分类号 ________________　　密级 ________________

UDC ________________

学 位 论 文

求解一类矩阵范数逼近问题的数值算法

（题名和副题名）

陈彩华

（作者姓名）

指导教师姓名、职务、职称、学位、单位名称及地址何炳生 教授，孙德锋 教授，卓

南京大学 数学系

申请学位级别　博士　专业名称　计算数学

论文提交日期 2012 年 6 月 17 日 论文答辩日期 2012 年 7 月 2 日

学位授予单位和日期 ________________

答辩委员会主席：________________

评阅人：________________

____

____

____

2012 年 6 月 17 日

南京大學

研 究 生 毕 业 论 文
（申请博士学位）

论 文 题 目 求解一类矩阵范数逼近问题的数值算法

作 者 姓 名 陈彩华

学科、专业方向 计算数学

指 导 教 师 何炳生 教授，孙德锋 教授，卓金权 教授

研 究 方 向 最优化理论与算法

2012 年 6 月

学　　　　号　：DG0921001

论文答辩日期　：2012 年 7 月 2 日

指 导 教 师　：　　　　　（签字）

Numerical Algorithms for a Class of Matrix Norm Approximation Problems

by

Caihua Chen

Directed by

Professor Bingsheng He	(Nanjing University)
Professor Defeng Sun	(National University of Singapore)
Professor Kim-Chuan Toh	(National University of Singapore)

Department of Mathematics

Nanjing University

June 2012

Submitted in partial fulfilment of the requirements

for the degree of PhD in Computational Mathematics

南京大学研究生毕业论文中文摘要首页用纸

毕业论文题目：求解一类矩阵范数逼近问题的数值算法

计算数学 专业 2009 级博士生姓名：陈彩华

指导教师（姓名、职称）：何炳生 教授，孙德锋 教授，卓金权 教授

摘 要

本文主要针对一类矩阵范数逼近问题提出快速有效而且稳定的算法。具体而言，矩阵范数逼近问题是在一族给定矩阵中寻找一个满足某些线性等式与不等式约束的线性组合，并且该组合与目标矩阵在谱范数意义下具有最近的距离。这类问题通常来源于数值代数，网络，控制，工程等领域，例如寻找矩阵的切比雪夫矩阵多项式以及求解最速收敛的混合马尔可夫链模型。

本文首先用目前流行的一阶交替方向法来求解此类问题，在算法的每步迭代中，由于子问题可以通过快速算法求解或者直接拥有解析解，故而比较容易实现。然而，交替方向法求解矩阵范数逼近问题的数值表现不稳定，对于某些算例特别是带有约束的问题，其无法在合理的时间内求得令人满意的解。

为了克服这个困难，我们引入一个不精确对偶邻近点算法来求解矩阵范数逼近问题。在每步迭代中，子问题可以被写成一个半光滑等式组，并可以用不精确半光滑牛顿法求解，其中牛顿方向利用预条件共轭梯度法求解。当子问题的原始约束非退化条件成立时，不精确半光滑牛顿法被证明有超线性的收敛速率。此外，我们对于该算法设计了高效的实现方法。各类问题的数值结果表明，半光滑牛顿共轭梯度对偶邻近点算法优于交替方向法，它可以稳定高效地得到矩阵范数逼近问题具有相对较高精度的解。

当矩阵范数问题的矩阵变为向量时，其可以等价的转换成一个二阶锥规划，可以被牛顿型方法如内点法求解，甚至对于大规模问题也是如此。受此启发，我们考虑用平方光滑牛顿法来求解给定矩阵行数远远小于列数的矩阵范数逼近问题。为此，我们首先提出了谱范数上图锥的投影算子的光滑函数，并且证明它在每个点至少具有γ阶的半光滑性，其中γ为某一正有理数。此外，原始对偶约束非退化条件在原始对偶最优点成立的情形下，该算法也具有超线性收敛性。初步的数值试验表明，该算法对于求解中小等规模问题非常稳定高效，它可以通过很少的迭代步数得到精度令人满意的解。

关键词： 矩阵范数逼近问题，交替方向法，邻近点算法，谱算子，半光滑牛顿法，共轭梯度法，约束非退化，平方光滑牛顿法。

南京大学研究生毕业论文英文摘要首页用纸

THESIS: Numerical Algorithms for a Class of Matrix Norm Approximation Problems

SPECIALIZATION: Computational Mathematics

POSTGRADUATE: Caihua Chen

MENTOR: Professor Bingsheng He (Nanjing University)
Professor Defeng Sun (National University of Singapore)
Professor Kim-Chuan Toh (National University of Singapore)

Abstract

This thesis focuses on designing robust and efficient algorithms for a class of matrix norm approximation (MNA) problems that are to find an affine combination of given matrices having the minimal spectral norm subject to some prescribed linear equality and inequality constraints. These problems arise often in numerical algebra, network, control, engineering and other areas, such as finding the Chebyshev polynomials of matrices and fastest mixing Markov chain models.

In this thesis, we first apply the popular first-order algorithm alternating direction method (ADM) to solve such problems. At each iteration of the algorithm, the subproblems involved can either be solved by a fast algorithm or admit closed form solutions, which allows us to implement the ADM easily and simply. Unfortunately, numerical experiments on MNA problems reveal that the ADM performs unstably, and it may fail to achieve satisfactory accuracy in reasonable cpu time for some tested examples, especially for the constrained cases.

To overcome this difficulty, we also introduce an inexact dual proximal point algorithm (in short SNDPPA) for solving the MNA problems. At each iteration, the inner problem, rewritten as a system of semismooth equations, is solved by an inexact semismooth Newton method using the preconditioned conjugate gradient method to compute the Newton directions. Furthermore, when the primal constraint nondegeneracy condition holds for the inner problems, our inexact semismooth Newton method is proven to have a suplinear convergence rate. We also design efficient implementation for the proposed algorithm to solve a variety of instances and compare its performance with that of ADM. Numerical results show that the semismooth Newton-CG dual proximal point algorithm substantially outperforms the alternating direction method, and it is able to solve the matrix norm approximation problems efficiently and stably to a relatively high accuracy.

When one restricts the matrices to vectors, then the matrix norm approximation problem can be converted into a second order cone (SOC) problem, which can be solved by Newton's method such as IPMs even for large scale problems. Motivated by this, we also consider a squared smoothing Newton method, to solve the MNA problems in which the matrix is of much more columns than rows (skinny ones) such as the vector case. For this purpose, we present an interior smoothing function for the metric projector over the epigraph cone of spectral norm and establish its γ order semismoothness everywhere. Moreover, suplinear convergence of the smoothing Newton method for solving the MNA problems is also shown to hold under the primal dual constraint nondegenerate conditions for the MNA problems and their dual at the primal dual optimal solution pairs. Preliminary numerical result demonstrate that the smoothing Newton is robust and efficient for the problem of small and moderate scale. Specifically, we can successfully find the solution with the desired accuracy in a few iterations.

Keywords: Matrix norm approximation, alternating direction method, proximal point algorithm, spectral operator, semismooth Newton method, conjugate gradient method, constraint nondegneracy, squared smoothing Newton method.

Contents

Chapter 1

Introduction

In this thesis, we focus on designing efficient algorithms for solving a special case of large scale matrix optimization problems. In particular, we are interested in a class of matrix norm approximation problems with linear equality and inequality constraints. Let $\Re^{m\times n}$ be the space of $m \times n$ matrices equipped with the standard inner product $\langle X, Y\rangle = \mathrm{Tr}(X^TY)$ for $X, Y \in \Re^{m\times n}$. Given a family of matrices $\{A_1, A_2, \dots, A_p\}$, define the linear operator $\mathcal{A}$ and its adjoint $\mathcal{A}^*$ by

$$\mathcal{A}(X) := [\langle A_1, X\rangle, \dots, \langle A_p, X\rangle]^T,\ \mathcal{A}^*(y) := \sum_{k=1}^{p} y_k A_k, \quad \forall\, X \in \Re^{m\times n},\ y \in \Re^p.$$

The matrix norm approximation (abbreviated as MNA) problem we consider in this thesis takes the following form

$$\min\Big\{\|A_0 - \mathcal{A}^*y\|_2 \mid By - b \in Q\Big\}, \tag{1.1}$$

where $A_0 \in \Re^{m\times n}$ and $B \in \Re^{(n_1+n_2)\times p}$ are given matrices, $b \in \Re^{n_1+n_2}$ is a vector and $Q = \{0\}^{n_1} \times \Re^{n_2}_+$ is a polyhedral cone. Without loss of generality, we assume that $m \le n$.

1.1 Motivating examples and related approaches

The MNA problems arise in numerical algebra, network, control, engineering and many other areas. An illustrative example is the problem of finding the degree t Chebyshev polynomial of a given matrix $A \in \Re^{N\times N}$. In this problem, one is interested in finding a degree t monic polynomial p_t^* which minimizes the spectral norm of $p_t(A)$, i.e.,

$$\min\{\|p_t(A)\|_2 \mid p_t \text{ is a monic polynomial of degree } t\}. \tag{1.2}$$

Problem (1.2) was firstly introduced in [36] under the name ideal Arboldi approximation problem and then extensively studied in [83] where the ideal Aronoli polynomial p_t^* is called the degree t Chebyshev polynomial of A, in analogy to the notion Chebyshev polynomial in approximation theory [44, 80], which is a monic polynomial that attains minimal essential-supremum on that set. Indeed, suppose A is a Hermitan, by the eigenvalue decomposition, the Chebyshev polynomials of A as defined by (1.2) collapses to the Chebyshev polynomial of the spectrum of A in the latter sense. Note that

$$p_t(A) = A^t - \sum_{i=0}^{t-1} y_i A^i \tag{1.3}$$

for some $y \in \Re^t$, the Chebyshev matrix approximation problem is actually a special case of (1.1).

Possibly due to its mathematical elegance, the Chebyshev polynomial matrices problem received much attention of the theoretical researchers, see [25, 26, 53, 96] and references therein. Nevertheless, with the exception of the early work in [83] concerning with algorithmic and computational results, no attention was paid on the numerical treatment of the Chebyshev matrix approximation problem. In [83], the model (1.2) is equivalently reformulated as the following semidefinite program problem

$$\begin{array}{ll} \min & -\lambda \\ \text{s.t.} & \sum_{k=1}^{t} y_k A_k + \lambda A_{t+1} + Z = A_0 \\ & Z \succeq 0, \end{array} \tag{1.4}$$

where

$$A_{t+1} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}, A_0 = \begin{pmatrix} 0 & B_0 \\ B_0^T & I \end{pmatrix}, A_0 = \begin{pmatrix} 0 & B_k \\ B_k^T & I \end{pmatrix}, B_0 = A^t$$

and $B_k = A^{k-1}$ for $k = 1, 2, \ldots, t$. Based on the semidefinite reformulation above, a primal-dual interior point method is proposed to compute the Chebyshev polynomials of matrices. In the implementation of such algorithm, the search direction is computed via a dense Schur complement equation even if the data is sparse and each iteration has a complexity $\mathcal{O}(tN^3) + \mathcal{O}(t^2N^2)$, which is reduced to $\mathcal{O}(tN^3)$ for the Cheyshev matrix approximation problems since $t < N$. Obviously, the high complexity may give rise to great difficulties for applying the interior point algorithms to solve the Chebyshev matrix norm approximation problems of large scale.

In contrast to the unconstrained example (1.2), some other problems may have prescribed linear constraints, for example, the fastest mixing Markov chain (FMMC) problem studied in [7–9]. Let $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ be an undirected connected graph with n

nodes. The FMMC problem is to find a symmetric stochastic matrix P with $P_{ij} = 0$ for $(i,j) \notin \mathcal{E}$ that minimizes $\mu(P)$, where

$$\mu(P) = \max_{i=2,\ldots,n} |\lambda_i(P)|$$

and $\lambda_i(P)$ is the ith largest eigenvalue of P in magnitude. Let the vector of transition probabilities on the edges (labeled by $l = 1, 2, \ldots, p$) be d, and let the matrix $B \in \Re^{n \times m}$ be defined by

$$B_{il} := \begin{cases} 1, & \text{if edge } l \text{ incident to vertex } i, \\ 0, & \text{otherwise.} \end{cases} \tag{1.5}$$

For any $l = 1, 2, \ldots, p$, write the matrix $E^{(l)}$ to denote

$$E^{(l)}_{ij} := \begin{cases} 1, & \text{if edge } l \text{ incident to vertex } i \text{ and } j,\ j \neq i \\ -1, & \text{if edge } l \text{ incident to vertex } i,\ j = i, \\ 0, & \text{otherwise.} \end{cases} \tag{1.6}$$

Then by the analysis in [8, 9], the FMMC problem can be written as a matrix norm approximation problem in terms of d as follows:

$$\min \left\{ \|I - (1/n)\mathbf{1}\mathbf{1}^T + \sum_{l=1}^{p} d_l E^{(l)}\|_2 \mid d \geq 0,\ Bd \leq \mathbf{1} \right\}. \tag{1.7}$$

Several simple heuristic methods, including the maximum-degree chain and Metroplis-Hasting chain, have been proposed to obtain the transition probability giving fast mixing. Let d_i be the degree of the vertex i, not counting the self-loop. Denote by $d_{\max}$ the maximum degree of the graph, i.e.,

$$d_{\max} = \max_{i \in \mathcal{N}} d_i.$$

Then the maximum-degree transition probability matrix P^{md} is given by

$$P^{\mathrm{md}}_{ij} = \begin{cases} 1/d_{\max}, & \text{if } (i,j) \in \mathcal{E} \text{ and } i \neq j, \\ 1 - d_i/d_{\max}, & \text{if } i = j, \\ 0, & \text{otherwise.} \end{cases}$$

Another typical heuristic is the Metropolis-Hasting chain, which is constructed based on the Metroplis-Hasting algorithm [3, 38, 60] applied to a random walk on a graph.

In this chain, the transition probability matrix P^{mh} is given in a symmetric form

$$P^{\mathrm{mh}}_{ij} = \begin{cases} \min\{1/d_i, 1/d_j\}, & \text{if } (i,j) \in \mathcal{E} \text{ and } i \neq j, \\ \sum_{(i,k)\in\mathcal{E}} \max\{0, 1/d_i - 1/d_j\}, & \text{if } i = j, \\ 0, & \text{otherwise.} \end{cases}$$

Other than the aforementioned methods, there are many other interesting works aiming at developing some heuristics to assign transition probabilities obtaining faster mixing Markov chain. Related materials can be found in [1, 19, 50, 75]. With the help of the semidefinite programming technique, Boyd et al. [9] proposed a primal-dual interior point algorithm to compute exactly the fastest mixing chain. In that paper, they expressed the FMMC problem as a SDP by introducing a scalar variable s:

$$\begin{aligned} \min \quad & s \\ \text{s.t.} \quad & \mathrm{Diag}\Big(P - (1/n)\mathbf{1}\mathbf{1}^T + sI,\ sI - P + (1/n)\mathbf{1}\mathbf{1}^T,\ \mathrm{vec}(P)\Big) \succeq 0, \\ & P\mathbf{1} = \mathbf{1},\ P = P^T, \\ & P_{ij} = 0,\ (i,j) \notin \mathcal{E}, \end{aligned} \tag{1.8}$$

where $\mathrm{Diag}(\cdot)$ forms a block diagonal matrix from its arguments, and $\mathrm{vec}(P)$ is a vector containing the $n(n+1)/2$ different coefficients in P. For graphs with up to a thousand or so edges, the resulting semidefinite programming can be solved efficiently by the standard interior point solvers. For larger problem, the authors suggested a projected subgradient method to solve the MNA formulation (1.7) of the FMMC problem. However, as pointed by the authors, the algorithm is relatively slow in terms of number of iterations and has no simple stopping criterion guaranteeing a certain level of suboptimality while compared to a primal-dual interior point method.

Another strong motivation for considering the model (1.1) comes from the fastest distributed linear averaging (FDLA) problem with symmetric weights. Let $\mathcal{G} = \{\mathcal{N}, \mathcal{E}\}$ as defined above be a connected graph with n nodes. In this problem, we aim at finding the symmetric weight matrix W, consistent with $\mathcal{G}$, that makes the convergence as fast as possible. Using [88, Theorem 1], we can formulate the FDLA problem as the following optimization problem

$$\min_d \ \rho\big(I - (1/n)\mathbf{1}\mathbf{1}^T + \sum_{l=1}^{p} d_l E^{(l)}\big), \tag{1.9}$$

where ρ stands for the spectral radius, p is the weight on the edges with different nodes, and B and $E^{(l)}$ are respectively defined by (1.5) and (1.6). Since the spectral radius of any symmetric matric coincides with its spectral norm, we know the FDLA problem can be stated as a MNA problem

$$\min_{d} \|I - (1/n)\mathbf{1}\mathbf{1}^T + \sum_{l=1}^{p} d_l E^{(l)}\|_2. \tag{1.10}$$

Similar to the FMMC problem, there exist some simple heuristics [88] for choosing the weight matrix W that gives reasonably fast convergence of distributed averaging. Define the Laplacian matrix L by

$$L_{ij} = \begin{cases} -1, & \text{if } (i,j) \in \mathcal{E} \text{ and } i \neq j, \\ d_i & \text{if } i = j, \\ 0, & \text{otherwise,} \end{cases}$$

where d_i is the degree of node i not counting the self-loop. The simplest approach is to set all the edge (with different nodes) weights to be a constant α; the self-weights are decided by the constraint $W\mathbf{1} = \mathbf{1}$. In the best constant weight graph, α is set to be $2/\big(\lambda_1(L) + \lambda_{n-1}(L)\big)$ where λ_1 and λ_{n-1} stand respectively for the largest and $n-1$-th eigenvalues of L. Additionally, one can also use the maximum-degree weight

$$\alpha^{\text{md}} = \frac{1}{d_{\max}},$$

provided that the graph is not bipartite. Another method is to assign the weight on an edge based on the larger degree of its two incident nodes:

$$W_{ij} = \frac{1}{\max(d_i, d_j)}, \quad \{i,j\} \in \mathcal{E}$$

and then determine W_{ii} using $W\mathbf{1} = \mathbf{1}$, which yields the so-called local-degree weights. By reformulating the FDLA equivalently as a SDP

$$\begin{aligned} \min \quad & s \\ \text{s.t.} \quad & \operatorname{Diag}\Big(W - (1/n)\mathbf{1}\mathbf{1}^T + sI,\, sI - W + (1/n)\mathbf{1}\mathbf{1}^T\Big) \succeq 0, \\ & W\mathbf{1} = \mathbf{1},\, W = W^T, \\ & W_{ij} = 0,\, (i,j) \notin \mathcal{E}, \end{aligned} \tag{1.11}$$

the authors show the interior-point method is able to solve efficiently the FDLA problem, for network with up to a thousand or so edges. A simple subgradient method, which suffers from slow convergence is also described in [88] to handle far larger problem.

The above examples serve to motivate the study of numerical algorithms for solving MNA problems.

1.2 Contributions of the thesis

By introducing a scalar variable t to bound the spectral norm $\|A_0 - A^*y\|_2$, the MNA problem (1.1) can be equivalently formulated as:

$$\begin{array}{ll} \min & t \\ \text{s.t.} & By \in b + Q, \\ & t \geq \|A_0 - A^*y\|_2, \end{array} \tag{1.12}$$

which is a natural generalization of the second order cone programming. It is well known that the matrix norm constraint $t \geq \|A_0 - A^*y\|_2$ is equivalent to a linear matrix inequality with block-arrow structure:

$$\begin{bmatrix} tI_m & A_0 - \mathcal{A}^*y \\ (A_0 - \mathcal{A}^*y)^T & tI_n \end{bmatrix} \succeq 0. \tag{1.13}$$

Therefore, the problem (1.12) can be expressed as a semidefinite programming problem:

$$\begin{array}{ll} \min & t \\ \text{s.t.} & \begin{bmatrix} tI_m & A_0 - \mathcal{A}^*y \\ (A_0 - \mathcal{A}^*y)^T & tI_n \end{bmatrix} \succeq 0, \\ & By \in b + Q, \end{array} \tag{1.14}$$

which falls into the applicable scope of standard SDP packages such as SDPT3 [82], SeDuMi [76], or SDPNAL [95]. When m or $n = 1$, the constraint $t \geq \|A_0 - A^*y\|_2$ reduces to a second order cone constraint. In this case, it is certainly not wise to solve the MNA problem (1.1) via (1.14). Instead one should deal with (1.1) or (1.12) directly since it is just a second-order cone problem, which requires far lower computational cost to solve compared to the SDP reformulation (1.14). In the case that $n > 1$, the block-arrow constraint $t \geq \|A_0 - A^*y\|_2$ is often handled via its SDP reformulation [2] because it can not be reduced to a second order cone constraint. This of course makes the MNA problem potentially very computationally expensive since one has to deal with $(m+n) \times (m+n)$ matrix variables instead of $m \times n$ matrices. The computational cost and memory requirement are especially high when we have large $m+n$, but $m \ll n$. For example, while applied the SDP (1.14) without linear equality and inequality constraints, the standard interior point methods require $\mathcal{O}\Big(p(m+n)^3 + p^2(m+n)^2 + p^3 + (m+n)^3\Big)$ flops at each iteration to solve the dense Schur complement equation and compute the search directions. The total memory requirement should be more than $8p^2 + \mathcal{O}(m+n)^2$.

Realizing the difficulties mentioned above, in this thesis we consider three different approaches to solve the MNA problem directly instead of via its SDP reformulation. The first idea to solve the matrix norm approximation problems is built on the classical alternating direction method [30, 34]. In the past several years, we have witnessed explosively increasing interests in ADM because of its effectiveness in diverse areas, such as image processing [90], compressive sensing [89], matrix completion [14], robust principle component analysis [81] and sparse matrix separation [74]. This provides an initial impetus for us to apply ADM to solve MNA problems by its variant which has a separable structure:

$$\begin{aligned} \min \quad & \|X\|_2 \\ \text{s.t.} \quad & \mathcal{A}^* y + X = A_0, \\ & By - b = z, \; z \in Q. \end{aligned} \tag{1.15}$$

At each iteration of the ADM, the subproblem involved can either be solved by a fast algorithm or it has a closed form solution, due to recent advances in [21].

