated sequence $\{(u^k, X^k)\}$ is bounded and converges to $(\bar{u}, \overline{X})$, where $(\bar{u}, \overline{X})$ is some optimal solution to problem (3.16), and $\{(\zeta^k, \xi^k)\}$ is asymptotically minimizing for problem (3.20) with*

$$\|C - \mathcal{A}^*(\zeta^{k+1}) - \mathcal{B}^*(\xi^{k+1}) - Z^{k+1}\| = \frac{1}{\sigma_k}\|X^{k+1} - X^k\| \to 0, \tag{3.43}$$

$$\text{asym sup}(D) - g_\rho(\zeta^k, \xi^k) \leq \frac{1}{2\sigma_k}\Big[\frac{1}{2}\varepsilon_k^2 + \|X^k\|^2 - \|X^{k+1}\|^2\Big], \tag{3.44}$$

where asym sup(D) *is the asymptotic supreme of the dual problem (3.20). If problem* (3.16) *satisfies the Slater condition (3.17), then the sequence $\{(\zeta^k, \xi^k)\}$ is also bounded, and all of its accumulation points are optimal solutions to the problem (3.20).*

Proof. Under the given assumption, we have from Proposition 3.4 that $0 \in \operatorname{int}\operatorname{Im}(\mathcal{T}_f)$. Moreover, we know from Proposition 3.3 that (3.36a) and (3.36b) implies the general stopping criterion (3.14a) for $\mathcal{T}_f$. It follows from [52, Theorem 1] that the sequence $\{(u^k, X^k)\}$ is bounded and any of its weak cluster point is an optimal solution to (3.16) and $X^k \to \overline{X}$. Since $f_\rho(u, X)$ is strongly convex with respect to u, the u-component of the optimal solution is uniquely determined, which implies that $\{u^k\} \to \bar{u}$. Thus the whole sequence $\{(u^k, X^k)\}$ converges to an optimal solution $(\bar{u}, \overline{X})$ of (3.16). The rest proof follows the similar discussion as in [108, Theorem 4]. Since

$$C - \mathcal{A}^*(\zeta^{k+1}) - \mathcal{B}^*(\xi^{k+1}) - Z^{k+1} = \frac{1}{\sigma_k}(X^{k+1} - X^k), \text{ and } X^{k+1} - X^k \to 0,$$

relation (3.43) holds. Observing that

$$\theta^\rho_{\sigma_k}(\zeta^{k+1}, \xi^{k+1}) = -\frac{1}{2}\|\zeta^{k+1}\|^2 + \langle b, \zeta^{k+1}\rangle + \langle d, \xi^{k+1}\rangle + \frac{1}{2\sigma_k}(\|X^k\|^2 - \|X^{k+1}\|^2),$$

we have

$$\theta^\rho_{\sigma_k}(\zeta^{k+1}, \xi^{k+1}) - g_\rho(\zeta^{k+1}, \xi^{k+1}) = \frac{1}{2\sigma_k}(\|X^k\|^2 - \|X^{k+1}\|^2). \tag{3.45}$$

From (3.39) in the proof of Proposition 3.3 we can also have that

$$\sup \theta^\rho_{\sigma_k}(\zeta, \xi) \geq f(P_{\sigma_k}(u^k, X^k)) \geq \min f(u, X). \tag{3.46}$$

Combining (3.45) and (3.46), we have

$$\min f(u, X) - g_\rho(\zeta^{k+1}, \xi^{k+1}) \tag{3.47}$$

$$\leq \sup \theta^\rho_{\sigma_k}(\zeta, \xi) - \theta^\rho_{\sigma_k}(\zeta^{k+1}, \xi^{k+1}) + \frac{1}{2\sigma_k}(\|X^k\|^2 - \|X^{k+1}\|^2) \tag{3.48}$$

$$\leq \frac{\varepsilon_k^2}{4\sigma_k} + \frac{1}{2\sigma_k}(\|X^k\|^2 - \|X^{k+1}\|^2) = \frac{1}{2\sigma_k}(\frac{1}{2}\varepsilon_k^2 + \|X^k\|^2 - \|X^{k+1}\|^2). \tag{3.49}$$

For every $(\zeta, \xi) \in \Re^m \times \mathcal{Q}^*$, we have

$$l(\bar{u}, \overline{X}; \zeta, \xi) \leq \sup_{\zeta \in \Re^m, \xi \in \Re^s} l(\bar{u}, \overline{X}; \zeta, \xi) = f(\bar{u}, \overline{X}) = \min f(u, X).$$

Since

$$g_\rho(\zeta, \xi) = \inf_{u \in \Re^m, X^{p \times q}} l(u, X; \zeta, \xi) \leq l(\bar{u}, \overline{X}; \zeta, \xi) \leq \min f(u, X),$$

we have asym $\sup(D) \leq \min f(u, X)$. Therefore we have from (3.47) that the relation (3.44) holds. If (3.16) satisfies the Slater condition (3.17), it follows from [106, Theorem 17 & 18] that all the level sets $\{(\zeta,\xi) \in \Re^m \times \Re^s \,|\, g_\rho(\zeta,\xi) \geq \beta,\, \beta \in \Re\}$ are bounded. Then the last part of the conclusion can be obtained from (3.43) and (3.44). □

Theorem 3.7. (**Local convergence**) *Suppose that the hypotheses in Proposition 3.4 are satisfied. Let the partial PPA be executed with the stopping criterion (3.36a), (3.36b), (3.36c) and (3.36d). If $\mathcal{T}_f^{-1}$ is Lipschitz continuous at the origin with modulus a_f, then $\{(u^k, X^k)\}$ converges to $(\bar{u}, \overline{X})$, where $(\bar{u}, \overline{X})$ is the unique optimal solution to problem* (3.16)*, and*

$$\|X^{k+1} - \overline{X}\| \leq \eta_k \|X^k - \overline{X}\|, \quad \textit{for all } k \textit{ sufficiently large}, \tag{3.50}$$

where

$$\eta_k = [a_f(a_f^2 + \sigma_k^2)^{-1/2} + \delta_k](1-\delta_k)^{-1} \to \eta_\infty = a_f(a_f^2 + \sigma_\infty^2)^{-1/2} < 1.$$

Moreover, the conclusions of Theorem 3.6 about $\{(\zeta^k, \xi^k)\}$ are valid.

If in addition to (3.36c), (3.36d) and the condition on $\mathcal{T}_f^{-1}$, one has (3.36e) and $\mathcal{T}_l^{-1}$ is Lipschitz continuous at the origin with modulus $a_l\,(\geq a_f)$, then $(\zeta^k, \xi^k) \to (\bar{\zeta}, \bar{\xi})$, where $(\bar{\zeta}, \bar{\xi})$ is the unique optimal solution to problem (3.20), and one has

$$\|(\zeta^{k+1}, \xi^{k+1}) - (\bar{\zeta}, \bar{\xi})\| \leq \eta'_k \|X^{k+1} - X^k\|, \quad \textit{for all } k \textit{ sufficiently large}, \tag{3.51}$$

where $\eta'_k = a_l(1+\delta'_k)/\sigma_k \to \eta'_\infty = a_l/\sigma_\infty$.

Proof. Since it follows from Proposition 3.3 that (3.36c) and (3.36d) implies the general stopping criterion (3.14b) and (3.14c), we can easily obtain the first part of the theorem from Theorem 3.6 and the general results in [52, Theorem 2]. The second part of theorem can be similarly obtained by following the discussion in [107, Theorem 5]. We omit it here. □

Remark 3.8. *At the moment, we do not study the characterization of the Lipschitz continuity of $\mathcal{T}_f^{-1}$ at the origin. But it is certainly an interesting problem to study.*

3.4 An inexact smoothing Newton method for inner subproblems

In this section, we will introduce an inexact smoothing Newton method for solving the inner subproblem (3.34).

3.4.1 Inner subproblems

For the convenience of subsequent discussion, we let

$$\widehat{\mathcal{A}} = \begin{pmatrix} \mathcal{A} \\ \mathcal{B} \end{pmatrix}, \ \hat{b} = (b; d) \in \Re^{m+s}, \ \mathcal{K} = \Re^m \times \mathcal{Q}^* \subseteq \Re^m \times \Re^s, \text{ and } y = (\zeta; \xi) \in \mathcal{K}. \quad (3.52)$$

In our proposed partial PPA, for some fixed $X \in \Re^{p \times q}$ and $\sigma > 0$, we need to solve the following form of inner subproblem

$$\min_{y \in \mathcal{K}} \Big\{ \varphi(y) := \frac{1}{2}\langle y, Ty \rangle + \frac{1}{2\sigma}\|D_{\rho\sigma}(W(y; X))\|^2 - \langle \hat{b}, y \rangle - \frac{1}{2\sigma}\|X\|^2 \Big\}, \quad (3.53)$$

where $T = [I_m, 0; 0, 0] \in \Re^{(m+s)\times(m+s)}, W(y; X) = X - \sigma(C - \widehat{\mathcal{A}}^* y)$ and $\widehat{\mathcal{A}}^* = (\mathcal{A}^*, \ \mathcal{B}^*)$ is the adjoint of $\widehat{\mathcal{A}}$. Note that $-\varphi(\cdot)$ is the objective function of the inner subproblem (3.34). The objective function $\varphi(\cdot)$ in (3.53) is continuously differentiable with

$$\nabla \varphi(y) = Ty + \widehat{\mathcal{A}} D_{\rho\sigma}(W(y; X)) - \hat{b}, \quad y \in \Re^{m+s}.$$

Since $\varphi(\cdot)$ is a convex function, $\bar{y} = (\bar{\zeta}; \bar{\xi}) \in \mathcal{K}$ solves problem (3.53) if and only if it satisfies the following variational inequality

$$\langle y - \bar{y}, \nabla \varphi(\bar{y}) \rangle \geq 0 \quad \forall y \in \mathcal{K}. \quad (3.54)$$

Define $F : \Re^{m+s} \to \Re^{m+s}$ by

$$F(y) \ := \ y - \Pi_{\mathcal{K}}(y - \nabla \varphi(y)), \quad y \in \Re^{m+s}. \quad (3.55)$$

Then one can easily obtain that $\bar{y} \in \mathcal{K}$ solves (3.54) if and only if $F(\bar{y}) = 0$ [34]. Thus, solving the inner problem (3.53) is equivalent to solving the following equation

$$F(y) = 0, \quad y \in \Re^{m+s}. \quad (3.56)$$

Since both $\Pi_{\mathcal{K}}(\cdot)$ and $D_{\rho\sigma}(\cdot)$ are globally Lipschitz continuous, F is globally Lipschitz continuous. For the purpose of introducing an inexact smoothing Newton method, we need to define a smoothing function for $F(\cdot)$.

The smoothing function for the soft threshold operator $D_{\rho\sigma}(\cdot)$ has been defined by (2.29) in which the threshold value is $\rho\sigma$. Next, we need to define the smoothing function for $\Pi_{\mathcal{K}}(\cdot)$. For simplicity, we shall use the smoothing function ϕ_H defined by (2.26). Let $\psi : \Re \times \Re^{m+s} \to \Re^{m+s}$ be defined by

$$\psi_i(\varepsilon, z) = \begin{cases} z_i & \text{if } 1 \leq i \leq m + s_1 \\ \phi_H(\varepsilon, z_i) & \text{if } m + s_1 + 1 \leq i \leq m + s \end{cases}, \quad (\varepsilon, z) \in \Re \times \Re^{m+s}. \quad (3.57)$$

The function ψ is obviously continuously differentiable around any $(\varepsilon, z) \in \Re\times\Re^{m+s}$ as long as $\varepsilon \neq 0$ and is strongly semismooth everywhere.

Now, we are ready to define a smoothing function for $F(\cdot)$. Let

$$\Upsilon(\varepsilon, y) := y - \psi(\varepsilon, y - (Ty + \widehat{\mathcal{A}}\Phi_{\rho\sigma}(\varepsilon, W(y;X)) - \hat{b})), \quad (\varepsilon, y) \in \Re \times \Re^{m+s}. \quad (3.58)$$

From the definitions of Υ, ψ, and $\Phi_{\rho\sigma}$, we have that $F(y) = \Upsilon(0, y)$ for any $y \in \Re^{m+s}$.

Proposition 3.9. *Let $\Upsilon : \Re \times \Re^{m+s} \to \Re^{m+s}$ be defined by (3.58). Let $y \in \Re^{m+s}$. Then Υ has the following properties:*

(i) Υ is globally Lipschitz continuous on $\Re \times \Re^{m+s}$.

(ii) Υ is continuously differentiable around (ε, y) when $\varepsilon \neq 0$. For any fixed $\varepsilon \in \Re$, $\Upsilon(\varepsilon, \cdot)$ is a P_0-function, i.e., for any $(y, z) \in \Re^{m+s} \times \Re^{m+s}$ with $y \neq z$,

$$\max_{y_i \neq z_i}(y_i - z_i)(\Upsilon_i(\varepsilon, y) - \Upsilon_i(\varepsilon, z)) \geq 0, \quad (3.59)$$

and thus for any fixed $\varepsilon \neq 0$, $\Upsilon'_y(\varepsilon, y)$ is a P_0-matrix (i.e., all its principal minors are nonnegative).

(iii) Υ is strongly semismooth at $(0, y)$. In particular, for any $\varepsilon \downarrow 0$ and $\Re^{m+s} \ni h \to 0$ we have

$$\Upsilon(\varepsilon, y + h) - \Upsilon(0, y) - \Upsilon'(\varepsilon, y + h)(\varepsilon, h) = O(\|(\varepsilon, h)\|^2).$$

(iv) For any $h \in \Re^{m+s}$,

$$\partial\Upsilon(0,y)(0,h) \subseteq h - \partial\psi(0, y - \nabla\varphi(y))(0, h - (Th + \sigma\widehat{\mathcal{A}}\partial\Phi_{\rho\sigma}(0, W(y;X))(0, \widehat{\mathcal{A}}^* h))).$$

Proof. (i) Since both ψ and $\Phi_{\rho\sigma}$ are globally Lipschitz continuous, Υ is also globally Lipschitz continuous.

(ii) By the definition of ψ and $\Phi_{\rho\sigma}$, we know that Υ is continuously differentiable around $(\varepsilon, y) \in \Re \times \Re^{m+s}$ when $\varepsilon \neq 0$. From part (i), we know Υ is continuous on $\Re \times \Re^{m+s}$, it is enough to show that for any $0 \neq \varepsilon \in \Re$, $\Upsilon(\varepsilon, \cdot)$ is a P_0-function.

For any fixed $\varepsilon \neq 0$. Define $g_\varepsilon : \Re^{m+s} \to \Re^{m+s}$ by

$$g_\varepsilon(y) = Ty + \widehat{\mathcal{A}}\Phi_{\rho\sigma}(\varepsilon, W(y;X)) - \hat{b}, \quad y \in \Re^{m+s}.$$

Then g_ε is continuously differentiable on $\Re^{m+s}$. For any $h \in \Re^{m+s}$, we have

$$\begin{aligned} \langle h, (g_\varepsilon)'(y)h \rangle &= \langle h, Th \rangle + \sigma \langle h, \widehat{\mathcal{A}}(\Phi_{\rho\sigma})'_W(\varepsilon, W)\widehat{\mathcal{A}}^* h \rangle \\ &= \langle h, Th \rangle + \sigma \langle \widehat{\mathcal{A}}^* h, (\Phi_{\rho\sigma})'_W(\varepsilon, W)\widehat{\mathcal{A}}^* h \rangle \geq 0, \end{aligned}$$

which implies that g_ε is a P_0-function on $\Re^{m+s}$. Let $(y, z) \in \Re^{m+s} \times \Re^{m+s}$ with $y \neq z$. Then there exists $i \in \{1, \ldots, m+s\}$ with $y_i \neq z_i$ such that

$$(y_i - z_i)((g_\varepsilon)_i(y) - (g_\varepsilon)_i(z)) \geq 0.$$

By noting that for any $h \in \Re^{m+s}$, $(\phi_H)'_{h_j}(\varepsilon, h_j) \in [0,1], j = 1, \ldots, m+s$, we have

$$(y_i - z_i)(\Upsilon_i(\varepsilon, y) - \Upsilon_i(\varepsilon, z)) \geq 0.$$

This shows that (3.59) holds. Hence, $\Upsilon'_y(\varepsilon, y)$ is P_0-matrix for any fixed $\varepsilon \neq 0$.

(iii) Since the composite of strongly semismooth functions is still strongly semismooth [38], Υ is strongly semismooth at $(0, y)$.

(iv) Let $\mathcal{N} = \{0\} \times \Re^{m+s}$ which has Lebesgue measure zero in $\Re \times \Re^{m+s}$ and

$$\partial_{\mathcal{N}}\Upsilon(0,y) := \Big\{ \lim_{k\to\infty} \Upsilon'(\varepsilon^k, y^k) \ : \ (\varepsilon^k, y^k) \to (0, y), \ \varepsilon^k \neq 0 \Big\}.$$

Then we have $\partial\Upsilon(0,y) = \mathrm{conv}(\partial_{\mathcal{N}}\Upsilon(0,y))$. Then it is enough to show that the inclusion is true where the term at the left-hand side is $\partial_{\mathcal{N}}\Upsilon(0,y)(0,h)$. Since both ψ and $\Phi_{\rho\sigma}$ are directionally differentiable, for any $(\varepsilon, y') \in \Re \times \Re^{m+s}$ with $\varepsilon \neq 0$,

$$\Upsilon'(\varepsilon, y')(0,h) = h - \psi'\Big((\varepsilon, z'); (0, h - (Th + \sigma\widehat{\mathcal{A}}\Phi'_{\rho\sigma}((\varepsilon, W); (0, \widehat{\mathcal{A}}^* h))))\Big),$$

where $z' = y' - (Ty' + \widehat{\mathcal{A}}\Phi_{\rho\sigma}(\varepsilon, W) - \hat{b})$, from which we can further have

$$\Upsilon'(\varepsilon, y')(0,h) \in h - \partial\psi(\varepsilon, z')(0, h - (Th + \sigma\widehat{\mathcal{A}}\partial\Phi_{\rho\sigma}(\varepsilon, W)(0, \widehat{\mathcal{A}}^* h))).$$

By taking $(\varepsilon, y') \to (0, y)$ in the above inclusion, the required result follows. □

3.4.2 An inexact smoothing Newton method

In this subsection we introduce an inexact smoothing Newton method, which was developed by Gao and Sun in [42], for solving the nonsmooth equation of the form (3.56). Let $\kappa \in (0, \infty)$ be a constant. Define $G : \Re \times \Re^{m+s} \to \Re^{m+s}$ by

$$G(\varepsilon, y) := \Upsilon(\varepsilon, y) + \kappa|\varepsilon|y, \quad (\varepsilon, y) \in \Re \times \Re^{m+s}, \tag{3.60}$$

where $\Upsilon : \Re \times \Re^{m+s} \to \Re^{m+s}$ is defined by (3.58). For any $(\varepsilon, y) \in \Re \times \Re^{m+s}$ with $\varepsilon \neq 0$, we have that $G'_y(\varepsilon, y)$ is a P-matrix (i.e., all its principal minors are positive), thus nonsingular, while by part (*ii*) of Proposition 3.9, $\Upsilon'_y(\varepsilon, y)$ is only a P_0-matrix which may be singular. Define $E : \Re \times \Re^{m+s} \to \Re \times \Re^{m+s}$ by

$$E(\varepsilon, y) := \begin{bmatrix} \varepsilon \\ G(\varepsilon, y) \end{bmatrix} = \begin{bmatrix} \varepsilon \\ \Upsilon(\varepsilon, y) + \kappa|\varepsilon|y \end{bmatrix}, \quad (\varepsilon, y) \in \Re \times \Re^{m+s}.$$

For any $(\varepsilon, y) \in \Re^{m+s}$ with $\varepsilon \neq 0, E'(\varepsilon, y)$ is a P-matrix, thus nonsingular. Then solving the nonsmooth equation $F(y) = 0$ is equivalent to solving the following smoothing-nonsmooth equation

$$E(\varepsilon, y) = 0.$$

Define the merit function $\widehat{\varphi} : \Re \times \Re^{m+s} \to \Re_+$ by

$$\widehat{\varphi}(\varepsilon, y) := \|E(\varepsilon, y)\|^2, \quad (\varepsilon, y) \in \Re \times \Re^{m+s}.$$

Then the inexact smoothing Newton method can be described as follows.

Algorithm 2: An inexact smoothing Newton method.

Step 0. Choose $r \in (0,1)$. Let $\hat{\varepsilon} \in (0,\infty)$ and $\eta \in (0,1)$ be such that $\delta := \sqrt{2}\max\{r\hat{\varepsilon}, \eta\} < 1$. Choose constants $\ell \in (0,1), \sigma \in (0,1/2), \tau \in (0,1)$, and $\hat{\tau} \in [1,\infty)$. Let $\varepsilon^0 := \hat{\varepsilon}$ and $y^0 \in \Re^{m+s}$ be an arbitrary starting point. Set $k := 0$.

Step 1. If $E(\varepsilon^k, y^k) = 0$, then stop. Otherwise, compute

$$\varsigma_k := r\min\{1, \widehat{\varphi}(\varepsilon^k, y^k)\} \quad \text{and} \quad \eta_k := \min\{\tau, \hat{\tau}\|E(\varepsilon^k, y^k)\|\}.$$

Step 2. Solve the following equation

$$E(\varepsilon^k, y^k) + E'(\varepsilon^k, y^k)\begin{bmatrix}\Delta\varepsilon^k \\ \Delta y^k\end{bmatrix} = \begin{bmatrix}\varsigma_k\hat{\varepsilon} \\ 0\end{bmatrix} \tag{3.61}$$

approximately such that

$$\|R_k\| \le \min\Big\{\eta_k\|G(\varepsilon^k, y^k) + G'_\varepsilon(\varepsilon^k, y^k)\Delta\varepsilon^k\|, \eta\|E(\varepsilon^k, y^k)\|\Big\}, \tag{3.62}$$

where $\Delta\varepsilon^k := -\varepsilon^k + \varsigma_k\hat{\varepsilon}$ and $R_k := G(\varepsilon^k, y^k) + G'(\varepsilon^k, y^k)\begin{bmatrix}\Delta\varepsilon^k \\ \Delta y^k\end{bmatrix}$.

Step 3. Let m_k be the smallest nonnegative integer m satisfying

$$\widehat{\varphi}(\varepsilon^k + \ell^m\Delta\varepsilon^k, y^k + \ell^m\Delta y^k) \le \Big[1 - 2\sigma(1-\delta)\ell^m\Big]\widehat{\varphi}(\varepsilon^k, y^k).$$

Set $(\varepsilon^{k+1}, y^{k+1}) = (\varepsilon^k + \ell^{m_k}\Delta\varepsilon^k, y^k + \ell^{m_k}\Delta y^k)$.

Step 4. Replace k by $k+1$ and go to **Step 1**.

Let

$$\mathcal{N} := \{(\varepsilon, y) \in \Re \times \Re^{m+s} \mid \varepsilon \ge \varsigma(\varepsilon, y)\hat{\varepsilon}\}.$$

From [42, Theorem 4.1 & Theorem 3.6], we have the following convergence result for the inexact smoothing Newton method. For more detailed discussion about inexact smoothing Newton method, see [42].

Theorem 3.10. *Algorithm 2 is well defined and generates an sequence* $\{(\varepsilon^k, y^k)\} \in \mathcal{N}$ *with the properties that any accumulation point* $(\bar{\varepsilon}, \bar{y})$ *of* $\{(\varepsilon^k, y^k)\}$ *is a solution of* $E(\varepsilon, y) = 0$ *and* $\lim_{k\to\infty} \widehat{\varphi}(\varepsilon^k, y^k) = 0$*. Additionally, if the Slater condition (3.17) holds, then* $\{(\varepsilon^k, y^k)\}$ *is bounded.*

Theorem 3.11. *Let* $(\bar{\varepsilon}, \bar{y})$ *be an accumulation point of the infinite sequence* $\{(\varepsilon^k, y^k)\}$ *generated by Algorithm 2. Suppose that* E *is strongly semismooth at* $(\bar{\varepsilon}, \bar{y})$ *and that all* $\mathcal{V} \in \partial E(\bar{\varepsilon}, \bar{y})$ *are nonsingular. Then the whole sequence* $\{(\varepsilon^k, y^k)\}$ *converges to* $(\bar{\varepsilon}, \bar{y})$ *quadratically, i.e.,*

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

3.4.3 Constraint nondegeneracy and quadratic convergence

Suppose that the Slater condition (3.17) holds. Let $(\bar{\varepsilon}, \bar{y})$ be an accumulation point of the sequence $\{(\varepsilon^k, y^k)\}$ generated by Algorithm 2. Then, we know that $\bar{\varepsilon} = 0$ and $F(\bar{y}) = 0$, which means that $\bar{y} = (\bar{\zeta}; \bar{\xi}) \in \mathcal{K}$ is an optimal solution to the inner subproblem (3.53). Let $\overline{X} := D_{\rho\sigma}(W(\bar{y}; X))$. Then $(\bar{\zeta}, \overline{X}) \in \Re^m \times \Re^{p\times q}$ is the unique optimal solution to problem (3.31).

In order to prove the quadratic convergence of Algorithm 2, we need the concept of constraint nondegeneracy which was initiated by Robinson [103] and later extensively studied by Bonnans and Shapiro [14]. For a given closed set $K \subseteq \mathcal{X}$, we denote $T_K(x)$ to be the tangent cone of K at $x \in K$ as in convex analysis [104]. The largest linear space contained in $T_K(x)$ is denoted by $\text{lin}(T_K(x))$, which is equal to $(-T_K(x)) \cap T_K(x)$. Define $g : \Re^{p\times q} \to \Re$ by $g(X) = \|X\|_*$. Let $K_{p,q}$ be the epigraph of g, i.e.,

$$K_{p,q} \;:=\; \text{epi}(g) = \{(X, t) \in \Re^{p\times q} \times \Re \mid g(X) \le t\},$$

which is a close convex cone. Let $\widehat{\mathcal{B}} := (\mathcal{B}, 0)$. Then the problem (3.15) can be rewritten in the following form:

$$\min\left\{\frac{1}{2}\|\mathcal{A}(X)-b\|^2+\rho\, t+\langle C,X\rangle\ :\ \widehat{\mathcal{B}}(X,t)\in d+\mathcal{Q},\ (X,t)\in K_{p,q}\right\}. \tag{3.63}$$

It is easy to see that $\overline{X}$ is an optimal solution to problem (3.15) if and only if $(\overline{X},\bar{t})$ is an optimal solution to (3.63) with $\bar{t}=\|\overline{X}\|_*$. Let $\mathcal{I}$ be the identity mapping from $\Re^{p\times q}\times\Re$ to $\Re^{p\times q}\times\Re$. Then the constraint nondegeneracy condition is said to hold at $(\overline{X},\bar{t})$ if

$$\begin{pmatrix}\widehat{\mathcal{B}}\\ \mathcal{I}\end{pmatrix}(\Re^{p\times q}\times\Re)+\begin{pmatrix}\mathrm{lin}(T_{\mathcal{Q}}(\widehat{\mathcal{B}}(\overline{X},\bar{t})-d))\\ \mathrm{lin}(T_{K_{p,q}}(\overline{X},\bar{t}))\end{pmatrix}=\begin{pmatrix}\Re^s\\ \Re^{p\times q}\times\Re\end{pmatrix}. \tag{3.64}$$

Note that $\mathrm{lin}(T_{\mathcal{Q}}(\widehat{\mathcal{B}}(\overline{X},\bar{t})-d))=\mathrm{lin}(T_{\mathcal{Q}}(\mathcal{B}(\overline{X})-d))$. Let $\mathcal{E}(\overline{X})$ denote the index set of active constraints at $\overline{X}$:

$$\mathcal{E}(\overline{X}):=\{i\mid\langle\mathcal{B}_i,\overline{X}\rangle=d_i,\ i=s_1+1,\ldots,s\},$$

and $l:=|\mathcal{E}(\overline{X})|$. Without loss of generality, we assume that

$$\mathcal{E}(\overline{X}):=\{s_1+1,\ldots,s_1+l\}.$$

Define $\widetilde{\mathcal{B}}:\Re^{p\times q}\to\Re^{s_1+l}$ by

$$\widetilde{\mathcal{B}}(X):=\left[\langle\mathcal{B}_1,X\rangle,\ldots,\langle\mathcal{B}_{s_1+l},X\rangle\right]^T,\quad X\in\Re^{p\times q}.$$

Let $\overline{\mathcal{B}}=(\widetilde{\mathcal{B}},\ 0)$. Since $\mathrm{lin}(T_{\mathcal{Q}}(\mathcal{B}(X)-d))$ can be computed directly as follows

$$\mathrm{lin}(T_{\mathcal{Q}}(\mathcal{B}(\overline{X})-d))=\{h\in\Re^s\mid h_i=0,i\in\{1,\ldots,s_1\}\cup\mathcal{E}(\overline{X})\},$$

we have that (3.64) can be reduced to

$$\begin{pmatrix}\overline{\mathcal{B}}\\ \mathcal{I}\end{pmatrix}(\Re^{p\times q}\times\Re)+\begin{pmatrix}\{0\}^{s_1+l}\\ \mathrm{lin}(T_{K_{p,q}}(\overline{X},\bar{t}))\end{pmatrix}=\begin{pmatrix}\Re^{s_1+l}\\ \Re^{p\times q}\times\Re\end{pmatrix},$$

which is equivalent to

$$\overline{\mathcal{B}}\Big(\mathrm{lin}(T_{K_{p,q}}(\overline{X},\bar{t}))\Big)=\Re^{s_1+l}. \tag{3.65}$$

Next, we shall characterize the linear space $\mathrm{lin}(T_{K_{p,q}}(\overline{X}, g(\overline{X})))$. Let $W(\bar{y}; X)$ admit the SVD as in (2.3). For the given threshold value $\rho\sigma$, decompose the index set $\alpha = \{1, \ldots, p\}$ into the following three subsets:

$$\alpha_1 := \{i \,|\, \sigma_i > \rho\sigma,\, i \in \alpha\},\ \alpha_2 := \{i \,|\, \sigma_i = \rho\sigma,\, i \in \alpha\},\ \alpha_3 := \{i \,|\, \sigma_i < \rho\sigma,\, i \in \alpha\},$$

where $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_p \geq 0$ are singular values of $W(\bar{y}; X)$ being arranged in non-increasing order. Then $U = [U_{\alpha_1}\ U_{\alpha_2}\ U_{\alpha_3}], V = [V_{\alpha_1}\ V_{\alpha_2}\ V_{\alpha_3}\ V_2]$, and $\overline{X} = D_{\rho\sigma}(W(\bar{y}; X))$ is of rank $|\alpha_1|$. For any $H \in \Re^{p\times q}$, by the results of Watson [127, Theorem 1], we can obtain that

$$g'(\overline{X}; H) = \begin{cases} \|H\|_* & \text{if } |\alpha_1| = 0, \\ \langle UV_1^T, H\rangle & \text{if } |\alpha_1| = p, \\ \langle U_{\alpha_1}V_{\alpha_1}^T, H\rangle + \|[U_{\alpha_2}\ U_{\alpha_3}]^T H [V_{\alpha_2}\ V_{\alpha_3}\ V_2]\|_* & \text{if } 0 < |\alpha_1| < p. \end{cases}$$

From [26, Proposition 2.3.6 & Theorem 2.4.9], we have

$$T_{K_{p,q}}(\overline{X}, g(\overline{X})) = \mathrm{epi}(g'(\overline{X}; \cdot)),$$

from which we can easily have that

$$T_{K_{p,q}}(\overline{X}, g(\overline{X})) = \{(H,t) \in \Re^{p\times q}\times\Re \mid \langle U_{\alpha_1}V_{\alpha_1}^T, H\rangle + \|[U_{\alpha_2}\ U_{\alpha_3}]^T H[V_{\alpha_2}\ V_{\alpha_3}\ V_2]\|_* \leq t\}.$$

Then its linearity space is as follows

$$\mathrm{lin}(T_{K_{p,q}}(\overline{X}, g(\overline{X}))) = \{(H,t) \in \Re^{p\times q}\times\Re \mid [U_{\alpha_2}\, U_{\alpha_3}]^T H[V_{\alpha_2}\, V_{\alpha_3}\, V_2] = 0, t = \langle U_{\alpha_1}V_{\alpha_1}^T, H\rangle\},$$

Let

$$\mathcal{T}(\overline{X}) := \{H \in \Re^{p\times q} \mid [U_{\alpha_2}\ U_{\alpha_3}]^T H[V_{\alpha_2}\ V_{\alpha_3}\ V_2] = 0\}, \tag{3.66}$$

which is a subspace of $\Re^{p\times q}$. The orthogonal complement of $\mathcal{T}(\overline{X})$ is given by

$$\mathcal{T}(\overline{X})^\perp = \{H \in \Re^{p\times q} \mid U_{\alpha_1}^T H = 0,\ HV_{\alpha_1} = 0\}. \tag{3.67}$$

Since $\overline{\mathcal{B}} = (\widetilde{\mathcal{B}}, 0)$, we have that (3.65) can be further reduced to

$$\boxed{\widetilde{\mathcal{B}}(\mathcal{T}(\overline{X})) = \Re^{s_1+l}.} \tag{3.68}$$

Lemma 3.12. *Let $W(\bar{y};X) = X - \sigma(C - \widehat{\mathcal{A}}^*\bar{y})$ admit the SVD as in (2.3). Then the constraint nondegeneracy condition (3.68) holds at $\overline{X} = D_{\rho\sigma}(W(\bar{y};X))$ if and only if for any $h \in \Re^{s_1+l}$,*

$$U_{\alpha_1}^T(\widetilde{\mathcal{B}}^*h) = 0 \quad \text{and} \quad (\widetilde{\mathcal{B}}^*h)V_{\alpha_1} = 0 \iff h = 0. \tag{3.69}$$

Proof. "$\Longrightarrow$" If $h = 0$, obviously we have that $U_{\alpha_1}^T(\widetilde{\mathcal{B}}^*h) = 0$ and $(\widetilde{\mathcal{B}}^*h)V_{\alpha_1} = 0$.

For any $h \in \Re^{s_1+l}$, if $U_{\alpha_1}^T(\widetilde{\mathcal{B}}^*h) = 0$ and $(\widetilde{\mathcal{B}}^*h)V_{\alpha_1} = 0$, since the constraint nondegenerate condition (3.68) holds at $\overline{X} = D_{\rho\sigma}(W(\bar{y};X))$, there exist $Z \in \mathcal{T}(\overline{X})$ such that $h = \widetilde{\mathcal{B}}(Z)$. Let $\bar{\alpha}_1 = \alpha_2 \cup \alpha_3$. Then we have

$$\langle h, h\rangle = \langle h, \widetilde{\mathcal{B}}(Z)\rangle = \langle \widetilde{\mathcal{B}}^*h, Z\rangle = \langle U^T(\widetilde{\mathcal{B}}^*h)V, U^TZV\rangle$$

$$= \left\langle \begin{bmatrix} 0 & 0 & 0 \\ 0 & U_{\bar{\alpha}_1}^T(\widetilde{\mathcal{B}}^*h)V_{\bar{\alpha}_1} & U_{\bar{\alpha}_1}^T(\widetilde{\mathcal{B}}^*h)V_2 \end{bmatrix}, \begin{bmatrix} U_{\alpha_1}^TZV_{\alpha_1} & U_{\alpha_1}^TZV_{\bar{\alpha}_1} & U_{\alpha_1}^TZV_2 \\ U_{\bar{\alpha}_1}^TZV_{\alpha_1} & 0 & 0 \end{bmatrix} \right\rangle = 0,$$

which means $h = 0$.

"$\Longleftarrow$" If the constraint nondegeneracy condition (3.68) does not hold at $\overline{X}$, then we have

$$\left[\widetilde{\mathcal{B}}(\mathcal{T}(\overline{X}))\right]^\perp \neq \{0\}.$$

Let $0 \neq h \in \left[\widetilde{\mathcal{B}}(\mathcal{T}(\overline{X}))\right]^\perp$. Then we have

$$0 = \langle h, \widetilde{\mathcal{B}}(Z)\rangle = \langle \widetilde{\mathcal{B}}^*h, Z\rangle \quad \forall\ Z \in \mathcal{T}(\overline{X}),$$

which means $\widetilde{\mathcal{B}}^*h \in \mathcal{T}(\overline{X})^\perp$. Thus, from (3.67), we have that

$$U_{\alpha_1}^T(\widetilde{\mathcal{B}}^*h) = 0 \quad \text{and } (\widetilde{\mathcal{B}}^*h)V_{\alpha_1} = 0,$$

from which we must have $h = 0$. This contradiction shows that the constraint nondegeneracy condition (3.68) holds at $\overline{X}$. □

Lemma 3.13. *Let $\widetilde{\mathcal{A}} = (\mathcal{A}; \widetilde{\mathcal{B}})$ and $\widetilde{\mathcal{A}}^* = (\mathcal{A}^*, \widetilde{\mathcal{B}}^*)$ be the adjoint of $\widetilde{\mathcal{A}}$. Let $\Phi_{\rho\sigma} : \Re \times \Re^{p\times q} \to \Re^{p\times q}$ be defined by (2.29). Assume that the constraint nondegeneracy condition (3.68) holds at $\overline{X}$. Then for any $\mathcal{V} \in \partial\Phi_{\rho\sigma}(0, W(\bar{y};X))$, we have*

$$\langle h, \widetilde{T}h + \sigma\widetilde{\mathcal{A}}\mathcal{V}(0, \widetilde{\mathcal{A}}^*h)\rangle > 0 \quad \forall\ 0 \neq h \in \Re^{m+s_1+l}, \tag{3.70}$$

where $\widetilde{T} = [I_m, 0; 0, 0]$ *is a matrix of size* $m + s_1 + l$.

Proof. For any $0 \neq h = (h_1; h_2) \in \Re^{m+s_1+l}$, where $h_1 \in \Re^m$ and $h_2 \in \Re^{s_1+l}$, we have

$$\langle h, \widetilde{T}h + \sigma\widetilde{\mathcal{A}}\mathcal{V}(0, \widetilde{\mathcal{A}}^*h)\rangle = \|h_1\|^2 + \sigma\langle h, \widetilde{\mathcal{A}}\mathcal{V}(0, \widetilde{\mathcal{A}}^*h)\rangle = \|h_1\|^2 + \sigma\langle \widetilde{\mathcal{A}}^*h, \mathcal{V}(0, \widetilde{\mathcal{A}}^*h)\rangle.$$

If $h_1 \neq 0$, since $\langle \widetilde{\mathcal{A}}^*h, \mathcal{V}(0, \widetilde{\mathcal{A}}^*h)\rangle \geq 0$, we have

$$\langle h, \widetilde{T}h + \sigma\widetilde{\mathcal{A}}\mathcal{V}(0, \widetilde{\mathcal{A}}^*h)\rangle > 0.$$

In the following proof, we assume $h_1 = 0$. For any $0 \neq h = (h_1; h_2) \in \Re^{m+s_1+l}$, we have $\widetilde{\mathcal{A}}^*h = \widetilde{\mathcal{B}}^*h_2$ and $\langle h, \widetilde{T}h + \sigma\widetilde{\mathcal{A}}\mathcal{V}(0, \widetilde{\mathcal{A}}^*h)\rangle = \sigma\langle h_2, \widetilde{\mathcal{B}}\mathcal{V}(0, \widetilde{\mathcal{B}}^*h_2)\rangle \geq 0$. Suppose that there exists $0 \neq h_2 \in \Re^{s_1+l}$ such that

$$\langle h_2, \widetilde{\mathcal{B}}\mathcal{V}(0, \widetilde{\mathcal{B}}^*h_2)\rangle = 0.$$

Let $H = \widetilde{\mathcal{B}}^*h_2, H_1 = U^T HV_1, H_2 = U^T HV_2$, and $\widetilde{H}_1 = \frac{1}{2}(H_1 + H_1^T)$. Then we have

$$\begin{aligned} 0 &= \langle H, \mathcal{V}(0, H)\rangle = \frac{1}{2}\Big\langle \begin{bmatrix} 0 & H \\ H^T & 0 \end{bmatrix}, \begin{bmatrix} 0 & \mathcal{V}(0, H) \\ (\mathcal{V}(0, H))^T & 0 \end{bmatrix} \Big\rangle \\ &= \frac{1}{2}\Big\langle Q^T \Xi(H) Q,\ Q^T \begin{bmatrix} 0 & \mathcal{V}(0, H) \\ (\mathcal{V}(0, H))^T & 0 \end{bmatrix} Q \Big\rangle, \end{aligned}$$

where $Q \in \mathcal{S}^{p+q}$ is of the form as in (2.5). From (2.42), (2.43) and Proposition 2.9, we know that there exists $\mathcal{V}_{|\alpha_2|} \in \partial(\Phi_\rho)_{|\alpha_2|}(0, \rho I_{|\alpha_2|})$ such that

$$Q^T \begin{bmatrix} 0 & \mathcal{V}(0, H) \\ (\mathcal{V}(0, H))^T & 0 \end{bmatrix} Q = \begin{bmatrix} M_{\alpha\alpha} & M_{\alpha\gamma} & M_{\alpha\beta} \\ M_{\alpha\gamma}^T & M_{\gamma\gamma} & M_{\gamma\beta} \\ M_{\alpha\beta}^T & M_{\gamma\beta}^T & M_{\beta\beta} \end{bmatrix},$$

where

$$M_{\alpha\alpha} = \begin{bmatrix} (\widetilde{H}_1)_{\alpha_1\alpha_1} & (\widetilde{H}_1)_{\alpha_1\alpha_2} & \Omega_{\alpha_1\alpha_3} \circ (\widetilde{H}_1)_{\alpha_1\alpha_3} \\ (\widetilde{H}_1)_{\alpha_1\alpha_2}^T & \mathcal{V}_{|\alpha_2|}(0, (\widetilde{H}_1)_{\alpha_2\alpha_2}) & 0 \\ \Omega_{\alpha_1\alpha_3}^T \circ (\widetilde{H}_1)_{\alpha_1\alpha_3}^T & 0 & 0 \end{bmatrix},$$

$$M_{\alpha\gamma} = \Gamma_{\alpha\gamma}\circ(\frac{H_1^T - H_1}{2}),\ M_{\alpha\beta} = \Gamma_{\alpha\beta}\circ(\frac{1}{\sqrt{2}}H_2),\ M_{\gamma\gamma} = -M_{\alpha\alpha},\ M_{\gamma\beta} = M_{\alpha\beta},\ M_{\beta\beta} = 0,$$

$\Omega_{\alpha_1\alpha_3}, \Gamma_{\alpha\gamma}$, and $\Gamma_{\alpha\beta}$ are of the forms as in (2.37), (2.38) and (2.39), respectively. Since $Q^T\Xi(H)Q$ is of the form in (2.13), we have that

$$\langle H, \mathcal{V}(0,H)\rangle = \langle \widetilde{H}_1, M_{\alpha\alpha}\rangle + \frac{1}{4}\langle H_1^T - H_1, \Gamma_{\alpha\gamma}\circ(H_1^T - H_1)\rangle + \langle H_2, \Gamma_{\alpha\beta}\circ H_2\rangle.$$

Since $\langle(\widetilde{H}_1)_{\alpha_2\alpha_2}, \mathcal{V}_{|\alpha_2|}(0,(\widetilde{H}_1)_{\alpha_2\alpha_2})\rangle \geq 0$, we obtain from $\langle H, \mathcal{V}(0,H)\rangle = 0$ that

$$(\widetilde{H}_1)_{\alpha_1\alpha} = 0,\ (\widetilde{H}_1)_{\alpha\alpha_1} = 0,\ (H_1^T - H_1)_{\alpha_1\alpha} = 0,\ (H_1^T - H_1)_{\alpha\alpha_1} = 0, \text{ and } (H_2)_{\alpha_1\bar{\beta}} = 0,$$

where $\bar{\beta} = \{1,\ldots,q-p\}$. Since $H_1 = \frac{1}{2}(H_1+H_1^T) - \frac{1}{2}(H_1^T - H_1) = \widetilde{H}_1 - \frac{1}{2}(H_1^T - H_1)$, we have that $(H_1)_{\alpha_1\alpha} = 0$ and $(H_1)_{\alpha\alpha_1} = 0$. Since $H_1 = [U_{\alpha_1}\, U_{\alpha_2}\, U_{\alpha_3}]^T H [V_{\alpha_1}\, V_{\alpha_2}\, V_{\alpha_3}]$ and $H_2 = [U_{\alpha_1}\, U_{\alpha_2}\, U_{\alpha_3}]^T H V_2$, we obtain that

$$U_{\alpha_1}^T H V_1 = 0, \quad U_{\alpha_1}^T H V_2 = 0, \quad \text{and} \quad U^T H V_{\alpha_1} = 0.$$

Since both U and $V = [V_1\ V_2]$ are orthogonal matrices, we have

$$U_{\alpha_1}^T H = 0 \quad \text{and} \quad H V_{\alpha_1} = 0,$$

which means

$$U_{\alpha_1}^T(\widetilde{\mathcal{B}}^* h_2) = 0 \quad \text{and} \quad (\widetilde{\mathcal{B}}^* h_2)V_{\alpha_1} = 0.$$

Since the constraint nondegeneracy condition (3.68) holds at $\overline{X}$, we have from Lemma 3.12 that $h_2 = 0$, which contradicts the assumption that $h_2 \neq 0$. This contradiction shows that for any $\mathcal{V} \in \partial\Phi_{\rho\sigma}(0, W(\bar{y}; X))$, (3.70) holds. □

Proposition 3.14. *Let $\Upsilon : \Re\times\Re^{m+s} \to \Re^{m+s}$ be defined by (3.58). Assume that the constraint nondegeneracy condition (3.68) holds at $\overline{X}$. Then for any $\mathcal{W} \in \partial\Upsilon(0,\bar{y})$, we have*

$$\max_i h_i(\mathcal{W}(0,h))_i > 0 \quad \forall\ 0 \neq h \in \Re^{m+s}. \tag{3.71}$$

Proof. Let $\mathcal{W} \in \partial\Upsilon(0,\bar{y})$. Suppose that there exists $0 \neq h \in \Re^{m+s}$ such that (3.71) does not hold, i.e.,

$$\max_i h_i(\mathcal{W}(0,h))_i \leq 0. \tag{3.72}$$

Then from part (iv) of Proposition 3.9, we know that there exist $\mathcal{D} \in \partial\psi(0, \bar{z})$ and $\mathcal{V} \in \partial\Phi_{\rho\sigma}(0, W(\bar{y}; X))$ such that

$$\mathcal{W}(0,h) = h - \mathcal{D}(0, h - (Th + \sigma\widehat{\mathcal{A}}\mathcal{V}(0, \widehat{\mathcal{A}}^* h))) = h - \mathcal{D}(0,h) + \mathcal{D}(0, Th + \sigma\widehat{\mathcal{A}}\mathcal{V}(0, \widehat{\mathcal{A}}^* h)),$$

where $\bar{z} = \bar{y} - (T\bar{y} + \widehat{\mathcal{A}}\Phi_{\rho\sigma}(0, W(\bar{y}; X)) - \hat{b})$. By simple calculations, we obtain that there exists a nonnegative vector $d \in \Re^{m+s}$ satisfying

$$d_i = \begin{cases} 1 & \text{if } 1 \le i \le m + s_1, \\ \in [0,1] & \text{if } m + s_1 + 1 \le i \le m + s_1 + l, \\ 0 & \text{if } m + s_1 + l + 1 \le i \le m + s, \end{cases}$$

such that for any $y \in \Re^{m+s}$,

$$(\mathcal{D}(0,y))_i = d_i y_i, \; i = 1, \ldots, m+s.$$

Then we have

$$h_i(\mathcal{W}(0,h))_i = h_i \left[h_i - d_i h_i + d_i \Big(Th + \sigma\widehat{\mathcal{A}}\mathcal{V}(0, \widehat{\mathcal{A}}^* h) \Big)_i \right], \; i = 1, \ldots, m+s.$$

This, together with (3.72), implies that

$$\begin{cases} h_i(Th + \sigma\widehat{\mathcal{A}}\mathcal{V}(0, \widehat{\mathcal{A}}^* h))_i \le 0 & \text{if } 1 \le i \le m + s_1, \\ h_i(Th + \sigma\widehat{\mathcal{A}}\mathcal{V}(0, \widehat{\mathcal{A}}^* h))_i \le 0 \text{ or } h_i = 0 & \text{if } m + s_1 + 1 \le i \le m + s_1 + l, \\ h_i = 0 & \text{if } m + s_1 + l + 1 \le i \le m + s. \end{cases}$$

Then we obtain that

$$\langle h, Th + \sigma\widehat{\mathcal{A}}\mathcal{V}(0, \widehat{\mathcal{A}}^* h)\rangle = \langle \tilde{h}, \widetilde{T}\tilde{h} + \sigma\widetilde{\mathcal{A}}\mathcal{V}(0, \widetilde{\mathcal{A}}^* \tilde{h})\rangle \le 0,$$

where $0 \neq \tilde{h} \in \Re^{m+s_1+l}$ is defined by $\tilde{h}_i = h_i, i = 1, \ldots, m + s_1 + l$. However, the above inequality contradicts (3.70) in Lemma 3.13. Hence, we have that (3.71) holds. □

Theorem 3.15. *Let $(\bar{\varepsilon}, \bar{y})$ be an accumulation point of the infinite sequence $\{(\varepsilon^k, y^k)\}$ generated by Algorithm 2. Assume that the constraint nondegeneracy condition (3.68) holds at $\overline{X}$. Then the whole sequence $\{(\varepsilon^k, y^k)\}$ converges to $(\bar{\varepsilon}, \bar{y})$ quadratically, i.e.,*

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

Proof. To prove the quadratic convergence of $\{(\varepsilon^k, y^k)\}$, by Theorem 3.11, it is enough to show that E is strongly semismooth at $(\bar{\varepsilon}, \bar{y})$ and all $\mathcal{V} \in \partial E(\bar{\varepsilon}, \bar{y})$ are nonsingular. The strong semismoothness of E at $(\bar{\varepsilon}, \bar{y})$ follows from part (iii) of Proposition 3.9 and the fact that the modulus function $|\cdot|$ is strongly semismooth everywhere on $\Re$.

Next, we show the nonsingularity of all elements in $\partial E(\bar{\varepsilon}, \bar{y})$. For any $\mathcal{V} \in \partial E(\bar{\varepsilon}, \bar{y})$, from Proposition 3.14 and the definition of E, we have that for any $0 \neq h \in \Re^{m+s+1}$, $\max_i h_i(\mathcal{V}d)_i > 0$, which implies that $\mathcal{V}$ is a P-matrix, and thus nonsingular [28, Theorem 3.3.4]. Hence we complete the proof of quadratic convergence of $\{(\varepsilon^k, y^k)\}$. □

3.5 Efficient implementation of the partial PPA

In this section, we introduce some techniques to improve the efficiency of our partial proximal point algorithm.

In our numerical implementation, we use the well-known alternating direction method of multipliers proposed by Gabay and Mercier [40], and Glowinski and Marrocco [45] to generate a good starting point for our partial PPA. To use the alternating direction method of multipliers, we introduce two auxiliary variables Y and v, and consider the following equivalent form of problem (3.15):

$$\min_{X\in\Re^{p\times q}, Y\in\Re^{p\times q}, v\in\mathcal{Q}} \left\{\frac{1}{2}\|\mathcal{A}(X)-b\|^2+\rho\|Y\|_*+\langle C, X\rangle \;:\; Y = X, \mathcal{B}(X)-v = d\right\}. \tag{3.73}$$

The augmented Lagrangian function for the problem (3.73) that corresponds to the linear equality constraints is defined as follows:

$$\begin{aligned} L_\beta(X, Y, v; Z, \lambda) \;=\; & \frac{1}{2}\|\mathcal{A}(X) - b\|^2 + \rho\|Y\|_* + \langle C, X\rangle + \langle Z, X - Y\rangle \\ & + \langle \lambda, d - \mathcal{B}(X) + v\rangle + \frac{\beta}{2}\|X - Y\|^2 + \frac{\beta}{2}\|d - \mathcal{B}(X) + v\|, \end{aligned}$$

where $Z \in \Re^{p\times q}$ and $\lambda \in \Re^s$ are the corresponding Lagrangian multipliers of the linear equality constraints and $\beta > 0$ is the penalty parameter for the violation of the linear equality constraints. Then, from a given starting point $(X^0, Y^0, v^0, Z^0, \lambda^0)$, the alternating direction method of multipliers generates new iterates according to the following procedure

$$X^{k+1} := \arg\min_{X \in \Re^{p\times q}} L_\beta(X, Y^k, v^k; Z^k, \lambda^k), \tag{3.74a}$$

$$Y^{k+1} := \arg\min_{Y \in \Re^{p\times q}} L_\beta(X^{k+1}, Y, v^k; Z^k, \lambda^k), \tag{3.74b}$$

$$v^{k+1} := \arg\min_{v \in \mathcal{Q}} L_\beta(X^{k+1}, Y^{k+1}, v; Z^k, \lambda^k), \tag{3.74c}$$

$$Z^{k+1} := Z^k + \gamma\beta(X^{k+1} - Y^{k+1}), \quad \lambda^{k+1} := \lambda^k + \gamma\beta(d - \mathcal{B}(X^{k+1}) + v^{k+1}), \tag{3.74d}$$

where $\gamma \in (0, (1+\sqrt{5})/2)$ is a given constant. Note that there is no theoretic convergence guarantee for the above procedure. This is only a heuristic approach for generating a good starting point for our partial proximal point algorithm. During our numerical implementation, we observe that the performance of the above procedure is very sensitive to the choice of the value of the penalty parameter β.

When applying Algorithm 2 to solve the inner subproblem (3.53), the most expensive step is in solving the linear system (3.61). In our numerical implementation, we first obtain $\Delta\varepsilon^k = -\varepsilon^k + \varsigma_k \hat{\varepsilon}$, and then apply the BiCGStab iterative solver of Van der Vost [124] to the following linear system

$$G'_y(\varepsilon^k, y^k)\Delta y^k = -G(\varepsilon^k, y^k) - G'_\varepsilon(\varepsilon^k, y^k)\Delta\varepsilon^k \tag{3.75}$$

to obtain a Δy^k satisfying condition (3.62). For the sake of convenience, we suppress the superscript k in our subsequent analysis. By noting that $G(\varepsilon, y)$ and $\Upsilon(\varepsilon, y)$ are defined by (3.60) and (3.58), respectively, we have that

$$\begin{aligned} G'_y(\varepsilon, y)\Delta y &= \Upsilon'_y(\varepsilon, y)\Delta y + \kappa\varepsilon\Delta y \\ &= (1+\kappa\varepsilon)\Delta y + \psi'_z(\varepsilon, z)\Big(T\Delta y + \sigma\widehat{\mathcal{A}}(\Phi_{\rho\sigma})'_W(\varepsilon, W)\widehat{\mathcal{A}}^*\Delta y - \Delta y\Big), \end{aligned} \tag{3.76}$$

where $z := y - (Ty + \widehat{\mathcal{A}}\Phi_{\rho\sigma}(\varepsilon, W) - \hat{b})$ and $W := X - \sigma(C - \widehat{\mathcal{A}}^* y)$. Let W admit the SVD as in (2.3). Then, by (2.34), we have

$$(\Phi_{\rho\sigma})'_W(\varepsilon, W)(\widehat{\mathcal{A}}^*\Delta y) = U\Big(\Lambda_{\alpha\alpha} \circ \frac{H_1 + H_1^T}{2} + \Lambda_{\alpha\gamma} \circ \frac{H_1 - H_1^T}{2}\Big)V_1^T + U(\Lambda_{\alpha\beta} \circ H_2)V_2^T, \tag{3.77}$$

where $\Lambda_{\alpha\alpha}, \Lambda_{\alpha\gamma}$ and $\Lambda_{\alpha\beta}$ are given by (2.30), $H_1 = U^T(\widehat{\mathcal{A}}^*\Delta y)V_1$, and $H_2 = U^T(\widehat{\mathcal{A}}^*\Delta y)V_2$. Note that the threshold value has been changed to $\rho\sigma$. When implementing the BiCGStab iterative method, one needs to repeatedly compute the matrix-vector multiplication $G'_y(\varepsilon, y)\Delta y$. From (3.77), it seems that a full SVD of W should be computed so that the matrix-vector multiplication $G'_y(\varepsilon, y)\Delta y$ can be evaluated. For a nonsymmetric matrix problem, in which p is moderate but q is large, computing the full SVD would incur huge memory space since the matrix $V \in \Re^{q\times q}$ is large and dense.

To over this difficulty, we first compute the economic form of the SVD of W, which is given by

$$W = U\Sigma V_1^T.$$

Then we construct V_2 via the QR factorization of V_1 with

$$V_1 = QR,$$

where $Q \in \Re^{q\times q}$ is orthogonal and $R \in \Re^{q\times p}$ is upper triangular. Decompose $Q \in \Re^{q\times q}$ into the form $Q = [Q_1 \ \ Q_2]$, where $Q_1 \in R^{q\times p}$ and $Q_2 \in \Re^{q\times(q-p)}$. From [49, Theorem 5.2.1], we know

$$\mathrm{range(Q_2)} = \mathrm{range(V_1)}^\perp = \mathrm{range(V_2)},$$

where range(Q$_2$) is the range space of Q_2. Since Q_2 has orthonormal columns which are orthogonal to those of V_1, Q_2 can be used in place of V_2. In our numerical implementation, Householder transformations are utilized to compute the QR factorization. Note that instead of storing the full Householder matrices, we only need to store the Householder vectors so as to compute the matrix-vector product involving V_2.

To achieve fast convergence for the BiCGStab method, we introduce an easy-to-compute diagonal preconditioner for the linear system (3.75). Since both $\psi'_z(\varepsilon, z)$ and T are diagonal matrices, we know from (3.76) that it is enough to find a good diagonal approximation of $\widehat{\mathcal{A}}(\Phi_{\rho\sigma})'_W(\varepsilon, W)\widehat{\mathcal{A}}^*$. Let

$$M := \widehat{\mathbf{A}}\mathbf{S}\widehat{\mathbf{A}}^T,$$

where $\widehat{\mathbf{A}}$ and $\mathbf{S}$ denote the matrix representation of the linear map $\widehat{\mathcal{A}}$ and $(\Phi_{\rho\sigma})'_W(\varepsilon, W)$ with respect to the standard bases in $\Re^{p\times q}$ and $\Re^{m+s}$, respectively. Let the standard basis in $\Re^{p\times q}$ be $\{E^{ij} \in \Re^{p\times q} : 1 \leq i \leq p, 1 \leq j \leq q\}$, where for each E^{ij}, its (i,j)-th entry is one and all the others are zero. Then the diagonal element of $\mathbf{S}$ with respect to the standard basis E^{ij} is given by

$$\mathbf{S}_{(i,j),(i,j)} = ((U \circ U)\widetilde{\Lambda}(V \circ V)^T)_{ij} + \frac{1}{2}\langle H_1^{ij} \circ (H_1^{ij})^T, \Lambda_{\alpha\alpha} - \Lambda_{\alpha\gamma}\rangle,$$

where $\widetilde{\Lambda} := [\frac{1}{2}(\Lambda_{\alpha\alpha} + \Lambda_{\alpha\gamma}), \Lambda_{\alpha\beta}]$ and $H_1^{ij} = U^T E^{ij} V_1$. Based on the above expression, the total cost of computing all the diagonal elements of $\mathbf{S}$ is equal to $2p(p+q)q + 3p^3 q$ flops, which is too expensive if $p^2 \gg p + q$. Fortunately, the first term

$$\mathbf{d}_{(ij)} = ((U \circ U)\widetilde{\Lambda}(V \circ V)^T)_{ij}$$

is usually a very good approximation of $\mathbf{S}_{(i,j),(i,j)}$, and the cost of computing all the elements $\mathbf{d}_{(ij)}$, for $1 \leq i \leq p, 1 \leq j \leq q$, is $2p(p + q)q$ flops since only the matrix product $(U \circ U)\widetilde{\Lambda}(V \circ V)^T$ is involved. Thus we propose the following diagonal preconditioner for the coefficient matrix $G'_y(\varepsilon, y)$:

$$M_G := (1 + \kappa\varepsilon)I + \psi'_z(\varepsilon, z)\Big(T + \sigma \mathrm{diag}(\widehat{\mathbf{A}}\mathrm{diag}(\mathbf{d})\widehat{\mathbf{A}}^T) - I\Big).$$

Finally, we should mention that the computational cost of either full or economic SVD can sometimes dominate the cost of the whole computation. In our implementation, we use the LAPACK routine `dgesdd.f`, which is based on the divide-and-conquer strategy, to compute either full or economic SVD of a matrix.

Chapter 4

A semismooth Newton-CG method for unconstrained inner subproblems

In this chapter, we consider the following nuclear norm regularized matrix least squares problem with linear equality constraints only:

$$\begin{array}{ll} \min\limits_{X\in\Re^{p\times q}} & \frac{1}{2}\|\mathcal{A}(X)-b\|^2+\rho\|X\|_*+\langle C,X\rangle \\ \text{s.t.} & \mathcal{B}(X)\ =\ d. \end{array} \tag{4.1}$$

where $\mathcal{A}:\Re^{p\times q}\to\Re^m$ and $\mathcal{B}:\Re^{p\times q}\to\Re^s$ are linear maps, $C\in\Re^{p\times q}, b\in\Re^m, d\in\Re^s$, ρ is a given positive parameter. It is easy to see that (4.1) can be rewritten as follows:

$$\begin{array}{ll} \min\limits_{u\in\Re^m, X\in\Re^{p\times q}} & f_\rho(u,X):=\frac{1}{2}\|u\|^2+\rho\|X\|_*+\langle C,X\rangle \\ \text{s.t.} & \mathcal{A}(X)+u=b, \\ & \mathcal{B}(X)=d. \end{array} \tag{4.2}$$

We apply the partial PPA introduced in Section 3 for solving (4.2). Since there is no inequality constraint in (4.2), the inner subproblems (3.34) in Algorithm 1 are unconstrained. Since the soft thresholding operator $D_{\rho\sigma_k}(\cdot)$ is strongly semismooth, we introduce a semismooth Newton-CG method, which is preferable to the inexact smoothing Newton method introduced in Section 3.4 for solving unconstrained inner subproblems (3.34). Throughout this chapter, the following Slater condition for (4.2) is assumed to hold:

$$\begin{cases} \mathcal{B} : \Re^{p\times q} \to \Re^{s} \text{ is onto}, \\ \exists X_0 \in \Re^{p\times q} \text{ such that } \mathcal{B}(X_0) \ = \ d. \end{cases} \tag{4.3}$$

4.1 A semismooth Newton-CG method

For the convenience of subsequent discussions, we let

$$\widehat{\mathcal{A}} = \begin{pmatrix} \mathcal{A} \\ \mathcal{B} \end{pmatrix}, \quad \widehat{b} = (b; d) \in \Re^{m+s}, \quad \text{and} \quad \widehat{y} = (\zeta; \xi) \in \Re^{m+s}. \tag{4.4}$$

In our proposed partial PPA, for some fixed $X \in \Re^{p\times q}$ and $\sigma > 0$, we need to solve the following form of inner subproblem

$$\min_{\widehat{y}\in\Re^{m+s}} \Big\{ \varphi(\widehat{y}) := \frac{1}{2}\langle \widehat{y},\, T\widehat{y}\rangle + \frac{1}{2\sigma}\|D_{\rho\sigma}(W(\widehat{y}; X))\|^2 - \langle \widehat{b},\, \widehat{y}\rangle - \frac{1}{2\sigma}\|X\|^2 \Big\}, \tag{4.5}$$

where $T = [I_m, 0; 0, 0] \in \Re^{(m+s)\times(m+s)}, W(\widehat{y}; X) = X - \sigma(C - \widehat{\mathcal{A}}^*\widehat{y})$ and $\widehat{\mathcal{A}}^* = (\mathcal{A}^*,\, \mathcal{B}^*)$ is the adjoint of $\widehat{\mathcal{A}}$. Note that $-\varphi(\cdot)$ is the objective function of the inner subproblem (3.34). The optimality condition for (4.5) is given by

$$\nabla\varphi(\widehat{y}) = \begin{bmatrix} I_m & \\ & 0 \end{bmatrix} \widehat{y} + \widehat{\mathcal{A}} D_{\rho\sigma}(W(\widehat{y}; X)) - \widehat{b} = 0. \tag{4.6}$$

Since the soft thresholding operator $D_{\rho\sigma}(\cdot)$ is Lipschitz continuous with modulus 1, the mapping $\nabla\varphi(\widehat{y})$ is Lipschitz continuous on $\Re^{m+s}$. Then for any $\widehat{y} \in \Re^{m+s}$, the generalized Hessian of $\varphi(\widehat{y})$ is well defined and it is defined as

$$\partial^2\varphi(\widehat{y}) \ := \ \partial(\nabla\varphi)(\widehat{y}), \tag{4.7}$$

where $\partial(\nabla\varphi)(\widehat{y})$ is the Clarke's generalized Jacobian of $\nabla\varphi$ at $\widehat{y}$ [26]. However, it is hard to express $\partial^2\varphi(\widehat{y})$ exactly, we define the following alternative for $\partial^2\varphi(\widehat{y})$

$$\hat{\partial}^2\varphi(\widehat{y}) \ := \ \begin{bmatrix} I_m & \\ & 0 \end{bmatrix} + \sigma\widehat{\mathcal{A}}\,\partial D_{\rho\sigma}(W(\widehat{y}; X))\widehat{\mathcal{A}}^*. \tag{4.8}$$

From [26, p.75], we have for $h \in \Re^{m+s}$,

$$\partial^2 \varphi(\widehat{y})h \subseteq \hat{\partial}^2 \varphi(\widehat{y})h, \tag{4.9}$$

which implies that if all elements in $\hat{\partial}^2\varphi(\widehat{y})$ are positive definite, so are those in $\partial^2\varphi(\widehat{y})$.