Recently, Zhao, Sun and Toh [95] designed a Newton-CG augmented Lagrangian (NAL) method to solve the standard SDP problems, which is essentially a proximal point algorithm applied to primal problem where the inner problems are solved by an inexact semi-smooth Newton method using a preconditioned conjugate gradient (PCG) solver. Their numerical results demonstrated the high efficiency and stability of the NAL method whenever the primal and dual constraint nondegeneracy conditions hold. This phenomenon can be partially explained by the theoretical results in [13, 77, 78] where it is shown that under the constraint nondegenerate conditions the augmented Lagrangian method can be locally regarded as an approximate generalized Newton's method applied to a semismooth equation. Shortly after, Wang, Sun and Toh extended this idea to solve a class of log-det problems, i.e., applying the proximal point algorithm to solve the primal problem where the innear problem is solved by a Newton-CG method. Extensive numerical experiments show that the resulting algorithm is shown to be approximately 2～20 times faster than the adaptive Nesterov's smoothing method [57]. Motivated by the stability and the effectiveness of the Newton-CG based PPA, we adopt the essential idea of the NAL method to propose a semismooth Newton-CG dual proximal point algorithm (SNDPPA) to solve the problem (1.1). As a starting point, we first derive an inexact dual proximal point algorithmic framework with checkable stopping criterions for the MNA problem. Based on the classical result of the proximal point algorithm in [72, 73], we analyze the global and local convergence of the PPA for solving (1.1). We note that the subproblem of the dual PPA in each iteration is an unconstrained minimization problem whose objective function is convex continuously differentiable though not twice continuously differentiable. However, since the corresponding gradient is strongly semismooth, we are able to apply the inexact semismooth Newton method to solve the unconstrained minimization subproblem with a fast convergence where at each iteration a preconditioned conjugate gradient method is used to compute approximately the Newton directions. Using the results of nonsmooth analysis in [16], we also clearly characterize the tangent cone and then the linearity space of the unit nuclear norm ball, and therefore introduce a constraint nondegeneracy condition for the subproblem. It turns out that the constraint nondegeneracy condition is equivalent to the nonsingularity of the generalized hessians of the subproblems and then ensures the quadratic convergence of our inexact semismooth Newton method. We also designed efficient implementation for our proposed algorithm to solve a variety of instances and compare its performance with the popular first order alternating direction method. The results show that our algorithm substantially outperforms the alternating direction method, especially for the constrained cases, and it is able to solve the matrix norm approximation problems efficiently to a relatively high accuracy.

In the last part of this thesis, we study a squared smoothing Newton method [47, 68, 79] for solving MNA problem (1.1), or equivalently (1.12). Assuming the strong duality holds for the problem (1.12) and its dual and there exists at least one saddle point.Then solving the MNA problem is equivalent to the following KKT syestem:

$$\begin{cases} (t, X) = \Pi_{\mathcal{K}}(t-1, X+Z) \\ \mathcal{A}Z + B^T w = 0, \\ \mathcal{A}^* y + X = A_0, \\ B_1 y = b_1, \\ w_2 = \Pi_{\Re^{n_2}_+}(w_2 - B_2 y + b_2), \end{cases} \tag{1.16}$$

where $\mathcal{K}$ is the epigraph cone of spectral norm, and $_1$ and $_2$ represent the indexes of equality and inequality constraints, respectively. In order to apply the smoothing Newton method to (6.5), we adopt the essential idea of [67, section 4] to provide a computable smoothing function for the metric projection onto the epigraph of l_1 norm, which, together with the recent developments of spectral operator [20], furnishs us a smoothing function G of $\Pi_{\mathcal{K}}$. It can be shown that $G(\cdot,\cdot,\cdot)$ is γ order semismooth at $(0,t,x)$ for any given $t \in \Re$ and $x \in \Re^m$. The search direction is computed by solving the Schur complement of the Newton system. Moreover, based on Clark's classical results on the tangent cone of convex sets [16], we also characterize the primal and dual constraint nondegeneracy and derive some equivalent conditions for the nondegeneracy. We are able to show when the primal and dual constraint nondegeneracy conditions of (1.12) and its dual hold, the proposed smoothing Newton method solves the MNA problem with a suplinear convergence rate. Preliminary numerical experiments demonstrate that the smoothing Newton is very robust and efficient for moderate and small scale problem. Specifically, we can successfully find the solution with the desired accuracy in a few iterations.

1.3 Organization of the thesis

The remaining parts of this thesis are organized as follows. In chapter 2, we list some preliminaries on the semismoothness mapping, spectral operator for nonsymmetric matrices, Moreau-Yosida regularization and smoothing functions. We give the closed form solution of the proximal point operator associated with the spectral function and establish its strongly semismoothness everywhere. We also discuss a computable smooth counterpart of the metric projection onto the epigraph of the spectral norm. With the help of the properties enjoyed by the spectral operator, we are able to show this smoothing function of $\Pi_{\mathcal{K}}$ is γ order semismooth at $(0,t,X)$ for given $t \in \Re$ and $X \in \Re^{m\times n}$. In chapter 3, we briefly review the history of the alternating direction method, develop some new results on the ADM and the proximal ADM and then discuss the details on the implementation of ADM for the MNA problem. In chapter 4, we introduce the framework of the inexact dual PPA for solving the MNA problem and establish its global and local convergence under certain conditions. The subproblems reformulated as a system of nonsmooth equations are solved by an inexact semismooth Newton method where the preconditioned conjugate gradient method is employed to compute the Newton directions. The suplinear convergence of our inexact semismooth Newton-CG method is established under the primal constraint nondegeneracy condition of subproblems, together with the strong semismoothness property of the metric projection onto the unit nuclear norm ball. In addition, some numerical issues pertaining to the efficient implementation of the semismooth Newton-CG method are also addressed in this chapter. In chapter 5, we implement the ADM and the SNDPPA to solve a variety of problems, including random matrix norm approximation problems, Chebyshev polynomial of matrices and FMMC/FDLA problems. Numerical results demonstrate that our SNDPPA is very efficient and robust to solve the MNA problem to a relatively high accuracy. In chapter 6, by using the smoothing function of Π_K introduced in chapter 2, we introduce the squared smoothing Newton method for the MNA problem where the Newton directions can be computed by solving the Schur complement of the Newton system. We conclude the thesis in chapter 7.

Notation. For any given positive integer m and n, we denote by I_n, $\mathbf{1}_{m\times n}$ and $\mathbf{0}_{m\times n}$ the $n \times n$ identity matrix, the $m \times n$ matrix of ones and zeros, respectively. We also use $\mathbf{1}_n$ and $\mathbf{0}_n$ to denote the vector of ones and zeros, respectively. We frequently drop m, n from the above notations when their size can be clear from the context. For any $x \in \Re^n$, $\mathrm{diag}(x)$ denotes the diagonal matrix with diagonal entries $x_i, i = 1, \ldots, n$, while for any $X \in \Re^{m\times n}$, $\mathrm{diag}(X)$ denotes the main diagonal of X. Let $\alpha \subseteq \{1, \ldots, n\}$ be an index set, we use $|\alpha|$ to represent the cardinality of α and X_α to denote the sub-matrix of X obtained by removing all the columns of X not in α. Let $\beta \subseteq \{1, \ldots, n\}$ be another index set, we use $X_{\alpha\beta}$ to denote the $|\alpha| \times |\beta|$ sub-matrix of X obtained by removing all the rows of $X \in \Re^{m\times n}$ not in α and all the columns of X not in β. The Hardamard product between matrices is denoted by "$\circ$", i.e., for any two matrices X and Y in $\Re^{m\times n}$, the (i, j)-th entry of $Z := X \circ Y$ is $Z_{ij} = X_{ij}Y_{ij}$.

Chapter 2

Preliminaries

In this chapter, we review and develop some results on the semismooth mappings, spectral operator, Moreau-Yosida regularization and smoothing functions, which are useful for our subsequent discussion.

2.1 Semismooth mapping

In this section, we briefly review the basic concepts B-subdifferential, Clark generalized Jacobian and semismooth functions.

Let $\mathcal{E}$ be a finite-dimensional real Hilbert space and $\mathcal{O}$ be an open set in $\mathcal{E}$. Let $\mathcal{E}'$ be another finite dimension Hilbert space. Suppose that $\Phi : \mathcal{O} \to \mathcal{E}'$ is a locally Lipschitz continuous function on the open set $\mathcal{O}$. By Rademacher's theorem, Φ is almost everywhere Féchet-differentiable in $\mathcal{E}$. Let Ω be the set of points where Φ is differentiable. For any $x \in \mathcal{E}$, the B-subdifferential of Φ is defined by

$$\partial_B \Phi(x) := \left\{ \lim_{\Omega \ni x^k \to x} \Phi'(x^k) \right\}$$

and the Clark's generalized Jacobian [16] of Φ at x is the convex hull of $\partial_B \Phi(x)$, i.e.,

$$\partial \Phi(x) = \mathrm{conv}\left\{ \partial_B \Phi(x) \right\}.$$

The concept semismooth was first introduced by Mifflin [61] for functionals and then extended to vector valued functions by Qi and Sun [69]. See also [27, 59].

Definition 2.1. Let $\Phi : \mathcal{O} \subset \mathcal{E} \to \mathcal{E}'$ be a locally Lipschitz continuous function on the open set $\mathcal{E}$. The function Φ is said to be G-semismooth at a point $x \in \mathcal{O}$ if for any $y \to x$ and $V \in \partial \Phi(y)$,

$$\Phi(y) - \Phi(x) - V(y - x) = o(\|y - x\|).$$

The function Φ is said to be γ order G-semismooth at x if for any $y \to x$ and $V \in \partial \Phi(y)$,

$$\Phi(y) - \Phi(x) - V(y - x) = O(\|y - x\|^{\gamma+1}).$$

If $\gamma = 1$, Φ is strongly G-semismooth at x. Furthermore, if the (strongly, γ order) G-semismooth function Φ is also directionally differentiable at x, then Φ is said to be (strongly, γ order) semismooth at x.

The (strong) semismoothness property plays crucial role in establishing the (quadratic) suplinear convergence of the semismooth Newton method for solving the nonlinear equations, as well as SC^1 unconstrained optimization problems. Many common functions such as convex functions and smooth functions can be verified to be semismooth everywhere. Piecewise linear functions and twice continuously differentiable functions are examples of strongly semismooth functions. In what follows, we provide a simple sufficient (not necessary) criteria to recognize the semismoothness of functions, which is based on the concept semialgebraic functions originally considered in the field of algebraic geometry [4].

Definition 2.2. A set in $\Re^n$ is semialgebraic if it is a finite union of sets of the form

$$\{x \in \Re^n : p_i(x) > 0,\, q_j(x) = 0, \quad i = 1, 2, \ldots, m,\, j = 1, 2, \ldots, k\},$$

where all $p_i(x)$, $q_j(x)$ are polynomials. A map $F : X \subset \Re^n \to \Re^m$ is called semialgebraic if its graph is a semialgebraic subset of $\Re^{n+m}$.

The following semismoothness result on semialgebraic functions is a special case of [5, Theorem 1] on tame functions. More characterizations of tame mappings can be found in [48].

Proposition 2.1. *Let a semialgebraic function $F : \Re^n \to \Re^m$ be locally Lipschitz. Then there exists a rational number $\gamma > 0$ such that F is γ-order semismooth.*

2.2 Spectral operator of matrices

In this section, we first list two useful results on the nonsymmetric matrices. The following inequality is called von Neumann's trace inequality [63].

Proposition 2.2. *Let Y and Z be two matrices in $\Re^{m\times n}$. Then*

$$\langle Y, Z\rangle \le \langle \sigma(Y), \sigma(Z)\rangle,$$

where $\sigma(Y)$ and $\sigma(Z)$ are the singular value vectors of Y and Z respectively.

Proposition 2.3. (c.f. [52]) *Suppose $X \in \Re^{m\times n}$ has the SVD*

$$X = U[\mathrm{Diag}(\sigma)\,0]V^T. \tag{2.1}$$

Then the orthogonal matrices P and W satisfy

$$P[\mathrm{Diag}(\sigma)\,0] = [\mathrm{Diag}(\sigma)\,0]W$$

if and only if there exist three orthogonal matrices $Q \in \Re^{r\times r}$, $Q' \in \Re^{(m-r)\times(m-r)}$ and $Q'' \in \Re^{(n-r)\times(n-r)}$ such that

$$P = \begin{bmatrix} Q & 0 \\ 0 & Q' \end{bmatrix} \quad \text{and} \quad W = \begin{bmatrix} Q & 0 \\ 0 & Q'' \end{bmatrix},$$

where r is the number of positive singular values of X and Q is a block diagonal matrix.

Let $\mathcal{X}$ be a Euclidean space which is Cartesian product of $\mathcal{S}^{m_1}$ and $\Re^{m\times n}$, i.e.,

$$\mathcal{X} := \mathcal{S}^{m_1} \times \Re^{m\times n}.$$

For any $X := (X_1, X_2)$, define $\kappa(X) \in \Re^{m_1+m}$ by

$$\kappa(X) := (\lambda(X_1), \sigma(X_2)).$$

Let w be a mapping from $\Re^{m_1+m}$ to $\Re^{m_1+m}$ which can be decomposed into the form

$$w := (h, g),$$

where $h : \Re^{m_1+m} \to \Re^{m_1}$ and $g : \Re^{m_1+m} \to \Re^{m}$. Suppose w is symmetric, that is for any perturbation matrix Q_1 and signed perturbation matrix Q_2,

$$w(x) = [Q_1\ Q_2]^T w(Q_1x_1\, Q_2x_2), \quad \forall x \in \Re^{m_1+m},$$

where $x_1 \in \Re^{m_1}, x_2 \in \Re^m$ and $x^T = [x_1^T,\, x_2^T]$.

Definition 2.3. [20] Let X_1 and X_2 have the following respective eigenvalue and singular value decomposition

$$X_1 = P\mathrm{Diag}(\lambda)P^T, \tag{2.2}$$

$$X_2 = U[\mathrm{Diag}(\sigma)\ 0]V^T. \tag{2.3}$$

The spectral operator $G : \mathcal{X} \to \mathcal{X}$ with respect to the symmetric function w is defined by

$$G(X) := (G_1(X), G_2(X)),\ \ X = [X_1,\ X_2] \in \mathcal{X},$$

where

$$G_k(X) := \begin{cases} P\mathrm{Diag}(h(\kappa))P^T, & \text{if } k = 1, \\ U[\mathrm{Diag}(g(\kappa))\ 0]V^T, & \text{if } k = 2, \end{cases}$$

and

$$\kappa = (\lambda(X_1), \sigma(X_2)).$$

Since w is symmetric, [20, Theorem 3.1] implies the spectral operator $G : \mathcal{X} \to \mathcal{X}$ is well defined. Before moving to introduce the properties of G, we first give some notations to simplify the subsequent discussion. Assume that g and h are F-differentiable (i.e., Féchet-differentiable). For any $x \in \Re^{m_1+m}$, rewrite h and g as the following form

$$h(x) := (h_1(x), h_2(x), \ldots, h_{m_1}(x)),$$

and

$$g(x) := (g_1(x), g_2(x), \ldots, g_m(x)).$$

Define the matrices $\mathcal{A}(\kappa), \Omega(\kappa), \Gamma(\kappa) \in \Re^{m\times m}$ and $\mathcal{F}(\kappa) \in \Re^{m\times(n-m)}$ by

$$[\mathcal{A}(\kappa)]_{ij} \ := \begin{cases} \dfrac{h_i(\kappa) - h_j(\kappa)}{\lambda_i - \lambda_j} & \text{if } \lambda_i \neq \lambda_j, \\ (h'(\kappa))_{ii} - (h'(\kappa))_{ij} & \text{if } \lambda_i = \lambda_j, i \neq j, \\ 0 & \text{otherwise}, \end{cases} \tag{2.4}$$

$$[\Omega(\kappa)]_{ij} \ := \begin{cases} \dfrac{g_i(\kappa) - g_j(\kappa)}{\sigma_i - \sigma_j} & \text{if } \sigma_i \neq \sigma_j, \\ (g'(\kappa))_{ii} - (g'(\kappa))_{ij} & \text{if } \sigma_i = \sigma_j, i \neq j, \\ 0 & \text{otherwise}, \end{cases} \tag{2.5}$$

$$[\Gamma(\kappa)]_{ij} \ := \begin{cases} \dfrac{g_i(\kappa) + g_j(\kappa)}{\sigma_i + \sigma_j} & \text{if } \sigma_i + \sigma_j \neq 0, \\ (g'(\kappa))_{ii} + (g'(\kappa))_{ij} & \text{otherwise}, \end{cases} \tag{2.6}$$

and

$$[\mathcal{F}(\kappa)]_{ij} \ := \begin{cases} \dfrac{g_i(\kappa)}{\sigma_i} & \text{if } \sigma_i \neq 0, \\ (g'(\sigma))_{ii} & \text{otherwise}. \end{cases} \tag{2.7}$$

Let the linear operators S and T be defined by

$$S(A) := \frac{1}{2}(A + A^T), \quad T(A) := \frac{1}{2}(A - A^T), \ \forall\, A \in \Re^{m\times m}.$$

Proposition 2.4. *[20] Let $X = [X_1\, X_2] \in \mathcal{X}$ be given. Suppose X_1 and X_2 have the eigenvalue decomposition* (2.2) *and the singular value decomposition* (2.3)*, respectively.*

(i) *The spectral operator G is F-differentiable at X if and only if the symmetric mapping w is F-differentiable at κ.*

(ii) *If G is F-differentiable at X, then its derivative is given as follows for any $H = (A, B) \in \mathcal{X}$,*

$$G'(X)H = \begin{pmatrix} P\Big[\mathcal{A}\circ \tilde{A} + \mathrm{Diag}(h'(\kappa)\mathrm{diag}(\tilde{H}))\Big]P^T, \\ U\Big[\Omega\circ S(\tilde{B}_1) + \Gamma\circ T(\tilde{B}_1) + \mathrm{Diag}(g'(\kappa)\mathrm{diag}(\tilde{H})),\ \mathcal{F}\circ\tilde{B}_2\Big]V^T \end{pmatrix},$$

where $\tilde{A} = P^TAP \in \mathcal{S}^{m_1}$; $\tilde{B}_1 \in \Re^{m\times m}$, $\tilde{B}_2 \in \Re^{m\times(n-m)}$ and $[\tilde{B}_1\ \tilde{B}_2] = U^TBV$; $\mathrm{diag}(\tilde{H}) = [\mathrm{diag}(\tilde{A}); \mathrm{diag}(\tilde{B}_1)]$.

(iii) *If w is locally Lipschitz continuous at κ, then the spectral operator G is (γ order, strongly) G-semismooth at X if and only w is (γ order, strongly) G-semismooth at κ.*

2.3 The Moreau-Yosida regularization

Let $f : \mathcal{E} \to (-\infty, +\infty]$ be a closed proper convex function, e.g. see [71]. The Moreau-Yosida regularization [62, 92] of f at $x \in \mathcal{E}$ is defined by

$$\psi_f^\beta(x) := \min_{y\in\mathcal{E}} f(y) + \frac{1}{2\beta}\|y - x\|^2. \tag{2.8}$$

The unique optimal solution of (2.8), denoted by $P_f^\beta(x)$, is called the proximal point of x associated with f.

Example 2.1. Let $C \subseteq \mathcal{E}$ be a closed convex set and δ_C be its indicator function. Then, for any $x \in \mathcal{E}$, the proximal point of x associated with δ_C reduces to the metric projection of x onto C by noting the fact that

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

For the Moreau-Yosida regularization, the following properties (see, e.g. [45, 46, 51]) are often very useful.

Proposition 2.5. *Let $f : \mathcal{E} \to (-\infty, +\infty]$ be a closed proper convex function, ψ_f^β be the Moreau-Yosida regularization of f, and P_f^β be the associated proximal point mapping. Then, the following properties hold:*

(i) *P_f^β is firmly non-expansive, i.e., $\forall\, x, y \in \mathcal{E}$,*

$$\|P_f^\beta(x) - P_f^\beta(y)\|^2 \le \langle P_f^\beta(x) - P_f^\beta(y), x - y\rangle. \tag{2.9}$$

Consequently, P_f^β is globally Lipschitz continuous with modulus 1.

(ii) *ψ_f^β is a continuously differentiable convex function, and*

$$\nabla\psi_f^\beta(x) = \frac{1}{\beta}(x - P_f^\beta(x)), \quad x \in \mathcal{E}. \tag{2.10}$$

A particular elegant and useful property on the Moreau-Yosida regularization is the so-called Moreau decomposition.

Theorem 2.6. *Let $f : \mathcal{E} \to (-\infty, +\infty]$ be a closed proper convex function and f^* be its conjugate. Define $g : \mathcal{E} \to (-\infty, +\infty]$ by*

$$g(x) = f^*(x/\beta), \quad \forall\, x \in \mathcal{E}.$$

Then any $x \in \mathcal{E}$ has the decomposition

$$x = P_f^\beta(x) + P_g^\beta(x). \tag{2.11}$$

Below, we state a well know result in convex analysis on the positive homogenous functions. See [71] for its proof.

Proposition 2.7. *Let $f : \mathcal{E} \to (-\infty, +\infty]$ be a proper convex function. Then f is positively homogeneous if and only if f^* is the indicator function of*

$$C = \{x^* \in \mathcal{E} : \langle x,\, x^*\rangle \le f(x),\ \forall\, x \in \mathcal{E}\}. \tag{2.12}$$

If $f(0) = 0$, in particular if f is closed, then $C = \partial f(0)$.

Combining this proposition with Moreau decomposition, we obtain the following corollary directly, which provides a powerful tool for us to calculate the proximal point mapping associated with the positive homogenous functions.