Since the soft thresholding operator $D_{\rho\sigma}(\cdot)$ is strongly semismooth, we consider solving (4.6) by a semismooth Newton-CG method for which the direction r at an iterate $\widehat{y}$ is computed from the following linear system:

$$\mathcal{V}\, r \;=\; -\nabla\varphi(\widehat{y}), \tag{4.10}$$

where

$$\mathcal{V} = \begin{bmatrix} I_m + \sigma\mathcal{A}\mathcal{W}\mathcal{A}^* & \sigma\mathcal{A}\mathcal{W}\mathcal{B}^* \\ \sigma\mathcal{B}\mathcal{W}\mathcal{A}^* & \sigma\mathcal{B}\mathcal{W}\mathcal{B}^* \end{bmatrix} = \begin{bmatrix} I_m & \\ & 0 \end{bmatrix} + \sigma\widehat{\mathcal{A}}\mathcal{W}\widehat{\mathcal{A}}^*. \tag{4.11}$$

Here $\mathcal{W}$ is an element in $\partial D_{\rho\sigma}(W(\widehat{y};X))$. Note that if $\mathcal{B} = \emptyset$, then $\mathcal{V}$ is always positive definite due to fact that all elements in $\partial D_{\rho\sigma}(\cdot)$ are positive semidefinite. To implement the above semismooth Newton-CG method, we need to compute an element $\mathcal{W}$ in $\partial D_{\rho\sigma}(W(\widehat{y};X))$. Define the operator $\mathcal{W}^0_{\widehat{y}} : \Re^{p\times q} \to \Re^{p\times q}$ as in (2.20), we can easily have that

$$\mathcal{V}^0_{\widehat{y}} = \begin{bmatrix} I_m & \\ & 0 \end{bmatrix} + \sigma\widehat{\mathcal{A}}\,\mathcal{W}^0_{\widehat{y}}\,\widehat{\mathcal{A}}^* \in \hat{\partial}^2\varphi(\widehat{y}). \tag{4.12}$$

Suppose that the Slater condition (4.3) holds and $\widetilde{y} = (\widetilde{\zeta};\widetilde{\xi}) \in \Re^{m+s}$ is the optimal solution to problem (4.5). Let $W(\widetilde{y};X) = X - \sigma(C - \widehat{\mathcal{A}}^*\widetilde{y})$ and $\overline{X} = D_{\rho\sigma}(W(\widetilde{y};X))$. Let $W(\widetilde{y};X)$ admit the SVD as in (2.3). For the given threshold value $\rho\sigma$, we decompose the index set $\alpha = \{1,\ldots,p\}$ into the following three subindex sets:

$$\alpha_1 := \{i\,|\,\sigma_i > \rho\sigma,\ i\in\alpha\},\quad \alpha_2 := \{i\,|\,\sigma_i = \rho\sigma,\ i\in\alpha\},\quad \alpha_3 := \{i\,|\,\sigma_i < \rho\sigma,\ i\in\alpha\}.$$

As discussed in Section 3.4.3, the constraint nondegeneracy condition is said to hold at $\overline{X}$ if

$$\mathcal{B}(\mathcal{T}(\overline{X})) = \Re^s, \tag{4.13}$$

where the subspace $\mathcal{T}(\overline{X})$ of $\Re^{p\times q}$ is defined as in (3.66)

$$\mathcal{T}(\overline{X}) := \Big\{H \in \Re^{p\times q} \mid [U_{\alpha_2}\ U_{\alpha_3}]^T H [V_{\alpha_2}\ V_{\alpha_3}\ V_2] = 0\Big\}, \tag{4.14}$$

and its orthogonal complement is given by

$$\mathcal{T}^{\perp}(\overline{X}) = \Big\{H \in \Re^{p\times q} \mid U_{\alpha_1}^T H = 0, \ \ HV_{\alpha_1} = 0\Big\}. \tag{4.15}$$

Lemma 4.1. *Let $W(\widetilde{y}; X)$ admit the SVD as in (2.3). For any $\mathcal{W} \in \partial D_{\rho\sigma}(W(\widetilde{y}; X))$ and $H \in \Re^{p\times q}$ such that $\mathcal{W}H = 0$, it holds that*

$$H \in \mathcal{T}^{\perp}(\overline{X}), \tag{4.16}$$

where $\mathcal{T}^{\perp}(\overline{X})$ is given by (4.15).

Proof. Let $\mathcal{W} \in \partial D_{\rho\sigma}(W(\widetilde{y}; X))$ and $H \in \Re^{p\times q}$ be such that $\mathcal{W}H = 0$. Then we have

$$\begin{aligned} 0 &= \langle H, \mathcal{W}H\rangle = \frac{1}{2}\langle \begin{bmatrix} 0 & H \\ H^T & 0 \end{bmatrix}, \begin{bmatrix} 0 & \mathcal{W}H \\ (\mathcal{W}H)^T & 0 \end{bmatrix} \rangle \\ &= \frac{1}{2}\langle \Xi(H), Q(\Gamma \circ (Q^T\Xi(H)Q))Q^T\rangle \\ &= \frac{1}{2}\langle Q^T\Xi(H)Q, \Gamma \circ (Q^T\Xi(H)Q)\rangle = \frac{1}{2}\langle \widetilde{H}, \Gamma \circ \widetilde{H}\rangle, \end{aligned}$$

where $\Gamma \in \mathcal{S}^{p+q}$ defined as in (2.16) and $\widetilde{H} = Q^T\Xi(H)Q$. From (2.13) and (2.19), we know

$$\langle \widetilde{H}, \Gamma \circ \widetilde{H}\rangle \geq \frac{1}{4}\sum_{i\in\alpha}\Big(\sum_{j\in\alpha}\Gamma_{ij}(H_1 + H_1^T)^2_{ij} + \sum_{j\in\gamma}\Gamma_{ij}(H_1^T - H_1)^2_{ij} + \sum_{j\in\beta}\Gamma_{ij}(\sqrt{2}H_2)^2_{ij}\Big),$$

where $H_1 = U^T HV_1$ and $H_2 = U^T HV_2$, which implies that

$$\sum_{i\in\alpha}\sum_{j\in\alpha}\Gamma_{ij}(H_1+H_1^T)^2_{ij} = 0, \quad \sum_{i\in\alpha}\sum_{j\in\gamma}\Gamma_{ij}(H_1^T-H_1)^2_{ij} = 0, \quad \sum_{i\in\alpha}\sum_{j\in\beta}\Gamma_{ij}(\sqrt{2}H_2)^2_{ij} = 0.$$

Then from (2.19) and the definition of the matrix Γ, we have

$$(H_1+H_1^T)_{\alpha_1\alpha} = 0, (H_1+H_1^T)_{\alpha\alpha_1} = 0, (H_1^T-H_1)_{\alpha_1\alpha} = 0, (H_1^T-H_1)_{\alpha\alpha_1} = 0, (H_2)_{\alpha_1\bar{\beta}} = 0.$$

Since $H_1 = \frac{1}{2}(H_1 + H_1^T) - \frac{1}{2}(H_1^T - H_1)$, we have $(H_1)_{\alpha_1\alpha} = 0$, $(H_1)_{\alpha\alpha_1} = 0$. Since $H_1 = [U_{\alpha_1}\, U_{\alpha_2}\, U_{\alpha_3}]^T H [V_{\alpha_1}\, V_{\alpha_2}\, V_{\alpha_3}]$, $H_2 = [U_{\alpha_1}\, U_{\alpha_2}\, U_{\alpha_3}]^T H V_2$, we obtain

$$U_{\alpha_1}^T H V_1 = 0, \quad U_{\alpha_1}^T H V_2 = 0, \quad \text{and} \quad U^T H V_{\alpha_1} = 0,$$

Since U and $V = [V_1\ V_2]$ are orthogonal matrices, we have

$$U_{\alpha_1}^T H = 0, \quad H V_{\alpha_1} = 0,$$

which means that $H \in \mathcal{T}^\perp(\overline{X})$. □

Proposition 4.2. *Suppose that the Slater condition (4.3) is satisfied. Let* $\widetilde{y} = (\widetilde{\zeta}; \widetilde{\xi}) \in \Re^{m+s}$ *be the optimal solution to problem (4.5),* $W(\widetilde{y}; X) = X - \sigma(C - \widehat{\mathcal{A}}^*\widetilde{y})$ *admit the SVD as in (2.3) and* $\overline{X} = D_{\rho\sigma}(W(\widetilde{y}; X))$*. Then the following conditions are equivalent:*

(a) The constraint nondegeneracy condition (4.13) holds at $\overline{X}$*.*

(b) Every $\mathcal{V}_{\widetilde{y}} \in \hat{\partial}^2\varphi(\widetilde{y})$ *is symmetric and positive definite.*

(c) $\mathcal{V}_{\widetilde{y}}^0 \in \hat{\partial}^2\varphi(\widetilde{y})$ *is symmetric and positive definite.*

Proof. "(a) ⇒ (b)". Let $\mathcal{V}_{\widetilde{y}}$ be an arbitrary element in $\hat{\partial}^2\varphi(\widetilde{y})$. Then there exists an element $\mathcal{W}_{\widetilde{y}} \in \partial D_{\rho\sigma}(W(\widetilde{y}; X))$ such that

$$\mathcal{V}_{\widetilde{y}} = \begin{bmatrix} I_m + \sigma\mathcal{A}\mathcal{W}_{\widetilde{y}}\mathcal{A}^* & \sigma\mathcal{A}\mathcal{W}_{\widetilde{y}}\mathcal{B}^* \\ \sigma\mathcal{B}\mathcal{W}_{\widetilde{y}}\mathcal{A}^* & \sigma\mathcal{B}\mathcal{W}_{\widetilde{y}}\mathcal{B}^* \end{bmatrix} = \begin{bmatrix} I_m & \\ & 0 \end{bmatrix} + \sigma\widehat{\mathcal{A}}\mathcal{W}_{\widetilde{y}}\widehat{\mathcal{A}}^*. \tag{4.17}$$

Since $\mathcal{W}_{\widetilde{y}}$ is self-adjoint and positive semidefinite, we have that $\mathcal{V}_{\widetilde{y}}$ is self-adjoint and positive semidefinite.

Next we show that $\mathcal{V}_{\widetilde{y}}$ is positive definite. From the structure (4.17) of $\mathcal{V}_{\widetilde{y}}$, we know that $\mathcal{V}_{\widetilde{y}}$ is positive definite if only if $\mathcal{B}\mathcal{W}_{\widetilde{y}}\mathcal{B}^*$ is positive definite. Hence, it it enough to show the positive definiteness of $\mathcal{B}\mathcal{W}_{\widetilde{y}}\mathcal{B}^*$. Let $h \in \mathcal{R}^s$ be such that $\mathcal{B}\mathcal{W}_{\widetilde{y}}\mathcal{B}^*h = 0$. Then, by (iii) of Proposition 2.7, we have

$$0 = \langle h, \mathcal{B}\mathcal{W}_{\widetilde{y}}\mathcal{B}^*h\rangle = \langle \mathcal{B}^*h, \mathcal{W}_{\widetilde{y}}\,\mathcal{B}^*h\rangle \geq \langle \mathcal{W}_{\widetilde{y}}\,\mathcal{B}^*h, \mathcal{W}_{\widetilde{y}}\,\mathcal{B}^*h\rangle,$$

which implies that $\mathcal{W}_{\tilde{y}}\,(\mathcal{B}^*h) = 0$. From Lemma 4.1, we have $\mathcal{B}^*h = \mathcal{T}(\overline{X})^\perp$. Since the constraint nondegeneracy condition holds at $\overline{X}$, there exists a $Y \in \mathcal{T}(\overline{X})$ such that $\mathcal{B}Y = h$. Then, we have

$$\langle h,\, h\rangle = \langle h,\, \mathcal{B}Y\rangle = \langle \mathcal{B}^*h,\, Y\rangle = 0.$$

Thus $h = 0$, which implies that $\mathcal{B}\mathcal{W}_{\tilde{y}}\mathcal{B}^*$ is positive definite. Hence, $\mathcal{V}_{\tilde{y}}$ is positive definite.

"(b) $\Rightarrow$ (c)". This is obviously true since $\mathcal{V}^0_{\tilde{y}} \in \hat{\partial}^2\varphi(\tilde{y})$.

"(c) $\Rightarrow$ (a)". Suppose that the constraint nondegeneracy condition (4.13) does not hold at $\overline{X}$. Then, we have

$$\Big[\mathcal{B}\mathcal{T}(\overline{X})\Big]^\perp \neq \{0\}.$$

Let $0 \neq h \in \Big[\mathcal{B}\mathcal{T}(\overline{X})\Big]^\perp$. Then, we have

$$0 = \langle h,\, \mathcal{B}Y\rangle = \langle H,\, Y\rangle \quad \forall\, Y \in \mathcal{T}(\overline{X}),$$

where $H = \mathcal{B}^*h$, which implies that $H \in \mathcal{T}(\overline{X})^\perp$. From (4.15), we have

$$U_{\alpha_1}^T H = 0 \quad \text{and} \quad HV_{\alpha_1} = 0.$$

Then it follows that

$$U_{\alpha_1}^T HV = U_{\alpha_1}^T H[V_1\ V_2] = 0 \quad \text{and} \quad U^T HV_{\alpha_1} = 0. \tag{4.18}$$

Since $H_1 = U^T HV_1$ and $H_2 = U^T HV_2$, we have from (4.18) that

$$(H_1)_{\alpha_1\alpha} = 0, \quad (H_1)_{\alpha\alpha_1} = 0, \quad \text{and } (H_2)_{\alpha_1\bar{\beta}} = 0,$$

where $\bar{\beta} = \beta - 2p = \{1, \ldots, q-p\}$, from which we further have that

$$(H_1 + H_1^T)_{\alpha_1\alpha} = 0,\ (H_1 + H_1^T)_{\alpha\alpha_1} = 0,\ (H_1^T - H_1)_{\alpha_1\alpha} = 0,\ \text{and } (H_1^T - H_1)_{\alpha\alpha_1} = 0.$$

Then we have

$$\Gamma^0_{\alpha\alpha} \circ (H_1 + H_1^T) = 0,\ \Gamma_{\alpha\gamma} \circ (H_1 - H_1^T) = 0,\ \text{and } \Gamma_{\alpha\beta} \circ H_2 = 0.$$

From the definition of $\mathcal{W}^0_{\tilde{y}}$ in (2.20), it follows that $\mathcal{W}^0_{\tilde{y}}(H) = 0$. Then we have

$$\langle h,\, \mathcal{B}\mathcal{W}^0_{\tilde{y}}\mathcal{B}^* h\rangle = \langle H,\, \mathcal{W}^0_{\tilde{y}}(H)\rangle = 0. \tag{4.19}$$

Since $\mathcal{V}^0_{\tilde{y}}$ is positive definite, it follows from (4.17) that $\mathcal{B}\mathcal{W}^0_{\tilde{y}}\mathcal{B}^*$ is also positive definite. Then (4.19) implies that $h = 0$, which contradicts the assumption that $h \neq 0$. Hence, we have that (a) holds. □

4.2 Convergence analysis

In this section, we state the semismooth Newton-CG algorithm (SSNCG) for solving (4.5) as follows.

SSNCG algorithm:

Given $\widehat{y}^0 \in \Re^{m+s}$, $c \in (0, 1/2)$, $\eta \in (0, 1)$, $\tau \in (0, 1]$, $\tau_1, \tau_2 \in (0, 1)$, and $\delta \in (0, 1)$. For $k = 0, 1, 2, \ldots$, obtain $\widehat{y}^{k+1}$ according to the following iteration:

Step 1. Compute

$$\eta_k := \min\{\eta, \|\nabla\varphi(\widehat{y}^k)\|^{1+\tau}\}.$$

Apply the CG method to find an approximation solution r^k to

$$(\mathcal{V}_k + \varepsilon_k I)\; r = -\nabla\varphi(\widehat{y}^k), \tag{4.20}$$

where $\mathcal{V}_k \in \hat{\partial}^2\varphi(\widehat{y}^k)$ is defined in (4.12) and $\varepsilon_k = \tau_1 \min\{\tau_2, \|\nabla\varphi(\widehat{y}^k)\|\}$, so that r^k satisfies the following condition:

$$\|(\mathcal{V}_k + \varepsilon_k I)r^k + \nabla\varphi(\widehat{y}^k)\| \leq \eta_k. \tag{4.21}$$

Step 2. Set $\alpha_k = \delta^{m_k}$, where m_k is the first nonnegative integer m for which

$$\varphi(\widehat{y}^k + \delta^m r^k) \;\leq\; \varphi(\widehat{y}^k) + c\delta^m\langle r^k,\, \nabla\varphi(\widehat{y}^k)\rangle.$$

Step 3. Set $\widehat{y}^{k+1} = \widehat{y}^k + \alpha_k r^k$.

In the SSNCG algorithm, since $\mathcal{V}_k$ is always positive semidefinite, the matrix $\mathcal{V}_k+\varepsilon_k I$ is positive definite as long as $\nabla\varphi(\widehat{y}^k) \neq 0$. To analyze the global convergence of the SSNCG algorithm, we assume that $\nabla\varphi(\widehat{y}^k) \neq 0$ for any $k \geq 0$. From [138, Lemma 3.1], we know that the generated search direction r^k is always a descent direction. The global convergence and the rate of convergence of the SSNCG algorithm can be derived similarly as studied in [138]. For details of the proof of the convergence analysis, see [138, Theorem 3.4 and Theorem 3.5].

Theorem 4.3. *Suppose that the Slater condition (4.3) holds. Then the SSNCG algorithm is well defined and any accumulation point $\widetilde{y}$ of $\{\widehat{y}^k\}$ generated by SSNCG algorithm is an optimal solution to the inner subproblem (4.5).*

Theorem 4.4. *Suppose that the Slater condition (4.3) holds. Let $\widetilde{y}$ be an accumulation point of the infinite sequence $\{\widehat{y}^k\}$ generated by SSNCG algorithm for solving the inner subproblem (4.5). Suppose that at each step $k \geq 0$, when the CG algorithm terminates, the tolerance η_k is achieved as in (4.21), i.e.,*

$$\|(\mathcal{V}_k + \varepsilon_k I)r^k + \nabla\varphi(\widehat{y}^k)\| \leq \eta_k. \tag{4.22}$$

Assume that the constraint nondegeneracy condition (4.13) holds at $\overline{X}$. Then the whole sequence $\{\widehat{y}^k\}$ convergence to $\widetilde{y}$ and

$$\|\widehat{y}^{k+1} - \widetilde{y}\| = O(\|\widehat{y}^k - \widetilde{y}\|^{1+\tau}). \tag{4.23}$$

4.3 Symmetric matrix problems

In this section, we will show that the partial PPA developed for solving (3.16) can be easily adapted for the symmetric matrix problems in which the matrix variable is symmetric and positive semidefinite. We consider the following regularized semidefinite matrix least squares problem:

$$\begin{array}{rl}\min\limits_{X\in\mathcal{S}^n} & \frac{1}{2}\|\mathcal{A}(X) - b\|^2 + \rho\langle I,\, X\rangle \\ \text{s.t.} & \mathcal{B}(X) \,=\, d, \\ & X \succeq\, 0,\end{array} \tag{4.24}$$

where $\mathcal{A} : \mathcal{S}^n \to \Re^m$ and $\mathcal{B} : \mathcal{S}^n \to \Re^s$ are linear maps, $b \in \Re^m, d \in \Re^s$, I is an identity matrix of size n and ρ is a given positive parameter. It is easy to see that (4.24) can be rewritten as follows:

$$\begin{array}{cl} \min\limits_{u\in\Re^m, X\in\mathcal{S}^n} & \frac{1}{2}\|u\|^2 + \langle C_\rho,\, X\rangle \\ \text{s.t.} & \mathcal{A}(X) + u \;=\; b, \\ & \mathcal{B}(X) \;=\; d, \\ & X \succeq\; 0,\; u \in \Re^m, \end{array} \tag{4.25}$$

where $C_\rho = \rho I$. The dual problem of (4.25) is given by:

$$\begin{array}{cl} \max\limits_{\zeta\in\Re^m, \xi\in\Re^s, Z\in\mathcal{S}^n} & -\frac{1}{2}\|\zeta\|^2 + \langle b,\, \zeta\rangle + \langle d,\, \xi\rangle \\ \text{s.t.} & \mathcal{A}^*(\zeta) + \mathcal{B}^*(\xi) + Z \;=\; C_\rho, \\ & Z \;\succeq\; 0,\; \zeta \in \Re^m,\; \xi \in \Re^s. \end{array} \tag{4.26}$$

For some fixed $X \in \mathcal{S}^n$ and $\sigma > 0$, the partial quadratic regularization of problem (4.25) is given by:

$$F_\sigma(X) = \min_{u\in\Re^m, Y\in\mathcal{S}^n} \; \frac{1}{2}\|u\|^2 + \langle C_\rho,\, Y\rangle + \frac{1}{2\sigma}\|Y - X\|^2$$

$$\mathcal{A}(Y) + u \;=\; b, \tag{4.27}$$

$$\mathcal{B}(Y) \;=\; d, \tag{4.28}$$

$$Y \succeq 0,\; u \in \Re^m.$$

The Lagrangian dual problem of (4.27) is

$$\max_{\zeta\in\Re^m, \xi\in\Re^s} \psi^\rho_\sigma(\zeta, \xi; X) := \inf_{u\in\Re^m, Y\succeq 0} L^\rho_\sigma(u, Y; \zeta, \xi, X)\,, \tag{4.29}$$

where

$$\begin{aligned} L^\rho_\sigma(u, Y; \zeta, \xi, X) \;&=\; \frac{1}{2}\|u\|^2 + \langle C_\rho,\, Y\rangle + \frac{1}{2\sigma}\|Y - X\|^2 + \langle \zeta,\, b - \mathcal{A}(Y) - u\rangle \\ &\quad + \langle \xi,\, d - \mathcal{B}(Y)\rangle \\ &=\; \frac{1}{2}\|u\|^2 - \langle \zeta, u\rangle + \langle b, \zeta\rangle + \langle d, \xi\rangle + \frac{1}{2\sigma}\|Y - W(\zeta, \xi; X)\|^2 \\ &\quad + \frac{1}{2\sigma}(\|X\|^2 - \|W(\zeta, \xi; X)\|^2), \end{aligned}$$

where $W(\zeta,\xi;X) = X - \sigma(C_\rho - \mathcal{A}^*\zeta - \mathcal{B}^*\xi)$. Then we have

$$\begin{aligned}\psi^\rho_\sigma(\zeta,\xi;X) &:= \inf_{u\in\Re^m, Y\succeq 0} L^\rho_\sigma(u,Y;\zeta,\xi,X) \\ &= -\frac{1}{2}\|\zeta\|^2 + \langle b,\zeta\rangle + \langle d,\xi\rangle + \frac{1}{2\sigma}\|X\|^2 - \frac{1}{2\sigma}\|\Pi_{S^n_+}(W(\zeta,\xi;X))\|^2. \quad (4.30)\end{aligned}$$

The optimality condition for (4.29) is given by

$$\begin{aligned}\nabla_\zeta\psi^\rho_\sigma(\zeta,\xi) &= b - \zeta - \mathcal{A}\Pi_{S^n_+}(W(\zeta,\xi;X)) = 0, \\ \nabla_\xi\psi^\rho_\sigma(\zeta,\xi) &= d - \mathcal{B}\Pi_{S^n_+}(W(\zeta,\xi;X)) = 0.\end{aligned} \quad (4.31)$$

Since $\Pi_{S^n_+}(\cdot)$ is strongly semismooth [115], (4.31) can be efficiently solved by the semismooth Newton-CG method developed in [138]. The convergence analysis of the semismooth Newton-CG method for solving (4.29) can be similarly derived as in [138].

Remark 4.5. Let $\sigma > 0$ be a given parameter. Consider the following function for (4.26):

$$\begin{aligned}\widetilde{L}(\zeta,\xi,Z;X) &= -\frac{1}{2}\|\zeta\|^2 + \langle b,\,\zeta\rangle + \langle d,\,\xi\rangle + \langle X,\,C_\rho - \mathcal{A}^*\zeta - \mathcal{B}^*\xi - Z\rangle \\ &\quad -\frac{\sigma}{2}\|C_\rho - \mathcal{A}^*\zeta - \mathcal{B}^*\xi - Z\|^2 \\ &= -\frac{1}{2}\|\zeta\|^2 + \langle b,\,\zeta\rangle + \langle d,\,\xi\rangle + \frac{1}{2\sigma}\|X\|^2 \\ &\quad -\frac{1}{2\sigma}\|\sigma Z + X - \sigma(C_\rho - \mathcal{A}^*\zeta - \mathcal{B}^*\xi)\|^2,\end{aligned}$$

the augmented Lagrangian function for the dual problem (4.26) is defined as follows:

$$\begin{aligned}\mathcal{L}^\rho_\sigma(\zeta,\xi;X) &:= \max_{Z\succeq 0}\widetilde{L}(\zeta,\xi,Z;X) \\ &= -\frac{1}{2}\|\zeta\|^2 + \langle b,\,\zeta\rangle + \langle d,\,\xi\rangle + \frac{1}{2\sigma}\|X\|^2 - \frac{1}{2\sigma}\|\Pi_{S^n_+}(W(\zeta,\xi;X))\|^2.\end{aligned}$$

Thus, the partial PPA for solving problem (4.25) is exactly the augmented Lagrangian method applied to the dual problem (4.26).

Chapter 5

An inexact APG method for linearly constrained convex SDP

In this chapter, we consider the following linearly constrained convex semidefinite programming problem:

$$
\begin{aligned}
(P)\quad \min_{x\in\mathcal{S}^n} \quad & f(x) \\
\text{s.t.} \quad & \mathcal{A}(x) \;=\; b, \\
& x \succeq 0,
\end{aligned}
$$

where f is a smooth convex function on $\mathcal{S}^n_+$, $\mathcal{A}:\,\mathcal{S}^n \to \mathcal{R}^m$ is a linear map, $b \in \mathcal{R}^m$, and $x \succeq 0$ means that $x \in \mathcal{S}^n_+$. Let $\mathcal{A}^*$ be the adjoint of $\mathcal{A}$. The dual problem associated with (P) is given by

$$
\begin{aligned}
(D)\quad \max \quad & f(x) - \langle \nabla f(x),\, x\rangle + \langle b,\, p\rangle \\
\text{s.t.} \quad & \nabla f(x) - \mathcal{A}^* p - z = 0, \\
& p \in \mathcal{R}^m,\; z \succeq 0,\; x \succeq 0.
\end{aligned}
$$

We assume that the linear map $\mathcal{A}$ is surjective, and that strong duality holds for (P) and (D). Let x_* be an optimal solution of (P) and (p_*, z_*) be an optimal solution of (D). Then, as a consequence of strong duality, they must satisfy the following

KKT conditions:

$$\mathcal{A}(x) = b, \quad \nabla f(x) - \mathcal{A}^* p - z = 0, \quad \langle x,\, z\rangle = 0, \quad x \succeq 0, \quad z \succeq 0.$$

The main purpose of this chapter is to design an efficient algorithm to solve the problem (P). The algorithm we propose here is based on the accelerated proximal gradient (APG) method of Beck and Teboulle [4] (the method is called FISTA in [4]), where in the kth iteration with iterate $\bar{x}_k$, a subproblem of the following form must be solved:

$$\min\Big\{\langle \nabla f(\bar{x}_k),\, x - \bar{x}_k\rangle + \frac{1}{2}\langle x - \bar{x}_k,\, \mathcal{H}_k(x - \bar{x}_k)\rangle : \mathcal{A}(x) = b,\ x \succeq 0\Big\}, \tag{5.1}$$

where $\mathcal{H}_k : \mathcal{S}^n \to \mathcal{S}^n$ is a given self-adjoint positive definite linear operator. In FISTA [4], $\mathcal{H}_k$ is restricted to $L\mathcal{I}$, where $\mathcal{I}$ denotes the identity map and L is a Lipschitz constant for ∇f. More significantly, for FISTA in [4], the subproblem (5.1) must be solved exactly to generate the next iterate x_{k+1}. In this chapter, we design an inexact APG method which overcomes the above mentioned two limitations. Specifically, in our inexact algorithm, the subproblem (5.1) is only solved approximately and $\mathcal{H}_k$ is not restricted to be a scalar multiple of $\mathcal{I}$. In addition, we are able to show that if the subproblem (5.1) is progressively solved with sufficient accuracy, then the number of iterations needed to achieve ε-optimality (in terms of the function value) is also proportional to $1/\sqrt{\varepsilon}$, just as in the exact algorithm.

5.1 An inexact accelerated proximal gradient method

For more generality, we consider the following minimization problem

$$\min\{F(x) := f(x) + g(x) : x \in \mathcal{X}\} \tag{5.2}$$

where $\mathcal{X}$ is a finite dimensional real Hilbert space with inner product $\langle\cdot,\cdot\rangle$ and its induced norm $\|\cdot\|$. The functions $f : \mathcal{X} \to \Re$, $g : \mathcal{X} \to \mathcal{R} \cup \{+\infty\}$ are proper, lower semi-continuous convex functions (possibly nonsmooth). We assume that $\mathrm{dom}(g) := \{x \in \mathcal{X} : g(x) < \infty\}$ is closed, f is continuously differentiable on $\mathcal{X}$ and its gradient ∇f is Lipschitz continuous with modulus L on $\mathcal{X}$, i.e.,

$$\|\nabla f(x) - \nabla f(y)\| \leq L\|x - y\| \quad \forall\, x, y \in \mathcal{X}.$$

We also assume that the problem (5.2) is solvable with an optimal solution $x_* \in \mathrm{dom}(g)$. The inexact APG algorithm we propose for solving (5.2) is described as follows.

Algorithm 3. Given a tolerance $\varepsilon > 0$. Input $y_1 = x_0 \in \mathrm{dom}(g)$, $t_1 = 1$. Set $k = 1$. Iterate the following steps.

Step 1. Find an approximate minimizer

$$x_k \approx \arg\min_{y\in\mathcal{X}} \Big\{ f(y_k) + \langle \nabla f(y_k),\, y - y_k\rangle + \frac{1}{2}\langle y - y_k,\, \mathcal{H}_k(y - y_k)\rangle + g(y) \Big\}, \quad (5.3)$$

where $\mathcal{H}_k$ is a self-adjoint positive definite linear operator that is chosen by the user.

Step 2. Compute $t_{k+1} = \dfrac{1 + \sqrt{1 + 4t_k^2}}{2}$.

Step 3. Compute $y_{k+1} = x_k + \Big(\dfrac{t_k - 1}{t_{k+1}}\Big)(x_k - x_{k-1})$.

Notice that Algorithm 3 is an inexact version of the algorithm FISTA in [4], where x_k need not be the exact minimizer of the subproblem (5.3). In addition, the quadratic term is not restricted to the form $\frac{L_k}{2}\|y - y_k\|^2$ where L_k is a positive scalar.

Given any positive definite linear operator $\mathcal{H}_j : \mathcal{X} \to \mathcal{X}$, and $y_j \in \mathcal{X}$, we define $q_j(\cdot) : \mathcal{X} \to \mathcal{R}$ by

$$q_j(x) = f(y_j) + \langle \nabla f(y_j),\, x - y_j\rangle + \frac{1}{2}\langle x - y_j,\, \mathcal{H}_j(x - y_j)\rangle. \quad (5.4)$$

Note that if we choose $\mathcal{H}_j = L\mathcal{I}$, then we have $f(x) \leq q_j(x)$ for all $x \in \mathrm{dom}(g)$.

Let $\{\xi_k\}, \{\epsilon_k\}$ be given convergent sequences of nonnegative numbers such that

$$\sum_{k=1}^{\infty} \xi_k < \infty \quad \text{and} \quad \sum_{k=1}^{\infty} \epsilon_k < \infty.$$

Suppose for each j, we have an approximate minimizer:

$$x_j \approx \arg\min\{q_j(x) + g(x) : x \in \mathcal{X}\} \tag{5.5}$$

that satisfies the conditions

$$F(x_j) \;\leq\; q_j(x_j) + g(x_j) + \frac{\xi_j}{2t_j^2} \tag{5.6}$$

$$\nabla f(y_j) + \mathcal{H}_j(x_j - y_j) + \gamma_j = \delta_j \quad \text{with } \|\mathcal{H}_j^{-1/2}\delta_j\| \leq \epsilon_j/(\sqrt{2}t_j) \tag{5.7}$$

where $\gamma_j \in \partial g(x_j; \frac{\xi_j}{2t_j^2})$ (the set of $\frac{\xi_j}{2t_j^2}$-subgradients of g at x_j). Note that for x_j to be an approximate minimizer, we must have $x_j \in \mathrm{dom}(g)$. We should mention that the condition (5.6) is usually easy to satisfy. For example, if $\mathcal{H}_j$ is chosen such that $f(x) \leq q_j(x)$ for all $x \in \mathrm{dom}(g)$, then (5.6) is automatically satisfied.

To establish the iteration complexity result analogous to the one in [4] for Algorithm 3, we need to establish a series lemmas whose proofs are extensions of those in [4] to account for the inexactness in x_k. We should note that although the ideas in the proofs are similar, but as the reader will notice later, the technical details become much more involved due to the error terms induced by the inexact solutions of the subproblems.

Lemma 5.1. *Given $y_j \in \mathcal{X}$ and a positive definite linear operator $\mathcal{H}_j$ on $\mathcal{X}$ such that the conditions (5.6) and (5.7) hold. Then for any $x \in \mathcal{X}$, we have*

$$\begin{aligned} F(x) - F(x_j) \;\geq\; & \frac{1}{2}\langle x_j - y_j,\, \mathcal{H}_j(x_j - y_j)\rangle + \langle y_j - x,\, \mathcal{H}_j(x_j - y_j)\rangle \\ & + \langle \delta_j,\, x - x_j\rangle - \frac{\xi_j}{t_j^2}. \end{aligned} \tag{5.8}$$

Proof. By condition (5.6), we have

$$\begin{aligned} F(x) - F(x_j) \;&\geq\; F(x) - q_j(x_j) - g(x_j) - \xi_j/(2t_j^2) \\ &= g(x) - g(x_j) + f(x) - f(y_j) - \langle \nabla f(y_j),\, x_j - y_j\rangle \\ &\quad - \frac{1}{2}\langle x_j - y_j,\, \mathcal{H}_j(x_j - y_j)\rangle - \frac{\xi_j}{2t_j^2}. \end{aligned}$$

By the convexity of f and the definition of γ_j, we have

$$f(x) \;\geq\; f(y_j) + \langle \nabla f(y_j),\, x - y_j\rangle, \quad g(x) \;\geq\; g(x_j) + \langle \gamma_j,\, x - x_j\rangle - \xi_j/(2t_j^2).$$

Hence

$$\begin{aligned} & F(x) - F(x_j) \\ \geq\; & \langle \gamma_j,\, x - x_j\rangle + \langle \nabla f(y_j),\, x - y_j\rangle - \langle \nabla f(y_j),\, x_j - y_j\rangle - \frac{1}{2}\langle x_j - y_j,\, \mathcal{H}_j(x_j - y_j)\rangle - \xi_j/t_j^2 \\ =\; & \langle \gamma_j + \nabla f(y_j),\, x - x_j\rangle - \frac{1}{2}\langle x_j - y_j,\, \mathcal{H}_j(x_j - y_j)\rangle - \xi_j/t_j^2 \\ =\; & \langle \delta_j - \mathcal{H}_j(x_j - y_j),\, x - x_j\rangle - \frac{1}{2}\langle x_j - y_j,\, \mathcal{H}_j(x_j - y_j)\rangle - \xi_j/t_j^2. \end{aligned}$$

From here, the required inequality (5.8) follows readily. □

For later purpose, we define the following quantities:

$$v_k = F(x_k) - F(x_*) \geq 0, \quad u_k = t_k x_k - (t_k - 1)x_{k-1} - x_*, \tag{5.9}$$

$$a_k = t_k^2 v_k \geq 0, \quad b_k = \tfrac{1}{2}\langle u_k,\, \mathcal{H}_k(u_k)\rangle \geq 0, \quad e_k = t_k\langle \delta_k,\, u_k\rangle \tag{5.10}$$

$$\tau = \tfrac{1}{2}\langle x_0 - x_*,\, \mathcal{H}_1(x_0 - x_*)\rangle, \quad \bar{\epsilon}_k = \textstyle\sum_{j=1}^k \epsilon_j, \quad \bar{\xi}_k = \sum_{j=1}^k (\xi_j + \epsilon_j^2). \tag{5.11}$$

Note that for the choice where $\epsilon_j = 1/j^\alpha = \xi_j$ for all $j \geq 1$, where $\alpha > 1$ is fixed, we have

$$\bar{\epsilon}_k \leq \frac{1}{\alpha - 1}, \quad \bar{\xi}_k \leq \frac{3}{2}\frac{1}{\alpha - 1} \;\forall\, k \geq 1.$$

Lemma 5.2. *Suppose that $\mathcal{H}_{k-1} \succeq \mathcal{H}_k \succ 0$ for all k. Then*

$$a_{k-1} + b_{k-1} \geq a_k + b_k - e_k - \xi_k. \tag{5.12}$$

Proof. By applying the inequality (5.8) to $x = x_{k-1} \in \mathrm{dom}(g)$ with $j = k$, we get

$$\begin{aligned} & v_{k-1} - v_k = F(x_{k-1}) - F(x_k) \qquad (5.13) \\ \geq\; & \frac{1}{2}\langle x_k - y_k,\, \mathcal{H}_k(x_k - y_k)\rangle + \langle y_k - x_{k-1},\, \mathcal{H}_k(x_k - y_k)\rangle + \langle \delta_k,\, x_{k-1} - x_k\rangle - \xi_k/t_k^2. \end{aligned}$$

Similarly, by applying the inequality (5.8) to $x = x_* \in \mathrm{dom}(g)$ with $j = k$, we get

$$\begin{aligned} -v_k \;=\; & F(x_*) - F(x_k) \\ \geq\; & \frac{1}{2}\langle x_k - y_k,\, \mathcal{H}_k(x_k - y_k)\rangle + \langle y_k - x_*,\, \mathcal{H}_k(x_k - y_k)\rangle \\ & + \langle \delta_k,\, x_* - x_k\rangle - \xi_k/t_k^2. \end{aligned} \tag{5.14}$$

By multiplying (5.13) throughout by $t_k - 1$ (note that $t_k \geq 1$ for all $k \geq 1$) and add that to (5.14), we get

$$\begin{aligned}
&(t_k-1)v_{k-1} - t_k v_k \\
\geq\ & \frac{t_k}{2}\langle x_k - y_k,\, \mathcal{H}_k(x_k-y_k)\rangle + \langle t_k y_k - (t_k-1)x_{k-1} - x_*,\, \mathcal{H}_k(x_k-y_k)\rangle \\
& - \langle \delta_k,\, t_k x_k - (t_k-1)x_{k-1} - x_*\rangle - \xi_k/t_k. \qquad (5.15)
\end{aligned}$$

Now, by multiplying (5.15) throughout by t_k and using the fact that $t_{k-1}^2 = t_k(t_k - 1)$, we have

$$\begin{aligned}
&a_{k-1} - a_k = t_{k-1}^2 v_{k-1} - t_k^2 v_k \\
\geq\ & \frac{t_k^2}{2}\langle x_k - y_k,\, \mathcal{H}_k(x_k-y_k)\rangle + t_k\langle t_k y_k - (t_k-1)x_{k-1} - x_*,\, \mathcal{H}_k(x_k-y_k)\rangle \\
& - \langle \delta_k,\, t_k^2 x_k - t_{k-1}^2 x_{k-1} - t_k x_*\rangle - \xi_k.
\end{aligned}$$

Let $\mathbf{a} = t_k y_k$, $\mathbf{b} = t_k x_k$, and $\mathbf{c} = (t_k - 1)x_{k-1} + x_*$. By using the fact that $\langle \mathbf{b} - \mathbf{a},\, \mathcal{H}_k(\mathbf{b}-\mathbf{a})\rangle + 2\langle \mathbf{a}-\mathbf{c},\, \mathcal{H}_k(\mathbf{b}-\mathbf{a})\rangle = \langle \mathbf{b}-\mathbf{c},\, \mathcal{H}_k(\mathbf{b}-\mathbf{c})\rangle - \langle \mathbf{a}-\mathbf{c},\, \mathcal{H}_k(\mathbf{a}-\mathbf{c})\rangle$, we get

$$\begin{aligned}
a_{k-1} - a_k \ \geq\ & \frac{1}{2}\langle \mathbf{b}-\mathbf{c},\, \mathcal{H}_k(\mathbf{b}-\mathbf{c})\rangle - \frac{1}{2}\langle \mathbf{a}-\mathbf{c},\, \mathcal{H}_k(\mathbf{a}-\mathbf{c})\rangle \\
& - \langle \delta_k,\, t_k^2 x_k - t_{k-1}^2 x_{k-1} - t_k x_*\rangle - \xi_k. \qquad (5.16)
\end{aligned}$$

Now $\mathbf{a} - \mathbf{c} = t_k y_k - \mathbf{c} = t_k x_{k-1} + (t_{k-1} - 1)(x_{k-1} - x_{k-2}) - \mathbf{c} = u_{k-1}$, $\mathbf{b} - \mathbf{c} = u_k$, and $t_k^2 x_k - t_{k-1}^2 x_{k-1} - t_k x_* = t_k u_k$. Thus (5.16) implies that

$$\begin{aligned}
a_{k-1} - a_k \ \geq\ & \frac{1}{2}\langle u_k,\, \mathcal{H}_k(u_k)\rangle - \frac{1}{2}\langle u_{k-1},\, \mathcal{H}_k(u_{k-1})\rangle - \langle \delta_k,\, t_k u_k\rangle - \xi_k \\
\geq\ & b_k - b_{k-1} - \langle \delta_k,\, t_k u_k\rangle - \xi_k. \qquad (5.17)
\end{aligned}$$

Note that in deriving (5.17), we have used the fact that $\mathcal{H}_{k-1} \succeq \mathcal{H}_k$. □

Lemma 5.3. *Suppose that $\mathcal{H}_{k-1} \succeq \mathcal{H}_k \succ 0$ for all k. Then*

$$a_k \ \leq\ (\sqrt{\tau} + \bar{\epsilon}_k)^2 + 2\bar{\xi}_k. \qquad (5.18)$$

Proof. Note that we have $|e_k| \leq \|\mathcal{H}_k^{-1/2}\delta_k\|\|\mathcal{H}_k^{1/2}u_k\|t_k \leq \epsilon_k \|\mathcal{H}_k^{1/2}u_k\|/\sqrt{2} = \epsilon_k\sqrt{b_k}$.

First, we show that $a_1 + b_1 \leq \tau + \epsilon_1\sqrt{b_1} + \xi_1$. Note that $a_1 = F(x_1) - F(x_*)$ and $b_1 = \frac{1}{2}\langle x_1 - x_*,\, \mathcal{H}_1(x_1 - x_*)\rangle$. By applying the inequality (5.8) to $x = x_*$ with $j = 1$, and noting that $y_1 = x_0$, we have that

$$\begin{aligned} -a_1 &\geq \frac{1}{2}\langle x_1 - y_1,\, \mathcal{H}_1(x_1 - y_1)\rangle + \langle y_1 - x_*,\, \mathcal{H}_1(x_1 - y_1)\rangle + \langle \delta_1,\, x_* - x_1\rangle - \xi_1 \\ &= \frac{1}{2}\langle x_1 - x_*,\, \mathcal{H}_1(x_1 - x_*)\rangle - \frac{1}{2}\langle y_1 - x_*,\, \mathcal{H}_1(y_1 - x_*)\rangle + \langle \delta_1,\, x_* - x_1\rangle - \xi_1 \\ &= b_1 - \frac{1}{2}\langle x_0 - x_*,\, \mathcal{H}_1(x_0 - x_*)\rangle + \langle \delta_1,\, x_* - x_1\rangle - \xi_1. \end{aligned}$$

Hence, by using the fact that $\|\mathcal{H}_1^{-1/2}\delta_1\| \leq \epsilon_1/\sqrt{2}$, we get

$$a_1 + b_1 \leq \frac{1}{2}\langle x_0 - x_*,\, \mathcal{H}_1(x_0 - x_*)\rangle - \langle \delta_1,\, x_* - x_1\rangle + \xi_1 \leq \tau + \epsilon_1\sqrt{b_1} + \xi_1. \quad (5.19)$$

Let

$$s_k = \epsilon_1\sqrt{b_1} + \cdots + \epsilon_k\sqrt{b_k} + \xi_1 + \cdots + \xi_k.$$

By Lemma 5.2, we have

$$\begin{aligned} \tau &\geq a_1 + b_1 - \epsilon_1\sqrt{b_1} - \xi_1 \geq a_2 + b_2 - \epsilon_1\sqrt{b_1} - \epsilon_2\sqrt{b_2} - \xi_1 - \xi_2 \\ &\geq \cdots \geq a_k + b_k - s_k. \end{aligned} \quad (5.20)$$

Thus we have $a_k + b_k \leq \tau + s_k$, and hence

$$s_k = s_{k-1} + \epsilon_k\sqrt{b_k} + \xi_k \leq s_{k-1} + \epsilon_k\sqrt{\tau + s_k} + \xi_k. \quad (5.21)$$

Note that since $\tau \geq b_1 - \epsilon_1\sqrt{b_1} - \xi_1$, we have $\sqrt{b_1} \leq \frac{1}{2}(\epsilon_1 + \sqrt{\epsilon_1^2 + 4(\tau + \xi_1)}) \leq \epsilon_1 + \sqrt{\tau + \xi_1}$. Hence $s_1 = \epsilon_1\sqrt{b_1} + \xi_1 \leq \epsilon_1(\epsilon_1 + \sqrt{\tau + \xi_1}) + \xi_1 \leq \epsilon_1^2 + \xi_1 + \epsilon_1(\sqrt{\tau} + \sqrt{\xi_1})$.

The inequality (5.21) implies that

$$(\tau + s_k) - \epsilon_k\sqrt{\tau + s_k} - (\tau + s_{k-1} + \xi_k) \leq 0.$$

Hence we must have

$$\sqrt{\tau + s_k} \leq \frac{1}{2}\left(\epsilon_k + \sqrt{\epsilon_k^2 + 4(\tau + s_{k-1} + \xi_k)}\right).$$

Consequently

$$\begin{aligned} s_k &\leq s_{k-1} + \frac{1}{2}\epsilon_k^2 + \xi_k + \frac{1}{2}\epsilon_k\sqrt{\epsilon_k^2 + 4(\tau + s_{k-1} + \xi_k)} \\ &\leq s_{k-1} + \epsilon_k^2 + \xi_k + \epsilon_k\Big(\sqrt{\tau} + \sqrt{s_{k-1}+\xi_k}\Big). \end{aligned}$$

This implies that

$$\begin{aligned} s_k &\leq s_1 + \sum_{j=2}^{k}\epsilon_j^2 + \sum_{j=2}^{k}\xi_j + \sqrt{\tau}\sum_{j=2}^{k}\epsilon_j + \sum_{j=2}^{k}\epsilon_j\sqrt{s_{j-1}+\xi_j} \\ &\leq \bar{\xi}_k + \sqrt{\tau}\,\bar{\epsilon}_k + \sum_{j=1}^{k}\epsilon_j\sqrt{s_j} \\ &\leq \bar{\xi}_k + \sqrt{\tau}\,\bar{\epsilon}_k + \sqrt{s_k}\,\bar{\epsilon}_k. \end{aligned} \tag{5.22}$$

In the last inequality, we used the fact that $s_{j-1} + \xi_j \leq s_j$ and $0 \leq s_1 \leq \cdots \leq s_k$. The inequality (5.22) implies that

$$\sqrt{s_k} \;\leq\; \frac{1}{2}\Big(\bar{\epsilon}_k + (\bar{\epsilon}_k^2 + 4\bar{\xi}_k + 4\bar{\epsilon}_k\sqrt{\tau})^{1/2}\Big).$$

From here, we get $s_k \leq \bar{\epsilon}_k^2 + 2\bar{\xi}_k + 2\bar{\epsilon}_k\sqrt{\tau}$, and the required result follows from the fact that $a_k \leq \tau + s_k$ in (5.20). □

Now we are ready to state the iteration complexity result for the inexact APG algorithm described in Algorithm 3.

Theorem 5.4. *Suppose the conditions (5.6) and (5.7) hold, and $\mathcal{H}_{k-1} \succeq \mathcal{H}_k \succ 0$ for all k. Then*

$$0 \leq F(x_k) - F(x_*) \leq \frac{4}{(k+1)^2}\Big((\sqrt{\tau} + \bar{\epsilon}_k)^2 + 2\bar{\xi}_k\Big). \tag{5.23}$$

Proof. By Lemma 5.3 and the fact that $t_k \geq (k+1)/2$, we have

$$F(x_k) - F(x_*) = a_k/t_k^2 \leq \frac{4}{(k+1)^2}((\sqrt{\tau} + \bar{\epsilon}_k)^2 + 2\bar{\xi}_k).$$

From the assumption on the sequences $\{\xi_k\}$ and $\{\epsilon_k\}$, we know that both $\{\bar{\epsilon}_k\}$ and $\{\bar{\xi}_k\}$ are bounded. Then the required convergent complexity result follows. □

Observe that in Theorem 5.4, we will recover the complexity result established in [4] if $\epsilon_j = 0 = \xi_j$ for all j.

5.1.1 Specialization to the case where $g = \delta(\cdot \,|\, \Omega)$

For the problem (P), it can be expressed in the form (5.2) with $g(x) = \delta(x \,|\, \Omega)$, where $\delta(\cdot \,|\, \Omega)$ denotes the indicator function on the set

$$\Omega = \{x \in \mathcal{S}^n : \mathcal{A}(x) = b,\ x \succeq 0\}. \tag{5.24}$$

The sub-problem (5.3), for a fixed y_k, then becomes the following constrained minimization problem:

$$\min \Big\{ \langle \nabla f(y_k),\, x - y_k \rangle + \frac{1}{2}\langle x - y_k,\, \mathcal{H}_k(x - y_k) \rangle \ :\ \mathcal{A}(x) = b, x \succeq 0 \Big\}. \tag{5.25}$$

Suppose we have an approximate solution (x_k, p_k, z_k) to the KKT optimality conditions for (5.25). More specifically,

$$\begin{aligned} \nabla f(y_k) + \mathcal{H}_k(x_k - y_k) - \mathcal{A}^* p_k - z_k =: \delta_k &\approx 0 \\ \mathcal{A}(x_k) - b &= 0 \\ \langle x_k,\, z_k \rangle =: \varepsilon_k &\approx 0, \quad x_k, z_k \succeq 0. \end{aligned} \tag{5.26}$$

To apply the complexity result established in Theorem 5.4, we need δ_k and ε_k to be sufficiently small so that the conditions (5.6) and (5.7) are satisfied. Observe that we need x_k to be contained in Ω in (5.26). Note that the first equation in (5.26) is the feasibility condition for the dual problem of (5.25), and it corresponds to the condition in (5.7) with $\gamma_k = -\mathcal{A}^* p_k - z_k$. Indeed, as we shall show next, γ_k is an ε_k-subgradient of g at $x_k \in \Omega$ if $z_k \succeq 0$. Now, given any $v \in \Omega$, we need to show that $g(v) \geq g(x_k) + \langle \gamma_k,\, v - x_k \rangle - \varepsilon_k$. We have $g(v) = 0$, $g(x_k) = 0$ since $v, x_k \in \Omega$, and

$$\begin{aligned} \langle \gamma_k,\, v - x_k \rangle &= \langle \mathcal{A}^* p_k + z_k,\, x_k - v \rangle = \langle p_k,\, \mathcal{A}(x_k) - \mathcal{A}(v) \rangle + \langle z_k,\, x_k \rangle - \langle z_k,\, v \rangle \\ &= \langle z_k,\, x_k \rangle - \langle z_k,\, v \rangle \ \leq\ \langle z_k,\, x_k \rangle \ =\ \varepsilon_k. \end{aligned} \tag{5.27}$$

Note that in deriving (5.27), we used the fact that $\langle z_k,\, v \rangle \geq 0$ since $v \succeq 0$ and $z_k \succeq 0$. Thus the condition (5.7) is satisfied if $\|\mathcal{H}_k^{-1/2}\delta_k\| \leq \epsilon_k/(\sqrt{2}t_k)$ and $\varepsilon_k \leq \xi_k/(2t_k^2)$.

As we have already noted in the last paragraph, the approximate solution x_k obtained by solving the sub-problem (5.25) should be feasible, i.e. $x_k \in \Omega$. In practice we can maintain the positive semidefiniteness of x_k by performing projection onto $\mathcal{S}^n_+$. But the residual vector $r_k := \mathcal{A}(x_k) - b$ is usually not exactly equal to 0. In the following paragraph, we will propose a strategy to find a feasible solution $\tilde{x}_k \in \Omega$ given an approximate solution x_k of (5.26) for which r_k is not necessarily 0, but (x_k, p_k, z_k) satisfies that conditions that $x_k \succeq 0$, $z_k \succeq 0$, and $\|\mathcal{H}_k^{-1/2}\delta_k\| \leq \frac{1}{2}\epsilon_k/(\sqrt{2}t_k)$ and $\varepsilon_k \leq \dfrac{1}{2}\xi_k/(2t_k^2)$.

Suppose that there exists $\bar{x} \succ 0$ such that $\mathcal{A}(\bar{x}) = b$. Since $\mathcal{A}$ is surjective, $\mathcal{A}\mathcal{A}^*$ is nonsingular. Let $\omega_k = -\mathcal{A}^*(\mathcal{A}\mathcal{A}^*)^{-1}(r_k)$. We note that $\|\omega_k\|_2 \leq \|r_k\|/\sigma_{\min}(\mathcal{A})$, and $\mathcal{A}(x_k + \omega_k) = b$, where $\|\cdot\|_2$ denotes the spectral norm. However, $x_k + \omega_k$ may not be positive semidefinite. Thus we consider the following iterate:

$$\tilde{x}_k = \lambda(x_k + \omega_k) + (1-\lambda)\bar{x} = \lambda x_k + (\lambda\omega_k + (1-\lambda)\bar{x}),$$

where $\lambda \in [0,1]$. It is clear that $\mathcal{A}\tilde{x}_k = b$. By choosing $\lambda = 1 - \|\omega_k\|_2/(\|\omega_k\|_2 + \lambda_{\min}(\bar{x}))$, we can guarantee that $\tilde{x}_k$ is positive semidefinite. For $\tilde{x}_k$, we have

$$\begin{aligned} 0 \leq \langle \tilde{x}_k,\, z_k\rangle &\leq \lambda\varepsilon_k + \lambda\sqrt{n}\|\omega_k\|_2\|z_k\| + \frac{\|\omega_k\|_2}{\|\omega_k\|_2 + \lambda_{\min}(\bar{x})}\sqrt{n}\lambda_{\max}(\bar{x})\|z_k\| \\ &\leq \varepsilon_k + \sqrt{n}\|\omega_k\|_2\|z_k\| + \sqrt{n}\frac{\|\omega_k\|_2}{\lambda_{\min}(\bar{x})}\lambda_{\max}(\bar{x})\|z_k\| \\ &\leq 2\varepsilon_k, \quad \text{if } \|\omega_k\|_2 \leq \frac{\varepsilon_k}{\sqrt{n}\|z_k\|}\Big(1 + \frac{\lambda_{\max}(\bar{x})}{\lambda_{\min}(\bar{x})}\Big)^{-1}. \end{aligned}$$

Moreover

$$\nabla f(y_k) + \mathcal{H}_k(\tilde{x}_k - y_k) - (\mathcal{A}^*p_k + z_k) = \delta_k + \mathcal{H}_k(\tilde{x}_k - x_k) =: \tilde{\delta}_k$$

Thus $\gamma_k = -\mathcal{A}^*p_k - z_k$ is an $2\varepsilon_k$-subgradient of g at $\tilde{x}_k \in \Omega$. Now $\|\mathcal{H}_k^{-1/2}\tilde{\delta}_k\| \leq \|\mathcal{H}_k^{-1/2}\delta_k\| + \|\mathcal{H}_k^{1/2}(\tilde{x}_k - x_k)\|$, and

$$\|\mathcal{H}_k^{1/2}(\tilde{x}_k - x_k)\|^2 = \langle \tilde{x}_k - x_k,\, \mathcal{H}_k(\tilde{x}_k - x_k)\rangle \leq n\|\omega_k\|_2^2\lambda_{\max}(H_1)\Big(1 + \frac{\|\bar{x} - x_k\|_2}{\lambda_{\min}(\bar{x})}\Big)^2.$$

Thus we have

$$\|\mathcal{H}_k^{-1/2}\tilde{\delta}_k\| \leq \epsilon_k/(\sqrt{2}t_k) \quad \text{if } \|w_k\|_2 \leq \frac{\epsilon_k}{2\sqrt{2n}\,t_k}(\lambda_{\max}(\mathcal{H}_1))^{-1/2}\Big(1 + \frac{\|\bar{x} - x_k\|_2}{\lambda_{\min}(\bar{x})}\Big)^{-1}.$$

To conclude $(\tilde{x}_k, p_k, z_k)$ would satisfy the condition (5.7) if

$$\|\omega_k\|_2 \le \min\Big\{\frac{\xi_k}{4t_k^2\sqrt{n}\|z_k\|}\Big(1+\frac{\lambda_{\max}(\bar{x})}{\lambda_{\min}(\bar{x})}\Big)^{-1},\ \frac{\epsilon_k}{2\sqrt{2n\lambda_{\max}(\mathcal{H}_1)}\,t_k}\Big(1+\frac{\|\bar{x}-x_k\|_2}{\lambda_{\min}(\bar{x})}\Big)^{-1}\Big\}. \tag{5.28}$$

We should note that even though we have succeeded in constructing a feasible $\tilde{x}_k$ in Ω. The accuracy requirement in (5.28) could be too stringent for computational efficiency. For example, when $\sigma_{\min}(\mathcal{A})$ is large, or $\|z_k\|$ is large, or $\bar{x}$ has a large condition number, or $\lambda_{\max}(\mathcal{H}_1)$ is large, we would expect that x_k must be computed to rather high accuracy so that (5.28) can be satisfied.

5.2 Analysis of an inexact APG method for (P)

To apply Algorithm 3 to solve the problem (P), the requirement that x_k must be primal feasible, i.e., $x_k \in \Omega$, can be restrictive as it limits our flexibility of choosing a non-primal feasible algorithm for solving (5.25). Even though the modification outlined in last paragraph of section 5.1.1 is able to produce a primal feasible $\tilde{x}_k$, the main drawback is that the residual norm $\|\omega_k\|$ must satisfy the stringent accuracy condition in (5.28). To overcome the drawbacks just mentioned, here we propose an inexact APG algorithm for solving (P) for which the iterate x_k need not be strictly contained in Ω. As the reader will observe later, the analysis of the iteration complexity of the proposed inexact APG becomes even more challenging than the analysis done in the previous section.

We let (x_*, p_*, z_*) be an optimal solution of (P) and (D). In the section, we let

$$q_k(x) = f(y_k) + \langle \nabla f(y_k),\, x - y_k\rangle + \frac{1}{2}\langle x - y_k,\, \mathcal{H}_k(x-y_k)\rangle, \quad x \in \mathcal{X}. \tag{5.29}$$

Note that $\mathcal{X} = \mathcal{S}^n$. The inexact APG algorithm we propose for solving (P) is given as follows.

Algorithm 4. Given a tolerance $\varepsilon > 0$. Input $y_1 = x_0 \in \mathcal{X}$, $t_1 = 1$. Set $k = 1$. Iterate the following steps.

Step 1. Find an approximate minimizer

$$x_k \approx \arg\min_{x\in\mathcal{X}} \Big\{ q_k(x) \; : \; x \in \Omega \Big\}, \tag{5.30}$$

where $\mathcal{H}_k$ is a self-adjoint positive definite operator that is chosen by the user, and x_k is allowed to be contained in a suitable enlargement Ω_k of Ω.

Step 2. Compute $t_{k+1} = \dfrac{1+\sqrt{1+4t_k^2}}{2}$.

Step 3. Compute $y_{k+1} = x_k + \Big(\dfrac{t_k - 1}{t_{k+1}}\Big)(x_k - x_{k-1})$.

Note that when $\Omega_k = \Omega$, the dual problem of (5.30) is given by

$$\max \Big\{ q_k(x) - \langle \nabla q_k(x),\, x\rangle + \langle b,\, p\rangle \mid \nabla q_k(x) - \mathcal{A}^* p - z = 0,\ z \succeq 0, x \succeq 0 \Big\}. \tag{5.31}$$

Let $\{\xi_k\}, \{\epsilon_k\}, \{\mu_k\}$ be given convergent sequences of nonnegative numbers such that

$$\sum_{k=1}^{\infty} \xi_k < \infty, \quad \sum_{k=1}^{\infty} \epsilon_k < \infty, \quad \text{and} \sum_{k=1}^{\infty} \mu_k < \infty,$$

and Δ is a given positive number. We assume that the approximate minimizer x_k in (5.30) has the property that x_k and its corresponding dual variables (p_k, z_k) satisfy the following conditions:

$$\begin{aligned}
f(x_k) \;&\le\; q_k(x_k) + \xi_k/(2t_k^2) \\
|\langle \nabla q_k(x_k),\, x_k\rangle &- \langle b,\, p_k\rangle| \le \Delta \\
\nabla q_k(x_k) - \mathcal{A}^* p_k - z_k \;&=\; \delta_k, \text{ with } \|H_k^{-1/2}\delta_k\| \le \epsilon_k/(\sqrt{2}t_k) \\
\|r_k\| \;&\le\; \mu_k/t_k^2 \\
\langle x_k,\, z_k\rangle \;&\le\; \xi_k/(2t_k^2) \\
x_k \succeq 0,&\ z_k \succeq 0,
\end{aligned} \tag{5.32}$$

where $r_k := \mathcal{A}(x_k) - b$. We assume that $\mu_k/t_k^2 \geq \mu_{k+1}/t_{k+1}^2$ and $\epsilon_k/t_k \geq \epsilon_{k+1}/t_{k+1}$ for all k. Observe that the last five conditions in (5.32) stipulate that (x_k, p_k, z_k) is an approximate optimal solution of (5.30) and (5.31).

Just as in the previous section, we need to establish a series of lemmas to analyse the iteration complexity of Algorithm 4. However, we should mention that the lack of feasibility in x_k (i.e., x_k may not be contained in Ω) introduces nontrivial technical difficulties in the proof of the complexity result for Algorithm 4. For example, $F(x_k) \geq F(x_*)$ no longer hold as in the feasible case when $x_k \in \Omega$.

Lemma 5.5. *Given $y_j \in \mathcal{X}$ and a positive definite linear operator $\mathcal{H}_j$ on $\mathcal{X}$ such that the conditions in (5.32) hold. Then for any $x \in \mathcal{S}_+^n$, we have*

$$\begin{aligned} f(x) - f(x_j) \;\; \geq \;\; & \frac{1}{2}\langle x_j - y_j,\, \mathcal{H}_j(x_j - y_j)\rangle + \langle y_j - x,\, \mathcal{H}_j(x_j - y_j)\rangle \\ & + \langle \delta_j + \mathcal{A}^* p_j,\, x - x_j\rangle - \xi_j/t_j^2. \end{aligned} \tag{5.33}$$

Proof. Since $f(x_j) \leq q_j(x_j) + \xi_j/(2t_j^2)$, we have

$$\begin{aligned} & f(x) - f(x_j) \geq f(x) - q_j(x_j) - \xi_j/(2t_j^2) \\ & = f(x) - f(y_j) - \langle \nabla f(y_j),\, x_j - y_j\rangle - \frac{1}{2}\langle x_j - y_j,\, \mathcal{H}_j(x_j - y_j)\rangle - \xi_j/(2t_j^2) \\ & \geq \langle \nabla f(y_j),\, x - x_j\rangle - \frac{1}{2}\langle x_j - y_j,\, \mathcal{H}_j(x_j - y_j)\rangle - \xi_j/(2t_j^2). \end{aligned}$$

Note that in the last inequality, we have used the fact that $f(x) - f(y_j) \geq \langle \nabla f(y_j),\, x - y_j\rangle$ for all $x \in \mathcal{X}$. Now, by using (5.32), we get

$$\begin{aligned} & f(x) - f(x_j) \geq \langle \nabla f(y_j),\, x - x_j\rangle - \frac{1}{2}\langle x_j - y_j,\, \mathcal{H}_j(x_j - y_j)\rangle - \xi_j/(2t_j^2) \\ & = \langle \delta_j + \mathcal{A}^* p_j - \mathcal{H}_j(x_j - y_j),\, x - x_j\rangle + \langle z_j,\, x - x_j\rangle - \frac{1}{2}\langle x_j - y_j,\, \mathcal{H}_j(x_j - y_j)\rangle - \frac{\xi_j}{2t_j^2} \\ & \geq \langle \delta_j + \mathcal{A}^* p_j - \mathcal{H}_j(x_j - y_j),\, x - x_j\rangle - \frac{1}{2}\langle x_j - y_j,\, \mathcal{H}_j(x_j - y_j)\rangle - \frac{\xi_j}{t_j^2}. \end{aligned}$$

From here, the required inequality (5.33) follows readily. Note that in deriving the last inequality, we have used the fact that $\langle z_j,\, x_j\rangle \leq \xi_j/(2t_j^2)$ and $\langle z_j,\, x\rangle \geq 0$. □

For later purpose, we define the following quantities for $k \geq 1$:

$$\begin{aligned} v_k &= f(x_k) - f(x_*), \quad u_k = t_k x_k - (t_k - 1)x_{k-1} - x_*, \\ a_k &= t_k^2 v_k, \quad b_k = \tfrac{1}{2}\langle u_k,\, \mathcal{H}_k(u_k)\rangle \geq 0, \quad e_k = t_k\langle \delta_k,\, u_k\rangle, \\ \eta_k &= \langle p_k,\, t_k^2 r_k - t_{k-1}^2 r_{k-1}\rangle, \text{ with } \eta_1 = \langle p_1,\, r_1\rangle, \\ \chi_k &= \|p_{k-1} - p_k\|\mu_{k-1}, \text{ with } \chi_1 = 0, \\ \bar{\epsilon}_k &= \textstyle\sum_{j=1}^k \epsilon_j, \quad \bar{\xi}_k = \sum_{j=1}^k (\xi_j + \epsilon_j^2), \quad \bar{\chi}_k = \sum_{j=1}^k \chi_j, \\ \tau &= \tfrac{1}{2}\langle x_0 - x_*,\, \mathcal{H}_1(x_0 - x_*)\rangle. \end{aligned} \tag{5.34}$$

Note that unlike the analysis in the previous section, a_k may be negative.