Corollary 2.8. *Suppose that the closed convex function* $f : \mathcal{E} \to (-\infty, +\infty]$ *is positively homogenous. Then for any* C*, we have*

$$P_f^{\beta}(x) = x - \Pi_{\beta C}(x),$$

where βC *represnets the set*

$$\beta C := \{\beta x \,|\, x \in C\},$$

and C *is defined by* (2.12)

In what follows, we use Corollary 2.8 to calculate the proximal point mapping associated with the spectral norm, which plays a crucial role in the efficient implementation of the ADM. For notational convenience, we write $\mathbb{B}_\beta := \{x \in \Re^m \,|\, \|x\|_1 \leq \beta\}$ and $\mathcal{B}_\beta := \{X \in \Re^{m\times n} \,|\, \|X\|_* \leq \beta\}$. If $\beta = 1$, we just use $\mathbb{B}$ and $\mathcal{B}$ to denote the unit l_1 norm ball and nuclear norm ball, respectively.

Proposition 2.9. *Let* $f(X) = \|X\|_2$ *be defined on* $\Re^{m\times n}$ *and* $\beta > 0$*. Suppose* X *has the following singular value decomposition (SVD)*

$$X = U[\mathrm{Diag}(\sigma)\ 0]V^T, \tag{2.13}$$

where $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_m)^T$ *with* $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_m \geq 0$*. Then it holds that*

$$P_f^{\beta}(X) = X - \Pi_{\mathcal{B}_\beta}(X), \tag{2.14}$$

where $\Pi_{\mathcal{B}_\beta}$ *is the projection onto* $\mathcal{B}_\beta$*, and it is given by*

$$\Pi_{\mathcal{B}_\beta}(X) = U[\mathrm{Diag}(\Pi_{\mathbb{B}_\beta}(\sigma))\ 0]V^T. \tag{2.15}$$

Proof. By directly computing the subdifferential of f , we have

$$\partial f(0) = \{X \in \Re^{m\times n} \,|\, \|X\|_* \leq 1\},$$

and then (2.14) follows immediately from Corollary 2.8. Note that $\Pi_{\mathcal{B}_\beta}(X)$ is the unique solution of the following optimization problem:

$$\begin{array}{ll} \min & \frac{1}{2}\|Y - X\|^2 \\ \text{s.t.} & \|Y\|_* \leq \beta. \end{array}$$

By the Von-Neumann's trace inequality [63], we know that

$$\|\sigma(X) - \sigma(Y)\| \leq \|X - Y\|_F.$$

Since the metric projection onto $\mathcal{B}_\beta$ is unique, we have

$$\Pi_{\mathcal{B}_\beta}(X) = U[\mathrm{Diag}\big(\sigma(\Pi_{\mathcal{B}_\beta}(X))\big)\ 0]V^T.$$

Therefore, $\sigma(\Pi_{\mathcal{B}_\beta}(X))$ is the optimal solution of

$$\begin{array}{ll} \min & \frac{1}{2}\|y - \sigma(X)\|^2 \\ \text{s.t.} & \|y\|_1 \leq \beta, \end{array}$$

which implies that

$$\sigma(\Pi_{\mathcal{B}_\beta}(X)) = \Pi_{\mathbb{B}_\beta}(\sigma).$$

This completes the proof of this assertion. □

Now, we are ready to give the exact expression of the projection $\Pi_{\mathbb{B}_\beta}$. Let x be a given vector in $\Re^m$. Define the vector $\zeta(x)$ to be

$$\zeta_i(x) = \frac{1}{i}\Big(\sum_{j=1}^{i} x_j - \beta\Big),\ i = 1, 2, \ldots, m.$$

Let $k_1(x)$ and $k_2(x)$ denote respectively the maximal indexes of the following two sets:

$$\{i :\ \sigma_i > \zeta_i(x),\ 1 \leq i \leq m\}, \quad \{i :\ \sigma_i \geq \zeta_i(x),\ 1 \leq i \leq m\}.$$

From the breakpoint search algorithm in [41, 42], it follows that

$$\Pi_{\mathbb{B}_\beta}(x) = \begin{cases} x, & \text{if } \|x\|_1 \leq 1, \\ \max(x - \zeta_{k_1(x)}(x), 0), & \text{otherwise.} \end{cases} \tag{2.16}$$

Also see [10, 12, 17] for breakpoint algorithm using medians. Let X have the singular value decomposition (2.1). According to the analysis above, we are able to express $\Pi_{\mathcal{B}_\beta}(X)$ analytically by

$$\Pi_{\mathcal{B}_\beta}(X) = \begin{cases} X, & \text{if } \|X\|_* \leq \beta, \\ U[\mathrm{Diag}\big(\max(\sigma - \zeta_{k_1(\sigma)}(\sigma), 0)\big)\ 0]V^T, & \text{otherwise.} \end{cases} \tag{2.17}$$

Remark 2.1. From (2.17), it follows that $\Pi_{\mathbb{B}_\beta}$ is differentiable at x if and only if x satisfies one of the following two conditions:

(i) $\|x\|_1 < \beta$;

(ii) $\|x\|_1 > \beta$ and $k_1(x) = k_2(x)$.

Also see [87] for a complete characterization of differentiability of metric projection onto vector k-norm ball. Let X be a matrix in $\Re^{m\times n}$ which admits the singular value decomposition (2.1). Combining Proposition 2.4 (iii) with the conditions (i) and (ii), we can easily deduce that $\Pi_{\mathcal{B}_\beta}$ is differentiable at X if and only if X satisfies either of the following two conditions:

(i) $\|X\|_* < \beta$;

(ii) $\|X\|_* > \beta$ and $k_1(\sigma) = k_2(\sigma)$.

Remark 2.2. Since $\Pi_{\mathbb{B}_\beta}$ is piecewise linear thus strongly semismooth, by part (iii) of Proposition 2.4, $\Pi_{\mathcal{B}_\beta}$ is strongly G-semismoothn at any $X \in \Re^{m\times n}$.

Since the proximal point mapping P_f^β is Lipschtiz continuous, it is differentiable almost everywhere on $\mathcal{E}$. Therefore, the B-subdifferential $\partial_B P_f^\beta$ and the Clarke generalized Jacobian ∂P_f^β of P_f^β are well defined.

Proposition 2.10. *Let f be a closed proper convex function on $\mathcal{E}$. For any $x \in \mathcal{E}$, $\partial P_f(x)$ has the following properties:*

(i) *Any $V \in \partial P_f(x)$ is self-adjoint.*

(ii) *$\langle Vd,\, d\rangle \geq \|Vd\|^2$ for any $V \in \partial P_f(x)$ and $d \in \mathcal{E}$.*

Proof. (i) Define $\phi : \mathcal{E} \to \Re$ by $\phi(y) := \frac{1}{2}\|y\|^2 - \psi_f(y)$, $y \in \mathcal{E}$. It follows from Proposition 2.5 that ϕ is continuously differentiable with $\nabla\phi(y) = P_f(y)$, $y \in \mathcal{E}$. Therefore, $(\nabla\phi)'(y)$ is self-adjoint if it exists. It follows that any element in $\partial_B P_f(x)$, and thus that in $\partial P_f(x) = \mathrm{conv}\,\partial_B P_f(x)$, is self-adjoint.

(ii) Let $d \in \mathcal{E}$ and $z \in \mathcal{D}_{P_f} := \{y \in \mathcal{E} : P_f \text{ is differentiable at } y\}$ be arbitrarily chosen. From Proposition 2.5, for any $t \geq 0$, we have $\langle P_f(z+td) - P_f(z),\, td\rangle \geq \|P_f(z+td) - P_f(z)\|^2$, from which it follows that

$$\left\langle t\,P_f'(z)\,d,\, t\,d\right\rangle + o(t^2) \geq \left\|t\,P_f'(z)\,d + o(t)\right\|^2. \tag{2.18}$$

By taking limits for $t \to 0$ in (2.18), we obtain

$$\langle P_f'(z)\,d,\, d\rangle \geq \|P_f'(z)\,d\|^2 \quad \forall z \in \mathcal{D}_{P_f}. \tag{2.19}$$

Let $V \in \partial P_f(x)$. Then there exists a positive integer $m > 0$, $V_i \in \partial_B P_f(x)$, $\lambda_i \geq 0$, $i = 1, 2, \ldots, m$, such that $\sum_{i=1}^m \lambda_i = 1$ and $V = \sum_{i=1}^m \lambda_i V_i$. For each $i = 1, \ldots, m$ and $k = 1, 2, \ldots$, there exists $x^{i_k} \in \mathcal{D}_{P_f}$ such that $\|x - x^{i_k}\| \leq 1/k$ and

$$\|P_f'(x^{i_k}) - V_i\| \leq 1/k.$$

By (2.19), we have $\langle P_f'(x^{i_k})\, d,\, d\rangle \geq \|P_f'(x^{i_k})\, d\|^2$. By taking limits, we get $\langle V_i\, d,\, d\rangle \geq \|V_i\, d\|^2$, from which it follows that

$$\langle V d,\, d\rangle = \sum_{i=1}^m \lambda_i \langle V_i\, d,\, d\rangle \geq \sum_{i=1}^m \lambda_i \|V_i\, d\|^2 \geq \left\| \sum_{i=1}^m \lambda_i V_i\, d \right\|^2 = \|V d\|^2, \tag{2.20}$$

where the second inequality follows from Jensen's inequality applied to the convex function $\theta(y) := \frac{1}{2}\|y\|^2$, $y \in \mathcal{E}$. Since d is arbitrarily chosen, part (ii) follows from (6.40). □

Remark 2.3. Let C be a closed convex set and f be its indicator function. In this case, P_f^β reduces to the metric projector onto C, and thus Proposition 2.10 recovers the positive semidefiniteness of $\partial \Pi_C$ established in [59].

2.4 Smoothing functions

Let $\|\cdot\|_\infty$ be the l_∞ norm, i.e., for each $x \in \Re^m$,

$$\|x\|_\infty = \max\{|x_i|\,|1 \leq i \leq m\}.$$

Denote the epigraph of $\|\cdot\|_\infty$ by epi_∞. For any $(t, x) \in \Re \times \Re^m$, let $\Pi_\infty(t, x)$ denote the metric projection of (t, x) over l_∞ norm, which is the unique solution of the following optimization problem:

$$\begin{array}{ll} \min & \frac{1}{2}\|y - x\|^2 + \frac{1}{2}(s - t)^2 \\ \text{s.t.} & \|y\|_\infty \leq s, \end{array}$$

or equivalently,

$$\begin{array}{ll} \min & \frac{1}{2}\|y - x\|^2 + \frac{1}{2}(s - t)^2 \\ \text{s.t.} & -s \leq y_i \leq s, 1 \leq i \leq m. \end{array} \tag{2.21}$$

Write $H(\varepsilon, t, x) = (s(\varepsilon, t, x), y(\varepsilon, t, x))$ to denote the unique optimal solution of the logarithmic penalty problem associated with (2.21), i.e.,

$$H(\varepsilon, t, x) = \arg\min \left\{ \frac{1}{2}\|y - x\|^2 + \frac{1}{2}(s - t)^2 - \varepsilon^2 \sum_{i=1}^m \log(s - y_i) - \varepsilon^2 \sum_{i=1}^m \log(s + y_i) \right\}. \tag{2.22}$$

Proposition 2.11. *Let (t, x) be a given vector in $\Re\times\Re^m$. Then the following statements are valid.*

(i) $H(\cdot,\cdot,\cdot)$ *is continuously differentiable on $\Re_{++}\times\Re\times\Re^m$, and for any $\varepsilon > 0$ and $(t, x)\times\Re^n\times\Re$,*

$$0 \prec \frac{\partial H(\varepsilon, t, x)}{\partial(t, x)} \prec I,$$

where $\prec$ means symmetric negative definiteness.

(ii) *For any $x_0 \in \Re^m$ and $t_0 \in \Re$,*

$$\lim_{\varepsilon\downarrow 0, t\to t_0, x\to x_0} H(\varepsilon, t, x) = \Pi_\infty(t_0, x_0).$$

(iii) $H(\cdot,\cdot,\cdot)$ *is γ order semismooth at $(0, t_0, x_0)$ for some rational number $\gamma > 0$. Furthermore, if $t_0 > -\|x_0\|_1$, then $H(\cdot,\cdot,\cdot)$ is strongly semismooth at $(0, t_0, x_0)$.*

Proof. i) By the definition of $H(\varepsilon, t, x)$, we know that

$$\begin{cases} y_i - x_i + \varepsilon^2 \dfrac{2y_i}{s^2 - y_i^2} = 0, \\ s - t - \varepsilon^2 \displaystyle\sum_{i=1}^m \dfrac{2s}{s^2 - y_i^2} = 0, \quad i = 1, 2, \dots, m \\ s > |y_i|, \end{cases} \tag{2.23}$$

For each $1 \le i \le m$, write

$$a_i = \frac{1}{(y_i + s)^2} - \frac{1}{(y_i - s)^2}$$

and

$$b_i = \frac{1}{(y_i + s)^2} + \frac{1}{(y_i - s)^2}.$$

Direct computation shows that

$$\begin{bmatrix} 1 + \varepsilon^2 \displaystyle\sum_{i=1}^m b_i & \varepsilon^2 a^T \\ \varepsilon^2 a & I + \varepsilon^2 \mathrm{Diag}(b) \end{bmatrix} \frac{\partial H(\varepsilon, t, x)}{\partial(t, x)} = I, \tag{2.24}$$

which implies the continuously differentiability of $H(\varepsilon,\cdot,\cdot)$. Moreover, by simple algebraic computation, one can easily establish that

$$\begin{bmatrix} 1 + \varepsilon^2 \displaystyle\sum_{i=1}^m b_i & \varepsilon^2 a^T \\ \varepsilon^2 a & I + \varepsilon^2 \mathrm{Diag}(b) \end{bmatrix} \succ I, \tag{2.25}$$

which, together with (2.24), completes the proof of (i).

ii) Since the Slater condition of (2.21) holds naturally, by [67, Proposition 4.1], it holds that

$$\lim_{\varepsilon\downarrow 0, t\to t_0, x\to x_0} H(\varepsilon, t, x) = \Pi_\infty(t_0, x_0)$$

for any $(t_0, x_0) \in \Re^{n+1}$.

iii) If $t_0 > -\|x_0\|_1$, then $\Pi_\infty(t_0, x_0) \neq (0, 0)$, which implies that the linear independence constraint qualification (LICQ) of (2.21) holds. By [67, Proposition 4.1], we can easily obtain that $H(\cdot,\cdot,\cdot)$ is strongly semismooth at $(0, t_0, x_0)$. Next, we show $H(\cdot,\cdot,\cdot)$ is locally Lipshcitz at $(0, t_0, x_0)$ for (t_0, x_0) satisfying $t_0 \le -\|x\|_1$. For any given (ε, t, x), we know from (2.23) that

$$\begin{aligned} 2n\varepsilon^2 &= \sum_{i=1}^m (s - y_i)\frac{\varepsilon^2}{s - y_i} + \sum_{i=1}^m (s + y_i)\frac{\varepsilon^2}{s + y_i} \\ &= [s\ y^T]\Big(\sum_{i=1}^m \frac{\varepsilon^2}{s - y_i}\begin{bmatrix} 1 \\ -e_i \end{bmatrix} + \sum_{i=1}^n \frac{\varepsilon^2}{s + y_i}\begin{bmatrix} 1 \\ e_i \end{bmatrix}\Big) \\ &= s^2 + \|y\|^2 - st - y^T x, \end{aligned} \tag{2.26}$$

where e_i is the usual ith base vector. Write $t = t_0 + \Delta t$ and $x = x_0 + \Delta x$. Then by direct computation applied to (2.26), we have

$$\begin{aligned} 2n\varepsilon^2 &= s^2 + \|y\|^2 - st_0 - y^T x_0 - s\Delta t - y^T \Delta x \\ &\ge s^2 + \|y\|^2 + s\|x_0\|_1 - \|y\|_\infty \|x_0\|_1 - s|\Delta t| - \|y\|_\infty \|\Delta x\|_1 \\ &\ge s^2 + \|y\|^2 - s(|\Delta t| + \|\Delta x\|_1) \\ &\ge s^2 - s(|\Delta t| + \sqrt{n}\|\Delta x\|). \end{aligned} \tag{2.27}$$

It therefore holds that

$$\|y\|_\infty \le s \le \|\Delta t\| + \sqrt{n}\|\Delta x\| + \sqrt{2n}|\varepsilon|,$$

which, together with $G(0, t_0, x_0) = \Pi_\infty(t_0, x_0) = (0, 0)$, implies the local Lipschtizness of $H(\cdot,\cdot,\cdot)$ at $(0, t_0, x_0)$. Since, for any $\varepsilon \neq 0$, $H(\varepsilon, t, x)$ is the unique solution of the fractional system (2.23) and $H(0, t, x) = \Pi_\infty(t, x)$ is a semialgebraic function with respect to (t, x), it can be checked directly that $H(\cdot,\cdot,\cdot)$ is semialgebraic. Therefore, by invoking Proposition 2.1, one can easily obtain the γ order semismoothness of $H(\cdot,\cdot,\cdot)$ at any $(0, t_0, x_0)$ for some rational number γ. $\square$

Proposition 2.12. *Let (t, x) be a given vector in $\Re \times \Re^m_+$.*

(i) *For any* $i = 1, 2, \ldots, m$, $0 \le y_i(\varepsilon, t, x) \le x_i$. *In particular, if* $x_i > 0$, *then* $0 < y_i(\varepsilon, t, x) < x_i$.

(ii) *If* $x_i < x_j$, *then* $0 < y_j(\varepsilon, t, x) - y_i(\varepsilon, t, x) < x_j - x_i$.

Proof. i) Note that $s(\varepsilon, t, x)^2 > y_i(\varepsilon, t, x)^2$ for any $1 \le i \le m$. Then one can easily deduce from the first equality of(2.23) that $y_i(\varepsilon, t, x) - x_i$ and $y_i(\varepsilon, t, x)$ have opposite signs, which implies the first assertion.
ii) For fixed (ε, x, s), it is easy to check that y_i and $\frac{\varepsilon^2 2y_i}{s^2 - y_i^2}$ are strictly increasing with respect to y_i for any i. It therefore follows from the first equality of (2.23) that $y_i < y_j$ when $x_i < x_j$ and thus $y_i - x_i > y_j - x_j$, which is exactly the second assertion of this proposition. □

Let Z be any given matrix in $\Re^{m\times n}$. We use $\sigma_1(Z) \ge \sigma_2(Z) \ge \ldots \ge \sigma_m(Z)$ to denote the singular values of Z (counting multiplicity). Let Z admit the following singular value decomposition (SVD)

$$Z = U[\text{diag}(\sigma(Z))\, 0]V^T,$$

where $\sigma(Z) = [\sigma_1(Z), \sigma_2(Z), \ldots, \sigma_m(Z)]^T$ and $U \in \Re^{m\times m}$ and $V \in \Re^{n\times}$ are orthogonal matrices. Define the spectral operator $G(\cdot,\cdot,\cdot) : \Re \times \Re^{m\times n} \times \Re \to \Re \times \Re^{m\times n}$ with respect to smoothing function $H(\cdot,\cdot,\cdot)$ as follows

$$G(\varepsilon, t, Z) := \begin{bmatrix} s(\varepsilon, t, \sigma(Z)) \\ U[\text{Diag}(y(\varepsilon, t, \sigma(Z))), 0]V^T \end{bmatrix} \tag{2.28}$$

for any $Z \in \Re^{m\times}$, $t \in \Re$ and $\varepsilon \in \Re$. The next proposition shows that $G(\cdot,\cdot,\cdot)$ is indeed the smoothing function of the metric projector over the epigraph of the spectral norm.

Proposition 2.13. *Let* $Z \in \Re^{m\times n}$ *and* $t \in \Re$.

(i) *The spectral operator* $G(\cdot,\cdot,\cdot)$ *is well-defined and, for any given* $\varepsilon > 0$, $t \in \Re$ *and* $Z \in \Re^{m\times n}$, $G(\cdot,\cdot,\cdot)$ *is continuously differentiable at* (ε, Z, t).

(ii) *For any* $X^0 \in \Re^{m\times n}$ *and* $t^0 \in \Re$,

$$\lim_{\varepsilon\downarrow 0, t\to t^0, X\to X^0} G(\varepsilon, t, X) = \Pi_2(t^0, X^0).$$

(iii) $G(\cdot,\cdot,\cdot)$ *is* γ *order G-semismooth at* $(0, t_0, X_0)$ *for any* (t_0, X_0) *in* $\Re \times \Re^{m\times n}$ *for some rational number* $\gamma > 0$. *Furthermore, if* $t_0 > -\|X_0\|_2$, *then* $G(\cdot,\cdot,\cdot)$ *is strongly G-semismooth at* $(0, t_0, X_0)$

Proof. i) Clearly, for given ε and t, $H(\varepsilon, \cdot, t)$ is absolutely symmetric with respect to $\Re^m$. Therefore, G is well defined. Since $H(\cdot, \cdot, \cdot)$ is continuously differentiable on $\Re\backslash\{0\} \times \Re \times \Re^m$, according to Proposition 2.4 (i), $G(\cdot, \cdot, \cdot)$ is continuously differentiable on $\Re\backslash\{0\} \times \Re \times \Re^m$.
ii) For given $\varepsilon \neq 0$ and $t \in \Re$, it follows from Von Neumann's trace inequality that

$$\|G(\varepsilon, t, X) - G(\varepsilon, t, Z)\| \leq \|H(\varepsilon, t, \sigma(X)) - H(\varepsilon, t, \sigma(Z))\|.$$

As stated in Proposition 2.11, $H(\cdot, \cdot, \cdot)$ is the smoothing function of $\Pi_\infty(\cdot, \cdot)$. Combining all these arguments, we deduce that

$$\lim_{\varepsilon \downarrow 0, t \to t^0, X \to X^0} G(\varepsilon, t, X) = \Pi_2(t^0, X^0).$$

iii) The last assertion follows directly from Proposition 2.11 (iii) and Proposition 2.4 (iii). We omit the details. $\square$

Next, we briefly review the CHKS smoothing functions for plus function. Let $h_u(\cdot, \cdot) : \Re^2 \to \Re$ be the CHKS function defined by

$$h_u(\varepsilon, w) = \frac{\sqrt{w^2 + 4\varepsilon^2} + w}{2}, \quad \forall\, (w, \varepsilon) \in \Re^2,$$

One can easily extend these smoothing functions for scalar plus function to vector-valued plus function. Indeed, define $H_u(\cdot, \cdot) : \Re^m \times \Re \to \Re^m$ by

$$H_u(\varepsilon, x) = \begin{pmatrix} h_u(\varepsilon, x_1) \\ h_u(\varepsilon, x_2) \\ \vdots \\ h_u(\varepsilon, x_m) \end{pmatrix}$$

for any $x \in \Re^m$ and $\varepsilon \in \Re$. It is easy to verify that H_u is smoothing functions for the vector plus function $\max(x, 0)$, which is continuously differentiable on $\Re_{++} \times \Re^m$ and enjoy the strongly semismooth property on $\{0\} \times \Re^m$.

Chapter 3

An alternating direction method

3.1 Introduction

Roughly speaking, alternating direction method (ADM) is an inexact implementation of the augmented Lagrangian method using the idea of Gauss-Seidel iteration. It solves problems of the form

$$\begin{array}{ll} \min & f(x) + g(y) \\ s.t. & Ax + By = b \end{array} \tag{3.1}$$

with variables $x \in \Re^n$ and $y \in \Re^m$, where $f : \Re^n \to \Re \cup \{+\infty\}$ and $g : \Re^m \to \Re \cup \{+\infty\}$ are proper closed convex functions, $A \in \Re^{p\times n}$, $B \in \Re^{p\times m}$, and b is vector in $\Re^p$. The augmented Lagrangian function of (3.1) is given by

$$\mathcal{L}_\beta(x, y, \lambda) = f(x) + g(y) - \langle \lambda,\, Ax + By - b\rangle + \frac{\beta}{2}\|Ax + By - b\|^2,$$

where $\lambda \in \Re^p$ is a Lagrangian multiplier and $\beta > 0$ is a penalty parameter. The classical augmented Lagrangian method [43, 64] consists of the iterations

$$\begin{cases} (x^{k+1}, y^{k+1}) & = \arg\min \mathcal{L}_\beta(x, y, \lambda^k), \\ \qquad \lambda^{k+1} & = \lambda^k - \gamma\beta(Ax^{k+1} + By^{k+1} - b), \end{cases} \tag{3.2}$$

where $\gamma \in (0, 2)$ guarantees the convergence. Seen clearly form (3.2), at each iteration the simple scheme of ALM involves a joint minimization with respect to x and y and therefore ignores the separable structure of (3.1). In contrast, the idea of ADM is to decompose the minimization task of $L_\beta(\cdot, \cdot, \lambda)$ into two easier and smaller subproblems such that the involved variables x and y can be minimized separately in the alternative order. The decomposed subproblems are usually much easier than the joint minimization task in (3.2) and even for some applications admit analytic solutions [14, 40, 54, 89, 93]. Given a couple (y^k, λ^k), the ADM applied to problem (3.2) yields the following iterative scheme

$$\begin{cases} x^{k+1} & = \arg\min \mathcal{L}_\beta(x, y^k, \lambda^k), \\ y^{k+1} & = \arg\min \mathcal{L}_\beta(x^{k+1}, y, \lambda^k), \\ \lambda^{k+1} & = \lambda^k - \beta(Ax^{k+1} - By^{k+1} - b). \end{cases} \tag{3.3}$$

Since its presence [30, 34] in the field of differential equation, ADM and its variants have been widely studied in many areas, such as convex programming and variational inequality. In [32, 33], the authors proposed an extension of the ADM (3.3), in which a stepsize $\gamma \in (0, \frac{\sqrt{5}+1}{2})$ is attached to the update of the Lagrangian multiplier λ. Specifically, the extended ADM iterates as

$$\begin{cases} x^{k+1} &= \arg\min \mathcal{L}_\beta(x, y^k, \lambda^k), \\ y^{k+1} &= \arg\min \mathcal{L}_\beta(x^{k+1}, y, \lambda^k), \\ \lambda^{k+1} &= \lambda^k - \gamma\beta(Ax^{k+1} - By^{k+1} - b). \end{cases} \tag{3.4}$$

Gabay [29] considered the ADM from the perspective of operator splitting and stated that the classical ADM is the tight Douglas-Rachford splitting method [22, 55] for finding a zero of the sum of two maximal monotone operators applied to the dual of (3.1). Sequently, in [24], it is shown that the Douglas-Rachford splitting method is a special implementation of the proximal point algorithm. Replacing the classical PPA by the relaxed PPA introduced in [35], Eckstein and Bertsekas obtained the following generalized alternating direction method:

$$\begin{cases} \|x^{k+1} - \arg\min\{f(x) - \langle\lambda^k, Ax\rangle + \frac{\beta}{2}\|Ax + By^k - b\|^2\}\| \le u_k, \\ \|y^{k+1} - \arg\min\{g(y) - \langle\lambda^k, By\rangle + \frac{\beta}{2}\|\rho_k(Ax^{k+1} - b) + B(y - (1-\rho_k)y^k)\|^2\}\| \le v_k, \\ \lambda^{k+1} = \lambda^k - \beta(\rho_k Ax^{k+1} - (1-\rho_k)By^k + By^{k+1}), \end{cases} \tag{3.5}$$

where $\rho_k \in (0, 2)$, $u_k > 0$, $v_k > 0$, $\sum_{i=1}^{\infty} u_k < \infty$ and $\sum_{i=1}^{\infty} v_k < \infty$. If $\rho_k = 1$, the generalized ADM (3.5) reduces to the original version of ADM. Suggested by Rockafellar, Eckstein [23] also considered a primal-dual saddle-point application of the Douglas-Rachford splitting to the separable convex programming. This resulting algorithm is known as the proximal alternating direction algorithm:

$$\begin{cases} x^{k+1} &= \arg\min\{f(x) - \langle\lambda^k, Ax\rangle + \frac{\beta}{2}\|Ax + By^k - b\|^2 + \frac{s}{2}\|x - x^k\|^2\}, \\ y^{k+1} &= \arg\min\{g(y) - \langle\lambda^k, By\rangle + \frac{\beta}{2}\|Ax^{k+1} + By - b\|^2 + \frac{t}{2}\|y - y^k\|^2\}, \\ \lambda^{k+1} &= \lambda^k - \beta(Ax^{k+1} + By^{k+1} - b). \end{cases} \tag{3.6}$$

By further investigating the contractive property of the proximal ADM, He et al. [39] presented a new variant of ADM in which x^{k+1} and y^{k+1} is produced by inexact minimizing the subproblems and the parameter β, s, t are replaced by positive definite matrices. More specifically,

$$\begin{cases} \|x^{k+1} - \arg\min\{f(x) - \langle \lambda^k, Ax\rangle + \frac{1}{2}\|Ax + By^k - b\|_{H_k}^2 + \frac{1}{2}\|x - x^k\|_{R_k}^2\|\} \le u_k, \\ \|y^{k+1} - \arg\min\{g(y) - \langle \lambda^k, By\rangle + \frac{1}{2}\|Ax^{k+1} + By - b\|_{H_k}^2 + \frac{1}{2}\|y - y^k\|_{S_k}^2\}\| \le v_k, \\ \lambda^{k+1} = \lambda^k - H_k(Ax^{k+1} + By^{k+1} - b). \end{cases} \tag{3.7}$$

where $u_k > 0$, $v_k > 0$, $\sum_{k=1}^{\infty} u_k < \infty$ and $\sum_{k=1}^{\infty} v_k < \infty$; $\{S_k\}$ and $\{T_k\}$ are sequences of both lower and upper bounded symmetric positive definite matrices; for a symmetric positive definite matrix G and a given vector x, $\|x\|_G = \sqrt{x^T G x}$; H_k is required to satisfy some technical condition introduced in [39]. From the perspective of contraction, Ye and Yuan [91] developed a variant of alternating direction method with an optimal stepsize. Given a couple of (y^k, λ^k), the new iterate of Ye-Yuan's algorithm is produced by

$$\begin{cases} x^{k+1} &= \arg\min \mathcal{L}_\beta(x, y^k, \lambda^k), \\ \tilde{y}^{k+1} &= \arg\min \mathcal{L}_\beta(x^{k+1}, y, \lambda^k), \\ \tilde{\lambda}^{k+1} &= \lambda^k - \beta(Ax^{k+1} + B\tilde{y}^{k+1} - b), \end{cases} \tag{3.8}$$

and

$$\begin{cases} y^{k+1} &= y^k - \gamma\alpha^*(y^k - \tilde{y}^k), \\ \lambda^{k+1} &= \lambda^k - \gamma\alpha^*(\lambda^k - \tilde{\lambda}^k), \end{cases} \tag{3.9}$$

where $\gamma \in (0, 2)$ and α^* is defined by

$$\alpha^* := \frac{1}{2}\Big[1 + \frac{\beta\|Ax^{k+1} + By^k - b\|^2}{\beta\|By^k - B\tilde{y}^k\|^2 + \frac{1}{\beta}\|\lambda^k - \tilde{\lambda}^k\|^2}\Big].$$

Numerical results demonstrated that an additional computation on the optimal size would improve the efficiency of the new variant of ADM. More recently, an ADM based relaxed customized proximal point aiming at accelerating the ADM was proposed by Cai et al. [11] for solving the separable convex programming (3.1). With the given couple (y^k, λ^k), this algorithm first generates a prediction point $(\tilde{x}^k, \tilde{y}^k, \tilde{\lambda}^k)$ by

$$\begin{cases} \tilde{x}^k &= \arg\min \mathcal{L}_\beta(x, y^k, \lambda^k), \\ \tilde{\lambda}^k &= \lambda^k - \beta(A\tilde{x}^k - By^k - b), \\ \tilde{y}^k &= \arg\min \mathcal{L}_\beta(\tilde{x}^k, y, \tilde{\lambda}^k), \end{cases} \tag{3.10}$$

and then updates (y^{k+1}, λ^{k+1}) according to the following rules

$$\begin{cases} y^{k+1} &= y^k - \gamma(y^k - \tilde{y}^k), \\ \lambda^{k+1} &= \lambda^k - \gamma(\lambda^k - \tilde{\lambda}^k), \end{cases} \tag{3.11}$$

where $\gamma \in (0, 2)$. As we will see later, the ADM based PPA is equivalent to a special implementation of Eckstein and Bertsekas' generalized ADM.

3.2 Equivalence of Eckstein-Bertseka's ADM and ADM based customized PPA

For ease of discussion and notations, we assume without generality that $A = M$ and $B = -I$. Then the separable convex programming collapses to:

$$\begin{aligned} \min \quad & f(x) + g(y) \\ \text{s.t.} \quad & Mx = y. \end{aligned}$$

and the exact implementation of Eckstein-Bertsekas's generalized ADM goes as follows:

$$\begin{cases} x^{k+1} & = \arg\min\{f(x) + \frac{\beta}{2}\|Mx - y^k + p^k\|^2\}, \\ w^{k+1} & = \arg\min\{g(y) + \frac{\beta}{2}\|\rho Mx^{k+1} + (1-\rho)y^k - y + p^k\|^2\}, \\ p^{k+1} & = p^k + \rho My^{k+1} + (1-\rho)y^k - y^{k+1}, \end{cases} \tag{3.12}$$

where $\rho \in (0, 2)$ is a constant relax factor. Given the couple (x^k, p^k), the next iteration of Cai et al.'s relaxed ADM is produced by

$$\begin{cases} \tilde{w}^k & = \arg\min\{g(y) + \frac{\beta}{2}\|Mx^k - y + p^k\|^2\}, \\ \tilde{p}^k & = p^k + Mx^k - \tilde{y}^k, \\ \tilde{x}^k & = \arg\min\{f(x) + \frac{\beta}{2}\|Mx - \tilde{y}^k + \tilde{p}^k\|^2\}, \\ x^{k+1} & = (1-\rho)x^k + \rho\tilde{x}^k, \\ p^{k+1} & = (1-\rho)p^k + \rho\tilde{p}^k. \end{cases} \tag{3.13}$$

Next we analyze Cai et al.'s algorithm in detail under the setting $(p^0, x^0) = (0, 0)$. Using the iterative formula in (3.13), it is easy to check that

$$Mx^{k+1} + p^{k+1} = (1-\rho)(Mx^k + p^k) + \rho(M\tilde{x}^k + \tilde{p}^k),$$

which means

$$Mx^{k+1} + p^{k+1} = \sum_{j=0}^{k} \rho(1-\rho)^j (M\tilde{x}^{k-j} + \tilde{p}^{k-j}). \tag{3.14}$$

Then by the above equality and an easy manipulation, the iterative scheme (3.13) can be reformulated as

$$\begin{cases} \tilde{w}^k & = \arg\min\{g(w) + \frac{\beta}{2}\|\sum_{j=0}^{k-1}\rho(1-\rho)^j(M\tilde{x}^{k-1-j} + \tilde{p}^{k-1-j}) - w\|^2\}, \\ \tilde{p}^k & = \sum_{j=0}^{k-1}\rho(1-\rho)^j(M\tilde{x}^{k-1-j} + \tilde{p}^{k-1-j}) - \tilde{w}^k, \\ \tilde{x}^k & = \arg\min\{f(x) + \frac{\beta}{2}\|Mx - \tilde{w}^k + \tilde{p}^k\|^2\}. \end{cases} \tag{3.15}$$

By changing the implementation order of w and p and replacing the notation $(\tilde{x}^k, \tilde{p}^k, \tilde{w}^k)$ by (x^k, p^k, w^k), we have an alternative form of (3.13):

$$\begin{cases} x^{k+1} & = \arg\min\{f(x) + \frac{\beta}{2}\|Mx - w^k + p^k\|^2\}, \\ w^{k+1} & = \arg\min\{g(w) + \frac{\beta}{2}\|\sum_{j=0}^{k}\rho(1-\rho)^j(Mx^{k+1-j} + p^{k-j}) - w\|^2\}, \\ p^{k+1} & = \sum_{j=0}^{k}\rho(1-\rho)^j(Mx^{k+1-j} + p^{k-j}) - w^{k+1}. \end{cases} \tag{3.16}$$

On the other hand, it follows from a direct computation that

$$\begin{aligned} \sum_{j=0}^{k}\rho(1-\rho)^j(Mx^{k+1-j} + p^{k-j}) = & \ \rho(Mx^{k+1} + p^k) + (1-\rho)\sum_{j=0}^{k-1}\rho(1-\rho)^j(Mx^{k-j} + p^{k-1-j}) \\ = & \ \rho(Mx^{k+1} + p^k) + (1-\rho)(p^k + w^k) \\ = & \ p^k + \rho Mx^{k+1} + (1-\rho)w^k, \end{aligned}$$

which implies

$$p^{k+1} = p^k + \rho Mx^{k+1} + (1-\rho)w^k - w^{k+1}.$$

Substituting the above equality into (3.16) yields nothing but Eckstein and Berteskas' generalized ADM.

3.3 Proximal alternating direction method

For the proximal alternating direction method (3.7), the positive definite matrix R_k, S_k and H_k are allowed to variate according to some particular rules. However, in many situations, H_k is set to be βI and the sequences $\{S_k\}$ and $\{T_k\}$ are constant matrices. In this case, we suppress the subindex and write S and T for S_k and T_k respectively. The convergence results provided in [39] need both S and T to be positive semidefinite, nevertheless this assumption may exclude some important applications (see [89, 94] for example). In [28], by slightly revising the proof in [39], the authors prove the convergence of the proximal alternating direction method under any of the following conditions:

(a) f and g are strongly convex;

(b) f is strongly convex and $B^TB + T$ is positive definite;

(c) g is strongly convex and $A^TA + S$ is positive definite;

(d) S is positive definite and B is injective;

(e) T is positive definite and A is injective;

(f) S and T are positive definite.

However, a moment's observation reveals that the conditions listed above don't cover the basic convergence result of the original ADM in which $S = 0, T = 0$ and A, B are required to be of full column rank. To fill the gap, we provide the following theorem which summarizes more general convergence results of the proximal ADM. Although the proof is a trivial extension of that in [39], we still include it here for the purpose of clarity and completeness.

Theorem 3.1. *Assume that problem* (3.1) *has at least a KKT point. Let* (x_k, y_k, λ_k) *be generated by the following proximal ADM:*

$$\begin{cases} x^{k+1} &= \arg\min\{f(x) - \langle \lambda^k, Ax\rangle + \frac{\beta}{2}\|Ax + By^k - b\|^2 + \frac{1}{2}\|x - x^k\|_S^2\}, \\ y^{k+1} &= \arg\min\{g(y) - \langle \lambda^k, By\rangle + \frac{\beta}{2}\|Ax^{k+1} + By - b\|^2 + \frac{1}{2}\|y - y^k\|_T^2\}, \\ \lambda^{k+1} &= \lambda^k - \beta(Ax^{k+1} + By^{k+1} - b). \end{cases} \tag{3.17}$$

where S and T are positive semidefinite. Then $\{(x_k, y_k)\}$ converges to an optimal solution to (3.5) *and $\{\lambda_k\}$ converges to an optimal solution to the dual of* (3.5) *if the following conditions hold:*

(a) *f is strongly convex or* $\begin{bmatrix} A \\ S \end{bmatrix}$ *has full column rank;*

(b) *g is strongly convex or* $\begin{bmatrix} B \\ T \end{bmatrix}$ *has full column rank.*

Proof. We first note that the positive semidefiniteness of A^TA+S (B^TB+T) is equivalent to the augmented matrices $[A^T\ S^T]^T$ ($[B^T\ T^T]^T$) has full column rank. Hence, it suffices to prove the convergence of the proximal ADM (3.17) under the condition that both $[A^T\ S^T]^T$ and $[B^T\ T^T]^T$ are of full column rank since other cases have been investigated in [28]. Let (x^*, y^*, λ^*) be a KKT point of (3.1). Then it holds that

$$\begin{cases} A^T\lambda^* \in \partial f(x^*), \\ B^T\lambda^* \in \partial g(y^*), \\ Ax^* + By^* = b. \end{cases} \tag{3.18}$$

Since x^{k+1} and y^{k+1} solve $L(\cdot, y^k, \beta) + \frac{1}{2}\|\cdot - x^k\|_S^2$ and $L(x^{k+1}, \cdot, \beta) + \frac{1}{2}\|\cdot - y^k\|_T^2$ respectively, we deduce from the first order optimality conditions that

$$\begin{cases} A^T[\lambda^k - \beta(Ax^{k+1} + By^k - b)] - S(x^{k+1} - x^k) \in \partial f(x^{k+1}), \\ B^T[\lambda^k - \beta(Ax^{k+1} + By^{k+1} - b)] - T(y^{k+1} - y^k) \in \partial g(y^{k+1}), \\ \lambda^{k+1} = \lambda^k - \beta(Ax^{k+1} + By^{k+1} - b). \end{cases} \tag{3.19}$$

By (3.19) and the monotonicity of $\partial f(\cdot)$ and $\partial g(\cdot)$, it is easily seen that

$$\begin{aligned} 0 \leq\ & \langle x^{k+1} - x^*,\, A^T[\lambda^k - \lambda^* - \beta(Ax^{k+1} + By^k - b)] - S(x^{k+1} - x^k)\rangle \\ & + \langle y^{k+1} - y^*,\, B^T[\lambda^k - \lambda^* - \beta(Ax^{k+1} + By^{k+1} - b)] - T(y^{k+1} - y^k)\rangle \\ & + \langle \lambda^{k+1} - \lambda^*,\, \frac{\lambda^k - \lambda^{k+1}}{\beta} - (Ax^{k+1} + By^{k+1} - b)\rangle, \\ =\ & \langle x^{k+1} - x^*,\, A^T[\lambda^{k+1} - \lambda^* - \beta B(y^k - y^{k+1})] - S(x^{k+1} - x^k)\rangle \\ & + \langle y^{k+1} - y^*,\, B^T[\lambda^{k+1} - \lambda^*] - T(y^{k+1} - y^k)\rangle \\ & + \langle \lambda^{k+1} - \lambda^*,\, \frac{\lambda^k - \lambda^{k+1}}{\beta} - [A(x^{k+1} - x^*) + B(y^{k+1} - y^*)]\rangle, \end{aligned} \tag{3.20}$$

which implies

$$\begin{aligned} & \langle \lambda^k - \lambda^{k+1},\, B(y^k - y^{k+1})\rangle + \langle B(y^{k+1} - y^*),\, \beta B(y^{k+1} - y^k)\rangle \\ + \ & \langle x^{k+1} - x^*, S(x^{k+1} - x^k)\rangle + \langle y^{k+1} - y^*,\, T(y^{k+1} - y^k)\rangle \\ + \ & \langle \lambda^{k+1} - \lambda^*,\, \frac{1}{\beta}(\lambda^{k+1} - \lambda^k)\rangle \leq 0. \end{aligned} \tag{3.21}$$

Using the elementary relationship $\langle u,\, v\rangle = \frac{1}{2}(\|u\|^2 + \|v\|^2 - \|u - v\|^2)$, we further obtain that

$$\begin{aligned} & \|x^{k+1} - x^*\|_S^2 + \|y^{k+1} - y^*\|_{T+\beta B^TB}^2 + \|\lambda^{k+1} - \lambda^*\|_{\frac{I}{\beta}}^2 \\ \leq\ & \|x^k - x^*\|_S^2 + \|y^k - y^*\|_{T+\beta B^TB}^2 + \|\lambda^k - \lambda^*\|_{\frac{I}{\beta}}^2 - 2\langle \lambda^k - \lambda^{k+1},\, B(y^k - y^{k+1})\rangle \\ & - \|x^k - x^{k+1}\|_S^2 - \|y^k - y^{k+1}\|_{T+\beta B^TB}^2 - \|\lambda^k - \lambda^{k+1}\|_{\frac{I}{\beta}}^2. \end{aligned} \tag{3.22}$$