Lemma 5.6. *Suppose that $\mathcal{H}_{k-1} \succeq \mathcal{H}_k \succ 0$ for all k. Then*

$$a_{k-1} + b_{k-1} \geq a_k + b_k - e_k - \xi_k - \eta_k. \tag{5.35}$$

Proof. By applying the inequality (5.33) to $x = x_{k-1} \succeq 0$ with $j = k$, we get

$$\begin{aligned} & v_{k-1} - v_k = f(x_{k-1}) - f(x_k) \qquad (5.36) \\ \geq & \frac{1}{2}\langle x_k - y_k,\, \mathcal{H}_k(x_k - y_k)\rangle + \langle y_k - x_{k-1},\, \mathcal{H}_k(x_k - y_k)\rangle + \langle \delta_k + \mathcal{A}^* p_k,\, x_{k-1} - x_k\rangle - \frac{\xi_k}{t_k^2}. \end{aligned}$$

Similarly, by applying the inequality (5.33) to $x = x_* \succeq 0$ with $j = k$, we get

$$\begin{aligned} & - v_k = f(x_*) - f(x_k) \qquad (5.37) \\ \geq & \frac{1}{2}\langle x_k - y_k,\, \mathcal{H}_k(x_k - y_k)\rangle + \langle y_k - x_*,\, \mathcal{H}_k(x_k - y_k)\rangle + \langle \delta_k + \mathcal{A}^* p_k,\, x_* - x_k\rangle - \frac{\xi_k}{t_k^2}. \end{aligned}$$

By multiplying (5.36) throughout by $t_k - 1$ (note that $t_k \geq 1$ for all $k \geq 1$) and add that to (5.37), we get

$$\begin{aligned} & (t_k - 1)v_{k-1} - t_k v_k \\ \geq \; & \frac{t_k}{2}\langle x_k - y_k,\, \mathcal{H}_k(x_k - y_k)\rangle + \langle t_k y_k - (t_k - 1)x_{k-1} - x_*,\, \mathcal{H}_k(x_k - y_k)\rangle \\ & - \langle \delta_k + \mathcal{A}^* p_k,\, t_k x_k - (t_k - 1)x_{k-1} - x_*\rangle - \xi_k / t_k. \end{aligned} \tag{5.38}$$

Now, by multiplying (5.15) throughout by t_k and using the fact that $t_{k-1}^2 = t_k(t_k-1)$, we get

$$\begin{aligned}
a_{k-1} - a_k &= t_{k-1}^2 v_{k-1} - t_k^2 v_k \\
&\geq \frac{t_k^2}{2}\langle x_k - y_k,\, \mathcal{H}_k(x_k - y_k)\rangle + t_k\langle t_k y_k - (t_k-1)x_{k-1} - x_*,\, \mathcal{H}_k(x_k - y_k)\rangle \\
&\quad - \langle \delta_k + \mathcal{A}^* p_k,\, t_k^2 x_k - t_{k-1}^2 x_{k-1} - t_k x_*\rangle - \xi_k.
\end{aligned}$$

Let $\mathbf{a} = t_k y_k$, $\mathbf{b} = t_k x_k$, and $\mathbf{c} = (t_k - 1)x_{k-1} + x_*$. By using the fact that $\langle \mathbf{b} - \mathbf{a},\, \mathcal{H}_k(\mathbf{b} - \mathbf{a})\rangle + 2\langle \mathbf{a} - \mathbf{c},\, \mathcal{H}_k(\mathbf{b} - \mathbf{a})\rangle = \langle \mathbf{b} - \mathbf{c},\, \mathcal{H}_k(\mathbf{b} - \mathbf{c})\rangle - \langle \mathbf{a} - \mathbf{c},\, \mathcal{H}_k(\mathbf{a} - \mathbf{c})\rangle$, we get

$$\begin{aligned}
a_{k-1} - a_k \;\geq\; & \frac{1}{2}\langle \mathbf{b} - \mathbf{c},\, \mathcal{H}_k(\mathbf{b} - \mathbf{c})\rangle - \frac{1}{2}\langle \mathbf{a} - \mathbf{c},\, \mathcal{H}_k(\mathbf{a} - \mathbf{c})\rangle \\
& - \langle \delta_k + \mathcal{A}^* p_k,\, t_k^2 x_k - t_{k-1}^2 x_{k-1} - t_k x_*\rangle - \xi_k. \qquad (5.39)
\end{aligned}$$

Now $\mathbf{a} - \mathbf{c} = t_k y_k - \mathbf{c} = t_k x_{k-1} + (t_{k-1} - 1)(x_{k-1} - x_{k-2}) - \mathbf{c} = u_{k-1}$, $\mathbf{b} - \mathbf{c} = u_k$, and $t_k^2 x_k - t_{k-1}^2 x_{k-1} - t_k x_* = t_k u_k$. Thus (5.39) implies that

$$\begin{aligned}
a_{k-1} - a_k \;&\geq\; \frac{1}{2}\langle u_k,\, \mathcal{H}_k(u_k)\rangle - \frac{1}{2}\langle u_{k-1},\, \mathcal{H}_k(u_{k-1})\rangle - \langle \delta_k + \mathcal{A}^* p_k,\, t_k u_k\rangle - \xi_k \\
&\geq\; b_k - b_{k-1} - \langle \delta_k,\, t_k u_k\rangle - \langle p_k,\, \mathcal{A}(t_k u_k)\rangle - \xi_k. \qquad (5.40)
\end{aligned}$$

Note that in deriving (5.40), we have used the fact that $\mathcal{H}_{k-1} \succeq \mathcal{H}_k$. Now

$$\langle p_k,\, \mathcal{A}(t_k u_k)\rangle = \langle p_k,\, t_k^2(\mathcal{A}x_k - b) - t_{k-1}^2(\mathcal{A}x_{k-1} - b)\rangle = \langle p_k,\, t_k^2 r_k - t_{k-1}^2 r_{k-1}\rangle.$$

From here, the required result is proved. $\square$

Lemma 5.7. *Suppose that $\mathcal{H}_{k-1} \succeq \mathcal{H}_k \succ 0$ and the conditions in (5.32) are satisfied for all k. Then*

$$a_k + b_k \;\leq\; (\sqrt{\tau} + \bar{\epsilon}_k)^2 + \|p_k\|\mu_k + 2(\bar{\xi}_k + \bar{\chi}_k + \omega_k) \qquad (5.41)$$

where $\omega_k = \sum_{j=1}^k \epsilon_j \sqrt{A_j}$, and

$$A_j = \|p_j\|\mu_j + a_j^-, \quad \textit{with } a_j^- = \max\{0, -a_j\}.$$

Proof. Note that we have $|e_k| \leq \|\mathcal{H}_k^{-1/2}\delta_k\| \|\mathcal{H}_k^{1/2}u_k\| t_k \leq \epsilon_k \, \|\mathcal{H}_k^{1/2}u_k\|/\sqrt{2} = \epsilon_k\sqrt{b_k}$.

First, we show that $a_1 + b_1 \leq \tau + |\langle p_1, r_1\rangle| + \epsilon_1\sqrt{b_1} + \xi_1$. Note that $a_1 = f(x_1) - f(x_*)$ and $b_1 = \frac{1}{2}\langle x_1 - x_*, \mathcal{H}_1(x_1 - x_*)\rangle$. By applying the inequality (5.33) to $x = x_*$ with $j = 1$, and noting that $y_1 = x_0$, we have that

$$
\begin{aligned}
-a_1 &\geq \frac{1}{2}\langle x_1 - y_1, \mathcal{H}_1(x_1 - y_1)\rangle + \langle y_1 - x_*, \mathcal{H}_1(x_1 - y_1)\rangle + \langle \delta_1 + \mathcal{A}^* p_1, x_* - x_1\rangle - \xi_1 \\
&= \frac{1}{2}\langle x_1 - x_*, \mathcal{H}_1(x_1 - x_*)\rangle - \frac{1}{2}\langle y_1 - x_*, \mathcal{H}_1(y_1 - x_*)\rangle + \langle \delta_1 + \mathcal{A}^* p_1, x_* - x_1\rangle - \xi_1 \\
&= b_1 - \frac{1}{2}\langle x_0 - x_*, \mathcal{H}_1(x_0 - x_*)\rangle + \langle \delta_1 + \mathcal{A}^* p_1, x_* - x_1\rangle - \xi_1.
\end{aligned}
$$

Hence, by using the fact that $\|\mathcal{H}_1^{-1/2}\delta_1\| \leq \epsilon_1/\sqrt{2}$, we get

$$
\begin{aligned}
a_1 + b_1 &\leq \frac{1}{2}\langle x_0 - x_*, \mathcal{H}_1(x_0 - x_*)\rangle - \langle \delta_1 + \mathcal{A}^* p_1, x_* - x_1\rangle + \xi_1 \\
&\leq \tau + \epsilon_1\sqrt{b_1} + \langle p_1, r_1\rangle + \xi_1 \;\leq\; \tau + |\langle p_1, r_1\rangle| + \epsilon_1\sqrt{b_1} + \xi_1.
\end{aligned}
$$

Let $s_1 = \epsilon_1\sqrt{b_1} + \xi_1$ and for $k \geq 2$,

$$
s_k = \sum_{j=1}^{k} \epsilon_j\sqrt{b_j} + \sum_{j=1}^{k} \xi_j + \sum_{j=1}^{k} \chi_j.
$$

By Lemma 5.6, we have

$$
\begin{aligned}
\tau &\geq a_1 + b_1 - \epsilon_1\sqrt{b_1} - \xi_1 - \eta_1 \\
&\geq a_2 + b_2 - \epsilon_2\sqrt{b_2} - \epsilon_1\sqrt{b_1} - \xi_1 - \xi_2 - \eta_1 - \eta_2 \\
&\geq \cdots \\
&\geq a_k + b_k - \sum_{j=1}^{k} \epsilon_j\sqrt{b_j} - \sum_{j=1}^{k+1} \xi_j - \sum_{j=1}^{k} \eta_j \\
&\geq a_k + b_k - \sum_{j=1}^{k} \epsilon_j\sqrt{b_j} - \sum_{j=1}^{k} \xi_j - |\langle p_k, t_k^2 r_k\rangle| - \sum_{j=1}^{k} \chi_j.
\end{aligned}
$$

Note that in the last inequality, we used the fact that

$$
\sum_{j=1}^{k} \eta_j \;=\; \langle p_k, t_k^2 r_k\rangle + \sum_{j=1}^{k-1} \langle p_j - p_{j+1}, t_j^2 r_j\rangle \;\leq\; |\langle p_k, t_k^2 r_k\rangle| + \sum_{j=1}^{k} \chi_j.
$$

Thus we have $a_k + b_k \leq \tau + |\langle p_k, t_k^2 r_k\rangle| + s_k$, and this implies that

$$
b_k \;\leq\; \tau_k + s_k \quad \text{where } \tau_k := \tau + |\langle p_k, t_k^2 r_k\rangle| - a_k \leq \tau + A_k. \tag{5.42}
$$

Hence

$$s_k = s_{k-1} + \epsilon_k \sqrt{b_k} + \xi_k + \chi_k \le s_{k-1} + \epsilon_k \sqrt{\tau_k + s_k} + \xi_k + \chi_k. \tag{5.43}$$

Note that since $\tau_1 \ge b_1 - \epsilon_1 \sqrt{b_1} - \xi_1$, we have $\sqrt{b_1} \le \frac{1}{2}(\epsilon_1 + \sqrt{\epsilon_1^2 + 4(\tau_1 + \xi_1)}) \le \epsilon_1 + \sqrt{\tau_1 + \xi_1}$. Hence $s_1 = \epsilon_1 \sqrt{b_1} + \xi_1 \le \epsilon_1(\epsilon_1 + \sqrt{\tau_1 + \xi_1}) + \xi_1 \le \epsilon_1^2 + \xi_1 + \epsilon_1(\sqrt{\tau_1} + \sqrt{\xi_1})$.

The inequality (5.43) implies that

$$(\tau_k + s_k) - \epsilon_k \sqrt{\tau_k + s_k} - \Big(\tau_k + s_{k-1} + \xi_k + \chi_k\Big) \le 0.$$

Hence we must have

$$\sqrt{\tau_k + s_k} \le \frac{1}{2}\left(\epsilon_k + \sqrt{\epsilon_k^2 + 4(\tau_k + s_{k-1} + \xi_k + \chi_k)}\right).$$

Consequently, by using the fact that $\sqrt{a+b} \le \sqrt{a} + \sqrt{b}$ for any $a, b \ge 0$, we have

$$\begin{aligned} s_k &\le s_{k-1} + \frac{1}{2}\epsilon_k^2 + \xi_k + \chi_k + \frac{1}{2}\epsilon_k \sqrt{\epsilon_k^2 + 4(\tau_k + s_{k-1} + \xi_k + \chi_k)} \\ &\le s_{k-1} + \frac{1}{2}\epsilon_k^2 + \xi_k + \chi_k + \frac{1}{2}\epsilon_k \sqrt{\epsilon_k^2 + 4(\tau + A_k + s_{k-1} + \xi_k + \chi_k)} \\ &\le s_{k-1} + \epsilon_k^2 + \xi_k + \chi_k + \epsilon_k \Big(\sqrt{\tau + A_k} + \sqrt{s_{k-1} + \xi_k + \chi_k}\Big). \end{aligned} \tag{5.44}$$

This implies that

$$\begin{aligned} s_k &\le s_1 + \sum_{j=2}^{k}(\epsilon_j^2 + \xi_j + \chi_j) + \sum_{j=2}^{k} \epsilon_j \sqrt{\tau + A_j} + \sum_{j=2}^{k} \epsilon_j \sqrt{s_{j-1} + \xi_j + \chi_j} \\ &\le \bar{\xi}_k + \bar{\chi}_k + \sum_{j=1}^{k} \epsilon_j \sqrt{\tau + A_j} + \sum_{j=1}^{k} \epsilon_j \sqrt{s_j} \\ &\le \bar{\xi}_k + \bar{\chi}_k + \omega_k + \bar{\epsilon}_k \sqrt{\tau} + \bar{\epsilon}_k \sqrt{s_k}. \end{aligned} \tag{5.45}$$

In the last inequality, we used the fact that $s_{j-1} + \xi_j + \chi_j \le s_j$, and $0 \le s_1 \le \cdots \le s_k$. The inequality (5.45) implies that

$$\sqrt{s_k} \le \frac{1}{2}\Big(\bar{\epsilon}_k + \sqrt{\bar{\epsilon}_k^2 + 4\theta_k}\Big), \tag{5.46}$$

where $\theta_k = \bar{\xi}_k + \bar{\chi}_k + \omega_k + \bar{\epsilon}_k \sqrt{\tau}$. From here, we get

$$s_k \le \bar{\epsilon}_k^2 + 2\theta_k. \tag{5.47}$$

The required result follows from (5.47) and the fact that $a_k + b_k \le \tau + s_k + |\langle p_k,\, t_k^2 r_k\rangle| \le \tau + s_k + \|p_k\|\mu_k$. □

Let

$$\Omega_k := \{x \in \mathcal{S}^n : \|\mathcal{A}(x) - b\| \le \mu_k/t_k^2,\ x \succeq 0\} \tag{5.48}$$

and

$$x_*^k \ := \ \operatorname{argmin}\{f(x) \ :\ x \in \Omega_k\}. \tag{5.49}$$

Since $x_*, x_k \in \Omega_k$, we have $f(x_*) \ge f(x_*^k)$ and $f(x_k) \ge f(x_*^k)$. Hence $v_k = f(x_k) - f(x_*) \le f(x_k) - f(x_*^k)$. Also, since $\mu_k/t_k^2 \ge \mu_{k+1}/t_{k+1}^2$, we have $f(x_*^{k+1}) \ge f(x_*^k)$ and $\Omega_{k+1} \subseteq \Omega_k$.

Lemma 5.8. *For all $k \ge 1$, we have*

$$0 \le f(x_*) - f(x_*^k) \le \|p_*\|\mu_k/t_k^2. \tag{5.50}$$

Proof. By the convexity of f, we have

$$\begin{aligned} f(x_*) - f(x_*^k) &\le \langle \nabla f(x_*),\, x_* - x_*^k\rangle = \langle \mathcal{A}^* p_* + z_*,\, x_* - x_*^k\rangle \\ &= \langle p_*,\, \mathcal{A}(x_*) - \mathcal{A}(x_*^k)\rangle + \langle z_*,\, x_*\rangle - \langle z_*,\, x_*^k\rangle \\ &\le \|p_*\|\|b - \mathcal{A}(x_*^k)\| \le \|p_*\|\mu_k/t_k^2. \end{aligned}$$

Note that in deriving the second last inequality, we have used the fact that $\langle z_*,\, x_*\rangle = 0$, $\langle z_*,\, x_*^k\rangle \ge 0$, and $\mathcal{A}(x_*) = b$. □

Theorem 5.9. *Suppose $M_k = \max_{1\le j\le k}\{\sqrt{(\|p_*\| + \|p_j\|)\mu_j}\}$. Then we have*

$$-\frac{4\|p_*\|\mu_k}{(k+1)^2} \le f(x_k) - f(x_*) \le \frac{4}{(k+1)^2}\Big((\sqrt{\tau} + \bar\epsilon_k)^2 + \|p_k\|\mu_k + 2\bar\epsilon_k M_k + 2(\bar\xi_k + \bar\chi_k)\Big). \tag{5.51}$$

Proof. The inequality on the left-hand side of (5.51) follows from Lemma 5.8 and the fact that $t_k \geq (k+1)/2$ and $f(x_*^k) - f(x_*) \leq f(x_k) - f(x_*)$. Next, we prove the inequality on the right-hand side of (5.51). By Lemma 5.7 and noting that $b_k \geq 0$, we have

$$t_k^2(f(x_k) - f(x_*)) = a_k \;\; \leq \;\; (\sqrt{\tau} + \bar{\epsilon}_k)^2 + \|p_k\|\mu_k + 2(\bar{\xi}_k + \bar{\chi}_k) + 2\omega_k.$$

Now $-a_j = t_j^2(f(x_*) - f(x_j)) \leq t_j^2(f(x_*) - f(x_*^j)) \leq \|p_*\|\mu_j$. Hence $a_j^- \leq \|p_*\|\mu_j$, and

$$\omega_k \;\leq\; \sum_{j=1}^{k} \epsilon_j \sqrt{\|p_j\|\mu_j + \|p_*\|\mu_j} \;\leq\; M_k \bar{\epsilon}_k. \tag{5.52}$$

From here, the required result follows. □

From the assumption on the sequences $\{\epsilon_k\}$, $\{\xi_k\}$, and $\{\mu_k\}$, we know that the sequences $\{\bar{\epsilon}_k\}$ and $\{\bar{\xi}_k\}$ are bounded. In order to show that the sequence of function values $f(x_k)$ converges to the optimal function value $f(x_*)$ with the convergent rate $O(1/k^2)$, it is enough to show that the sequence $\{\|p_k\|\}$ is bounded under certain conditions, from which we can also have the boundedness of $\{M_k\}$ and $\{\bar{\chi}_k\}$. Then the desired convergent rate of $O(1/k^2)$ for our inexact APG method follows.

5.2.1 Boundedness of $\{p_k\}$

In this subsection, we consider sufficient conditions to ensure the boundedness of the sequence $\{p_k\}$.

Lemma 5.10. *Suppose that there exists $(\bar{x}, \bar{p}, \bar{z})$ such that*

$$\mathcal{A}(\bar{x}) = b, \; \bar{x} \succeq 0, \quad \nabla f(\bar{x}) = \mathcal{A}^*\bar{p} + \bar{z}, \; \bar{z} \succ 0. \tag{5.53}$$

If the sequence $\{f(x_k)\}$ is bounded from above, then the sequence $\{x_k\}$ is bounded.

Proof. By using the convexity of f, we have

$$\begin{aligned}
& f(\bar{x}) - f(x_k) \leq \langle \nabla f(\bar{x}),\, \bar{x} - x_k \rangle = \langle \mathcal{A}^* \bar{p} + \bar{z},\, \bar{x} - x_k \rangle \\
= \ & \langle \bar{p},\, \mathcal{A}(\bar{x}) - \mathcal{A}(x_k) \rangle + \langle \bar{z},\, \bar{x} \rangle - \langle \bar{z},\, x_k \rangle \\
\leq \ & \|\bar{p}\| \|b - \mathcal{A}(x_k)\| + \langle \bar{z},\, \bar{x} \rangle - \langle \bar{z},\, x_k \rangle \\
\leq \ & \|\bar{p}\| \mu_k / t_k^2 + \langle \bar{z},\, \bar{x} \rangle - \langle \bar{z},\, x_k \rangle \leq \|\bar{p}\| \mu_1 + \langle \bar{z},\, \bar{x} \rangle - \langle \bar{z},\, x_k \rangle.
\end{aligned}$$

Thus

$$\lambda_{\min}(\bar{z}) \mathrm{Tr}(x_k) \leq \langle \bar{z},\, x_k \rangle \leq \|\bar{p}\| \mu_1 + \langle \bar{z},\, \bar{x} \rangle - f(\bar{x}) + f(x_k). \tag{5.54}$$

From here, the required result is proved. □

Remark 5.11. The condition that $\{f(x_k)\}$ is bounded from above appears to be fairly weak. But unfortunately we are not able to prove that this condition holds true. In many cases of interest, such as the nearest correlation matrix problem (1.14), the condition that $\{f(x_k)\}$ is bounded above or that $\{x_k\}$ is bounded can be ensured since Ω_1 is bounded.

Lemma 5.12. *Suppose that $\{x_k\}$ is bounded and there exists $\hat{x}$ such that*

$$\mathcal{A}(\hat{x}) = b, \ \ \hat{x} \succ 0.$$

Then the sequence $\{z_k\}$ is bounded. In addition, the sequence $\{p_k\}$ is also bounded.

Proof. From (5.32), we have

$$\begin{aligned}
& \lambda_{\min}(\hat{x}) \mathrm{Tr}(z_k) \ \leq \ \langle \hat{x},\, z_k \rangle \ = \ \langle \hat{x},\, \nabla q_k(x_k) - \mathcal{A}^* p_k - \delta_k \rangle \\
= \ & -\langle b,\, p_k \rangle + \langle \hat{x},\, \nabla q_k(x_k) \rangle - \langle \hat{x},\, \delta_k \rangle \\
\leq \ & \Delta + \langle \hat{x} - x_k,\, \nabla q_k(x_k) \rangle - \langle \hat{x},\, \delta_k \rangle \\
= \ & \Delta + \langle \hat{x} - x_k,\, \nabla f(y_k) + \mathcal{H}_k(x_k - y_k) \rangle - \langle \hat{x},\, \delta_k \rangle \\
\leq \ & \Delta + \|\hat{x} - x_k\| \|\nabla f(y_k) + \mathcal{H}_k(x_k - y_k)\| + \|\mathcal{H}_k^{1/2} \hat{x}\| \epsilon_k / (\sqrt{2} t_k) \\
\leq \ & \Delta + \|\hat{x} - x_k\| \|\nabla f(y_k) + \mathcal{H}_k(x_k - y_k)\| + \|\mathcal{H}_1^{1/2} \hat{x}\| \epsilon_1 / \sqrt{2} \qquad (5.55)
\end{aligned}$$

Since $\{x_k\}$ is bounded, it is clear that $\{y_k\}$ is also bounded. By the continuity of ∇f and that fact that $0 \preceq \mathcal{H}_k \preceq \mathcal{H}_1$, the sequence $\{\|\nabla f(y_k) + \mathcal{H}_k(x_k - y_k)\|\}$ is also bounded. From (5.55), we can now conclude that $\{z_k\}$ is bounded.

Next, we show that $\{p_k\}$ is bounded. Let $\mathcal{A}^\dagger = (\mathcal{A}\mathcal{A}^*)^{-1}\mathcal{A}$. Note that the matrix $\mathcal{A}\mathcal{A}^*$ is nonsingular since $\mathcal{A}$ is assumed to be surjective. From (5.32), we have $p_k = \mathcal{A}^\dagger(\nabla q_k(x_k) - z_k - \delta_k)$, and hence

$$\|p_k\| \leq \|\mathcal{A}^\dagger\|\|\nabla q_k(x_k) - z_k - \delta_k\| \leq \|\mathcal{A}^\dagger\|\Big(\|\nabla f(y_k) + \mathcal{H}_k(x_k - y_k)\| + \|z_k\| + \|\delta_k\|\Big).$$

Since $\mathcal{H}_k \preceq \mathcal{H}_1 \preceq \lambda_{\max}(\mathcal{H}_1)I$, we have $\|\delta_k\| \leq \sqrt{\lambda_{\max}(\mathcal{H}_1)}\|\mathcal{H}_k^{-1/2}\delta_k\| \leq \sqrt{\lambda_{\max}(\mathcal{H}_1)}\frac{\epsilon_1}{\sqrt{2}}$. By using the fact that the sequences $\{\|\nabla f(y_k)+\mathcal{H}_k(x_k-y_k)\|\}$ and $\{z_k\}$ are bounded, the boundedness of $\{p_k\}$ follows. □

5.2.2 A semismooth Newton-CG method

In Section 5.2, an inexact APG method (Algorithm 4) was presented for solving (P) with the desired convergent rate of $O(1/k^2)$. However, an important issue on how to efficiently solve the inner subproblem (5.30) has not been addressed.

In this section, we propose the use of a semismooth Newton-CG (SSNCG) method to solve (5.30) with warm-start using the iterate from the previous iteration. Note that the self-adjoint positive definite linear operator $\mathcal{H}_k$ can be chosen by the user. Suppose that at each iteration we are able to choose a linear operator of the form:

$$\mathcal{H}_k := w_k \circledast w_k, \quad \text{where } w_k \in \mathcal{S}^n_{++},$$

such that $f(x) \leq q_k(x)$ for all $x \in \Omega$. (Note that we can always choose $w_k = \sqrt{L}I$ if there is no other better choice.) Then the objective function $q_k(\cdot)$ in (5.30) can equivalently be written as

$$q_k(x) = \frac{1}{2}\|w_k^{1/2}(x - u_k)w_k^{1/2}\|^2 + f(y_k) - \frac{1}{2}\|w_k^{-1/2}\nabla f(y_k)w_k^{-1/2}\|^2,$$

where $u_k = y_k - w_k^{-1}\nabla f(y_k)w_k^{-1}$. By dropping the last two constant terms in the above equation, we can equivalently write (5.30) as the following well-studied W-weighted semidefinite least squares problem

$$\min\Big\{\frac{1}{2}\|w_k^{1/2}(x-u_k)w_k^{1/2}\|^2 \;:\; \mathcal{A}(x)=b, x\succeq 0\Big\}. \tag{5.56}$$

Let $\bar{x}=w_k^{1/2}xw_k^{1/2}$, $\bar{u}_k=w_k^{1/2}u_kw_k^{1/2}$, and define the linear map $\bar{\mathcal{A}}:\mathcal{S}^n\to\Re^m$ by

$$\bar{\mathcal{A}}(\bar{x})=\mathcal{A}(w_k^{-1/2}\,\bar{x}\,w_k^{-1/2}).$$

Then (5.56) can equivalently be written as

$$\min\Big\{\frac{1}{2}\|\bar{x}-\bar{u}_k\|^2 \;:\; \bar{\mathcal{A}}(\bar{x})=b,\ \bar{x}\succeq 0\Big\}, \tag{5.57}$$

whose Lagrangian dual problem is given by

$$\max\Big\{\theta(p):=b^Tp-\frac{1}{2}\|\Pi_{\mathcal{S}^n_+}(\bar{u}_k+\bar{\mathcal{A}}^*p)\|^2 \mid p\in\Re^m\Big\} \tag{5.58}$$

where $\Pi_{\mathcal{S}^n_+}(u)$ denotes the metric projection of $u\in\mathcal{S}^n$ onto $\mathcal{S}^n_+$. The problem (5.58) is an unconstrained continuously differentiable convex optimization problem, and it can be efficiently solved by the SSNCG method developed in [97]. Note that once an approximate solution p_k is computed from (5.58), an approximate solution for (5.56) can be computed by $x_k=\Pi_{\mathcal{S}^n_+}(\bar{u}_k+\bar{\mathcal{A}}^*p_k)$ and its complementary dual slack variable is $z_k=\bar{u}_k+\bar{\mathcal{A}}^*p_k-x_k$.

Note that the problem (5.58) is an unconstrained continuously differentiable convex optimization problem which can also be solved by a gradient ascent method. In our numerical implementation, we use a gradient method to solve (5.58) during the initial phase of Algorithm 4 when high accuracy solutions are not required. When the gradient method encounters difficulty in solving the subproblem to the required accuracy or becomes inefficient, we switch to the SSNCG method to solve (5.58). We should note that approximate solution computed for the current subproblem can be used to warm start the SSNCG and gradient methods for solving the next subproblem. In fact, the strategy of solving a semidefinite least squares subproblem (5.30) in each iteration of our inexact APG algorithm is practically viable precisely because we are able to warm start the SSNCG or gradient methods when solving the subproblems. In our numerical experience, the SSNCG method would typically take less than 5 Newton steps to solve each subproblem with warm start.

To successfully apply the SSNCG method to solve (5.30), we must find a suitable symmetrized Kronecker product approximation of $\mathcal{Q}$. Note that for the H-weighted nearest correlation matrix problem (1.14) where $\mathcal{Q}$ is a diagonal operator defined by $\mathcal{Q}(x) = (H \circ H) \circ x$, a positive definite symmetrized Kronecker product approximation for $\mathcal{Q}$ can be derived as follows. Consider a rank-one approximation dd^T of $H \circ H$, then $\mathrm{Diag}(d) \circledast \mathrm{Diag}(d)$ is a symmetrized Kronecker product approximation of $\mathcal{Q}$. For the vector $d \in \mathcal{R}^n$, one can simply take

$$d_j = \max\Big\{\epsilon, \max_{1\le i\le n}\{H_{ij}\}\Big\}, \; j = 1, \dots, n. \tag{5.59}$$

where $\epsilon > 0$ is a small positive number.

For the convex QSDP problem (1.12) where the linear operator $\mathcal{Q}$ is defined by

$$\mathcal{Q}(x) = B \circledast I(x) = (Bx + xB)/2, \tag{5.60}$$

where $B \in \mathcal{S}^n_+$, we propose the following strategy for constructing a suitable symmetrized Kronecker product approximation of $\mathcal{Q} = B \circledast I$. Suppose we have the eigenvalue decomposition $B = P\Lambda P^T$, where $\Lambda = \mathrm{diag}(\lambda)$ and $\lambda = (\lambda_1, \dots, \lambda_n)^T$ is the vector of eigenvalues of B. Then

$$\langle x, B \circledast I(x)\rangle = \frac{1}{2}\langle \hat{x}, \Lambda\hat{x} + \hat{x}\Lambda\rangle = \frac{1}{2}\sum_{i=1}^n\sum_{j=1}^n \hat{x}_{ij}^2(\lambda_i + \lambda_j) = \sum_{i=1}^n\sum_{j=1}^n \hat{x}_{ij}^2 M_{ij},$$

where $\hat{x} = P^T x P$ and $M = \frac{1}{2}(\lambda e^T + e\lambda^T)$ with $e \in \mathcal{R}^n$ being the vector of all ones. For the choice of w_k, one may simply choose $w_k = \sqrt{\max(M)}I$, where $\max(M)$ is the largest element of M. However, if the matrix B is ill-conditioned, this choice of w_k may not work very well in practice since $\max(M)I \circledast I$ may not be a good approximation of $\mathcal{Q} = B \circledast I$. To find a better approximation of $\mathcal{Q}$, we propose to consider the following nonconvex minimization problem:

$$\min\Big\{\sum_{i=1}^n\sum_{j=1}^n h_i h_j \mid h_i h_j - M_{ij} \ge 0 \;\forall\, i, j = 1, \dots, n, \; h \in \mathcal{R}^n_+\Big\}. \tag{5.61}$$

Thus if $\hat{h}$ is a feasible solution to the above problem, then we have

$$\langle x,\, B \circledast I(x)\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} \hat{x}_{ij}^2 M_{ij} \leq \sum_{i=1}^{n}\sum_{j=1}^{n} \hat{x}_{ij}^2 \hat{h}_i \hat{h}_j = \langle x, w_k x w_k\rangle \quad \forall\, x \in \mathcal{S}^n$$

with $w_k = P\mathrm{Diag}(\hat{h})P^T$. Note that the above strategy can also be used to get a suitable symmetrized Kronecker product approximation of the form $\mathrm{Diag}(d) \circledast \mathrm{Diag}(d)$ when $\mathcal{Q}$ is a diagonal operator.