By the above inequality, it follows that the sequences $\{Sx^k\}, \{(T+\beta B^TB)y^k\}$ and $\{\lambda^k/\beta\}$ are bounded. Since $T+\beta B^TB$ and $\frac{I}{\beta}$ are positive semidefinite, we deduce that $\{y^k\}$ and $\{\lambda^k\}$ are also bounded. Recall that

$$Ax^{k+1} = \lambda^k - \lambda^{k+1} - By^{k+1} + b. \tag{3.23}$$

Hence, $\{(S+\beta A^TA)x^k\}$ is bounded and this together with the positive definiteness of $S+\beta A^TA$ implies the boundedness of $\{x^k\}$. Moreover, from (3.22), we see immediately that

$$\begin{aligned} &\sum_{k=1}^{\infty} \|x^k - x^{k+1}\|_S^2 + \|y^k - y^{k+1}\|_{T+\beta B^TB}^2 + \|\lambda^k - \lambda^{k+1}\|_{\frac{I}{\beta}}^2 \\ + \quad & 2\sum_{k=1}^{\infty} \langle \lambda^k - \lambda^{k+1},\, B(y^k - y^{k+1})\rangle < +\infty. \end{aligned} \tag{3.24}$$

One the other hand, by the monotonicity of $\partial g(\cdot)$ combined with (3.19), we have

$$\begin{aligned} & \langle y^k - y^{k+1},\, B^T(\lambda^k - \lambda^{k+1})\rangle \\ \geq\ & \|y^k - y^{k+1}\|_T^2 + \langle y^k - y^{k+1}, T(y^k - y^{k-1})\rangle \\ \geq\ & \|y^k - y^{k+1}\|_T^2 - \frac{1}{2}\|y^k - y^{k+1}\|_T^2 - \frac{1}{2}\|y^{k-1} - y^k\|_T^2 \\ =\ & \frac{1}{2}\|y^k - y^{k+1}\|_T^2 - \frac{1}{2}\|y^{k-1} - y^k\|_T^2, \end{aligned} \tag{3.25}$$

which , by the boundedness of $\{y^k\}$, implies that

$$\sum_{k=1}^{\infty} \langle y^k - y^{k+1},\, B^T(\lambda^k - \lambda^{k+1})\rangle > -\infty.$$

It therefore holds

$$\sum_{k=1}^{\infty} \|x^k - x^{k+1}\|_S^2 < +\infty,\ \sum_{k=1}^{\infty} \|y^k - y^{k+1}\|_{\beta B^TB+T} < +\infty$$

and

$$\sum_{k=1}^{\infty} \|\lambda^k - \lambda^{k+1}\|_{\frac{I}{\beta}}^2 < +\infty.$$

This together with (3.23) shows

$$\lim_{k\to\infty} \|x^k - x^{k+1}\| = 0,\ \lim_{k\to\infty} \|y^k - y^{k+1}\| = 0,\ \lim_{k\to\infty} \|\lambda^k - \lambda^{k+1}\| = 0. \tag{3.26}$$

Since the sequences $\{x^k\}$, $\{y^k\}$ and $\{\lambda^k\}$ are bounded, there exists a triple $(x^\infty, y^\infty, \lambda^\infty)$ and a subsequence n_k such that

$$\lim_{k\to\infty} x_{n_k} = x^\infty,\ \lim_{k\to\infty} y_{n_k} = y^\infty,\ \lim_{k\to\infty} \lambda_{n_k} = \lambda^\infty.$$

Then, by taking the limits on the both sides of (3.19), using (3.26) and invoking the upper semicontinuous of $\partial g(\cdot)$ [71], one can immediately write

$$\begin{cases} A^T\lambda^\infty \in \partial f(x^\infty), \\ B^T\lambda^\infty \in \partial g(y^\infty), \\ Ax^\infty + By^\infty = b, \end{cases} \tag{3.27}$$

which means $(x^\infty, y^\infty, \lambda^\infty)$ is a KKT point of (3.5). Hence, the inequality (3.22) is also valid if $(x^\infty, y^\infty, \lambda^\infty)$ is replaced by (x^*, y^*, λ^*). Then it holds that

$$\begin{aligned} & \|x^{k+1} - x^\infty\|_S^2 + \|y^{k+1} - y^\infty\|_{T+\beta B^T B}^2 + \|\lambda^{k+1} - \lambda^\infty\|_{\frac{I}{\beta}}^2 \\ \leq \; & \|x^k - x^\infty\|_S^2 + \|y^k - y^\infty\|_{T+\beta B^T B}^2 + \|\lambda^k - \lambda^\infty\|_{\frac{I}{\beta}}^2. \end{aligned} \tag{3.28}$$

Since $T + \beta B^T B$ is positive definite, we deduce from (3.28) that

$$\lim_{k\to\infty} \|x^k - x^\infty\|_S^2 = 0, \tag{3.29}$$

and

$$\lim_{k\to\infty} y^k = y^\infty, \; \lim_{k\to\infty} \lambda^k = \lambda^\infty.$$

By the relationship (3.23) and $Ax^\infty + By^\infty = b$, it is easy to see

$$\lim_{k\to\infty} Ax^k = Ax^\infty,$$

which together with (3.29) and the positive definiteness of $S + \beta B^T B$ implies

$$\lim_{k\to\infty} x^k = x^\infty$$

Therefore, we have shown that the whole sequence $\{(x^k, y^k, \lambda^k)\}$ converges to $(x^\infty, y^\infty, \lambda^\infty)$ under the assumption of this theorem. □

Remark 3.1. Theorem (3.1) provides more general conditions for the convergence of the proximal ADM. It includes all the conditions in [28, Theorem 8.1] as its special case. For $S = 0$ and $T = 0$, the proximal alternating direction method reduces to the original ADM whose convergence can be established under the condition A, B have full column rank. This basic convergence result is also included in Theorem 3.1 while not contained in the six conditions provided in [28].

Corollary 3.2. *Assume that problem* (3.5) *has at least a KKT point. Let* (x_k, y_k, λ_k) *be generated by the following linearized alternating direction method:*

$$\begin{cases} x^{k+1} &= \arg\min\{f(x) - \langle \lambda^k, Ax\rangle + \beta\langle A^T(Ax^k + By^k - b),\, x - x^k\rangle + \frac{r}{2}\|x - x^k\|^2, \\ y^{k+1} &= \arg\min\{g(y) - \langle \lambda^k, By\rangle + \beta\langle B^T(Ax^{k+1} + By^k - b),\, y - y^k\rangle + \frac{s}{2}\|y - y^k\|^2, \\ \lambda^{k+1} &= \lambda^k - \beta(Ax^{k+1} + By^{k+1} - b), \end{cases} \tag{3.30}$$

where $r \geq \beta\|A^TA\|_2 > 0$ *and* $s \geq \beta\|B^TB\|_2 > 0$ *. Then* $\{(x_k, y_k)\}$ *converges to an optimal solution to* (3.5) *and* $\{\lambda^k\}$ *converges to an optimal solution to the dual of* (3.5).

Proof. The linearized ADM (3.30) is a special case of the proximal ADM (3.17) where S is taken as $rI - \beta A^TA$ and T is $sI - \beta B^TB$. Note that the conditions $r \geq \beta\|A^TA\|_2 > 0$ and $s \geq \beta\|B^TB\|_2 > 0$ ensure the full column rank property of $[A^T\, S^T]^T$ and $[B^T\, T^T]^T$. Therefore, this corollary follows immediately from Theorem 3.1. □

3.4 ADM for the matrix norm approximation problem

In this section, we employ the ADM to solve the matrix norm approximation problem. Note that problem (1.1) can be expressed in the following equivalent form:

$$\begin{aligned} \min \quad & \|X\|_2 \\ \text{s.t.} \quad & \mathcal{A}^*y + X = A_0, \\ & By - b = z, \quad z \in \mathcal{Q}. \end{aligned} \tag{3.31}$$

The augmented Lagrangian function associated with (3.31) is given by

$$\begin{aligned} \mathcal{L}_\beta(y, X, z; Z, w) := \|X\|_2 - \langle Z, \mathcal{A}^*y + X - A_0\rangle - \langle w, By - b - z\rangle \\ + \frac{\beta}{2}\|\mathcal{A}^*y + X - A_0\|_F^2 + \frac{\beta}{2}\|By - b - z\|^2, \end{aligned} \tag{3.32}$$

where Z and w are Lagrangian multipliers, and $\beta > 0$ is the penalty parameter. Given $X^0, Z^0 \in \Re^{m\times n}$, $z^0, w^0 \in \Re^{n_1+n_2}$, and $\beta_0 > 0$, the ADM for problem (3.31) at k-th iteration can be described as follows:

$$\begin{cases} y^{k+1} &= \arg\min\{\mathcal{L}_{\beta_k}(y, X^k, z^k; Z^k, w^k) | y \in \Re^p\}, \\ (X^{k+1}, z^{k+1}) &= \arg\min\{\mathcal{L}_{\beta_k}(y^{k+1}, X, z; Z^k, w^k) | (X, z) \in \Re^{m\times n} \times \mathcal{Q}\}, \\ Z^{k+1} &= Z^k - \varrho\beta_k(\mathcal{A}^* y^{k+1} + X^{k+1} - A_0), \\ w^{k+1} &= w^k - \varrho\beta_k(By^{k+1} - b - z^{k+1}), \\ \beta_{k+1} &= r\beta_k, r \geq 1 \end{cases}$$

where $\varrho \in (0, \frac{1+\sqrt{5}}{2})$. It is easy to see that the minimizer y^{k+1} is the solution of the following linear system of equations:

$$(\mathcal{A}\mathcal{A}^* + B^T B)y^{k+1} = \mathcal{A}(A_0 - X^k + Z^k/\beta_k) + B^T(b + z^k + w^k/\beta_k). \tag{3.33}$$

Since $\mathcal{L}_{\beta_k}(y^{k+1}, X, z; Z^k, w^k)$ is separable in X and z, simple algebraic manipulations then give

$$X^{k+1} = A_0 - \mathcal{A}^* y^{k+1} + Z^k/\beta_k - \Pi_{\mathcal{B}_\beta}(A_0 - \mathcal{A}^* y^{k+1} + Z^k/\beta_k),$$
$$z^{k+1} = \Pi_{\mathcal{Q}}(By^k - b - w^k/\beta_k).$$

As analyzed in the previous chapter, X^{k+1} can be computed analytically. Moreover, z^{k+1} is just a simple projection over $\mathcal{Q}$. Specifically, for any given $x \in \Re^{n_1+n_2}$,

$$(\Pi_{\mathcal{Q}}(x))_i = \begin{cases} 0 & \text{if} 1 \leq i \leq n_1 \\ \max(0, x_i) & \text{if } n_1 + 1 \leq i \leq n_1 + n_2 \end{cases}.$$

Remark 3.2. In the implementation of ADM, the subproblem (3.33) is solved by a direct solver using the Cholesky decomposition where the number of matrices p is small or medium. For larger p, we employ the conjugate gradient method with a diagonal preconditioner to obtain an approximate solution of (3.33). In addition, one can completely avoid solving the linear system by using a linearized technique in the ADM. Thus, y^{k+1} is given by

$$\begin{aligned} y^{k+1} := \ & \arg\min_y\{-\langle \mathcal{A}^*\xi_1^k + B\xi_2^k, y\rangle + \beta\langle \mathcal{A}(\mathcal{A}^* y^k + X^{k+1} - A_0), y - y^k\rangle + \frac{r}{2}\|y - y^k\|^2 \\ & +\beta\langle B^T(By^k - b - z^{k+1}), y - y^k\rangle + \frac{s}{2}\|y - y^k\|^2\}. \end{aligned} \tag{3.34}$$

Direct calculation yields the closed form y^{k+1} satisfying

$$y^{k+1} = y^k + \frac{1}{r+s}[\mathcal{A}^*\xi_1^k + B\xi_2^k - \beta\mathcal{A}(\mathcal{A}^* y^k + X^{k+1} - A_0) - \beta B^T(By^k - b - z^{k+1})],$$

where $r \geq \beta\|\mathcal{A}\mathcal{A}^*\|_2 > 0$ and $s \geq \beta\|B^T B\|_2 > 0$ guarantee the convergence.

Remark 3.3. Some more general ADM-based methods in the literature can be easily extended to solve (1.1). For example, the ADM-based descent method developed in [91] and the ADM based customized PPA [11]. We here omit details of these general ADM type methods for succinctness.

Chapter 4

A semismooth Newton-CG dual proximal point algorithm

4.1 A dual proximal point algorithm framework

In this section, we shall introduce the framework of the inexact dual PPA for solving the MNA problem and establish its global and local convergence.

4.1.1 Proximal point algorithm

Let H be a real Hilbert space with inner product $\langle \cdot, \cdot \rangle$. A multifunction $T : H \to H$ is a monotone operator if

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

It is said to be maximal monotone if, in addition, the graph

$$G(T) = \{(z, w) \in H \times H \,|\, w \in T(z)\}$$

is not strictly contained in the graph of any other monotone operator $T' : H \to H$. In various fields of applied mathematics, many problems can be equivalently formulated as a maximal monotone inclusion problem, that is, given a, possibly multi-valued, maximal monotone operator $\mathcal{T} : \mathcal{H} \to \mathcal{H}$, it is to find a $x \in \mathcal{X}$ such that

$$0 \in \mathcal{T}(x).$$

For example, let $f : H \to (-\infty, \infty]$ be a proper lower semicontinuous convex functions. Then $T = \partial f(\cdot)$ is a maximal operator (see) and $0 \in T(z)$ means $f(z) = \min f(x)$.

The proximal point method, initiated by [58] and later investigated extensively by [72, 73]. The PPA [73] applied to the maximal monotone inclusion problem takes the following scheme

$$x^{k+1} \approx p_{\lambda_k}(x^k) := (I + \lambda_k \mathcal{T})^{-1}(x^k),$$

where $\lambda_k > 0$ is bounded away from zero. In [73], Rockafellar suggested computing x^{k+1} only approximately to satisfy the following accuracy criteria:

$$\|x^{k+1} - p_{\lambda_k}(x^k)\| \leq \varepsilon_k, \quad \varepsilon_k > 0, \sum_{k=1}^{\infty} \varepsilon_k < \infty, \tag{4.1}$$

$$\|x^{k+1} - p_{\lambda_k}(x^k)\| \leq \delta_k \|x^{k+1} - x^k\|, \quad \delta_k > 0, \sum_{k=1}^{\infty} \delta_k < \infty. \tag{4.2}$$

In that paper, he also showed that the sequence generated above converges (in the weak topology) to a zero point of $\mathcal{T}$, if it exists. Moreover, if $\lambda_k \uparrow \lambda_\infty \leq \infty$ and $\mathcal{T}^{-1}$ is Lipschitz continuous at 0, then condition (4.2) ensures that the local convergence is linear and the rate is approximately proportional to $1/\lambda_\infty$. If in addition $\lambda_\infty = \infty$, then the convergence becomes superlinear. A problem of particular importance is the convex minimization $\min f(x)$ where f is assumed to be proper, lower semicontinuous and convex. In this case, the above inexact PPA reduces to

$$x_{k+1} \approx \arg\min f(x) + \frac{1}{2}\|x - x^k\|^2. \tag{4.3}$$

The attractive feature of this approach is that the objective function in (4.3) is strongly, which motivates us to apply an indirect method for solving (4.3) based on the duality theory for convex programming.

Possibly due to its versatility and effectiveness, the proximal point algorithm receives continuous attention from numerous researchers and is well accepted as a powerful tool for solving various classes of optimization problems, see, e.g.[37, 56, 72, 84, 95]. In this section, we consider the dual proximal point algorithm, i.e., applying the idea to the maximal monotone operator associated with the dual problem. By rewriting (1.1) as

$$\min \Big\{ \|X\|_2 \ \big| \ \mathcal{A}^* y + X = A_0, \ By - b \in \mathcal{Q} \Big\}, \tag{4.4}$$

we can easily derive the following explicit form of its dual

$$\begin{aligned} \min \quad & -\langle A_0, Z\rangle - \langle b, w\rangle \\ \text{s.t.} \quad & \mathcal{A}Z + B^T w = 0, \\ & \|Z\|_* \leq 1, \ \ w \in \mathcal{Q}^*, \end{aligned} \tag{4.5}$$

where $\|\cdot\|_*$ denotes the nuclear norm of a matrix which is defined as the sum of its singular values and $\mathcal{Q}^*$ is the dual cone of $\mathcal{Q}$. For the convergence analysis later, we assume that the Slater condition for (4.5) holds, i.e., there exists $(Z, w) \in \Re^{m\times n} \times \Re^{n_1+n_2}$ such that

$$\begin{cases} \|Z\|_* < 1, \\ w_i > 0,\ i = n_1+1, \ldots, n_1+n_2, \\ \mathcal{A}Z + B^T w = 0. \end{cases} \tag{4.6}$$

Write

$$\mathcal{T}_l(X, y) = -\begin{pmatrix} \partial\|X\|_2 \\ 0 \end{pmatrix} + \partial\chi_{\mathcal{F}_1}(X, y), \quad \forall\, X \in \Re^{m\times n},\ y \in \Re^p,$$

$$\mathcal{T}_g(Z, w) = -\begin{pmatrix} A_0 \\ b \end{pmatrix} + \partial\chi_{\mathcal{F}_2}(Z, w), \quad \forall\, Z \in \Re^{m\times n},\ w \in \Re^{n_1+n_2},$$

and

$$p_\lambda(Z, w) = (I + \lambda\mathcal{T}_g)^{-1}(Z, w),$$

where $\mathcal{F}_1$ and $\mathcal{F}_2$ are the feasible sets of (4.4) and (4.5) respectively.

For any given $Z^k \in \Re^{m\times n}$, $w^k \in \Re^{n_1+n_2}$ and $\lambda_k > 0$, it is easy to see that $p_{\lambda_k}(Z^k, w^k)$ is the unique solution of the following minimization problem

$$\begin{array}{ll} \min & -\langle A_0, Z\rangle - \langle b, w\rangle + \dfrac{1}{2\lambda_k}\|Z - Z^k\|^2 + \dfrac{1}{2\lambda_k}\|w - w^k\|^2 \\ \text{s.t.} & \mathcal{A}Z + B^T w = 0, \\ & \|Z\|_* \le 1, \quad w \in \mathcal{Q}^*. \end{array} \tag{4.7}$$

By attaching a lagrangian y to the equality constraint, the dual of (4.7) is of the following form

$$\begin{array}{ll} \max\limits_{y\in\Re^p} & \{-\langle A_0, Z\rangle - \langle b, w\rangle + \dfrac{1}{2\lambda_k}\|Z - Z^k\|^2 + \dfrac{1}{2\lambda_k}\|w - w^k\|^2 \\ & -\langle y, \mathcal{A}Z + B^T w\rangle \mid \|Z\|_* \le 1, w \in \mathcal{Q}^*\}. \end{array} \tag{4.8}$$

Simple calculation shows that the dual of (4.7) can be further expressed as

$$\begin{array}{rl} \max\limits_{y\in\Re^p} \theta_k(y) := & \dfrac{1}{2\lambda_k}\|\Pi_{\mathcal{B}}(Z^k - \lambda_k(\mathcal{A}^*y - A_0)) - (Z^k - \lambda_k(\mathcal{A}^*y - A_0))\|^2 \\ & +\dfrac{1}{2\lambda_k}\Big(\|Z^k\|^2 - \|Z^k - \lambda_k(\mathcal{A}^*y - A_0)\|^2\Big) \\ & +\dfrac{1}{2\lambda_k}\Big(\|w^k\|^2 - \|\Pi_{\mathcal{Q}^*}[w^k - \lambda_k(By - b)]\|^2\Big). \end{array} \tag{4.9}$$

Clearly, the Slater condition (4.6) asserts that the optimal solution set of (4.9) is nonempty. Let y^{k+1} be a minimizer of (4.9). Then from the relationship between the primal and dual variables, we have

$$p_{\lambda_k}(Z^k, w^k) = \begin{bmatrix} \Pi_{\mathcal{B}}(Z^k - \lambda_k(\mathcal{A}^* y^{k+1} - A_0)) \\ \Pi_{\mathcal{Q}^*}(w^k - \lambda_k(B y^{k+1} - b)) \end{bmatrix}. \tag{4.10}$$

Therefore, to implement the proximal point algorithm, one need to solve (4.9) and then update the variable (Z, w) by

$$(Z^{k+1}, w^{k+1}) \approx p_{\lambda_k}(Z^k, w^k).$$

In view of (4.10), we are able to present the inexact dual PPA framework:

Algorithm 4.1 (An inexact dual PPA framework) Given (Z^0, w^0, y^0) and $\lambda_0 > 0$, at the k-th iteration, do the following steps:

Step 1**.** For fixed Z^k, w^k and y^k, compute an approximate maximizer

$$y^{k+1} \approx \arg\max_{y \in \Re^p} \theta_k(y),$$

where θ_k is defined in (4.9).

Step 2**.** Update the variables Z^{k+1}, X^{k+1} and w^{k+1} via

$$\begin{aligned} Z^{k+1} &= \Pi_{\mathcal{B}_{\frac{1}{\lambda^k}}}(Z^k - \lambda_k(\mathcal{A}^* y^{k+1} - A_0)), \\ w^{k+1} &= \Pi_{\mathcal{Q}^*}(w^k - \lambda_k(B y^{k+1} - b)), \\ X^{k+1} &= (Z^k - \lambda_k(\mathcal{A}^* y^{k+1} - A_0) - Z^{k+1})/\lambda_k. \end{aligned}$$

Step 3**.** If $\max\{\|A_0 - \mathcal{A}^* y^{k+1} - X^{k+1}\|_F, \|\Pi_{\mathcal{Q}^*}(b - B y^{k+1})\|\} \leq \varepsilon$, stop; else, update λ_k to λ_{k+1}, end.

4.1.2 Convergence analysis

In Step 1 of the dual PPA, we use the following stopping criteria:

$$\max \theta_k(y) - \theta_k(y^{k+1}) \leq \tfrac{\varepsilon_k^2}{2\lambda_k}, \quad \varepsilon_k > 0, \ \textstyle\sum_{k=1}^{\infty} \varepsilon_k < \infty \tag{4.11}$$

$$\max \theta_k(y) - \theta_k(y^{k+1}) \leq \tfrac{\delta_k^2}{2\lambda_k}(\|Z^{k+1} - Z^k\|^2 + \|w^{k+1} - w^k\|^2), \ \delta_k > 0, \ \textstyle\sum_{k=1}^{\infty} \delta_k < \infty \tag{4.12}$$

$$\|\nabla_y \theta_k(y^{k+1})\| \leq \frac{\delta_k'}{\lambda_k} \left\| \begin{pmatrix} Z^{k+1} - Z^k \\ w^{k+1} - w^k \end{pmatrix} \right\|, \ 0 \leq \delta_k' \to 0 \tag{4.13}$$

to terminate our proposed dual PPA. For the constrained minimization (4.4), the augmented Lagrangian is

$$\begin{aligned} L_\beta(X, y, Z, w, \lambda) &= \min_{z \in \Re_+^p} \{ \|X\|_2 + \frac{\lambda}{2}\|\mathcal{A}^* y + X - A_0 - \frac{Z}{\lambda}\|_F^2 \\ &\qquad + \frac{\lambda}{2}\|By - b - z - \frac{w^k}{\lambda}\| - \frac{\|w^k\|^2}{\lambda}\}, \\ &= \|X\|_2 + \frac{\lambda}{2}\|\mathcal{A}^* y + X - A_0 - \frac{Z}{\lambda}\|_F^2 \\ &\qquad + \frac{1}{2\lambda}\|\Pi_{\Re_+^n}(w - \lambda(By - b))\|^2 - \frac{\|w^k\|^2}{\lambda}. \end{aligned} \tag{4.14}$$

For $k = 1, 2, \ldots$, let $\phi_k(X, y) = L_\beta(X, y, Z^k, w^k, \lambda^k)$. By the construction of X^k at each k, we know

$$X^k = \arg\min_X \phi_k(X, y^k) \tag{4.15}$$

and

$$\theta_k(y^{k+1}) = -\phi_k(X^{k+1}, y^{k+1}).$$

Combining [72, Theorem 4-5] with the above preparation, we present below two results on the global and local convergence of the dual PPA.

Theorem 4.1 (Global Convergence)**.** *Let the inexact PPA be executed with stopping criterion* (4.11)*. Suppose that the primal problem* (4.4) *satisfies the Slater condition. Then the sequence* $\{(Z^{k+1}, w^{k+1})\} \subset \mathcal{B} \times \mathcal{Q}^*$ *generated by the inexact PPA is bounded and it converges to an optimal solution of* (4.5)*. Moreover, the sequence* $\{y^k\}$ *is also bounded and any of its accumulation point is an optimal solution of* (1.1)*.*

Proof. Observing from the definition $\theta_k(y)$ and $\phi(X, y)$, we obtain by direct computation that

$$\min_{X,y} \phi_k(X, y) = -\max_y \theta_k(y).$$

Therefore, the stopping criterion (4.11) can be written as

$$\phi_k(X^{k+1}, y^{k+1}) - \min_{X,y} \phi_k(X, y) \le \frac{\varepsilon_k^2}{2\lambda_k},\ \varepsilon_k > 0,\ \sum_{k=1}^{\infty} \varepsilon_k < \infty,$$

which is the criterion (A'') in [72]. Then by directly invoking [72, Theorem 4], we can complete the proof of this theorem. □

Theorem 4.2 (Local Convergence). *Let the dual PPA be executed with stopping criteria* (4.11) *and* (4.12). *Suppose that the Slater condition holds for* (4.4). *If $\mathcal{T}_g^{-1}$ is Lipschitz continuous at the origin with the modulus a_g, then $\{(Z^{k+1}, w^{k+1})\}$ converges to an optimal solution $(\overline{Z}, \bar{w})$ of* (4.5)*, and*

$$\left\| \begin{pmatrix} Z^{k+1} - \overline{Z} \\ w^{k+1} - \bar{w} \end{pmatrix} \right\| \le \nu_k \left\| \begin{pmatrix} Z^{k} - \overline{Z} \\ w^{k} - \bar{w} \end{pmatrix} \right\|, \quad \text{for all } k \text{ sufficiently large}, \tag{4.16}$$

where $\nu_k = [a_g(a_g + \lambda_k^2)^{-1/2} + \delta_k](1 - \delta_k)^{-1} \to a_g(a_g^2 + \lambda_\infty^2)^{-1/2} < 1$. Moreover, the conclusion about $\{y^k\}$ in Theorem 4.1 is valid.

If in addition to (4.12) *and the condition on $\mathcal{T}_g^{-1}$, one also has* (4.13) *and that $\mathcal{T}_l^{-1}$ is Lipschitz continuous at the origin with modulus a_l $(\ge a_g)$, then $\{y^{k+1}\}$ converges to the unique optimal solution $\bar{y}$ of* (1.1)*, and*

$$\left\| \begin{pmatrix} X^{k+1} - \overline{X} \\ y^{k+1} - \bar{y} \end{pmatrix} \right\| \le \nu_k' \left\| \begin{pmatrix} Z^{k+1} - Z^{k} \\ w^{k+1} - w^{k} \end{pmatrix} \right\|, \quad \text{for all } k \text{ sufficiently large,}$$

where $\overline{X} = A_0 - \mathcal{A}^*\bar{y}$, and $\nu_k' = a_l(1 + \delta_k')/\lambda_k \to a_l/\lambda_\infty$.

Proof. Since

$$\partial_{(X,y)}\phi_k(X^{k+1}, y^{k+1}) = \begin{bmatrix} \partial_X \phi_k(X^{k+1,y^{k+1}}) \\ \nabla_y \phi_k(X^{k+1,y^{k+1}}) \end{bmatrix},$$

by using the first order optimality condition of (4.15), we deduce that

$$\text{dist}(0, \partial_{(X,y)}\phi_k(X^{k+1}, y^{k+1})) = \|\nabla_y \phi_k(X^{k+1}, y^{k+1})\|.$$

It can be established by direct computation that

$$\nabla_y \phi_k(X^{k+1}, y^{k+1}) = -\nabla_y \theta(y^{k+1})$$

and therefore the criterion (4.13) is equivalent to

$$\text{dist}(0, \partial_{(X,y)}\phi_k(X^{k+1,y^{k+1}})) \le \frac{\delta_k'}{\lambda_k} \left\| \begin{pmatrix} Z^{k+1} - Z^{k} \\ w^{k+1} - w^{k} \end{pmatrix} \right\|,\ 0 \le \delta_k' \to 0\,.$$

Note that

$$\min_{X,y} \phi_k(X,y) = -\max_y \theta_k(y).$$

and

$$\theta_k(y^{k+1}) = -\phi_k(X^{k+1}, y^{k+1}).$$

Then [72, Theorem 5] is applicable to the inexact proximal point algorithm and this claim follows immediately. □

Remark 4.1. In practical implementation, a proximal term $-\frac{1}{2\beta_k}\|y - y^k\|^2$ can be added to the dual objective function θ_k. This corresponds to the proximal method of multiplier considered in [72, Section 5] whose convergence analysis can be conducted in a parallel way for algorithm 4.1. Let $\hat{\theta}_k(y)$ be the resulted new dual objetive function, i.e.,

$$\hat{\theta}_k(y) = \theta_k(y) - \frac{1}{2\beta_k}\|y - y^k\|^2.$$

By the strong convexity of the new objective function, it holds that:

$$\sup \hat{\theta}_k(y) - \hat{\theta}_k(y^{k+1}) \leq \frac{1}{2\beta}\|\nabla\hat{\theta}_k(y^{k+1})\|^2.$$

Therefore, the stopping criterias (4.11) and (4.12) can be modified into the following practical conditions respectively:

$$\|\nabla\hat{\theta}_k(y^{k+1})\| \leq \varepsilon_k,\ \epsilon_k \geq 0,\ \sum_{k=0}^{\infty} \epsilon_k < \infty$$

$$\|\nabla\hat{\theta}_k(y^{k+1})\| \leq \delta_k\sqrt{\|Z^{k+1} - Z^k\|^2 + \|w^{k+1} - w^k\|^2},\ \delta_k > 0,\ \sum_{k=1}^{\infty} \delta_k < \infty.$$

4.2 A semismooth Newton-CG method for the inner problem

In this section, we will apply the well-known inexact semismooth Newton method to approximately solve the unconstrained subproblem (4.9). Using Proposition 2.5 (ii), we know that the first order optimality condition for (4.9) is given by

$$0 = \nabla\theta_k(y) := \mathcal{A}\Pi_{\mathcal{B}}\big[Z^k - \lambda_k(\mathcal{A}^*y - A_0)\big] + B^T\Pi_{\mathcal{Q}^*}\big[w^k - \lambda_k(By - b)\big].$$

Since $\Pi_{\mathcal{B}}(\cdot)$ and $\Pi_{\mathcal{Q}^*}(\cdot)$ are Lipschitz continuous, $\nabla\theta_k(\cdot)$ is also Lipschitz continuous. Hence the Clarke's generalized Jacobian of $\nabla\theta_k$ (which is the generalized Hessian of θ_k and we denote it by $\partial^2\theta_k$) is well defined. Since it is difficult to derive the exact characterization of $\partial^2\theta_k$, we will slightly modify the classical semismooth Newton method by selecting elements in a larger differentiable set $\hat{\partial}^2\theta_k$ instead of $\partial^2\theta_k$, where $\hat{\partial}^2\theta_k$ is a set-valued mapping defined by

$$\hat{\partial}^2\theta_k(y) := -\lambda_k\big[\mathcal{A}\partial\Pi_{\mathcal{B}}(Z^k-\lambda_k(\mathcal{A}^*y-A_0))\mathcal{A}^*+B^T\partial\Pi_{\mathcal{Q}^*}(w^k-\lambda_k(By-b))B\big],\ \forall\, y\in\Re^p.$$

due to the fact [16, p.75]

$$\partial^2\theta_k(y)H \subseteq \hat{\partial}^2\theta_k(y)H$$

for any $H\in\Re^{m\times n}$.

4.2.1 Characterization of $\hat{\partial}^2\theta_k$

To obtain the explicit expression of $\hat{\partial}^2\theta_k$, it suffices to characterize $\partial\Pi_{\mathcal{B}}(\cdot)$ and $\partial\Pi_{\mathcal{Q}^*}(\cdot)$. For a given $Y\in\Re^{m\times n}$, suppose that it has the following SVD:

$$Y = U[\mathrm{diag}(\sigma)\ 0]V^T,$$

where $\sigma=(\sigma_1,\sigma_2,\ldots,\sigma_m)^T$ with $\sigma_1\geq\sigma_2\geq\ldots\geq\sigma_r>0=\sigma_{r+1}=\ldots=\sigma_m$.

Let $\{Y^i\}_{i\geq1}$ be a sequence converging to Y such that every element $Y^i\in\mathcal{E}^0$, where $\mathcal{E}^0$ is defined by

$$\mathcal{E}^0 := \{Y\in\mathcal{D}_{\Pi_{\mathcal{B}}}\,|\,\sigma_1(Y)>\sigma_2(Y)>\ldots>\sigma_m(Y)>0\},$$

where $\mathcal{D}_{\Pi_{\mathcal{B}}}$ is the collection of the points at which $\Pi_{\mathcal{B}}(\cdot)$ is differentiable. This implies $\|Y^i\|_*\neq1$ for each $i\geq1$. Indeed, the expression (2.17) clearly shows that $\Pi_{\mathbb{B}}(\cdot)$ is non-differentiable at any $x\in\Re^m$ with $\|x\|_1=1$. Then, by Proposition 2.4 (i), we know that $\Pi_{\mathcal{B}}(\cdot)$ is also non-differentiable at any point $Z\in\Re^{m\times n}$ satisfying $\|Z\|_*=1$. Let

$$\partial_{\mathcal{E}^0}\Pi_{\mathcal{B}}(Y) := \left\{\lim_{\mathcal{E}^0\ni Y^j\to Y}\Pi'_{\mathcal{B}}(Y^j)\right\}.$$

Let the SVD of Y^j be $Y^i=U^i[\mathrm{diag}(\sigma^i),0](V^i)^T$. We consider 3 cases.

Case 1: $\|Y\|_*<1$.
In this case, $\Pi_{\mathcal{B}}$ is continuously differentiable at Y and its generalized Jacobian is a singleton set consisting of the identity operator $\mathcal{I}$ from $\Re^{m\times n}$ to $\Re^{m\times n}$.

Case 2: $\|Y\|_*=1$.
In this case, a quick computation yields that $k_1(\sigma)=r$ and $k_2(\sigma)=m$.

Since Y can be approximated by a sequence in the interior of $\mathcal{B}$, it follows that the identity operator is always an element of $\partial_{\mathcal{E}^0}\Pi_{\mathcal{B}}(Y)$. To obtain other elements, we consider the case in which $\{Y^i\}$ has an infinite subsequence outside $\mathcal{B}$. Without loss of generality, we assume that $\|Y^i\|_* > 1$ for all i. By passing through a subsequence if necessary, we know that there exists a positive integer $N \in [r,\, m]$ such that $N = k_1(\sigma^i)$ for each i. Therefore, one has

$$(\Pi_{\mathbb{B}}(\sigma^i))_k \;=\; \begin{cases} \sigma_k^i - \frac{1}{N}\Big(\sum_{j=1}^N \sigma_j^i - 1\Big), & 1 \le k \le N, \\ 0, & \text{otherwise} \end{cases}$$

and

$$\Pi'_{\mathbb{B}}(\sigma^i) = \begin{bmatrix} I_N & 0 \\ 0 & 0 \end{bmatrix} - \frac{1}{N}\begin{bmatrix} \mathbf{1}_{N\times N} & 0 \\ 0 & 0 \end{bmatrix}.$$

where I_N and $\mathbf{1}_{N\times N}$ are the $N \times N$ identity matrix and the $m \times n$ matrix respectively. For notational simplicity, we write $\Pi_{\mathbb{B}}(\sigma^i)$ as

$$\Pi_{\mathbb{B}}(\sigma^i) =: g^i(\sigma^i) = (g_1^i(\sigma^i), \dots, g_N^i(\sigma^i), 0, \dots, 0)^T$$

and define the following four index sets

$$\begin{aligned} \alpha_1 &:= \{1, 2, \dots, r\}, & \alpha_2 &:= \{r+1, r+2, \dots, N\}, \\ \alpha_3 &:= \{N+1, N+2, \dots, m\}, & \alpha_4 &:= \{m+1, m+2, \dots, n\}. \end{aligned} \tag{4.17}$$

Let Ω^i and Γ^i be the following $m \times m$ symmetric matrices

$$\Omega^i = \begin{bmatrix} \multicolumn{2}{c}{\mathbf{1}_{N\times N}} & \begin{matrix}\Omega^i_{\alpha_1\alpha_3} \\ \Omega^i_{\alpha_2\alpha_3}\end{matrix} \\ (\Omega^i_{\alpha_1\alpha_3})^T & (\Omega^i_{\alpha_2\alpha_3})^T & 0 \end{bmatrix}, \tag{4.18}$$

$$\Gamma^i = \begin{bmatrix} \Gamma^i_{\alpha_1\alpha_1} & \Gamma^i_{\alpha_1\alpha_2} & \Gamma^i_{\alpha_1\alpha_3} \\ (\Gamma^i_{\alpha_1\alpha_2})^T & \Gamma^i_{\alpha_2\alpha_2} & \Gamma^i_{\alpha_2\alpha_3} \\ (\Gamma^i_{\alpha_1\alpha_3})^T & (\Gamma^i_{\alpha_2\alpha_3})^T & 0 \end{bmatrix}, \tag{4.19}$$

where

$$(\Omega^i_{\alpha_1\alpha_3})_{kj} = \frac{g^i_k}{\sigma^i_k - \sigma^i_{j+N}}, \quad \text{for } k \in \alpha_1,\ j \in \alpha_3 - N,$$

$$(\Omega^i_{\alpha_2\alpha_3})_{kj} = \frac{g^i_{k+r}}{\sigma^i_{k+r} - \sigma^i_{j+N}}, \quad \text{for } k \in \alpha_2 - r,\ j \in \alpha_3 - N,$$

$$(\Gamma^i_{\alpha_1\alpha_1})_{kj} = \frac{g^i_k + g^j_k}{\sigma^i_k + \sigma^i_j}, \quad \text{for } k \in \alpha_1,\ j \in \alpha_1,$$

$$(\Gamma^i_{\alpha_1\alpha_2})_{kj} = \frac{g^i_k + g^i_{j+r}}{\sigma^i_k + \sigma^i_{j+r}}, \quad \text{for } k \in \alpha_1,\ j \in \alpha_2 - r,$$

$$(\Gamma^i_{\alpha_1\alpha_3})_{kj} = \frac{g^i_k}{\sigma^i_k + \sigma^i_{j+N}}, \quad \text{for } k \in \alpha_1,\ j \in \alpha_3 - N,$$

$$(\Gamma^i_{\alpha_2\alpha_2})_{kj} = \frac{g^i_{k+r} + g^i_{j+r}}{\sigma^i_{k+r} + \sigma^i_{j+r}}, \quad \text{for } k \in \alpha_2 - r,\ j \in \alpha_2 - r,$$

$$(\Gamma^i_{\alpha_2\alpha_3})_{kj} = \frac{g^i_{k+r}}{\sigma^i_{k+r} + \sigma^i_{j+N}}, \quad \text{for } k \in \alpha_2 - r,\ j \in \alpha_3 - N.$$

To simplify notation, we also write

$$\Upsilon^i = \begin{bmatrix} \Upsilon^i_{\alpha_1} \\ \Upsilon^i_{\alpha_2} \end{bmatrix}, \quad \text{with } \Upsilon^i_k := \frac{g^i_k}{\sigma^i_k},\ k = 1, 2, \ldots, N.$$

Now from Proposition 2.4, we know that for any given $H \in \Re^{m\times n}$,

$$\Pi'_{\mathcal{B}}(Y^i)H = U^i \begin{bmatrix} W^i_1 & W^i_2 \end{bmatrix} (V^i)^T, \tag{4.20}$$

where the matrices $W^i_1 \in \Re^{m\times m}$ and $W^i_2 \in \Re^{m\times(n-m)}$ are defined by

$$W^i_1 = \Omega^i \circ S(\widetilde{H}^i_1) + \Gamma^i \circ T(\widetilde{H}^i_1) - \frac{\mathrm{Tr}(\widetilde{H}^i_{11})}{N} \begin{bmatrix} I_N & 0 \\ 0 & 0 \end{bmatrix},$$

and

$$W^i_2 = \begin{bmatrix} \Upsilon^i \mathbf{1}^T_{n-m} \\ 0 \end{bmatrix} \circ \widetilde{H}^i_2,$$

with $\widetilde{H}^i_1 \in \Re^{m\times m}$, $\widetilde{H}^i_2 \in \Re^{m\times(n-m)}$, $[\widetilde{H}^i_1\ \widetilde{H}^i_2] = (U^i)^T H V^i$ and $\widetilde{H}^i_{11}$ being the matrix extracted from the first N columns and rows of $\widetilde{H}^i_1$. By simple algebraic computation, we are able to show

$$\lim_{i\to\infty} \Omega^i_{\alpha_1\alpha_3} = \mathbf{1}_{r\times(m-N)},$$

$$\lim_{i\to\infty} \Gamma^i_{\alpha_1\alpha_1} = \mathbf{1}_{r\times r},$$

$$\lim_{i\to\infty} \Gamma^i_{\alpha_1\alpha_2} = \mathbf{1}_{r\times(N-r)},$$

$$\lim_{i\to\infty} \Omega^i_{\alpha_1\alpha_3} = \mathbf{1}_{N\times(m-N)},$$

$$\lim_{i\to\infty} \Upsilon^i_{\alpha_1} = \mathbf{1}_r.$$

Let $\mathcal{S}_N$ be the set of cluster points of $\{(\Omega^i_{\alpha_2\alpha_3}, \Gamma^i_{\alpha_2\alpha_2}, \Gamma^i_{\alpha_2\alpha_3}, \Upsilon^i_{\alpha_2})\}_{i\geq 1}$. By taking limits on both sides of (4.20), we are able to establish the conclusion that $\mathcal{V}(\neq \mathcal{I})$ is an element of $\partial_{\mathcal{E}^0}\Pi_{\mathcal{B}}(Y)$ if and only if there exist an integer $N \in [r,\, m]$, $\{(\Omega^\infty_{\alpha_2\alpha_3}, \Gamma^\infty_{\alpha_2\alpha_2}, \Gamma^\infty_{\alpha_2\alpha_3}, \Upsilon^\infty_{\alpha_2})\} \in \mathcal{S}_N$ and singular vector matrices U^∞, V^∞ of Y such that for any $H \in \Re^{m\times n}$,

$$\mathcal{V}H = U^\infty \left[W_1^\infty \quad W_2^\infty\right] (V^\infty)^T, \tag{4.21}$$

where the matrices $W_1^\infty \in \Re^{m\times m}$ and $W_2^\infty \in \Re^{m\times(n-m)}$ are defined by

$$W_1^\infty = \Omega^\infty \circ S(\widetilde{H}_1) + \Gamma^\infty \circ T(\widetilde{H}_1) - \frac{\mathrm{Tr}(\widetilde{H}_{11})}{N}\begin{bmatrix} I_N & 0 \\ 0 & 0 \end{bmatrix}$$

and

$$W_2^\infty = \begin{bmatrix} \mathbf{1}_{r\times(n-m)} \\ \Upsilon^\infty_{\alpha_2}\mathbf{1}^T_{n-m} \\ 0 \end{bmatrix} \circ \widetilde{H}_2,$$

here, $\widetilde{H}_1 \in \Re^{m\times m}, \widetilde{H}_2 \in \Re^{m\times(n-m)}$, $[\widetilde{H}_1\ \widetilde{H}_2] = (U^\infty)^T H V^\infty$ and $\widetilde{H}_{11}$ is the matrix extracted from the first N columns and rows of $\widetilde{H}_1$, and

$$\Omega^\infty = \begin{bmatrix} \mathbf{1}_{N\times N} & \begin{matrix}\mathbf{1}_{r\times(m-N)} \\ \Omega^\infty_{\alpha_2\alpha_3}\end{matrix} \\ \mathbf{1}_{(m-N)\times r}\ (\Omega^\infty_{\alpha_2\alpha_3})^T & 0 \end{bmatrix}$$

and

$$\Gamma^\infty = \begin{bmatrix} \mathbf{1}_{r\times r} & \mathbf{1}_{r\times(N-r)} & \mathbf{1}_{r\times(m-N)} \\ \mathbf{1}_{(N-r)\times r} & \Gamma^\infty_{\alpha_2\alpha_2} & \Gamma^\infty_{\alpha_2\alpha_3} \\ \mathbf{1}_{(m-N)\times r} & (\Gamma^\infty_{\alpha_2\alpha_3})^T & 0 \end{bmatrix}.$$

By taking a convex hull of such $\mathcal{V}$ described above and the identity operator $\mathcal{I}$, we can obtain the generalized Jacobian of $\Pi_{\mathcal{B}}(\cdot)$ at Y since it is indifference to sets of zero measure [85].