To find a good feasible solution for (5.61), we consider the following strategy. Suppose we are given an initial vector $u \in \mathcal{R}^n_+$ such that $uu^T - M \geq 0$. For example, we may take $u = \sqrt{\max(M)}e$. Our purpose is to find a correction vector $\xi \in \mathcal{R}^n_+$ such that $h := u - \xi$ satisfies the constraint in (5.61) while the objective value is reduced. Note that we have

$$h_i h_j - M_{ij} = u_i u_j - M_{ij} - (u_i \xi_j + u_j \xi_i) + \xi_i \xi_j \geq u_i u_j - M_{ij} - (u_i \xi_j + u_j \xi_i).$$

Thus the constraints in (5.61) are satisfied if $\xi \leq u$ and

$$u_i \xi_j + u_j \xi_i \leq u_i u_j - M_{ij} \quad \forall\, i, j = 1, \ldots, n.$$

Since $\sum_{i=1}^n \sum_{j=1}^n h_i h_j = (e^T\xi)^2 - 2(e^T u)(e^T \xi) + (e^T u)^2$, and noting that $0 \leq e^T\xi \leq e^T u$, the objective value in (5.61) is minimized if $e^T\xi$ is maximized. Thus we propose to consider the following LP:

$$\max\Big\{ e^T\xi \mid u_i\xi_j + u_j\xi_i \leq u_i u_j - M_{ij}\, \forall\, i, j = 1, \ldots, n,\ 0 \leq \xi \leq u \Big\}. \qquad (5.62)$$

Observe that the correction vector ξ depends on the given vector u. Thus if necessary, after a new u is obtained, one may repeat the process by solving the LP associated with the new u.

Chapter 6

Numerical Results

In this chapter, we conduct a variety of large scale numerical experiments to evaluate the performance of our proposed algorithms. In section 6.1, we present numerical results of the partial PPA for solving nuclear norm regularized matrix least squares problems. In section 6.2, we present numerical results of the inexact APG method for solving large scale linear constrainted convex QSDP problems, including the H-weighted nearest correlation problem.

6.1 Numerical Results for nuclear norm minimization problems

In this section, we report some numerical results to demonstrate the efficiency of our partial proximal point algorithm.

In order to measure the infeasibilities of the primal problem (3.16), we define two linear operators $\mathcal{B}_e : \Re^{p\times q} \to \Re^{s_1}$ and $\overline{\mathcal{B}}_e : \Re^{p\times q} \to \Re^{s_2}$, respectively, as follows:

$$\begin{cases} (\mathcal{B}_e(X))_i := \langle \mathcal{B}_i, X\rangle, \text{ for } i = 1,\dots,s_1, \\ (\overline{\mathcal{B}}_e(X))_i := \langle \mathcal{B}_i, X\rangle, \text{ for } i = s_1+1,\dots,s. \end{cases}$$

Let $d = (d_{s_1}; d_{s_2})$ where $d_{s_1} \in \Re^{s_1}$ and $d_{s_2} \in \Re^{s_2}$. We measure the infeasibilities and optimality for the primal problem (3.16) and the dual problem (3.20) as follows:

$$\begin{aligned} R_P &= \frac{\|(b-\zeta-\mathcal{A}(X); d_{s_1}-\mathcal{B}_e(X); \max(0, d_{s_2}-\overline{\mathcal{B}}_e(X)))\|}{1+\|\hat{b}\|}, \\ R_D &= \frac{\|C-\widehat{\mathcal{A}}^*y-Z\|}{1+\|\widehat{\mathcal{A}}^*\|}, \quad \text{relgap} = \frac{f_\rho(\zeta,X)-g_\rho(\zeta,\xi)}{1+|f_\rho(\zeta,X)|+|g_\rho(\zeta,\xi)|}, \end{aligned}$$

where $y=(\zeta;\xi), Z=(D_{\rho\sigma}(W)-W)/\sigma$ with $W=X-\sigma(C-\widehat{\mathcal{A}}^*y)$, and $f_\rho(\zeta,X)$ and $g_\rho(\zeta,\xi)$ are the objective functions of the primal problem (3.16) and the dual problem (3.20), respectively. The infeasibility of the condition $\|Z\|_2 \leq \rho$ is not checked since it is satisfied up to machine precision throughout the algorithm. In our numerical experiments, we stop the partial PPA when

$$\max\{R_P, R_D\} \leq \texttt{Tol}, \tag{6.1}$$

where `Tol` is a pre-specified accuracy tolerance. We choose the initial iterate $X^0 = 0, y^0 = 0$, and $\sigma_0 = 1$. Unless otherwise specified, we set $\texttt{Tol} = 10^{-6}$ as the default. For examples from 1 to 4, where the inner subproblem (3.34) is solved by the inexact smoothing Newton method, besides (6.1), we impose another stopping criterion $|\text{relgap}| \leq 10^{-5}$ for stoping the partial PPA. The parameter ρ in (3.16) is set to be $\rho = 10^{-3}\|\mathcal{A}^*b\|_2$ if the data is not contaminated by noise; otherwise, the parameter ρ is set to be $\rho = 5 \times 10^{-3}\|\mathcal{A}^*b\|_2$. For examples from 5 to 8, where the inner subproblem (3.34) is solved by the semismooth Newton-CG method, we use the default stopping criterion (6.1) and the parameter ρ in (3.16) is set to be $\rho = 10^{-3}\|\mathcal{A}^*b\|_2$ for all cases.

Example 1

We consider the nearest matrix approximation problem which was discussed by Golub, Hoffman and Stewart in [48], where the classic Eckart-Young[35]-Mirsky[85] theorem was extended to obtain the nearest lower-rank approximation while certain specified columns of the matrix are fixed. The Eckart-Young-Mirsky theorem has the drawback that the approximation generally differs from the original matrix in all its entries. This is not suitable for application where some columns of the original matrix may be fixed. For example, in statistics the regression matrix for the multiple regression model with a constant term has a column of all ones, and this column should not be perturbed. For each triplet (p, q, r), where r is the predetermined rank, we generate a random matrix $M \in \Re^{p\times q}$ of rank r by setting $M = M_1 M_2^T$ where both $M_1 \in \Re^{p\times r}$ and $M_2 \in \Re^{q\times r}$ have i.i.d. standard uniform entries in $(0, 1)$. As observed entries in practice are rarely exact, we corrupt the entries of M by Gaussian noises to simulate the situation where the observed data may be noisy as follows. First we generate a random matrix $N \in \Re^{p\times q}$ with i.i.d Gaussian entries. Then we assume that the observed data is given by $\widetilde{M} = M + \tau N\|M\|/\|N\|$, where τ is the noise factor. In our numerical experiments, we choose the parameter $\tau = 0.1$. We assume that the first column of M should be fixed. Then the minimization problem can be stated as follows:

$$\min_{X\in\Re^{p\times q}} \Big\{ \frac{1}{2}\|X - \widetilde{M}\|^2 + \rho\|X\|_* \, : \, Xe_1 \, = \, Me_1, \, X \geq 0 \Big\}, \tag{6.2}$$

where e_1 is the first column of the q-by-q identity matrix. Here we impose an extra constraint $X \geq 0$ since the original matrix M is nonnegative. Note that the approximation derived in [48] generally is not nonnegative.

For each p, q, r and τ, we repeat the above procedure 5 times. In Table 6.1, we report the total number of constraints $(m + s)$ in (3.16), the average number of outer iterations, the average total number of inner iterations, the average number of BiCGStab steps taken to solve (3.75), the average infeasibilities in (3.16) and (3.20), respectively, the average relative gap between (3.16) and (3.20), the average relative mean square error MSE $:= \|X - M\|/\|M\|$ (where M is the original matrix), the mean value of the rank (#sv) of X, and the average CPU time taken (in the format hours:minutes:seconds). We may observe from the table that the partial PPA is very efficient for solving (6.2). For the nonsquare matrix problem, where p is moderate but q is large, e.g., $p = 100$ and $q = 20000$, we only need to compute the economic form of the SVD. Thus we use the technique introduced in section 3.5 to compute

V_2 via the QR factorization of V_1. It takes about three and a half minutes to solve the last instance for achieving the tolerance 10^{-6} while the MSE is reasonably small.

$p \times q$	r	$m+s$	it.|itsub|bicg	R_p | R_D | relgap	MSE	#sv	time
500×500	10	500500	5.0 | 13.0 | 4.0	5.57e-7 | 7.29e-7 | 3.61e-7	3.33e-2	169 (10)	38
500×500	50	500500	3.0 | 8.0 | 3.3	1.46e-7 | 4.58e-7 | 1.30e-7	4.01e-2	177 (NA)	26
500×500	100	500500	3.0 | 7.8 | 3.4	3.63e-7 | 6.68e-7 | 1.88e-7	3.72e-2	177 (NA)	26
1000×1000	10	2001000	7.0 | 16.4 | 4.6	1.03e-7 | 5.33e-7 | -3.15e-6	2.09e-2	121 (10)	3:14
1000×1000	50	2001000	9.0 | 16.2 | 3.0	6.13e-9 | 1.90e-8 | -3.75e-6	3.31e-2	138 (NA)	2:27
1000×1000	100	2001000	9.0 | 15.8 | 2.7	9.45e-9 | 1.92e-8 | -3.82e-6	3.10e-2	143 (NA)	2:20
1500×1500	10	4501500	9.0 | 19.0 | 4.4	3.47e-8 | 4.01e-8 | 9.14e-6	1.85e-2	22 (10)	9:18
1500×1500	50	4501500	9.0 | 16.2 | 3.2	1.06e-8 | 2.48e-8 | -3.71e-6	3.26e-2	54 (50)	6:54
1500×1500	100	4501500	8.0 | 15.2 | 3.2	1.14e-8 | 1.52e-8 | -4.50e-6	3.19e-2	67 (NA)	6:41
100×5000	10	1000100	10.0 | 12.8 | 1.5	2.67e-8 | 4.10e-8 | 3.85e-6	5.72e-2	100 (10)	46
100×10000	10	2000100	10.0 | 12.4 | 1.4	2.00e-8 | 4.09e-8 | 4.05e-6	5.70e-2	100 (10)	1:40
100×20000	10	4000100	10.4 | 13.0 | 1.4	1.89e-8 | 4.13e-8 | 3.90e-6	5.70e-2	100 (10)	3:32

Table 6.1: Numerical results of the partial PPA on (6.2).

In the numerical implementation, we observe that when the generated matrix M is of small rank, e.g., $r = 10$, the singular values of the computed solution X are separated into two clusters with the first cluster having much larger mean value than that of the second cluster (see, e.g., Figure 6.1). We may view the number of singular values in the first cluster as a good estimate of the rank of the optimal solution, while the smaller positive singular values in the second cluster may be attributed to the presence of noise in the given data. When the matrix M is of high rank, e.g., $r = 50$, the singular values of X are usually not separated into two clusters (see, e.g., Figure 6.1), excluding the largest singular value of X. In Table 6.1, when the singular values of X are separated into two clusters, we also report the number of singular values in the first cluster in parenthesis next to #sv. In the table, “NA” means that the singular values of X are not separated into two clusters.

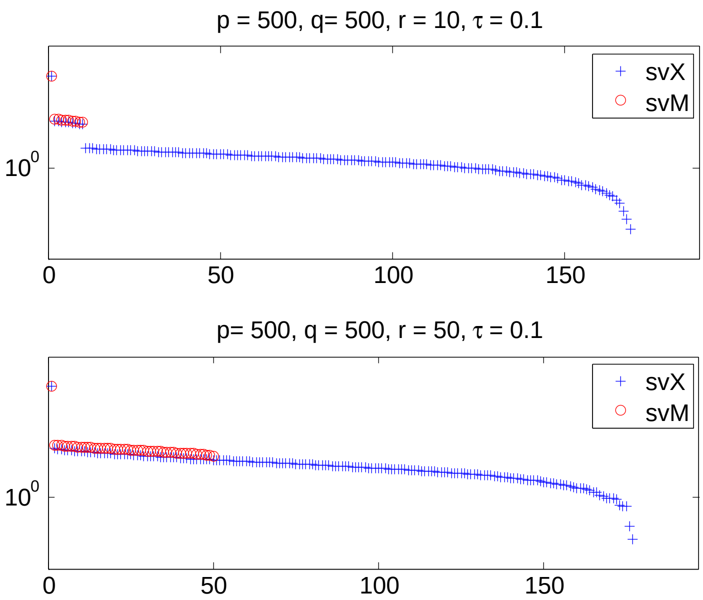

Figure 6.1: Distribution of singular values of X and M.

Example 2

Recently Lin [70] proposed the Latent Markov Analysis (LMA) approach for finding the reduced rank approximations of transition matrices. The LMA is applied to clustering based on pairwise similarities such that the inferred cluster relationships can be described probabilistically by the reduced-rank transition matrix. In [6], Benczúr, Csalogány and Sarlós considered the problem of finding the low rank approximation of the transition matrix for computing the personalized PageRank, which describes the backlink-based page quality around user-selected pages.

In this example, we evaluate the performance of our partial PPA for finding the nearest transition matrix of lower rank. Consider the set of n web pages as a directed graph whose nodes are the web pages and whose edges are all the links between pages. Let $\deg(i)$ be the outdegreee of the page i, i.e., the number of pages which can be reached by a direct link from page i. Note that all the self-referential links in the web graph are excluded. Let $P \in \Re^{n\times n}$ be the matrix which describes the transition probability between the page i and j, where $P_{ij} = 1/\deg(i)$ if $\deg(i) > 0$ and there is a link form page i to page j. For some page i having no outlink (dangling pages), we assume $P_{ij} = 1/n$ for $j = 1, \ldots, n$, i.e., the user will make a random choice with uniform distribution $1/n$. Since the matrix P for the web graph generally is reducible, P may have several eigenvalues on the unit circle, which could cause convergence problems to the power method for computing the PageRank [65]. The standard way of ensuring irreducibility is that we replace P by the matrix

$$P_c = cM + (1-c)ev^T,$$

where $c \in (0,1)$, $e \in \Re^n$ is a vector of all ones, and $v \in \Re^n$ is a probability vector such that $v \geq 0$ and $e^T v = 1$. We generate a random matrix $N \in \Re^{n\times n}$ with i.i.d Gaussian entries. Then we assume that the observed data is given by $\widetilde{P}_c = P_c + \tau N \|P_c\|/\|N\|$, where τ is the noise factor. In our numerical experiments, we choose the parameter $\tau = 0.1$, $c = 0.85$ which is a typical value used by Google, and $v_i = 1/n$, for $i = 1, \ldots, n$. Then the minimization problem that we finally solve can be stated as follows:

$$\min_{X\in\Re^{n\times n}} \left\{ \frac{1}{2}\|X - \widetilde{P}_c\|^2 + \rho\|X\|_* \,:\, Xe = e,\ X \geq 0 \right\}. \tag{6.3}$$

We use the data `Harvard500.mat` generated by Cleve Moler's MATLAB program `surfer.m` to evaluate the performance of our algorithm. The data and program are available at `http://www.mathworks.com/moler/ncmfilelist.html`. We also use the M file `surfer.m` to generate three adjacency graphs of a portion of web pages starting at the root page "`http://www.nus.edu.sg`". We also apply our algorithm to the data sets * collected by Panayiotis Tsaparas on querying the Google search engine about four topics: "automobile industries", "computational complexity", "computational geometry", and "randomized algorithms". Table 6.2 reports the average numerical results of PPA for solving (6.3) over 5 runs, where r denotes the

*Datasets are available at: `http://www.cs.toronto.edu/~tsap/experiments/datasets/index.html` and `http://www.cs.toronto.edu/~tsap/experiments/download/download.html` rank of P_c for each data set. We can observe from the table that the partial PPA is very efficient for solving (6.3) when applied to the real web graph data sets.

Problem	n	r	$m+s$	it.|itsub|bicg	R_p | R_D | relgap	MSE	#sv	time
Harvard500	500	218	500500	6.0 | 14.6 | 7.8	7.58e-8 | 4.92e-9 | -7.48e-6	5.87e-2	366	1:01
NUS500	500	225	500500	6.2 | 12.4 | 5.9	4.51e-8 | 1.60e-9 | -5.22e-6	5.70e-2	382	47
NUS1000	1000	466	2001000	5.4 | 14.2 | 7.7	3.35e-7 | 4.67e-9 | -6.46e-6	5.62e-2	658	5:19
NUS1500	1500	807	4501500	5.0 | 15.0 | 8.8	3.34e-7 | 5.68e-9 | -7.00e-6	6.35e-2	957	17:21
RandomAlg	742	216	1101870	7.0 | 16.0 | 7.7	3.37e-7 | 3.19e-9 | -7.02e-6	4.48e-2	631	2:48
Complexity	884	255	1563796	7.0 | 16.2 | 7.7	6.75e-8 | 1.75e-9 | -4.65e-6	4.77e-2	712	4:22
Automobile	1196	206	2862028	6.0 | 16.4 | 8.7	2.02e-7 | 5.91e-9 | -8.80e-6	4.05e-2	844	10:14
Geometry	1226	416	3007378	7.0 | 17.2 | 8.0	7.22e-8 | 1.99e-9 | -4.18e-6	4.67e-2	1018	11:01

Table 6.2: Numerical results of the partial PPA on (6.3).

Example 3

We consider the problem of finding a low rank doubly stochastic matrix with a prescribed entry. A matrix $M \in \Re^{n\times n}$ is called doubly stochastic if it is nonnegative and all its row and column sums are equal to one. This problem arose from numerical simulation of large circuit networks. In order to reduce the complexity of the simulation of the whole system, the Padé approximation with Krylov subspace method, such as the Lanczos algorithm, is a useful tool for generating a lower order approximation to the linear system matrix which describes the large linear network [3]. The tridiagonal matrix $M \in \Re^{n\times n}$ produced by the Lanczos algorithm is generally not doubly stochastic. If the original system matrix is doubly stochastic, then we need to find a low rank approximation of M, which is doubly stochastic and matches the maximal moments. In our numerical experiments, we will not restrict the matrix M to be tridiagonal.

For each pair (n, r), we generate a positive matrix $\overline{M} \in \Re^{n\times n}$ with rank r by the same method as in Example 1. Then we use the Sinkhorn-Knopp algorithm [113] to find two diagonal matrices $D_1 \in \Re^{n\times n}$ and $D_2 \in \Re^{n\times n}$, where all the diagonal entries of D_1 and D_2 are positive, such that $M = D_1\overline{M}D_2$ is a doubly stochastic matrix of rank r. We sample a subset $\mathcal{E}$ of m entries of M uniformly at random, and generate a random matrix $N_{\mathcal{E}} \in \Re^{p\times q}$ with sparsity pattern $\mathcal{E}$ and i.i.d standard Gaussian random entries. Then we assume that the observed data is given by $\widetilde{M}_{\mathcal{E}} = M_{\mathcal{E}} + \tau N_{\mathcal{E}}\|M_{\mathcal{E}}\|/\|N_{\mathcal{E}}\|$, where τ is the noise factor. Then the problem for matching the first moment of M can be stated as follows:

$$\min_{X\in\Re^{n\times n}} \Big\{\frac{1}{2}\|X_{\mathcal{E}} - \widetilde{M}_{\mathcal{E}}\|^2 + \rho\|X\|_* \,:\, Xe = e,\, X^Te = e,\, X_{11} = M_{11},\, X \geq 0\Big\}. \quad (6.4)$$

In our numerical experiments, we set $\tau = 0, 0.1$, and the number of sampled entries to be $m = 10dr$, where $dr = r(2n-r)$ is the value of the degrees of freedom in an $n \times n$ matrix of rank r. In Table 6.3, we report the average numerical results for solving (6.4) on randomly generated matrices over 5 runs, where m is the average number of sampled entries, and $m + s$ is the average number of total constraints in (3.16). For problems with noise, if the singular values of X are separated into two clusters, we report the number of singular values in the first cluster in parenthesis next to #sv, and we use "NA" to denote that the singular values of X are not separated into two clusters. We can observe from the table that the partial PPA can solve (6.4) very efficiently for all the instances with or without Gaussian noise.

n/τ	r	m	$m+s$	it.|itsub|bicg	R_p | R_D | relgap	MSE	#sv	time
500/0.0	10	99148	350148	7.0 | 15.4 | 3.3	5.71e-7 | 6.88e-8 | -4.30e-6	3.53e-3	10	26
	50	250000	501000	6.0 | 8.2 | 1.5	2.03e-7 | 8.45e-8 | -3.45e-6	7.07e-3	50	09
	100	250000	501000	5.0 | 7.0 | 1.4	1.02e-7 | 1.50e-7 | -3.62e-6	1.01e-2	100	08
1000/0.0	10	199034	1201034	9.0 | 20.0 | 4.0	6.80e-7 | 5.66e-8 | -8.17e-6	4.07e-3	10	2:56
	50	974915	1976915	6.0 | 12.0 | 2.7	2.88e-7 | 4.68e-8 | -3.69e-6	7.11e-3	50	1:24
	100	1000000	2002000	5.0 | 7.0 | 1.4	3.63e-8 | 7.41e-8 | -3.52e-6	1.01e-2	100	39
1500/0.0	10	299194	2552194	10.0 | 23.0 | 4.0	5.81e-7 | 3.94e-8 | -8.79e-6	4.41e-3	10	9:29
	50	1474481	3727481	7.0 | 13.8 | 2.6	1.41e-7 | 4.40e-8 | -4.70e-6	7.54e-3	50	4:12
	100	2250000	4503000	5.0 | 7.0 | 1.4	1.34e-8 | 4.92e-8 | -3.50e-6	1.01e-2	100	2:02

n/τ	r	m	$m+s$	it.|itsub|bicg	R_p | R_D | relgap	MSE	#sv	time
500/0.1	10	99148	350148	7.0 | 16.0 | 3.2	1.97e-7 | 1.93e-7 | -6.27e-6	5.42e-2	174 (10)	26
	50	250000	501000	5.0 | 9.2 | 2.0	1.65e-7 | 2.31e-7 | -8.58e-6	3.97e-2	177 (NA)	12
	100	250000	501000	5.0 | 9.0 | 2.1	1.11e-7 | 1.83e-7 | -5.37e-6	3.65e-2	177 (NA)	12
1000/0.1	10	199034	1201034	8.0 | 18.8 | 3.6	1.45e-7 | 9.18e-8 | -9.31e-6	5.50e-2	234 (10)	2:41
	50	974915	1976915	5.0 | 10.0 | 2.7	7.25e-7 | 7.91e-8 | -3.93e-6	3.30e-2	145 (NA)	1:13
	100	1000000	2002000	3.0 | 6.6 | 2.1	4.43e-7 | 3.32e-7 | -7.58e-6	3.07e-2	143 (NA)	45
1500/0.1	10	299194	2552194	9.0 | 22.2 | 3.9	1.69e-7 | 3.84e-8 | -5.68e-6	5.49e-2	275 (10)	8:56
	50	1474481	3727481	5.0 | 11.0 | 2.7	4.76e-7 | 1.11e-7 | -6.87e-6	3.41e-2	194 (NA)	3:36
	100	2250000	4503000	2.0 | 5.2 | 3.1	2.11e-7 | 2.71e-7 | -3.26e-6	3.19e-2	68 (NA)	1:55

Table 6.3: Numerical results of the partial PPA on (6.4). In the table, $m = 10dr$ and $dr = r(2n-r)$.

We may also consider a generalized version of problem (6.4), where we want to find a low rank doubly stochastic matrix with k prescribed entries of M. The problem could be stated as follows:

$$\begin{array}{ll} \min & \frac{1}{2}\|X_{\mathcal{E}} - \widetilde{M}_{\mathcal{E}}\|^2 + \rho\|X\|_* \\ \text{s.t.} & Xe \;=\; e,\ X^T e = e, \\ & e_{i_t}^T X e_{j_t} \;=\; e_{i_t}^T M e_{j_t}, \quad 1 \le t \le k, \\ & X \ge 0,\ X \in \Re^{n\times n}, \end{array} \tag{6.5}$$

where $(i_1, j_1), \ldots, (i_k, j_k)$ are distinct pairs and e_i is the i-th column of the n-by-n identity matrix. In our numerical experiments, we set $k = \lceil 10^{-3} n^2 \rceil$, which is the number of prescribed entries selected uniformly at random. Table 6.4 presents the average numerical results for solving (6.5) on randomly generated matrices over 5 runs. For problems with noise, if the singular values of X are separated into two clusters, we also report the number of singular values in the first cluster in parenthesis next to #sv, and we use "NA" to denote that the singular values of X are not separated into two clusters.

n/τ	r	m	$m+s$	it.|itsub|bicg	R_p | R_D | relgap	MSE	#sv	time
500/0.0	10	99148	350148	7.0 | 18.2 | 6.4	5.62e-7 | 6.89e-8 | -4.27e-6	3.47e-3	10	39
	50	250000	501000	6.0 | 10.8 | 3.0	3.71e-7 | 8.39e-8 | -3.39e-6	7.05e-3	50	15
	100	250000	501000	5.0 | 11.0 | 2.9	4.78e-7 | 1.49e-7 | -3.64e-6	1.00e-2	100	14
1000/0.0	10	199034	1201034	9.0 | 21.0 | 5.3	6.56e-7 | 5.78e-8 | -8.35e-6	3.94e-3	10	3:32
	50	974915	1976915	6.0 | 12.4 | 4.3	3.06e-7 | 4.65e-8 | -3.64e-6	7.06e-3	50	1:47
	100	1000000	2002000	5.0 | 12.0 | 3.4	3.76e-7 | 7.31e-8 | -3.47e-6	1.00e-2	100	1:27
1500/0.0	10	299194	2552194	10.0 | 25.8 | 6.3	6.68e-7 | 3.94e-8 | -8.85e-6	4.20e-3	11	12:54
	50	1474481	3727481	6.8 | 16.4 | 4.9	3.53e-7 | 5.39e-8 | -5.64e-6	7.50e-3	50	6:36
	100	2250000	4503000	5.0 | 12.0 | 4.3	3.39e-7 | 4.82e-8 | -3.42e-6	1.00e-2	100	4:41
500/0.1	10	99148	350148	7.0 | 16.0 | 3.6	2.10e-7 | 1.92e-7 | -6.26e-6	5.41e-2	174 (10)	28
	50	250000	501000	5.0 | 11.0 | 2.9	5.43e-7 | 2.29e-7 | -8.51e-6	3.97e-2	177 (NA)	17
	100	250000	501000	5.0 | 11.0 | 3.0	4.18e-7 | 1.81e-7 | -5.35e-6	3.65e-2	177 (NA)	17
1000/0.1	10	199034	1201034	8.0 | 19.0 | 4.1	1.61e-7 | 9.14e-8 | -9.28e-6	5.47e-2	234 (10)	2:58
	50	974915	1976915	5.0 | 13.2 | 4.6	6.18e-7 | 7.67e-8 | -3.90e-6	3.29e-2	151 (NA)	2:06
	100	1000000	2002000	3.0 | 11.2 | 5.5	1.17e-7 | 3.18e-7 | -7.13e-6	3.06e-2	151 (NA)	1:57
1500/0.1	10	299194	2552194	9.0 | 22.0 | 4.5	1.35e-7 | 3.81e-8 | -5.64e-6	5.45e-2	276 (10)	9:43
	50	1474481	3727481	5.0 | 13.0 | 4.6	6.26e-7 | 1.13e-7 | -6.75e-6	3.39e-2	203 (NA)	5:23
	100	2250000	4503000	2.0 | 10.2 | 8.1	2.07e-7 | 3.50e-7 | -6.19e-6	3.11e-2	119 (NA)	5:55

Table 6.4: Numerical results of the partial PPA on (6.5). In the table, $m = 10dr$ and $dr = r(2n - r)$.

Example 4

We consider the problem of finding a low rank nonnegative approximation which preserves the left and right principal eigenvectors of a square positive matrix. This problem was suggested by Ho and Dooren in [56]. Let $M \in \Re^{n\times n}$ be a positive matrix, i.e., all entries of M are positive. By the well-known Perron [95]-Frobenius [39] theorem, M has a positive eigenvalue λ which is simple and has the largest magnitude among all the eigenvalues of M. Moreover, there exist two positive eigenvectors $v \in \Re^n$ and $w \in \Re^n$ such that $Mv = \lambda v$ and $M^T w = \lambda w$. As suggested by Bonacich [13], the principal eigenvector could be used to measure the network centrality, where the i-th component of the eigenvector gives the centrality of the i-th node in the network. For example, the well-known Google's PageRank [65] is a variant of the eigenvector centrality for ranking web pages.

For each pair (n, r), we generate a positive matrix $M \in \Re^{n\times n}$ of rank r by the same method as in Example 1. Suppose that we sample a subset $\mathcal{E}$ of m entries of M that are possibly corrupted by Gaussian noise as in Example 3. Given the largest positive eigenvalue λ and the left and right principal eigenvectors v and w of M, we want to find a low rank approximation of M while preserving the left and right principal eigenvectors of M. Then the problem can be stated as follows:

$$\min_{X\in\Re^{n\times n}} \Big\{\frac{1}{2}\|X_{\mathcal{E}} - \widetilde{M}_{\mathcal{E}}\|^2 + \rho\|X\|_* \,:\, Xv = \lambda v, X^T w = \lambda w, X \geq 0\Big\}. \tag{6.6}$$

In our numerical experiments, we set the noise factor $\tau = 0, 0.1$. Table 6.5 reports the average numerical results of the partial PPA for solving (6.6) over 5 runs. For problems with noise, if the singular values of X are separated into two clusters, we report the number of singular values in the first cluster in parenthesis next to #sv, and we use "NA" to denote that the singular values of X are not separated into two clusters.

n/τ	r	m	$m+s$	it.|itsub|bicg	R_p | R_D | relgap	MSE	#sv	time
500/0.0	10	99157	350157	6.0 | 14.8 | 2.5	1.27e-7 | 1.45e-7 | -7.75e-6	3.54e-3	10	24
	50	250000	501000	3.0 | 7.0 | 2.0	3.95e-7 | 5.25e-7 | 3.36e-6	6.97e-3	50	09
	100	250000	501000	3.0 | 7.0 | 2.0	2.41e-7 | 5.46e-7 | -1.68e-6	1.01e-2	100	11
1000/0.0	10	199029	1201029	7.0 | 17.2 | 2.6	2.15e-7 | 7.20e-8 | 5.32e-6	2.93e-3	10	2:12
	50	974912	1976912	2.4 | 7.2 | 2.3	7.32e-8 | 6.53e-7 | -6.72e-6	7.39e-3	50	56
	100	1000000	2002000	2.0 | 7.0 | 2.1	2.13e-7 | 8.02e-7 | -9.34e-7	1.03e-2	100	52
1500/0.0	10	299187	2552187	8.0 | 22.4 | 2.7	5.26e-7 | 3.90e-8 | 4.89e-6	2.86e-3	10	8:23
	50	1474471	3727471	3.0 | 9.0 | 2.2	6.12e-7 | 8.03e-8 | -5.41e-6	7.54e-3	50	3:51
	100	2250000	4503000	2.0 | 7.0 | 2.1	1.38e-7 | 4.92e-7 | 2.85e-7	1.02e-2	100	2:41

n/τ	r	m	$m+s$	it.|itsub|bicg	R_p | R_D | relgap	MSE	#sv	time
500/0.1	10	99157	350157	2.0 | 5.8 | 2.2	1.85e-7 | 4.88e-7 | -4.58e-6	5.38e-2	170 (10)	16
	50	250000	501000	1.6 | 5.6 | 2.1	4.68e-7 | 8.07e-9 | -4.63e-7	3.94e-2	177 (NA)	18
	100	250000	501000	1.8 | 6.2 | 2.1	3.35e-7 | 9.18e-9 | -2.35e-7	3.64e-2	176 (NA)	17
1000/0.1	10	199029	1201029	2.0 | 5.2 | 1.9	6.13e-7 | 2.54e-7 | -2.08e-6	5.28e-2	230 (10)	1:20
	50	974912	1976912	2.0 | 6.8 | 2.4	9.95e-8 | 1.61e-8 | -5.18e-8	3.27e-2	145 (NA)	1:18
	100	1000000	2002000	2.0 | 6.0 | 2.2	9.21e-7 | 1.73e-7 | -2.64e-6	3.04e-2	142 (NA)	1:13
1500/0.1	10	299187	2552187	2.0 | 5.0 | 1.8	4.56e-7 | 1.83e-7 | 2.16e-6	5.22e-2	278 (10)	3:53
	50	1474471	3727471	2.0 | 5.6 | 2.4	3.95e-7 | 2.93e-8 | 1.75e-7	3.35e-2	192 (NA)	3:36
	100	2250000	4503000	2.0 | 7.4 | 2.2	6.33e-8 | 4.31e-8 | -5.30e-7	3.14e-2	67 (NA)	3:40

Table 6.5: Numerical results of the partial PPA on (6.6). In the table, $m = 10dr$ and $dr = r(2n - r)$.