Case 3: $\|Y\|_* > 1$.

In this case, it is easily seen that $k_1(\sigma) \leq k_2(\sigma) \leq r$. By taking a subsequence if necessary, there exists a positive integer $N \in [k_1(\sigma), k_2(\sigma)]$ such that $N = k_1(\sigma^i)$ for each i. Therefore,

$$\Pi_{\mathbb{B}}(\sigma^i) = \begin{cases} \sigma^i_k - \frac{1}{N}\Big(\sum_{j=1}^N \sigma^i_j - 1\Big), & 1 \leq k \leq N, \\ 0, & \text{otherwise}, \end{cases}$$

and

$$\Pi'_{\mathbb{B}}(\sigma^i) = \begin{bmatrix} I_N & 0 \\ 0 & 0 \end{bmatrix} - \frac{1}{N}\begin{bmatrix} \mathbf{1}_{N\times N} & 0 \\ 0 & 0 \end{bmatrix}.$$

For later discussion, we partition the set $\{1, 2, \ldots, m\}$ into the following five subsets:

$$\beta_1 := \{1, 2, \ldots, k_1(\sigma)\}, \qquad \beta_2 := \{k_1(\sigma)+1, k_1(\sigma)+2, \ldots, N\}$$
$$\beta_3 := \{N+1, N+2, \ldots, k_2(\sigma)\}, \quad \beta_4 := \{k_2(\sigma)+1, k_2(\sigma)+2, \ldots, r\}$$
$$\beta_5 := \{r+1, r+2, \ldots, m\}$$

and write

$$\gamma_1 := \beta_1 \cup \beta_2, \quad \gamma_2 := \beta_3 \cup \beta_4.$$

For each i, redefine Ω^i, Γ^i and Υ^i by

$$\Omega^i := \left[\begin{array}{cc|ccc} \multicolumn{2}{c|}{\mathbf{1}_{N\times N}} & \begin{matrix}\Omega^i_{\beta_1\beta_3} \\ \Omega^i_{\beta_2\beta_3}\end{matrix} & \Omega^i_{\gamma_1\beta_4} & \Omega^i_{\gamma_1\beta_5} \\ (\Omega^i_{\beta_1\beta_3})^T & (\Omega^i_{\beta_2\beta_3})^T & & & \\ \multicolumn{2}{c|}{(\Omega^i_{\gamma_1\beta_4})^T} & \multicolumn{3}{c}{\mathbf{0}_{(m-N)\times(m-N)}} \\ \multicolumn{2}{c|}{(\Omega^i_{\gamma_1\beta_5})^T} & & & \end{array}\right], \tag{4.22}$$

$$\Gamma^i =: \begin{bmatrix} \Gamma^i_{\gamma_1\gamma_1} & \Gamma^i_{\gamma_1\gamma_2} & \Gamma^i_{\gamma_1\beta_5} \\ (\Gamma^i_{\gamma_1\gamma_2})^T & \multicolumn{2}{c}{\multirow{2}{*}{$\mathbf{0}_{(m-N)\times(m-N)}$}} \\ (\Gamma^i_{\gamma_1\beta_5})^T & & \end{bmatrix},$$

and

$$\Upsilon^i := \Big(\frac{g^i_1}{\sigma^i_1}, \frac{g^i_2}{\sigma^i_2}, \ldots, \frac{g^i_N}{\sigma^i_N}\Big)^T,$$

where

$$(\Omega^i_{\beta_1\beta_3})_{kj} = \tfrac{g^i_k}{\sigma^i_k - \sigma^i_{j+N}}, \qquad \text{for } k \in \beta_1,\ j \in \beta_3 - N,$$

$$(\Omega^i_{\beta_2\beta_3})_{kj} = \tfrac{g^i_{k+k_1(\sigma)}}{\sigma^i_{k+k_1(\sigma)} - \sigma^i_{j+N}}, \quad \text{for } k \in \beta_2 - k_1(\sigma),\ j \in \beta_3 - N,$$

$$(\Omega^i_{\gamma_1\beta_4})_{kj} = \tfrac{g^i_k}{\sigma^i_k - \sigma^i_{j+k_2(\sigma)}}, \qquad \text{for } k \in \gamma_1,\ j \in \beta_4 - k_2(\sigma),$$

$$(\Omega^i_{\gamma_1\beta_5})_{kj} = \tfrac{g^i_k}{\sigma^i_k - \sigma^i_{j+r}}, \qquad \text{for } k \in \gamma_1,\ j \in \beta_5 - r,$$

$$(\Gamma^i_{\gamma_1\gamma_1})_{kj} = \tfrac{g^i_k + g^i_j}{\sigma^i_k + \sigma^i_j}, \qquad \text{for } k \in \gamma_1,\ j \in \gamma_1,$$

$$(\Gamma^i_{\gamma_1\gamma_2})_{kj} = \tfrac{g^i_k}{\sigma^i_k + \sigma^i_{j+N}}, \qquad \text{for } k \in \gamma_1,\ j \in \gamma_2 - N,$$

$$(\Gamma^i_{\gamma_1\beta_5})_{kj} = \tfrac{g^i_k}{\sigma^i_k + \sigma^i_{j+r}}, \qquad \text{for } k \in \gamma_1,\ j \in \beta_5 - r.$$

Then the equality (4.20) is also valid. Simple calculation shows that

$$
\begin{aligned}
&\lim_{i\to\infty}(\Omega^i_{\beta_1\beta_3})_{kj} = \tfrac{g_k}{\sigma_k-\sigma_{j+N}} := (\Omega^\infty_{\beta_1\beta_3})_{kj}, && \text{for } k\in\beta_1,\ j\in\beta_3-N,\\
&\lim_{i\to\infty}(\Omega^i_{\gamma_1\beta_4})_{kj} = \tfrac{g_k}{\sigma_k-\sigma_{j+k_2(\sigma)}} := (\Omega^\infty_{\gamma_1\beta_4})_{kj}, && \text{for } k\in\gamma_1,\ j\in\beta_4-k_2(\sigma),\\
&\lim_{i\to\infty}(\Omega^i_{\gamma_1\beta_5})_{kj} = \tfrac{g_k}{\sigma_k} := (\Omega^\infty_{\gamma_1\beta_5})_{kj}, && \text{for } k\in\gamma_1,\ j\in\beta_5-r,\\
&\lim_{i\to\infty}(\Gamma^i_{\gamma_1\gamma_1})_{kj} = \tfrac{g_k+g_j}{\sigma_k+\sigma_j} := (\Gamma^\infty_{\gamma_1\gamma_1})_{kj}, && \text{for } k\in\gamma_1,\ j\in\gamma_1,\\
&\lim_{i\to\infty}(\Gamma^i_{\gamma_1\gamma_2})_{kj} = \tfrac{g_k}{\sigma_k+\sigma_{j+N}} := (\Gamma^\infty_{\gamma_1\gamma_2})_{kj}, && \text{for } k\in\gamma_1,\ j\in\gamma_2-N,\\
&\lim_{i\to\infty}(\Gamma^i_{\gamma_1\beta_5})_{kj} = \tfrac{g_k}{\sigma_k} := (\Gamma^\infty_{\gamma_1\beta_5})_{kj}, && \text{for } k\in\gamma_1,\ j\in\beta_5-r,\\
&\lim_{i\to\infty}(\Upsilon^i)_{k} = \tfrac{g_k}{\sigma_k} := (\Upsilon^\infty)_{k}, && \text{for } k=1,2,\ldots,N.
\end{aligned}
$$

Redefine $\mathcal{S}_N$ to be the set of limit points of $\{\Omega^i_{\beta_2\beta_3}\}$. By taking limits on both sides of (4.20), we have the conclusion that $\mathcal{V}$ is an element of $\partial_{\mathcal{E}^0}\Pi_{\mathcal{B}}(Y)$ if and only if there exist an integer $N\in[k_1(\sigma),\ k_2(\sigma)]$, $\Omega^\infty_{\beta_2\beta_3}\in\mathcal{S}_N$ and singular vector matrices U^∞, V^∞ of Y such that for any $H\in\Re^{m\times n}$,

$$
\mathcal{V}H = U^\infty\,[W_1^\infty \quad W_2^\infty]\,(V^\infty)^T, \tag{4.23}
$$

where the matrices $W_1^\infty\in\Re^{m\times m}$ and $W_2^\infty\in\Re^{m\times(n-m)}$ are defined by

$$
W_1^\infty = \Omega^\infty\circ S(\widetilde{H}_1) + \Gamma^\infty\circ T(\widetilde{H}_1) - \frac{\mathrm{Tr}(\widetilde{H}_{11})}{N}\begin{bmatrix} I_N & 0\\ 0 & 0\end{bmatrix},
$$

$$
W_2^\infty = \begin{bmatrix}\Upsilon^\infty\mathbf{1}^T_{n-m}\\ 0\end{bmatrix}\circ\widetilde{H}_2,
$$

here, $\widetilde{H}_1\in\Re^{m\times m}, \widetilde{H}_2\in\Re^{m\times(n-m)}$, $[\widetilde{H}_1\ \widetilde{H}_2] = (U^\infty)^T H V^\infty$ and $\widetilde{H}_{11}$ is the matrix extracted from the first N columns and rows of $\widetilde{H}_1$, and the matrices Ω^∞ and Γ^∞ are defined by

$$
\Omega^\infty = \begin{bmatrix} \mathbf{1}_{N\times N} & \begin{matrix}\Omega^\infty_{\beta_1\beta_3}\\ \Omega^\infty_{\beta_2\beta_3}\end{matrix} \ \ \Omega^\infty_{\gamma_1\beta_4}\ \ \Omega^\infty_{\gamma_1\beta_5}\\ \begin{matrix}(\Omega^\infty_{\beta_1\beta_3})^T\ (\Omega^\infty_{\beta_2\beta_3})^T\\ (\Omega^\infty_{\gamma_1\beta_4})^T\\ (\Omega^\infty_{\gamma_1\beta_5})^T\end{matrix} & \mathbf{0}_{(m-N)\times(m-N)}\end{bmatrix},
$$

$$\Gamma^\infty = \begin{bmatrix} \Gamma^\infty_{\gamma_1\gamma_1} & \Gamma^\infty_{\gamma_1\gamma_2} \ \Gamma^\infty_{\gamma_1\beta_5} \\ (\Gamma^\infty_{\gamma_1\gamma_2})^T & \\ (\Gamma^\infty_{\gamma_1\beta_5})^T & \mathbf{0}_{(m-N)\times(m-N)} \end{bmatrix}.$$

By taking a convex hull of those $\mathcal{V}$ described above, we can obtain the generalized Jacobian of $\Pi_{\mathcal{B}}(\cdot)$ at Y since it is blind to sets of zero measure [85].

In order to characterize $\partial\Pi_{\mathcal{Q}^*}(\cdot)$ at $z \in \Re^{n_1+n_2}$, we define the following three index sets:

$$\mathcal{J}_1 := \{i : z_i > 0,\ n_1+1 \le i \le n_1+n_2\} \cup \{1,2,\ldots,n_1\},$$
$$\mathcal{J}_2 := \{i : z_i = 0,\ n_1+1 \le i \le n_1+n_2\},$$
$$\mathcal{J}_3 := \{i : z_i < 0,\ n_1+1 \le i \le n_1+n_2\}.$$

By direct calculation, it follows that $\mathcal{V}$ is an element of $\partial\Pi_{\mathcal{Q}^*}(z)$ if and only if there exists a vector $a \in [0,1]^{|\mathcal{J}_2|}$ such that

$$\mathcal{V}h = \begin{bmatrix} h_{\mathcal{J}_1} \\ a \circ h_{\mathcal{J}_2} \\ 0 \end{bmatrix}, \ \forall\, h \in \Re^{n_1+n_2}. \tag{4.24}$$

Remark 4.2. In the implementation of our inexact semismooth Newton-CG method, we need to select an element $\mathcal{V}_1^0 \in \partial\Pi_{\mathcal{B}}(Y)$ and an element $\mathcal{V}_2^0 \in \partial\Pi_{\mathcal{Q}^*}(z)$. If $\|Y\|_* \le 1$, $\mathcal{V}_1^0$ is chosen as the identity operator from $\Re^{m\times n}$ to $\Re^{m\times n}$. For the case where Y is outside of $\mathcal{B}$, we take $U^\infty = U$, $V^\infty = V$ and $N = k_1(\sigma)$ in (4.23). Thus $\beta_2 = \emptyset$ and for any $H \in \Re^{m\times n}$,

$$\mathcal{V}_1^0 H = U\,[W_1 \quad W_2]\,V^T, \tag{4.25}$$

where the matrices $W_1 \in \Re^{m\times m}$ and $W_2 \in \Re^{m\times(n-m)}$ are defined by

$$W_1 = \Omega^\infty \circ S(\widetilde{H}_1) + \Gamma^\infty \circ T(\widetilde{H}_1) - \frac{\mathrm{Tr}(\widetilde{H}_{11})}{N}\begin{bmatrix} I_N & 0 \\ 0 & 0 \end{bmatrix},$$

$$W_2 = \begin{bmatrix} \Upsilon^\infty \mathbf{1}^T_{n-m} \\ 0 \end{bmatrix} \circ \widetilde{H}_2,$$

with $\widetilde{H}_1 \in \Re^{m\times m}, \widetilde{H}_2 \in \Re^{m\times(n-m)}$, $[\widetilde{H}_1\ \widetilde{H}_2] = U^T H V$ and $\widetilde{H}_{11}$ being the matrix extracted from the first N columns and rows of $\widetilde{H}_1$, and

$$\Omega^\infty = \begin{bmatrix} \mathbf{1}_{N\times N} & \Omega^\infty_{\beta_1\gamma_2} \ \Omega^\infty_{\beta_1\beta_5} \\ (\Omega^\infty_{\beta_1\gamma_2})^T & \\ (\Omega^\infty_{(\beta_1\beta_5})^T & \mathbf{0}_{(m-N)\times(m-N)} \end{bmatrix},$$

$$\Gamma^{\infty} = \begin{bmatrix} \Gamma^{\infty}_{\beta_1\gamma_1} & \Gamma^{\infty}_{\beta_1\gamma_2} \; \Gamma^{\infty}_{\beta_1\beta_5} \\ (\Gamma^{\infty}_{\beta_1\gamma_2})^T & \\ (\Gamma^{\infty}_{\beta_1\beta_5})^T & \mathbf{0}_{(m-N)\times(m-N)} \end{bmatrix}.$$

As to the selection of $\mathcal{V}_2^0$, we take $a = 0$ in (4.24) and

$$\mathcal{V}_2^0 h = \begin{bmatrix} h_{\mathcal{J}_1} \\ 0 \end{bmatrix}, \; \forall\, h \in \Re^{n_1+n_2}. \tag{4.26}$$

4.2.2 Constraint nondegeneracy

For the convergence analysis of the semismooth Newton method, we need the concept of constraint nondegeneracy which is originally introduced by Robinson [70] and extended by Bonnans and Shapiro [6]. Let $\mathcal{X}$ and $\mathcal{Y}$ be two finite dimensional space, $\Phi : \mathcal{X} \to \mathcal{Y}$ be a continuously differentiable function and $\mathcal{C}$ be a closed convex set. We use $T_{\mathcal{C}}(x)$ and $\mathrm{lin}(T_{\mathcal{C}}(x))$ to denote the tangent cone of $\mathcal{C}$ at x and its linearity space, respectively. A feasible point $\bar{x}$ to the feasibility problem $\{\Phi(x) \in \mathcal{C},\ x \in \mathcal{X}\}$ is constraint nondegenerate if

$$\Phi'(\bar{x})\mathcal{X} + \mathrm{lin}(T_{\mathcal{C}}(\Phi(\bar{x}))) = \mathcal{Y}.$$

Thus the constraint nondegeneracy condition associated with the minimizer $(\widehat{Z}, \hat{w})$ of (4.7) has the form

$$\begin{bmatrix} \mathcal{A} & B^T \\ \mathcal{I} & 0 \\ 0 & \mathcal{I} \end{bmatrix} \begin{pmatrix} \Re^{m\times n} \\ \Re^{n_1+n_2} \end{pmatrix} + \begin{bmatrix} \{0\}^p \\ \mathrm{lin}(T_{\mathcal{B}}(\widehat{Z})) \\ \mathrm{lin}(T_{\mathcal{Q}^*}(\hat{w})) \end{bmatrix} = \begin{bmatrix} \Re^p \\ \Re^{m\times n} \\ \Re^{n_1+n_2} \end{bmatrix}, \tag{4.27}$$

or equivalently,

$$\mathcal{A}\mathrm{lin}(T_{\mathcal{B}}(\widehat{Z})) + B^T \mathrm{lin}(T_{\mathcal{Q}^*}(\hat{w})) = \Re^p. \tag{4.28}$$

Proposition 4.3. *Let $(\widehat{Z}, \hat{w})$ be the unique solution pair of* (4.7)*. Let $\widehat{Z}$ have the following SVD:*

$$\widehat{Z} = U[\mathrm{diag}(\sigma(\widehat{Z}))\ 0]V^T = [U_1\ U_2][\mathrm{diag}(\sigma(\widehat{Z}))\ 0][V_1\ V_2]^T,$$

where $\sigma_1(\widehat{Z}) \geq \ldots \geq \sigma_r(\widehat{Z}) > 0 = \sigma_{r+1}(\widehat{Z}) = \ldots = \sigma_m(\widehat{Z})$, $\sigma(\widehat{Z}) := (\sigma_1(\widehat{Z}), \ldots, \sigma_m(\widehat{Z}))^T$, and $U = [U_1\ U_2]$ and $V = [V_1\ V_2]$ with $U_1 \in \Re^{m\times r}, U_2 \in \Re^{m\times(m-r)}, V_1 \in \Re^{n\times r}, V_2 \in$

$\Re^{n\times(n-r)}$. *Define the following two index sets* κ_1 *and* κ_2 *by*

$$\kappa_1 := \{1, 2, \ldots, n_1\} \cup \{i|\, \hat{w}_i > 0,\, n_1 + 1 \leq i \leq n_1 + n_2\},$$
$$\kappa_2 := \{i|\, \hat{w}_i = 0,\, n_1 + 1 \leq i \leq n_1 + n_2\}.$$

Then it holds that:

(i) *if* $\|\widehat{Z}\|_* < 1$*, the constraint nondegeneracy holds at* $(\widehat{Z}, \hat{w})$ *if and only if*

$$\begin{cases} B_{\kappa_1} y = 0 \\ \mathcal{A}^* y = 0 \end{cases} \Longrightarrow y = 0. \tag{4.29}$$

(ii) *if* $\|\widehat{Z}\|_* = 1$*, the constraint nondegeneracy holds at* $(\widehat{Z}, \hat{w})$ *if and only if, for any given* $k \in \Re$*,*

$$\begin{cases} B_{\kappa_1} y = 0 \\ (U_1)^T(\mathcal{A}^* y)V_1 = kI_r \\ (U_1)^T(\mathcal{A}^* y)V_2 = 0 \\ (U_2)^T(\mathcal{A}^* y)V_1 = 0 \end{cases} \Longrightarrow y = 0. \tag{4.30}$$

Proof. (i) Under the condition that $\|\widehat{Z}\|_* < 1$, it is easy to see that

$$\text{lin}(T_{\mathcal{B}}(\widehat{Z})) = \Re^{m\times n}$$

and

$$\text{lin}(T_{\mathcal{Q}^*}(\hat{w})) = \begin{bmatrix} \Re^{|\kappa_1|} \\ \{0\}^{|\kappa_2|} \end{bmatrix}.$$

Thus the constraints nondegeneracy condition (4.28) is reduced to

$$\mathcal{A}\Re^{m\times n} + B_{\kappa_1}^T \Re^{|\kappa_1|} = \Re^p, \tag{4.31}$$

which, by taking orthogonal complement, is equivalent to (4.29).