Example 5

We consider the random matrix completion problem discussed in [19]. For each triplet (p, q, r), we first generate a random matrix $M \in \Re^{p\times q}$ by setting $M = M_1 M_2^T$ where $M_1 \in \Re^{p\times r}$, $M_2 \in \Re^{q\times r}$ each has i.i.d. Gaussian entries. Then we sample a subset $\mathcal{E}$ of m entries uniformly at random. As observed entries in practice are rarely exact, we corrupt the entries of $M_{\mathcal{E}}$ by Gaussian noises to simulate the situation where the observed data may be noisy as follows. First we generate a random matrix $N_{\mathcal{E}} \in \Re^{p\times q}$ that has sparsity pattern $\mathcal{E}$ and i.i.d Gaussian entries. Then we assume that the observed data is given by $\widetilde{M}_{\mathcal{E}} = M_{\mathcal{E}} + \tau N_{\mathcal{E}} \|M_{\mathcal{E}}\|_F / \|N_{\mathcal{E}}\|_F$, where τ is the noise factor. The minimization problem which we finally solve is the following:

$$\min \left\{ \frac{1}{2} \|X_{\mathcal{E}} - \widetilde{M}_{\mathcal{E}}\|_F^2 + \rho \|X\|_* \, : \, X \in \Re^{p\times q} \right\}. \tag{6.7}$$

In our numerical experiments, we set the noise level $\tau = 0$, 0.1 and the number of entries to sample $m = 10dr$, where $dr = r(p + q - r)$ is the value of degrees of freedom in an $p \times q$ matrix of rank r.

For each triplet $(p, q, r), m$ and τ, we repeat the above procedure 5 times. Table 6.6 presents the average numerical results of the partial PPA for solving the randomly generated matrix completion problem (6.7) over 5 runs. In the table, we report the number of sampled entries m, the average number of outer iterations, the average total number of inner iterations, the average number of CG steps taken to solve (4.20), the average infeasibilities in (3.16) and (3.20), respectively, the average relative gap between (3.16) and (3.20), the average relative mean square error MSE $:= \|X - M\|/\|M\|$ (where M is the original matrix), the mean value of the rank (#sv) of X, and the average CPU time taken. Here we report the numerical rank of X defined as follows:

$$\#\mathrm{sv}(X) := \max\{k \,:\, \sigma_k(X) \geq \max\{10^{-8}, \tau\}\sigma_1(X)\}. \tag{6.8}$$

We can observe that from Table 6.6 that the partial PPA is able to recover the original data rather accurately. In the numerical experiments in which the sampled entries are corrupted by 10% Gaussian noise, the errors (MSE) are all smaller than the noise factor ($\tau = 0.1$) in the given data. The errors are smaller than the theoretical result established in [18].

Example 6

We consider the positive semidefinite random matrix completion problem. For each pair (n, r), we generate a positive semidefinite matrix $M \in \mathcal{S}^n$ of rank r by setting $M = M_1 M_1^T$ where $M_1 \in \Re^{n\times r}$ is a random matrix with i.i.d Gaussian entries. Then we sample a subset $\mathcal{E}$ of m entries uniformly at random from the upper triangular part of M. The observed data is set to be $\widetilde{M}_{\mathcal{E}} = M_{\mathcal{E}} + \tau N_{\mathcal{E}}\|M_{\mathcal{E}}\|_F/\|N_{\mathcal{E}}\|_F$, where $N_{\mathcal{E}} \in \mathcal{S}^n$ is generated in a similar fashion as in Example 5 and τ is the noise factor. Then the minimization problem that we finally solve is given by

$$\min\Big\{\frac{1}{2}\|X_{\mathcal{E}} - \widetilde{M}_{\mathcal{E}}\|_F^2 + \rho\langle I,\, X\rangle \,:\, X \succeq 0\Big\}. \tag{6.9}$$

In our numerical experiments, we set the noise level $\tau = 0$, 0.1, and the number of entries to sample $m = 10dr$, where $dr = nr - r(r-1)/2$ is the value of degrees of

$p/q/\tau$	r	m	it.|itsub|cg	R_p | R_D | relgap	MSE	#sv	time
500/500/0.0	10	99189	11.4 | 16.6 | 7.0	5.36e-8 | 6.14e-7 | -9.85e-5	1.36e-3	10	19
	50	250000	10.0 | 10.0 | 4.0	2.71e-15 | 2.81e-7 | -2.77e-5	1.43e-3	50	12
	100	250000	9.4 | 9.4 | 4.0	3.06e-15 | 2.20e-7 | -1.36e-5	1.65e-3	100	13
1000/1000/0.0	10	199104	13.8 | 26.6 | 8.0	1.54e-7 | 6.18e-7 | -3.86e-5	1.36e-3	10	2:54
	50	974891	9.6 | 9.8 | 4.4	1.62e-8 | 2.93e-7 | -2.67e-5	1.32e-3	50	1:14
	100	1000000	10.0 | 10.0 | 4.0	3.05e-15 | 4.34e-7 | -3.07e-5	1.45e-3	100	1:15
1500/1500/0.0	10	299272	13.0 | 29.4 | 10.4	4.88e-7 | 1.96e-7 | -1.54e-5	1.36e-3	10	11:45
	50	1474562	11.0 | 14.0 | 6.9	1.86e-8 | 2.86e-7 | -2.08e-5	1.35e-3	50	5:47
	100	2250000	10.0 | 10.0 | 4.0	3.56e-15 | 2.93e-7 | -2.14e-5	1.37e-3	100	4:15
500/500/0.1	10	99189	22.0 | 43.0 | 7.1	1.97e-7 | 7.60e-7 | -4.47e-5	8.33e-2	10	50
	50	250000	11.0 | 11.0 | 4.0	3.12e-15 | 2.68e-7 | -1.04e-5	9.68e-2	50	15
	100	250000	13.2 | 13.2 | 4.0	3.28e-15 | 6.04e-7 | -2.13e-5	9.77e-2	100	21
1000/1000/0.1	10	199104	22.0 | 44.4 | 8.2	5.90e-7 | 7.55e-7 | -9.35e-5	7.78e-2	10	5:21
	50	974891	21.2 | 22.2 | 5.8	2.99e-7 | 6.67e-7 | -1.69e-6	9.51e-2	50	3:08
	100	1000000	14.0 | 14.0 | 4.0	3.46e-15 | 5.89e-7 | -1.91e-5	9.67e-2	100	2:14
1500/1500/0.1	10	299272	23.6 | 49.0 | 8.6	6.64e-7 | 5.44e-7 | -9.45e-5	7.54e-2	10	18:42
	50	1474562	22.0 | 39.8 | 8.2	5.47e-7 | 7.03e-7 | -2.63e-6	8.91e-2	50	17:39
	100	2250000	11.0 | 11.0 | 4.0	3.98e-15 | 2.41e-7 | -8.33e-6	9.60e-2	100	4:49
500/1000/0.0	10	149363	13.2 | 20.2 | 7.2	2.49e-8 | 6.24e-7 | -6.25e-5	1.39e-3	10	50
500/1000/0.1	10	149363	23.0 | 44.0 | 7.1	7.04e-7 | 6.40e-7 | -5.27e-5	7.97e-2	10	1:58
1000/2000/0.0	20	596592	12.0 | 18.0 | 7.9	5.22e-8 | 7.25e-7 | -9.31e-5	1.35e-3	20	5:15
1000/2000/0.1	20	596592	23.0 | 43.6 | 7.7	6.13e-7 | 8.06e-7 | -4.51e-5	7.94e-2	20	13:52
500/5000/0.0	25	1368353	9.0 | 17.0 | 9.8	4.22e-8 | 2.56e-7 | -3.26e-5	1.34e-3	25	9:12
500/5000/0.1	25	1368353	17.2 | 23.2 | 5.7	6.79e-7 | 7.68e-7 | -3.70e-5	7.99e-2	25	11:24

Table 6.6: Numerical results of the partial PPA for solving the randomly generated matrix completion problem (6.7). In the table, $m = 10dr$ and $dr = r(p + q - r)$.

freedom in an $n \times n$ matrix of rank r. For each pair $(n, r), m$ and τ, we repeat the above procedure 5 times.

n/τ	r	m	it.|itsub|cg	R_p | R_D | relgap	MSE	#sv	time
500/0.0	10	49637	12.4 | 17.4 | 6.1	3.34e-8 | 7.40e-7 | -1.02e-4	1.45e-3	10	17
	50	125250	9.0 | 9.0 | 3.2	2.47e-8 | 5.03e-7 | -3.93e-5	1.55e-3	50	9
	100	125250	9.0 | 9.0 | 3.3	4.65e-8 | 3.16e-7 | -6.18e-6	1.72e-3	100	9
1000/0.0	10	99359	14.2 | 23.4 | 7.0	2.14e-7 | 5.04e-7 | -6.00e-5	1.41e-3	10	1:20
	50	487724	10.0 | 10.0 | 3.5	3.46e-8 | 3.87e-7 | -3.18e-5	1.43e-3	50	49
	100	500500	9.0 | 9.0 | 3.2	1.24e-8 | 8.05e-7 | -4.42e-5	1.57e-3	100	45
1500/0.0	10	149545	14.8 | 25.6 | 8.9	3.88e-7 | 4.07e-7 | -9.77e-5	1.41e-3	10	4:52
	50	737608	11.0 | 14.0 | 5.9	9.58e-9 | 7.20e-7 | -4.56e-5	1.45e-3	50	2:40
	100	1125750	10.0 | 10.0 | 3.0	1.12e-9 | 2.33e-7 | -1.40e-5	1.48e-3	100	2:00
500/0.1	10	49637	25.4 | 45.0 | 9.5	2.49e-7 | 6.35e-7 | -1.12e-5	1.17e-1	10	43
	50	125250	9.0 | 11.0 | 5.2	9.46e-8 | 4.50e-7 | -6.71e-6	6.84e-2	50	12
	100	125250	9.0 | 11.6 | 6.0	5.24e-8 | 1.64e-7 | -1.26e-6	8.04e-2	100	13
1000/0.1	10	99359	25.4 | 48.2 | 10.6	7.65e-7 | 7.87e-7 | -4.78e-5	1.01e-1	10	3:22
	50	487724	25.8 | 27.2 | 6.4	3.49e-7 | 6.89e-7 | -8.02e-7	7.20e-2	50	2:26
	100	500500	9.2 | 10.4 | 5.7	7.75e-8 | 5.57e-7 | -4.79e-6	6.83e-2	100	1:04
1500/0.1	10	149545	25.0 | 47.8 | 11.4	8.39e-7 | 8.88e-7 | -1.05e-4	9.17e-2	10	9:51
	50	737608	25.6 | 32.8 | 9.8	4.91e-7 | 6.29e-7 | -1.03e-6	1.17e-1	50	9:14
	100	1125750	10.0 | 11.0 | 5.2	3.62e-8 | 3.29e-7 | -2.41e-6	6.33e-2	100	2:48

Table 6.7: Numerical results of the partial PPA on positive semidefinite random matrix completion problems. In the table, $m = 10dr$ and $dr = nr - r(r-1)/2$.

Table 6.7 presents the average numerical results of the partial PPA for solving (6.9) over 5 runs, where #sv is the numerical rank of X defined in (6.8). We can observe that from Table 6.7 that the partial PPA performed very well on randomly generated positive semidefinite matrix completion problems and it is able to recover the original data rather accurately.

Example 7

We consider matrix completion problems based on some real data sets including the Jester joke data set [46] and the MovieLens data set. The Jester joke data set contains 4.1 million ratings for 100 jokes from 73421 users and is available on the website `http://www.ieor.berkeley.edu/~goldberg/jester-data/`. The whole data is stored in three excel files with the following characteristics.

1. jester-1: 24983 users who have rated 36 or more jokes;
2. jester-2: 23500 users who have rated 36 or more jokes;
3. jester-3: 24938 users who have rated between 15 and 35 jokes.

For each data set, we let M be the original incomplete data matrix such that the i-th row of M corresponds to the ratings given by the i-th user on the jokes. For convenience, let Γ be the set of indices for which M_{ij} is given. We tested the jester joke data sets in the same way as in [121]. For each user, we randomly choose 10 ratings. Thus we select a subset Ω randomly from Γ. Since some of the entries in M are missing, we cannot compute the relative error of the estimated matrix X as we did for the randomly generated matrices. Instead, we computed the Normalized Mean Absolute Error (NMAE) as in [46]. The Mean Absolute Error (MAE) is defined as

$$\text{MAE} = \frac{1}{|\Gamma \backslash \Omega|} \sum_{(i,j)\in\Gamma\backslash\Omega} |M_{ij} - X_{ij}|, \tag{6.10}$$

where M_{ij} and X_{ij} are the original and computed ratings of joke j given by user i respectively. The normalized MAE is defined as

$$\text{NMAE} = \frac{\text{MAE}}{r_{\max} - r_{\min}}, \tag{6.11}$$

where $r_{\min}$ and $r_{\max}$ are lower and upper bounds for the ratings, For the jester joke data sets, all ratings are scaled to the range $[-10, 10]$, so we have $r_{\min} = -10, r_{\max} = 10$.

The MovieLens data set is from the GroupLens Research Group. This data set consists of 100,000 ratings on 1682 movies given by 943 users and is available on the website `http://www.grouplens.org`. Each user has rated at least 20 movies with scores from the range 1 to 5, and 5 is the highest score. In this data set we have $r_{\min} = 1, r_{\max} = 5$. For the MovieLens data sets, the matrices M is very sparse. In our experiments, we randomly select about 50% of the ratings given by each user, i.e., $|\Omega|/|\Gamma| = 50\%$.

In this example, we set `Tol` $= 10^{-5}$. We repeat the above precedure 5 times for each data set. Table 6.8 reports the average number of outer iterations, the average total number of inner iterations, the average number of CG steps taken to solve (4.20), the average infeasibilities in (3.16) and (3.20), respectively, the average relative gap between (3.16) and (3.20), the average NMAE value, the mean value of the numerical rank (#sv) of X defined by #sv := $\max\{k : \sigma_k(X) \geq 10^{-8}\sigma_1(X)\}$, and the average CPU time taken. We can observe from the table that the partial PPA performed very well on real matrix completion problems based on the jester joke and MovieLens data sets.

problem	p/q	N	$\|\Omega\|/N$	it.|itsub|cg	R_p | R_D | relgap	NMAE	#sv	time
jester-1	24983/100	1.81e+6	1.60e-1	18.0 | 24.0 | 5.9	3.69e-6 | 9.26e-6 | -5.57e-4	1.89e-1	99	9:56
jester-2	23500/100	1.71e+6	1.60e-1	18.0 | 24.2 | 6.2	4.05e-6 | 9.44e-6 | -5.63e-4	1.90e-1	98	9:43
jester-3	24938/100	6.17e+5	6.78e-1	21.4 | 37.6 | 20.0	1.99e-6 | 7.03e-6 | -1.69e-4	1.94e-1	71	40:14
jester-4	73421/100	4.14e+6	2.16e-1	18.0 | 23.6 | 5.6	4.24e-6 | 6.10e-6 | -3.45e-4	1.89e-1	100	27:08
movie	943/1682	1.00e+5	9.91e-1	23.0 |57.0 | 22.2	3.26e-6 | 5.50e-6 | -6.14e-4	2.05e-1	140	20:26

Table 6.8: Numerical results on the real matrix completion problems.

Example 8

In the Euclidean metric embedding problem, we are given an incomplete, possibly noisy, dissimilarity matrix $B \in \mathcal{S}^n$ with $\mathrm{Diag}(B) = 0$ and sparsity pattern specified by the set of indices $\mathcal{E}$. The goal is to find an Euclidean distance matrix [1] that is nearest to B. If the measure of nearness is in the Frobenius norm, then the mathematical formulation of the problem is as follows:

$$\min\Big\{\frac{1}{2}\sum_{(i,j)\in\mathcal{E}} W_{ij}(D_{ij}-B_{ij})^2+\frac{\rho}{2n}\langle E,\,D\rangle\ :\ D \text{ is an Euclidean distance matrix}\Big\}, \tag{6.12}$$

where $W_{ij}>0, (i,j)\in\mathcal{E}$, are given weights, $E\in\mathcal{S}^n$ is a matrix of all ones and ρ is a positive regularization parameter. Here we added the term $\frac{\rho}{2n}\langle E,\,D\rangle$ to encourage a sparse solution. Recall that a standard characterization [1] of an Euclidean distance matrix D is that $D=\mathrm{Diag}(X)e^T+e\,\mathrm{Diag}(X)^T-2X$ for some $X\succeq 0$ with $Xe=0$, where $e\in\Re^n$ is a vector of all ones. Thus the problem (6.12) can be rewritten as:

$$\min\Big\{\frac{1}{2}\sum_{(i,j)\in\mathcal{E}} W_{ij}(\langle A_{ij},\,X\rangle-B_{ij})^2+\rho\langle I,\,X\rangle\ :\ \langle E,\,X\rangle=0,\ X\succeq 0\Big\}, \tag{6.13}$$

where $A_{ij}=e_{ij}e_{ij}^T$ with $e_{ij}=e_i-e_j$. Note that under the condition $X\succeq 0$, the constraint $Xe=0$ is equivalent to $\langle E,\,X\rangle=0$. It is interesting to note that desiring sparsity in the Euclidean distance matrix D leads to the regularization term $\rho\langle I,\,X\rangle$, which is a proxy for desiring a low-rank X.

The Euclidean metric problem arises in many applications. For the regularized kernel estimation (RKE) problem in statistics [74], we are given a set of n objects and dissimilarity measures d_{ij} for certain object pairs $(i,j)\in\mathcal{E}$. The goal is to estimate a positive semidefinite kernel matrix $X\in\mathcal{S}^n_+$ such that the fitted squared distances between objects induced by X satisfy

$$X_{ii}+X_{jj}-2X_{ij}=\langle A_{ij},\,X\rangle\ \approx\ d_{ij}^2\quad \forall\,(i,j)\in\mathcal{E},$$

where $A_{ij}=(e_i-e_j)(e_i-e_j)^T$. Formally, one version of the RKE problem proposed in [74] is to solve the SDP problem (6.13).

In our numerical experiments, the data d_{ij} are normalized to be in the interval [0, 1], and $\mathcal{E}=\{(i,j): 1\le i<j,\, 1\le j\le 630\}$. We set $W_{ij}=1$ for all $(i,j)\in\mathcal{E}$ and `Tol` $=10^{-6}$. In [74], due to the computational difficulties encountered by the

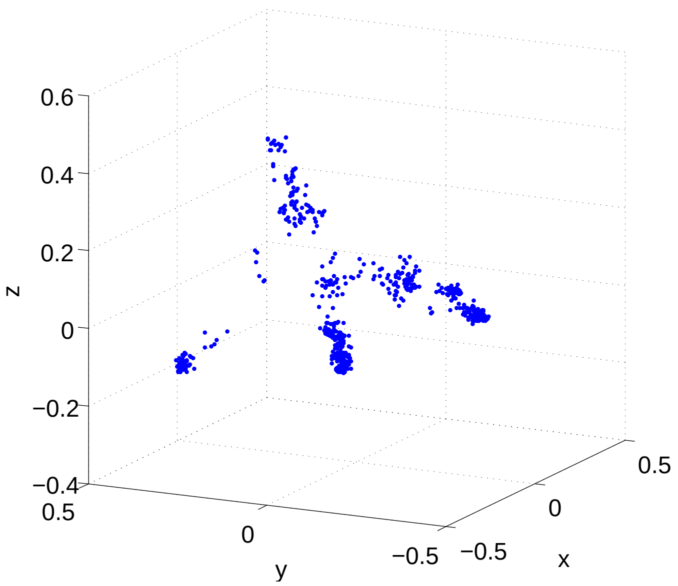

Figure 6.2: A 3D representation of the sequence space for 630 proteins.

interior-point method used to solve (6.13), a subset of 280 globin proteins were selected from the entire set of 630 proteins. And for each of the selected proteins, 55 dissimilarities were randomly selected out of the total of 280. Here we are able to consider the entire set of 630 proteins and the dissimilarities among all the pairs of proteins.

As mentioned in [74], the RKE methodology can provide an efficient way to represent each protein sequence by a feature vector in a chosen coordinate system using the pairwise dissimilarity between protein sequences, and the projection of the computed solution $\overline{X}$ on to a 3D space, which corresponds to the largest three eigenvalues, is quite informative. Figure 6.2 displays a 3D representation of the sequence space for 630 proteins from the globin family. There are at leat 4 classes visually identifiable in the data set of 630 proteins, which is consistent with the observations in [74]. The numerical results for solving (6.13) are reported in Table 6.9, where #sv is the number of positive eigenvalues of $\overline{X}$. For the obtained solution $\overline{X}$, we have $\langle \overline{X}, E\rangle = 1.09 \times 10^{-14}$ and $\langle \overline{X}, I\rangle = 1.85 \times 10^{2}$.

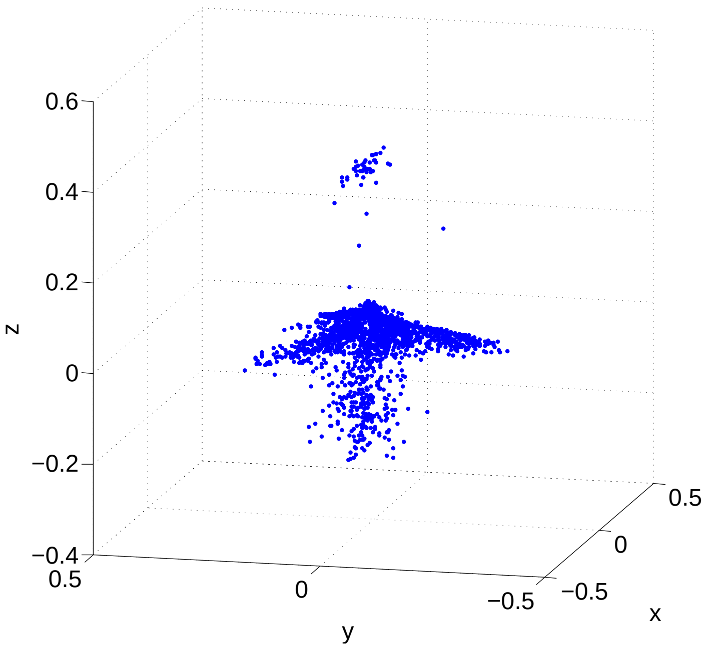

Figure 6.3: A 3D representation of the protein structure space for 1874 proteins.

problem	n	m	ρ	it.|itsub|cg	R_p | R_D | relgap	#sv	time
RKE630	630	198136	5.07e-1	6 | 36| 24.6	1.07e-7 | 2.42e-8 | -1.81e-6	388	1:59
PDB25	1898	1646031	1.84e+0	18 | 55 | 55.8	4.89e-7 | 4.78e-6 | -1.46e-5	1388	1:19:11

Table 6.9: Numerical results on the RKE problem arising from protein clustering.

We also conducted numerical experiments on a much larger protein data set to evaluate the performance of our algorithm. We used the `PDB_SELECT 25` data set, a representative subset of the Protein Data Bank database [8], which contains 1898 protein chains. In our numerical implementation, we set $\texttt{Tol} = 5 \times 10^{-6}$. Figure 6.3 displays a 3D representation of the structure space for 1898 proteins, which is consistent with the protein structure space studied in [58]. The numerical results for the `PDB_SELECT 25` data set are reported in Table 6.9. For the obtained solution $\overline{X}$, we have $\langle \overline{X}, E \rangle = 3.43 \times 10^{-14}$ and $\langle \overline{X}, I \rangle = 8.76 \times 10^{2}$.

6.2 Numerical Results for linearly constrained QSDP problems

In this section, we report the numerical performance of the inexact APG algorithm (Algorithm 4) for large scale linearly constrained QSDP problems.

We measure the infeasibilities for the primal and dual problems (1.12) and (1.13) as follows:

$$R_P = \frac{\|b - \mathcal{A}(X)\|}{1 + \|b\|}, \quad R_D = \frac{\|\mathcal{Q}(X) + C - \mathcal{A}^* p - Z\|}{1 + \|C\|}, \tag{6.14}$$

where X, p, Z are computed from (5.58). In our numerical experiments, we stop the inexact APG algorithm when

$$\max\{R_P, R_D\} \leq \texttt{Tol}, \tag{6.15}$$

where Tol is a pre-specified accuracy tolerance. Unless otherwise specified, we set $\texttt{Tol} = 10^{-6}$ as the default tolerance. When solving the subproblem (5.58) at iteration k of our inexact APG method, we stop the SSNCG or gradient method when $\|\nabla\theta(p_k)\|/(1 + \|b\|) < \min\{1/t_k^{3.1}, 0.2\|\nabla f(X_{k-1}) - \mathcal{A}^* p_{k-1} - Z_{k-1}\|/(1 + \|C\|)\}$.

Example 9

In this example, we consider the following H-weighted nearest correlation matrix problem

$$\min\Big\{\frac{1}{2}\|H \circ (X - G)\|^2 \mid \mathrm{Diag}(X) = e, X \succeq 0\Big\}. \tag{6.16}$$

We compare the performance of our inexact APG (IAPG) method and the augmented Lagrangian dual method (AL) studied by Qi and Sun in [98], whose MATLAB codes are available at `http://www.math.nus.edu.sg/∼matsundf`. We consider the gene correlation matrices $\widehat{G}$ from [69]. For testing purpose we perturb $\widehat{G}$ to

$$G := (1 - \alpha)\widehat{G} + \alpha E,$$

where $\alpha \in (0, 1)$ and E is a randomly generated symmetric matrix with entries in $[-1, 1]$. We also set $G_{ii} = 1, i = 1, \ldots, n$. The weight matrix H is a sparse random symmetric matrix with about 50% nonzero entries. The MATLAB code for generating H and E is as follows:

```
H = sprand(n,n,0.5); H = triu(H) + triu(H,1)'; H = (H + H')/2;
E = 2*rand(n,n)-1; E = triu(E) + triu(E,1)'; E = (E + E')/2.
```

In order to generate a good initial point, we use the SSNCG method in [97] to solve the following unweighted nearest correlation matrix problem

$$\min\left\{\frac{1}{2}\|X-G\|^2 \mid \operatorname{Diag}(X)=e, X\succeq 0\right\}. \tag{6.17}$$

Due to the difference in stopping criteria for different algorithms, we set different accuracy tolerances for the IAPG and augmented Lagrangian methods. For the IAPG method, we set `Tol` $=10^{-6}$. For the augmented Lagrangian method, its stopping criteria depends on a tolerance parameter `Tol1` which controls the three conditions in the KKT system (5.26). We set `Tol1` $=10^{-4}$.

Table 6.10 presents the numerical results obtained by the IAPG method and the augmented Lagrangian dual method (AL) for various instances of Example 1. We use the primal infeasibility, primal objective value and computing time to compare the performance of the two algorithms. For each instance in the table, we report the matrix dimension (n), the noise level (α), the number of outer iterations (iter), the total number of Newton systems solved (newt) the primal infeasibility (R_P), the dual infeasibility (R_D), the primal objective value (pobj) in (6.16), as well as the computation time (in the format hours:minutes:seconds) and the rank of the computed solution (sv). We may observe from the table that the IAPG method can solve (6.16) very efficiently. For each instance, the IAPG method can achieve nearly the same primal objective value as the augmented Lagrangian method, and the former can achieve much better primal infeasibility while taking less than 50% of the time needed by the augmented Lagrangian method.

Algo.	problem	n	α	iter/newt	R_P	R_D	pobj	time	sv
IAPG	Lymph	587	0.1	169/177	3.09e-11	1.75e-6	5.04289799e+0	2:35	179
			0.05	300/327	1.31e-10	2.04e-6	2.53103607e-1	4:16	207
AL	Lymph	587	0.1	12	4.13e-7	9.96e-7	5.04289558e+0	5:39	179
			0.05	12	2.96e-7	1.07e-6	2.53101698e-1	30:58	207
IAPG	ER	692	0.1	137/142	2.27e-10	2.43e-6	1.26095534e+1	3:10	189
			0.05	187/207	3.93e-11	9.54e-7	1.14555927e+0	3:40	220
AL	ER	692	0.1	12	3.73e-7	4.63e-7	1.26095561e+1	9:28	189
			0.05	12	3.21e-7	1.02e-6	1.14555886e+0	14:14	220
IAPG	Arabidopsis	834	0.1	115/123	3.28e-10	1.78e-6	3.46252363e+1	3:53	191
			0.05	131/148	2.41e-10	9.75e-7	5.50148194e+0	4:09	220
AL	Arabidopsis	834	0.1	13	2.28e-7	7.54e-7	3.46252429e+1	12:35	191
			0.05	12	2.96e-8	1.01e-6	5.50148169e+0	22:49	220
IAPG	Leukemia	1255	0.1	104/111	5.35e-10	7.97e-7	1.08939600e+2	9:24	254
			0.05	96/104	4.81e-10	9.31e-7	2.20789464e+1	8:35	276
AL	Leukemia	1255	0.1	12	3.06e-7	2.74e-7	1.08939601e+2	22:04	254
			0.05	11	2.90e-7	8.57e-7	2.20789454e+1	28:37	276
IAPG	hereditarybc	1869	0.1	67/87	2.96e-10	8.68e-7	4.57244497e+2	17:56	233
			0.05	64/85	9.58e-10	7.04e-7	1.13171325e+2	17:32	236
AL	hereditarybc	1869	0.1	13	2.31e-7	3.55e-7	4.57244525e+2	38:35	233
			0.05	11	2.51e-7	6.29e-7	1.13171335e+2	36:31	236

Table 6.10: Comparison of the inexact APG (IAPG) and augmented Lagrangian dual (AL) algorithms on (6.16) using sample correlation matrix from gene data sets. The weight matrix H is a sparse random matrix with about 50% nonzero entries.

Example 10

We consider the same problem as in Example 9, but the weight matrix H is generated from a weight matrix H_0 used by a hedge fund company. The matrix H_0 is a 93×93 symmetric matrix with all positive entries. It has about 24% of the entries equal to 10^{-5} and the rest are distributed in the interval $[2, 1.28 \times 10^3]$. It has 28 eigenvalues in the interval $[-520, -0.04]$, 11 eigenvalues in the interval $[-5 \times 10^{-13}, 2 \times 10^{-13}]$, and the rest of 54 eigenvalues in the interval $[10^{-4}, 2 \times 10^4]$. The MATLAB code for generating the matrix H is given by `tmp = kron(ones(25,25),H0); H = tmp([1:n],[1:n]); H = (H + H')/2.`

We use the same implementation techniques as in Example 9. The stopping tolerance for the IAPG method is set to `Tol` $= 10^{-6}$ while the tolerance for the augmented Lagrangian method is set to a less demanding value with `Tol1` $= 10^{-2}$. Table 6.11 presents the numerical results obtained by the IAPG and augmented Lagrangian dual (AL) methods. In the table, " $*$ " means that the augmented Lagrangian method cannot achieve the required tolerance of 10^{-2} in 24 hours. As we can see from Table 6.11, the IAPG method is much more efficient than the augmented Lagrangian method, and it can achieve much better primal infeasibility. For the last gene correlation matrix of size 1869, the IAPG method can find a good approximate solution within half an hour. For the augmented Lagrangian method, because the map $\mathcal{Q}$ associated with the weight matrix H is highly ill-conditioned, the CG method has great difficulty in solving the ill-conditioned linear system of equations obtained by the semismooth Newton method.

Example 11

In this example, we report the performance of the inexact APG on the linearly constrained QSDP problem (1.12). The linear operator $\mathcal{Q}$ is given by

$$\mathcal{Q}(X) = \frac{1}{2}(BX + XB)$$

Algo.	problem	n	α	iter/newt	R_P	R_D	pobj	time	sv
IAPG	Lymph	587	0.1	72/159	1.76e-8	9.90e-7	8.92431024e+6	1:50	238
			0.05	60/148	3.81e-8	9.75e-7	1.69947194e+6	1:41	278
AL	Lymph	587	0.1	14	2.64e-5	1.06e-5	8.92425480e+6	56:07	260
			0.05	12	1.69e-4	4.41e-5	1.69925778e+6	29:15	286
IAPG	ER	692	0.1	62/156	2.48e-9	9.72e-7	1.51144194e+7	2:33	254
			0.05	56/145	3.58e-9	9.55e-7	3.01128282e+6	2:22	295
AL	ER	692	0.1	16	1.22e-5	5.80e-6	1.51144456e+7	2:05:38	288
			0.05	12	3.11e-5	6.29e-6	3.01123631e+6	53:15	309
IAPG	Arabidopsis	834	0.1	61/159	6.75e-9	9.98e-7	2.69548461e+7	4:01	254
			0.05	54/145	1.06e-8	9.82e-7	5.87047119e+6	3:41	286
AL	Arabidopsis	834	0.1	19	3.04e-6	3.94e-6	2.69548769e+7	4:49:00	308
			0.05	13	1.69e-5	6.76e-6	5.87044318e+6	1:28:59	328
IAPG	Leukemia	1255	0.1	65/158	8.43e-9	9.86e-7	7.17192454e+7	11:32	321
			0.05	55/143	1.19e-7	9.80e-7	1.70092540e+7	10:18	340
AL	Leukemia	1255	0.1	*	*	*	*	*	*
			0.05	13	3.19e-5	5.15e-6	1.70091646e+7	5:55:21	432
IAPG	hereditarybc	1869	0.1	48/156	2.08e-8	9.16e-7	2.05907938e+8	29:07	294
			0.05	49/136	6.39e-8	9.61e-7	5.13121563e+7	26:16	297
AL	hereditarybc	1869	0.1	*	*	*	*	*	*
			0.05	*	*	*	*	*	*

Table 6.11: Same as Table 6.10, but with a “bad” weight matrix H.

n	m	cond(B)	iter/newt	R_P	R_D	pobj	dobj	time
500	500	9.21e+0	9/9	3.24e-10	9.70e-7	-4.09219187e+4	-4.09219188e+4	13
1000	1000	9.43e+0	9/9	3.68e-10	9.28e-7	-8.41240999e+4	-8.41241006e+4	1:13
2000	2000	9.28e+0	9/9	3.16e-10	8.53e-7	-1.65502323e+5	-1.65502325e+5	8:49
2500	2500	9.34e+0	9/9	3.32e-10	8.57e-7	-2.07906307e+5	-2.07906309e+5	16:15
3000	3000	9.34e+0	9/9	2.98e-10	8.13e-7	-2.49907743e+5	-2.49907745e+5	29:02

Table 6.12: Numerical results of the inexact APG algorithm on (1.12), where the positive definite matrix B for the linear operator $\mathcal{Q}$ is well-conditioned.

for a given $B \succ 0$, and the linear map $\mathcal{A}$ is given by $\mathcal{A}(X) = \mathrm{Diag}(X)$. We generate a positive definite matrix X and set $b = \mathcal{A}(X)$. Similarly we can generate a random vector $p \in \Re^m$ and a positive definite matrix Z and set $C = \mathcal{A}^*(p) + Z - \mathcal{Q}(X)$. The MATLAB code for generating the matrix B is given by `randvec = 1+ 9*rand(n,1); tmp = randn(n,ceil(n/4)); B = diag(randvec)+(tmp*tmp')/n; B = (B+B')/2`. Note that the matrix B generated is rather well conditioned.

As discussed in section 5.2, we are able to find a good symmetrized Kronecker product approximation $W \circledast W$ of $\mathcal{Q}$. By noting that

$$\frac{1}{2}\langle X, W \circledast W(X)\rangle + \langle C, X\rangle = \frac{1}{2}\|W^{1/2}(X-U)W^{1/2}\|^2 - \frac{1}{2}\|W^{-1/2}CW^{-1/2}\|^2,$$

where $U = -W^{-1}CW^{-1}$, and dropping the constant term, we propose to solve the following problem to generate a good initial point for the inexact APG method:

$$\min\left\{\frac{1}{2}\|W^{1/2}(X-U)W^{1/2}\|^2 \;\mid\; \mathcal{A}(X) = b,\; X \succeq 0\right\},$$

which can be efficiently solved by the the SSNCG method in [97].

The performance results of our IAPG method on convex QSDP problems are given in Table 6.12, where "pobj" and "dobj" are the primal and dual objective values for QSDP, respectively. We may see from the table that the IAPG method can solve all the five instances of QSDP problems very efficiently with very good primal infeasibility.

n	m	cond(B)	iter/newt	R_P	R_D	pobj	dobj	time
500	10000	2.67e+5	51/102	3.02e-8	9.79e-7	-9.19583895e+3	-9.19584894e+3	1:29
1000	50000	1.07e+6	62/115	2.43e-8	9.71e-7	-1.74777588e+4	-1.74776690e+4	11:46
2000	100000	4.32e+6	76/94	5.24e-9	5.28e-7	-3.78101950e+4	-3.78101705e+4	1:14:04
2500	100000	6.76e+6	80/96	4.62e-9	5.64e-7	-4.79637904e+4	-4.79637879e+4	2:11:01

Table 6.13: Same as Table 6.12, but the matrix B for the linear operator $\mathcal{Q}$ is ill-conditioned and the linear map $\mathcal{A}$ is randomly generated as in [79].

Example 12

We consider the same problem as Example 11 but the linear map $\mathcal{A}$ is generated by using the first generator in [79] with order $p = 3$. The positive definite matrix B is generated by using MATLAB's built-in function: `B = gallery('lehmer',n)`. The condition number cond(B) of the generated Lehmer matrix B is within the range $[n, 4n^2]$. For this example, the simple choice of $W = \sqrt{\lambda_{\max}(B)}I$ in the symmetrized Kronecker product $W \circledast W$ for approximating $\mathcal{Q}$ does not work well. In our numerical implementation, we employ the strategy described in section 3.2 to find a suitable W.

Table 6.13 presents the numerical results of our IAPG method on convex QSDP problems where the matrix B is very ill-conditioned. As observed from the table, the condition numbers of B are large. We may see from the table that the IAPG method can solve the problem very efficiently with very accurate approximate optimal solution.

Chapter 7

Conclusions

In this thesis, we designed algorithms for solving large scale nuclear norm minimization and convex quadratic semidefinite programming (QSDP) problems. We introduced a partial proximal point algorithm for solving nuclear norm regularized matrix least squares problems with linear equality and inequality constraints. Based on the results of the general partial proximal point algorithm, we analyzed the global and local convergence of our proposed algorithm. The inner subproblems, due to the presence of inequality constraints, were reformulated as a system of semismooth equations, which are solved by an inexact smoothing Newton method. The quadratic convergence of the inexact smoothing Newton method was proved under the constraint nondegeneracy condition, together with the strong semismoothness property of the soft thresholding operator. When the nuclear norm regularized matrix least squares problem has equality constraints only, we proposed a semismooth Newton-CG method to solve the unconstrained inner subproblem in each iteration. The quadratic convergence of the semismooth Newton-CG method was also established.

In order to efficiently solve large scale convex QSDP problems, we extended the APG algorithm to the inexact setting where the subproblem in each iteration was only solved approximately. We showed that the inexact APG enjoys the same superior worst-case iteration complexity as the exact version. Numerical experiments conducted on a variety of large scale nuclear norm minimization and convex QSDP problems demonstrated that our proposed algorithms are very efficient and robust.

There are still many interesting problems that will lead to further development of algorithms for solving large scale structured matrix optimization problems. The current theorectical guarantees of using the nuclear norm $\|\cdot\|_*$ as a surrogate for the rank function for matrix completion problems require that the entries of the matrix are uniformly sampled [19, 20, 102]. However, if the entries of the matrix are non-uniformly sampled, the nuclear norm regularizer may perform very poorly for matrix completion problems [110]. A weighted nuclear norm function, i.e., $\|W_1(\cdot)W_2\|_*$, where $W_1 \in \Re^{p\times q}$ and $W_2 \in \Re^{p\times q}$ are given weight matrices, was suggested in [110] to deal with matrix completion problems with non-uniformly sampled entries. It will be worthwhile to develop an efficient and robust algorithm for solving weighted nuclear norm regularized matrix least squares problems. In many applications [29, 93, 128], based on the available prior information about the target matrix, we may use the operator norm or the Ky Fan k-norm which is defined as the sum of k largest singular values as a regularizer for obtaining certain desired structures.

Based on the high efficiency and robustness of the inexact APG algorithm for large scale convex QSDP problems, it will be very attractive to design an inexact APG algorithm for solving (weighted) nuclear norm regularized matrix least squares problems with equality and inequality constraints and second order cone constraints.

Bibliography

[1] A. Alfakih, A. Khandani, and H. Wolkowicz, *Solving euclidean distance matrix completion problems via semidefinite programming*, Computational optimization and applications, 12 (1999), pp. 13–30.

[2] B. Ames and S. Vavasis, *Nuclear norm minimization for the planted clique and biclique problems*, Mathematical Programming, (2009), pp. 1–21.

[3] Z. Bai and R. Freund, *A partial padé-via-lanczos method for reduced-order modeling*, Linear Algebra and its Applications, 332 (2001), pp. 139–164.

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

[5] M. Bee, *Estimating rating transition probabilites with missing data*, Statistical Methods and Applications, 14 (2005), pp. 127–141.

[6] A. Benczúr, K. Csalogány, and T. Sarlós, *On the feasibility of low-rank approximation for personalized pagerank*, in Special interest tracks and posters of the 14th international conference on World Wide Web, ACM, 2005, pp. 972–973.

[7] A. Berman and R. Plemmons, *Nonnegative matrices in the mathematical sciences*, Classics in Applied Mathematics. SIAM, Philadelphia, 1994.

[8] H. Berman, T. Battistuz, T. Bhat, W. Bluhm, P. Bourne, K. Burkhardt, Z. Feng, G. Gilliland, L. Iype, S. Jain, P. Fagan, J. Marvin, D. Padilla, V. Ravichandran, B. Schneider, N. Thanki, H. Weissig, J. Westbrook, and C. Zardecki, *The protein data bank*, Acta Crystallographica Section D: Biological Crystallography, 58 (2002), pp. 899–907.

[9] D. Bertsekas, *Constrained optimization and lagrange multiplier methods*, Computer Science and Applied Mathematics, Boston: Academic Press, 1982, 1 (1982).

[10] ——, *Nonlinear programming*, Athena Scientific Belmont, MA, 1999.

[11] D. Bertsekas and P. Tseng, *Partial proximal minimization algorithms for convex programming*, SIAM Journal on Optimization, 4 (1994), pp. 551–572.

[12] R. Bhatia, *Matrix analysis*, Springer Verlag, NewYork, 1997.

[13] P. Bonacich, *Factoring and weighting approaches to status scores and clique identification*, Journal of Mathematical Sociology, 2 (1972), pp. 113–120.

[14] J. Bonnans and A. Shapiro, *Perturbation analysis of optimization problems*, Springer Verlag, 2000.

[15] S. Burer and R. Monteiro, *A nonlinear programming algorithm for solving semidefinite programs via low-rank factorization*, Mathematical Programming, 95 (2003), pp. 329–357.

[16] ——, *Local minima and convergence in low-rank semidefinite programming*, Mathematical Programming, 103 (2005), pp. 427–444.

[17] J. Cai, E. Candes, and Z. Shen, *A singular value thresholding algorithm for matrix completion*, SIAM Journal on Optimization, 20 (2010), pp. 1956–1982.

[18] E. Candes and Y. Plan, *Matrix completion with noise*, Proceedings of the IEEE, 98 (2010), pp. 925–936.

[19] E. Candès and B. Recht, *Exact matrix completion via convex optimization*, Foundations of Computational Mathematics, 9 (2009), pp. 717–772.

[20] E. Candès, J. Romberg, and T. Tao, *Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information*, Information Theory, IEEE Transactions on, 52 (2006), pp. 489–509.

[21] E. Candes and T. Tao, *Near-optimal signal recovery from random projections: Universal encoding strategies?*, Information Theory, IEEE Transactions on, 52 (2006), pp. 5406–5425.

[22] V. Chandrasekaran, B. Recht, P. Parrilo, and A. Willsky, *The convex geometry of linear inverse problems.* preprint, 2010.

[23] S. Chen, D. Donoho, and M. Saunders, *Atomic decomposition by basis pursuit*, SIAM journal on scientific computing, 20 (1999), pp. 33–61.

[24] S. Chennubhotla, *Spectral methods for multi-scale feature extraction and data clustering*, PhD thesis, University of Toronto, 2004.

[25] M. T. Chu, R. E. Funderlic, and R. J. Plemmons, *Structured low rank approximation*, Linear algebra and its applications, 366 (2003), pp. 157–172.

[26] F. Clarke, *Optimization and Nonsmooth Analysis*, John Wiley and Sons, New York, 1983.

[27] R. Corless, P. Gianni, B. Trager, and S. Watt, *The singular value decomposition for polynomial systems*, in Proceedings of the International Symposium on Symbolic and Algebraic Computation, Montreal, Canada, 1995, pp. 195–207.

[28] R. Cottle, J. Pang, and R. Stone, *The linear complementarity problem*, Academic Press, Boston, 1992.

[29] A. d'Aspremont, *Subsampling algorithms for semidefinite programming.* Preprint available at `http://www.cmap.polytechnique.fr/~aspremon/subsampsdp.html`.

[30] O. Devolder, F. Glineur, and Y. Nesterov, *First-order methods of smooth convex optimization with inexact oracle*, CORE Discussion Papers, Université catholique de Louvain, 2011.

[31] C. Ding, D. Sun, and K. Toh, *An introduction to a class of matrix cone programming*, Technical Report, Department of Mathematics, National University of Singapore, 2010.

[32] V. Dolezal, *Monotone operators and applications in Control and Network Theory*, vol. 2, Elsevier Scientific Pub. Co.,(Amsterdam and New York and New York), 1979.

[33] D. Donoho, *Compressed sensing*, Information Theory, IEEE Transactions on, 52 (2006), pp. 1289–1306.

[34] B. Eaves, *On the basic theorem of complementarity*, Mathematical Programming, 1 (1971), pp. 68–75.

[35] C. Eckart and G. Young, *The approximation of one matrix by another of lower rank*, Psychometrika, 1 (1936), pp. 211–218.

[36] M. Fazel, *Matrix rank minimization with applications*, PhD thesis, Stanford University, 2002.

[37] M. Fazel, H. Hindi, and S. Boyd, *A rank minimization heuristic with application to minimum order system approximation*, in American Control Conference, 2001. Proceedings of the 2001, vol. 6, 2001, pp. 4734–4739.

[38] A. Fischer, *Solution of monotone complementarity problems with locally lipschitzian functions*, Mathematical Programming, 76 (1997), pp. 513–532.

[39] F. Frobenius, *Über Matrizen aus nicht negativen Elementen*, Königliche Akademie der Wissenschaften, 1912.

[40] D. Gabay and B. Mercier, *A dual algorithm for the solution of nonlinear variational problems via finite element approximation*, Computers & Mathematics with Applications, 2 (1976), pp. 17–40.

[41] K. Gabriel and S. Zamir, *Lower rank approximation of matrices by least squares with any choice of weights*, Technometrics, (1979), pp. 489–498.

[42] Y. Gao and D. Sun, *Calibrating least squares covariance matrix problems with equality and inequality constraints*, SIAM Journal on Matrix Analysis and Applications, 31 (2009), pp. 1432–1457.

[43] Y. Gao and D. Sun, *A majorized penalty approach for calibrating rank constrained correlation matrix problems*, Technical Report, Department of Mathematics, National University of Singapore, 2010.

[44] L. E. Ghaoui and P. Gahinet, *Rank minimization under lmi constraints: A framework for output feedback problems*, in Proceedings of the European Control Conference, 1993.

[45] R. Glowinski and A. Marrocco, *Sur l'approximation, par éléments finis d'ordre un, et la résolution, par pénalisation-dualité, d'une classe de problèmes de Dirichlet non linéaires*, vol. 2, LABORIA, 1975.

[46] K. Goldberg, T. Roeder, D. Gupta, and C. Perkins, *Eigentaste: A constant time collaborative filtering algorithm*, Information Retrieval, 4 (2001), pp. 133–151.

[47] A. Goldstein, *Convex programming in hilbert space*, Bull. Amer. Math. Soc, 70 (1964), pp. 709–710.

[48] G. Golub, A. Hoffman, and G. Stewart, *A generalization of the eckart-young-mirsky matrix approximation theorem*, Linear Algebra and Its Applications, 88 (1987), pp. 317–327.

[49] G. H. Golub and C. F. Van Loan, *Matrix computations*, Johns Hopkins Studies in the Mathematical Sciences, Johns Hopkins University Press, Baltimore, MD, third ed., 1996.

[50] T. Graepel, *Kernel matrix completion by semidefinite programming*, in Proceedings of the International Conference on Artificial Neural Networks, J. R. Dorronsoro, ed., Springer-Verlag, 2002, pp. 694–699.

[51] D. Gross, *Recovering low-rank matrices from few coefficients in any basis*, IEEE Transactions on Information Theory, 57 (2011), pp. 1548–1566.

[52] C. D. Ha, *A generalization of the proximal point algorithm*, SIAM Journal on Control and Optimization, 28 (1990), pp. 503–512.

[53] E. T. Hale, W. Yin, and Y. Zhang, *A fixed-point continuation method for l1-regularized minimization with applications to compressed sensing*, Technical Report, Department of Computational and Applied Mathematics, Rice University, 2007.

[54] M. R. Hestenes and E. Stiefel, *Methods of conjugate gradients for solving linear systems*, Journal of Research of the National Bureau of Standards, 49 (1952), pp. 409–436.

[55] N. J. Higham, *Computing the nearest correlation matrix—a problem from finance*, IMA J. Numer. Anal., 22 (2002), pp. 329–343.

[56] N. D. Ho and P. Van Dooren, *Non-negative matrix factorization with fixed row and column sums*, Linear Algebra and its Applications, 429 (2008), pp. 1020–1025.

[57] R. A. Horn and C. R. Johnson, *Topics in matrix analysis*, Cambridge University Press, Cambridge, 1994.

[58] J. Hou, S. R. Jun, C. Zhang, and S. H. Kim, *Global mapping of the protein structure space and application in structure-based inference of protein function*, Proceedings of the National Academy of Sciences of the United States of America, 102 (2005), pp. 3651–3656.

[59] Y. T. Hu, R. Kiesel, and W. Perraudin, *The estimation of transition matrices for sovereign credit ratings*, Journal of Banking & Finance, 26 (2002), pp. 1383–1406.

[60] S. Ibaraki and M. Fukushima, *Partial proximal method of multipliers for convex programming problems*, Journal of the Operations Research Society of Japan-Keiei Kagaku, 39 (1996), pp. 213–229.

[61] S. Ji and J. Ye, *An accelerated gradient method for trace norm minimization*, in Proceedings of the 26th Annual International Conference on Machine Learning, ACM, 2009, pp. 457–464.

[62] N. K. Karmarkar and Y. N. Lakshman, *On approximate gcds of univariate polynomials*, Journal of Symbolic Computation, 26 (1998), pp. 653–666.

[63] R. H. Keshavan, A. Montanari, and S. Oh, *Matrix completion from a few entries*, IEEE Transactions on Information Theory, 56 (2010), pp. 2980–2998.

[64] B. Kummer, *Newton's method for non-differentiable functions*, Mathematical research, 45 (1988), pp. 114–125.

[65] P. Lawrence, B. Sergey, M. Rajeev, and W. Terry, *The pagerank citation ranking: Bringing order to the web*, tech. report, Technical Report, Stanford Digital Library Technologies Project, 1999.

[66] D. D. Lee and H. S. Seung, *Learning the parts of objects by non-negative matrix factorization*, Nature, 401 (1999), pp. 788–791.

[67] E. S. Levitin and B. T. Polyak, *Constrained minimization problems*, USSR Computational Mathematics and Mathematical Physics, 6 (1966), p. 50.

[68] H. Li, P. Stoica, and J. Li, *Computationally efficient maximum likelihood estimation of structured covariance matrices*, IEEE Transactions on Signal Processing, 47 (1999), pp. 1314–1323.

[69] L. Li and K. C. Toh, *An inexact interior point method for l 1-regularized sparse covariance selection*, Mathematical Programming Computation, (2010), pp. 1–25.

[70] J. Lin, *Reduced rank approximations of transition matrices*, in Proceedings of the Ninth International Workshop on Artificial Intelligence and Statistics, 2003.

[71] Y. J. Liu, D. Sun, and K. C. Toh, *An implementable proximal point algorithmic framework for nuclear norm minimization*, Mathematical Programming, (2009), pp. 1–38.

[72] Z. Liu and L. Vandenberghe, *Interior-point method for nuclear norm approximation with application to system identification*, SIAM Journal on Matrix Analysis and Applications, 31 (2009), pp. 1235–1256.

[73] K. Löwner, *Über monotone matrixfunktionen*, Mathematische Zeitschrift, 38 (1934), pp. 177–216.

[74] F. Lu, S. Keleş, S. J. Wright, and G. Wahba, *Framework for kernel regularization with application to protein clustering*, Proceedings of the National Academy of Sciences of the United States of America, 102 (2005), pp. 12332–12337.

[75] Z. Lu, R. D. C. Monteiro, and M. Yuan, *Convex optimization methods for dimension reduction and coefficient estimation in multivariate linear regression*, Mathematical Programming, 123 (2010).

[76] Y. Luo, *A smoothing newton-bicgstab method for least squares matrix nuclear norm problems*, master's thesis, Master thesis, Department of Mathematics, National University of Singapore, 2010.

[77] S. Ma, D. Goldfarb, and L. Chen, *Fixed point and bregman iterative methods for matrix rank minimization*, Mathematical Programming, (2009), pp. 1–33.

[78] J. Malick, *A dual approach to semidefinite least-squares problems*, SIAM Journal on Matrix Analysis and Applications, 26 (2004), pp. 272–284.

[79] J. Malick, J. Povh, F. Rendl, and A. Wiegele, *Regularization methods for semidefinite programming*, SIAM Journal on Optimization, 20 (2009), pp. 336–356.

[80] B. Martinet, *Régularisation d'inéquations variationnelles par approximations successives*, Rev. Française Informat. Recherche Opérationnelle, 4 (1970), pp. 154–158.

[81] F. Meng, D. Sun, and G. Zhao, *Semismoothness of solutions to generalized equations and the moreau-yosida regularization*, Mathematical programming, 104 (2005), pp. 561–581.

[82] M. Mesbahi and G. Papavassilopoulos, *On the rank minimization problem over a positive semidefinite linear matrix inequality*, IEEE Transactions on Automatic Control, 42 (1997), pp. 239–243.

[83] R. Mifflin, *Semismooth and semiconvex functions in constrained optimization*, SIAM Journal on Control and Optimization, 15 (1977), pp. 959–972.

[84] G. J. Minty, *Monotone (nonlinear) operators in hilbert space*, Duke Mathematical Journal, 29 (1962), pp. 341–346.

[85] L. Mirsky, *Symmetric gauge functions and unitarily invariant norms*, The Quarterly Journal of Mathematics, 11 (1960), pp. 50–59.

[86] J. Moré and Z. Wu, *Issues in large scale global molecular optimization*, Large-Scale Optimization with Applications, (1997), pp. 99–121.

[87] J. J. Moré and Z. Wu, *Distance geometry optimization for protein structures*, Journal of Global Optimization, 15 (1999), pp. 219–234.

[88] J. J. Moreau, *Proximité et dualité dans un espace hilbertien*, Bulletin de la Societe Mathematique de France, 93 (1965), pp. 273–299.

[89] A. S. Nemirovsky and D. B. Yudin, *Problem complexity and method efficiency in optimization.*, Wiley, New York, 1983.

[90] Y. Nesterov, *A method of solving a convex programming problem with convergence rate* $O(1/k^2)$, Soviet Mathematics Doklady, 27 (1983), pp. 372–376.

[91] ——, *Introductory Lectures on Convex Optimization: A Basic Course*, Kluwer Academic Publishers, Boston, MA, 2004.

[92] J. Nocedal and S. J. Wright, *Numerical Optimization*, Springer verlag, 1999.

[93] M. L. Overton and R. S. Womersley, *Optimality conditions and duality theory for minimizing sums of the largest eigenvalues of symmetric matrices*, Mathematical Programming, 62 (1993), pp. 321–357.

[94] J. S. Pang, D. Sun, and J. Sun, *Semismooth homeomorphisms and strong stability of semidefinite and lorentz complementarity problems*, Mathematics of Operations Research, (2003), pp. 39–63.

[95] O. Perron, *Zur theorie der matrices*, Mathematische Annalen, 64 (1907), pp. 248–263.

[96] T. K. Pong, P. Tseng, S. Ji, and J. Ye, *Trace norm regularization: Reformulations, algorithms, and multi-task learning*, SIAM Journal on Optimization, 20 (2010), pp. 3465–3489.

[97] H. Qi and D. Sun, *A quadratically convergent newton method for computing the nearest correlation matrix*, SIAM Journal on Matrix Analysis and Applications, 28 (2007), pp. 360–385.

[98] ——, *An augmented lagrangian dual approach for the h-weighted nearest correlation matrix problem*, IMA Journal of Numerical Analysis, 31 (2011), pp. 491–511.

[99] L. Qi, D. Sun, and G. Zhou, *A new look at smoothing newton methods for nonlinear complementarity problems and box constrained variational inequalities*, Mathematical Programming, 87 (2000), pp. 1–35.

[100] L. Qi and J. Sun, *A nonsmooth version of newton's method*, Mathematical Programming, 58 (1993), pp. 353–367.

[101] B. Recht, *A simpler approach to matrix completion.* to appear in Journal of Machine Learning Research. Preprint available at `http://pages.cs.wisc.edu/~recht/publications.html`.

[102] B. Recht, M. Fazel, and P. Parrilo, *Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization*, SIAM Review, 52 (2010), pp. 471–501.

[103] S. M. Robinson, *Local structure of feasible sets in nonlinear programming, part ii: Nondegeneracy*, Mathematical Programming Study, 22 (1984), pp. 217–230.

[104] R. T. Rockafellar, *Convex analysis*, Princeton University Press, Princeton, 1970.

[105] ——, *On the maximality of sums of nonlinear monotone operators*, Transactions of the American Mathematical Society, 149 (1970), pp. 75–88.

[106] ——, *Conjugate duality and optimization*, vol. 16, Regional Conferences Series in Applied Mathematics, SIAM, Philadelphia, 1974.

[107] ——, *Augmented lagrangians and applications of the proximal point algorithm in convex programming*, Mathematics of Operations Research, 1 (1976), pp. 97–116.

[108] ——, *Monotone operators and the proximal point algorithm*, SIAM Journal on Control and Optimization, 14 (1976), pp. 877–898.

[109] R. T. Rockafellar and J. B. W. Roger, *Variational Analysis*, Springer-Verlag, New York, 1998.

[110] R. Salakhutdinov and N. Srebro, *Collaborative filtering in a non-uniform world: Learning with the weighted trace norm*, in Advances in Neural Information Processing Systems 23, J. Lafferty, C. K. I. Williams, J. Shawe-Taylor, R. Zemel, and A. Culotta, eds., The MIT Press, 2010, pp. 2056–2064.

[111] T. Schulz, *Penalized maximum-likelihood estimation of covariance matrices with linear structure*, IEEE Transactions on Signal Processing, 45 (1997), pp. 3027–3038.

[112] A. K. Shaw and R. Kumaresan, *Some structured matrix approximation problems*, in IEEE International Conference on Acoustics, Speech, and Signal Processing, 1988, pp. 2324–2327.

[113] R. Sinkhorn and P. Knopp, *Concerning nonnegative matrices and doubly stochastic matrices*, Pacific Journal of Mathematics, 21 (1967), pp. 343–348.

[114] J. F. Sturm, *Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones*, Optimization Methods and Software, 11/12 (1999), pp. 625–653.

[115] D. Sun and J. Sun, *Semismooth matrix-valued functions*, Mathematics of Operations Research, 27 (2002), pp. 150–169.

[116] J. Sun, D. Sun, and L. Qi, *A squared smoothing newton method for nonsmooth matrix equations and its applications in semidefinite optimization problems*, SIAM Journal on Optimization, 14 (2004), pp. 783–806.

[117] J. Sun and S. Zhang, *A modified alternating direction method for convex quadratically constrained quadratic semidefinite programs*, European Journal of Operational Research, 207 (2010), pp. 1210–1220.

[118] K. C. Toh, *An inexact primal–dual path following algorithm for convex quadratic sdp*, Mathematical Programming, 112 (2008), pp. 221–254.

[119] K. C. Toh, M. J. Todd, and R. H. Tütüncü, *SDPT3—a MATLAB software package for semidefinite programming, version 1.3*, Optimization Methods and Software, 11 (1999), pp. 545–581.

[120] K. C. Toh, R. H. Tütüncü, and M. J. Todd, *Inexact primal-dual path-following algorithms for a special class of convex quadratic sdp and related problems*, Pacific Journal of Optimization, 3 (2007), pp. 135–164.

[121] K. C. Toh and S. Yun, *An accelerated proximal gradient algorithm for nuclear norm regularized linear least squares problems*, Pacific Journal of Optimization, 6 (2010), pp. 615–640.

[122] M. W. Trosset, *Distance matrix completion by numerical optimization*, Computational Optimization and Applications, 17 (2000), pp. 11–22.

[123] P. Tseng, *On accelerated proximal gradient methods for convex-concave optimization*, submitted to SIAM Journal on Optimization, (2008).

[124] H. A. Van der Vorst, *Bi-cgstab: A fast and smoothly converging variant of bi-cg for the solution of nonsymmetric linear systems*, SIAM Journal on scientific and Statistical Computing, 13 (1992), p. 631.

[125] C. Wang, D. Sun, and K. C. Toh, *Solving log-determinant optimization problems by a newton-cg proximal point algorithm*, SIAM Journal on Optimization, 20 (2010), pp. 2994–3013.

[126] J. Warga, *Fat homeomorphisms and unbounded derivate containers*, Journal of Mathematical Analysis and Applications, 81 (1981), pp. 545–560.

[127] G. A. Watson, *Characterization of the subdifferential of some matrix norms*, Linear algebra and its applications, 170 (1992), pp. 33–45.

[128] ——, *On matrix approximation problems with ky fan k norms*, Numerical Algorithms, 5 (1993), pp. 263–272.

[129] M. Wax and T. Kailath, *Detection of signals by information theoretic criteria*, IEEE Transactions on Acoustics, Speech and Signal Processing, 33 (1985), pp. 387–392.

[130] K. G. Woodgate, *Efficient stiffness matrix estimation for elastic structures*, Computers and structures, 69 (1998), pp. 79–84.

[131] Z. Yang, *A study on nonsymmetric matrix-valued functions*, master's thesis, Department of Mathematics, National University of Singapore, 2009.

[132] K. Yosida, *Functional Analysis*, Springer-Verlag, Berlin, 1964.

[133] M. Yuan, A. Ekici, Z. Lu, and R. Monteiro, *Dimension reduction and coefficient estimation in multivariate linear regression*, Journal of the Royal Statistical Society: Series B (Statistical Methodology), 69 (2007), pp. 329–346.

[134] E. H. Zarantonello, *Projections on convex sets in hilbert space and spectral theory*, Contributions to Nonlinear Functional Analysis (E.H. Zarantonello, ed.), (1971), pp. 237–424.

[135] S. Zhang, *On alternating direction methods for monotropic semidefinite programming*, PhD thesis, PhD thesis, Department of Mathematics, National University of Singapore, 2009.

[136] S. Zhang, J. Ang, and J. Sun, *An alternating direction method for solving convex nonlinear semidefinite programming problems*, to appear in Optimization, (2011).

[137] X. Y. Zhao, *A semismooth Newton-CG augmented Lagrangian method for large scale linear and convex quadratic SDPs*, PhD thesis, PhD thesis, Department of Mathematics, National University of Singapore, 2009.

[138] X. Y. Zhao, D. Sun, and K. C. Toh, *A newton-cg augmented lagrangian method for semidefinite programming*, SIAM Journal on Optimization, 20 (2010), pp. 1737–1765.

[139] K. Zię̧tak, *On the characterization of the extremal points of the unit sphere of matrices*, Linear Algebra and its Applications, 106 (1988), pp. 57–75.

ALGORITHMS FOR LARGE SCALE NUCLEAR NORM MINIMIZATION AND CONVEX QUADRATIC SEMIDEFINITE PROGRAMMING PROBLEMS

JIANG KAIFENG

NATIONAL UNIVERSITY OF SINGAPORE

2011

Algorithms for large scale nuclear norm minimization and convex quadratic semidefinite programming problems **Jiang Kaifeng**

2011