(ii) Since

$$\text{lin}(T_{\mathcal{Q}^*}(\hat{w})) = \begin{bmatrix} \Re^{|\kappa_1|} \\ \{0\}^{|\kappa_2|} \end{bmatrix},$$

the constraints nondegeneracy condition (4.28) is reduced to

$$\mathcal{A}\text{lin}(T_{\mathcal{B}}(\widehat{Z})) + B_{\kappa_1}^T \Re^{|\kappa_1|} = \Re^p, \tag{4.32}$$

which is equivalent to

$$\begin{cases} B_{\kappa_1} y = 0 \\ \mathcal{A}^* y \in \mathrm{lin}(T_{\mathcal{B}}(\widehat{Z}))^{\perp} \end{cases} \Longrightarrow y = 0. \tag{4.33}$$

Since from [86, Example 2] that

$$\partial \|X\|_* = \{U_1 V_1^T + U_2 T V_2^T \,|\, T \in \Re^{(m-r)\times(n-r)},\ \|T\|_2 \le 1\},$$

it follows that

$$\begin{aligned} \|\cdot\|_*(X;H) &= \max_{V\in\partial\|X\|_*} \langle V,\, H\rangle \\ &= \mathrm{Tr}((U_1)^T H V_1) + \max_T\{\langle U_2 T V_2^T, H\rangle \,|\, \|T\|_2 \le 1\,\} \\ &= \mathrm{Tr}((U_1)^T H V_1) + \max_T\{\langle U_2 T V_2^T, H\rangle \,|\, \|T\|_2 \le 1\,\} \\ &= \mathrm{Tr}((U_1)^T H V_1) + \max_T\{\langle T, U_2^T H V_2\rangle \,|\, \|T\|_2 \le 1\,\} \\ &= \mathrm{Tr}((U_1)^T H V_1) + \|U_2^T H V_2\|_*. \end{aligned} \tag{4.34}$$

Then by [16, Proposition 2.3.6,Theorem 2.4.9], one can establish that

$$T_{\mathcal{B}}(\widehat{Z}) = \{H \in \Re^{m\times n} |\, \mathrm{Tr}((U_1)^T H V_1) + \|(U_2)^T H V_2\|_* \le 0\},$$

and therefore obtain

$$\begin{aligned} \mathrm{lin}(T_{\mathcal{B}}(\widehat{Z})) &= T_{\mathcal{B}}(\widehat{Z}) \cap \{-T_{\mathcal{B}}(\widehat{Z})\} \\ &= \{H \in \Re^{m\times n} |\, H V_1 \in (U_1)^{\perp},\ (U_2)^T H V_2 = 0\}. \end{aligned}$$

This implies

$$\begin{aligned} \mathrm{lin}(T_{\mathcal{B}}(\widehat{Z}))^{\perp} &= \{Y \in \Re^{m\times n} \,|\, \langle Y,\, H\rangle = 0,\ \forall\, H \in \mathrm{lin} T_{\mathcal{B}}\}, \\ &= \{Y \in \Re^{m\times n} \,|\, \langle U^T Y V,\, U^T H V\rangle = 0,\ \forall\, H \in \mathrm{lin} T_{\mathcal{B}}\}, \\ &= \{Y \in \Re^{m\times n} \,|\, \langle U^T Y V, \begin{bmatrix} U_1^T H V_1 & U_1^T H V_2 \\ U_2^T H V_1 & 0 \end{bmatrix} = 0,\ H V_1 \in (U_1)^{\perp}\,\} \\ &= \{Y \in \Re^{m\times n} \,|\, \exists k \in \Re,\ (U_1)^T Y V_1 = k I_r, (U_1)^T Y V_2 = 0, (U_2)^T Y V_1 = 0\}, \end{aligned} \tag{4.35}$$

which, together with (6.37), completes the proof. □

With the above proposition, we next establish a result which exploits the close relationship between the constraint nondegeneracy of the optimal solution of (4.7) and the negative definiteness of the elements of $\hat{\partial}^2 \theta_k$.

Proposition 4.4. *Suppose that the problem* (4.7) *satisfies the Slater condition* (4.6). *Let* $(\widehat{Z}, \hat{w})$ *and* $\hat{y}$ *denote, respectively, the optimal solutions of* (4.7) *and* (4.9). *Then the following conditions are equivalent:*

(i) *The constraint nondegeneracy condition* (4.28) *holds at* $(\widehat{Z}, \hat{w})$.

(ii) *Every element in* $\hat{\partial}^2\theta_k(\hat{y})$ *is symmetric and negative definite.*

(iii) *The operator*

$$\mathcal{V}_0 = -\lambda_k(\mathcal{A}\mathcal{V}_1^1\mathcal{A}^* + B^T\mathcal{V}_2^0 B)$$

is symmetric and negative definite, where $\mathcal{V}_1^1$ *is the same as* $\mathcal{V}_1^0$ *in* (4.25) *except when in the case of* $\|\widehat{W}\|_* = 1$, *the operator is defined by* (4.23) *with* N *being the rank of* $\widehat{W}$, *where* $\widehat{W} := Z^k - \lambda_k(\mathcal{A}^*\hat{y} - A_0)$.

Proof. Assume that the SVD of $\widehat{Z}$ and the index sets κ_1, κ_2 are given in Proposition 6.2.

"(i) $\Rightarrow$ (ii)". Let $\mathcal{V}$ be an arbitrary element of $\hat{\partial}^2\theta_k(\hat{y})$. Then there exist $\mathcal{V}_1$ and $\mathcal{V}_2$ in $\partial\Pi_{\mathcal{B}}[Z^k - \lambda_k(\mathcal{A}^*\hat{y} - A_0)]$ and $\partial\Pi_{\mathcal{Q}^*}[w^k - \lambda_k(B\hat{y} - b)]$, respectively, such that

$$\mathcal{V} = -\lambda_k[\mathcal{A}\mathcal{V}_1\mathcal{A}^* + B^T\mathcal{V}_2 B].$$

Therefore, $\mathcal{V}$ is self-adjoint. Moreover, it follows from [59, Proposition 1] that for any $h \in \Re^p$,

$$\begin{aligned}
\langle h,\, \mathcal{V}h\rangle &= -\lambda_k\langle h,\, \mathcal{A}\mathcal{V}_1\mathcal{A}^*h\rangle - \lambda_k\langle h,\, B^T\mathcal{V}_2Bh\rangle \\
&= -\lambda_k\langle \mathcal{A}^*h,\, \mathcal{V}_1\mathcal{A}^*h\rangle - \lambda_k\langle Bh,\, \mathcal{V}_2Bh\rangle \\
&\le -\lambda_k\langle \mathcal{V}_1\mathcal{A}^*h,\, \mathcal{V}_1\mathcal{A}^*h\rangle - \lambda_k\langle \mathcal{V}_2Bh,\, \mathcal{V}_2Bh\rangle \qquad (4.36)\\
&\le 0,
\end{aligned}$$

which implies that $\mathcal{V}$ is negative semidefinite. To complete the proof of this part, it suffices to show that $\mathcal{V}$ is nonsingular. Consider the following linear system

$$\mathcal{V}h = 0, \quad \text{or equivalently,} \quad \begin{cases} \mathcal{V}_1\mathcal{A}^*h = 0, \\ \mathcal{V}_2Bh = 0. \end{cases} \qquad (4.37)$$

Now we proceed to prove that $h = 0$ by considering the following two cases.

Case 1: $\|\widehat{W}\|_* < 1$.
In this case, $\|\widehat{Z}\|_* < 1$ and $\mathcal{V}_1 = \mathcal{I}$. Then, it follows from (4.37) that

$$\begin{cases} \mathcal{A}^* h = 0, \\ B_{\kappa_1} h = 0, \end{cases} \tag{4.38}$$

which, together with the constraint nondegeneracy assumption (4.29), implies that $h = 0$.

Case 2: $\|\widehat{W}\|_* \geq 1$ and thus $\|\widehat{Z}\|_* = 1$.
We first show the nonsingularity of $\mathcal{V}$ for the choice that $\mathcal{V}_1 = \mathcal{I}$. In this situation, (4.38) still holds and hence by taking $k = 0$ in (4.30), we know that $\mathcal{V}$ is negative definite. Next, we turn to the case in which $\mathcal{V}_1$ is another element selected from $\partial_{\mathcal{E}^0}\Pi_{\mathcal{B}}(\widehat{W})$. We consider two sub-cases.

Case 2.1: $\|\widehat{W}\|_* = 1$. Let $H = \mathcal{A}^* h$. In view of the analysis in the previous subsection, we know from $\mathcal{V}_1(H) = 0$ that

$$\begin{aligned} 0 \;=\; & U^\infty \left[\left[\begin{array}{ccc} \multicolumn{2}{c}{\mathbf{1}_{N\times N}} & \begin{array}{c}\mathbf{1}_{r\times(m-N)} \\ \Omega^\infty_{\alpha_2\alpha_3}\end{array} \\ \mathbf{1}_{(m-N)\times r} & (\Omega^\infty_{\alpha_2\alpha_3})^T & 0 \end{array} \right] \circ S(\widetilde{H}_1) \quad \left[\begin{array}{c} \mathbf{1}_{r\times(n-m)} \\ \Upsilon^\infty_{\alpha_2}\mathbf{1}^T_{n-m} \\ 0 \end{array} \right] \circ \widetilde{H}_2 \right] (V^\infty)^T \\ & + \; U^\infty \left(\left[\begin{array}{ccc} \mathbf{1}_{r\times r} & \mathbf{1}_{r\times(N-r)} & \mathbf{1}_{r\times(m-N)} \\ \mathbf{1}_{(N-r)\times r} & \Gamma^\infty_{\alpha_2\alpha_2} & \Gamma^\infty_{\alpha_2\alpha_3} \\ \mathbf{1}_{(m-N)\times r} & (\Gamma^\infty_{\alpha_2\alpha_3})^T & 0 \end{array} \right] \circ T(\widetilde{H}_1) - \frac{\mathrm{Tr}(\widetilde{H}_{11})}{N} \left[\begin{array}{cc} I_N & 0 \\ 0 & 0 \end{array} \right] \right) (V_1^\infty)^T, \end{aligned}$$

where $V_1^\infty \in \Re^{m\times m}$, $V_2^\infty \in \Re^{m\times(n-m)}$ and $V^\infty := [V_1^\infty \; V_2^\infty]$, the index sets $\alpha_1, \alpha_2, \alpha_3$ and α_4 are defined as in (4.17). This implies that

$$\begin{cases} (U^\infty_{\alpha_1})^T (H) V^\infty_{\alpha_1} = \dfrac{1}{N} I_r, \\ (U^\infty_{\alpha_1})^T (H) V^\infty_{\alpha_2\cup\alpha_3\cup\alpha_4} = 0, \\ (U^\infty_{\alpha_2\cup\alpha_3})^T (H) V^\infty_{\alpha_1} = 0. \end{cases} \tag{4.39}$$

By Proposition 2.3, there exist orthogonal matrices $Q \in \Re^{r\times r}$, $Q' \in \Re^{(m-r)\times(m-r)}$ and $Q'' \in \Re^{(n-r)\times(n-r)}$ such that

$$\begin{cases} U^\infty_{\alpha_1} = U_{\alpha_1} Q, \\ U^\infty_{\alpha_2\cup\alpha_3} = U_{\alpha_2\cup\alpha_3} Q', \\ V^\infty_{\alpha_1} = Q V_{\alpha_1}, \\ V^\infty_{\alpha_2\cup\alpha_3\cup\alpha_4} = Q'' V_{\alpha_2\cup\alpha_3\cup\alpha_4}. \end{cases} \tag{4.40}$$

Moreover, we know from $\mathcal{V}_2 h = 0$ that

$$B_{\kappa_1} h = 0. \tag{4.41}$$

Combining (4.30), (4.41), (4.39) with (6.56), we deduce $h = 0$ and hence $\mathcal{V}$ is negative definite.

Case 2.2: $\|\widehat{W}\|_* > 1$. The proof of the negative definiteness of $\mathcal{V}$ is similar to that of Case 2.1, with the equality (4.23) replacing (4.21).

By taking the convex hull of $\partial_{\mathcal{E}^0}\Pi_{\mathcal{B}}(\widehat{W})$, we complete the proof of the first part.

"(ii) $\Rightarrow$ (iii)". This is trivial since $\mathcal{V}_0 \in \hat{\partial}^2\theta_k(\hat{y})$.

"(iii) $\Rightarrow$ (i)". Assume the contrary that the constraint nondegeneracy condition fails to hold at $(\widehat{Z}, \hat{w})$. Again, we consider two cases.

Case 1: $\|\widehat{W}\|_* < 1$ and hence $\|\widehat{Z}\|_* < 1$.

By assumption, there exists a $z \neq 0$ such that

$$\begin{cases} \mathcal{A}^* z = 0, \\ B_{\kappa_1} z = 0. \end{cases}.$$

namely,

$$\mathcal{V}_0 z = 0.$$

This means that $\mathcal{V}_0$ is singular, which contradicts to (iii).

Case 2: $\|\widehat{W}\|_* \geq 1$ and hence $\|\widehat{Z}\|_* = 1$.

By assumption, there exist $k \in \Re$ and $z \neq 0$ such that

$$\begin{cases} B_{\kappa_1} z = 0, \\ (U_1)^T(\mathcal{A}^* z)V_1 = kI_r, \\ (U_1)^T(\mathcal{A}^* z)V_2 = 0, \\ (U_2)^T(\mathcal{A}^* z)V_1 = 0. \end{cases} \tag{4.42}$$

Using the equalities above, simple computation yields $\mathcal{V}_0 z = 0$. This contradicts to the statement (iii). The proof is completed. □

4.2.3 A semismooth Newton-CG algorithm

In this subsection, we briefly describe the semismooth Newton-CG algorithm for solving (4.9). The basic template of the algorithm is given as follows. For simplicity, we drop the outer iteration index k.

Algorithm 4.2 (An inexact semismooth Newton-CG method)

Step 0. Given $\varsigma \in (0, 0.5), \eta \in (0, 1), \tau_1, \tau_2 \in (0, 1)$ and $\rho \in (0, 1)$. Choose $y^0 \in \Re^p$.

Step 1. For $j = 0, 1, 2, \ldots$,

Compute

$$\eta_j := \min\{\eta, \|\nabla_y \theta_k(y^j)\|^{1+\tau}.$$

Step 1.1. Apply the PCG method to find an approximation solution d^j to

$$(\mathcal{V}_j - \epsilon_j I)d = -\nabla_y \theta_k(y^j), \tag{4.43}$$

where $\mathcal{V}_j = -\lambda_k(\mathcal{A}\mathcal{V}_1^0\mathcal{A}^* + B^T\mathcal{V}_2^0 B)$ and $\epsilon_j = \tau_1 \min\{\tau_2, \|\nabla_y \theta_k(y^j)\|\}$, such that d^j satisfies the following condition:

$$\|(\mathcal{V}_j - \epsilon_j I)d + \nabla_y \theta_k(y^j)\| \leq \eta_j.$$

Step 1.2. Let m_j be the smallest nonnegative integer m satisfying

$$\theta_k(y^j + \rho^m d^j) - \theta_k(y^j) \geq \varsigma\rho^m \langle \nabla_y \theta_k(y^j), d^j \rangle.$$

Set $\alpha_j := \rho^{m_j}$ and $y^{j+1} := y^j + \alpha_j d^j$.

From the structures of $\mathcal{V}_1^0$ and $\mathcal{V}_2^0$, we know that $\mathcal{V}$ is always negative semidefinite. Hence $\mathcal{V} - \epsilon_j I$ is always negative definite as long as $\nabla_y \theta_k(y^j) \neq 0$. So, it is reasonable for us to apply the PCG method to solve (4.43). Furthermore, by noting the strong semismoothness of $\Pi_{\mathcal{B}}(\cdot)$ and $\Pi_{\mathcal{Q}^*}(\cdot)$, and using a theorem similar to [95, Theorem 3.4], we can easily derive the following convergence result for Algorithm 4.2.

Theorem 4.5. *Suppose that the Slater condition holds for* (4.7). *Then the inexact semismooth Newton-CG algorithm* 4.2 *is well defined and any accumulation point* $\hat{y}$ *of* $\{y^j\}$ *generated by algorithm* 4.2 *is an optimal solution to the innear subproblem* (4.9).

Theorem 4.6. *Assume that the Slater condition holds for* (4.7). *Let* $\hat{y}$ *be an accumulation point of the infinite sequence* $\{y^j\}$ *generated by the Newton-CG algorithm for solving* (4.9). *Suppose that at each step* $j \geq 0$, *when the PCG algorithm terminates, the tolerance* η_j *is achieved, i.e.,*

$$\|(\mathcal{V}_j - \epsilon_j I)d + \nabla_y \theta_k(y^j)\| \leq \eta_j.$$

Assume that the constraint nondegeneracy condition (4.28) *holds at* $Z^k - \lambda_k(\mathcal{A}^*(\hat{y}) - A_0)$. *Then the whole sequence* $\{y^j\}$ *converges to* $\hat{y}$ *with* η *order convergence rate, i.e.,*

$$\|y^{j+1} - \hat{y}\| = O(\|y^j - \hat{y}\|^{1+\eta}).$$

4.3 Numerical issues

In applying the Newton-CG method to solve the innear problem (4.9), the most expensive step is to compute the direction from the linear equation (4.43). As is well known, the basic operation in implementing the PCG method is to calculate the multiplication $V_0 y$ for any given $y \in \Re^p$. With the exception of the trivial case in which $\|Z^k - \lambda_k(\mathcal{A}^* y - A_0)\|_* \leq 1$, we know from the analysis in subsection 4.2.1 that a full SVD appears to be necessary.

For a problem in which m is moderate but n is large, the full SVD computation would be expensive and huge memory space is also needed to store the large and dense matrix V. However, as explained in [49], this can be done indirectly via an economical SVD and a QR factorization. First, we can compute the economical SVD of $Z^k - \lambda_k(\mathcal{A}^* y - A_0)$ as

$$Z^k - \lambda_k(\mathcal{A}^* y - A_0) = U\Sigma V_1^T$$

and construct $V_2 \in \Re^{n\times(n-r)}$ by computing the QR factorization of V_1

$$V_1 = QR := [Q_1 \ V_2]R,$$

where $Q \in \Re^{n\times n}$ is an orthogonal matrix and $R \in \Re^{n\times r}$ is upper triangular. In the numerical implementation, Householder transformations are used to compute the QR factorization and only the Householder vectors are stored to compute the matrix-vector product involving V_2. After the SVD is done, one can easily calculate $\mathcal{V}_0^1 H$ via (4.25) for any given $H \in \Re^{m\times n}$ in about $4n(mk_1(\sigma) + nk_1(\sigma) + m^2 - k_1(\sigma)^2)$ flops. The above computational complexity shows that our algorithm is able to utilize any low rank or flat rectangular structure of a matrix to reduce the computational cost.

In fact, one can completely avoid the computation of V_2 by carefully analyzing the structure of (4.25). The part W_2 in (4.25) is given as follows:

$$\begin{aligned}
[W_2 V_2^T]_{1:k_1(\sigma)} &:= \Big((\Upsilon^\infty \mathbf{1}_{n-m}^T) \circ (U_{\beta_1}^T H V_2)\Big) V_2^T \\
&= \mathrm{diag}(\Upsilon^\infty) U_{\beta_1}^T H (V_2 V_2^T) \\
&= \mathrm{diag}(\Upsilon^\infty) U_{\beta_1}^T H (I - V_1 V_1^T). \qquad (4.44)
\end{aligned}$$

From the above equation, it is obvious that one need not compute V_2 in order to evaluate W_2 in (4.25). After the SVD is done, one can easily calculate $\mathcal{V}_0^1 H$ via (4.25) for any given $H \in \Re^{m\times n}$ in about $O(k_1(\sigma)mn)$ flops.

Next, we introduce two diagonal preconditioners to accelerate the convergence of the CG method applied to solve the linear system (4.43). Let $\mathbf{A}$ and $\mathbf{V}$ be the matrix representation of the linear mappings $\mathcal{A}$ and $\mathcal{V}_1^0$, respectively. Then the coefficient matrix in (4.43) has the following form

$$W = -\lambda \mathbf{A}\mathbf{V}\mathbf{A}^T - \lambda B_{\mathcal{J}_1}^T B_{\mathcal{J}_1} - \epsilon I.$$

Note that we have omitted the iteration index for brevity. Let the standard basis in $\Re^{m\times n}$ be $\{E_{ij} \in \Re^{m\times n} : 1 \leq i \leq m,\, 1 \leq n\}$, where E_{ij} is the matrix whose (i,j)-th entry is one and zero otherwise. The diagonal element of $\mathcal{V}_1^0$ with respect to the standard basis is given by

$$\begin{aligned} \mathbf{V}_{(i,j),(i,j)} = \;& \langle \mathcal{V}_1^0 E_{ij},\, E_{ij}\rangle \\ = \;& ((U\circ U)\Lambda^\infty (V\circ V))_{ij} - \frac{1}{k_1(\sigma)}((U_1' V_1'^T)\circ(U_1' V_1'^T))_{ij} \\ & + \frac{1}{2}\langle H_{ij}\circ H_{ij}^T,\, \Omega^\infty - \Gamma^\infty\rangle, \end{aligned} \tag{4.45}$$

where

$$\Lambda^\infty = \begin{bmatrix} \frac{1}{2}(\Omega^\infty + \Gamma^\infty) & \Upsilon^\infty \mathbf{1}_{k_1(\sigma)\times(n-m)}^T \\ & \mathbf{0}_{(m-k_1(\sigma))\times(n-k_2(\sigma))} \end{bmatrix}, \quad H_{ij} = U^T E_{ij} V_1,$$

and U_1' and V_1' are the matrices formed by the first $k_1(\sigma)$ columns of U and V, respectively. To avoid excessive computational cost, we only calculate the first two terms on the right side of (4.45)

$$\mathbf{D}_{(i,j),(i,j)} = ((U\circ U)\Lambda^\infty (V\circ V))_{ij} - \frac{1}{k_1(\sigma)}((U_1' V_1'^T)\circ(U_1' V_1'^T))_{ij}, \tag{4.46}$$

as a good approximation of (4.45). Thus we propose the following diagonal preconditioner for the coefficient matrix

$$M = \lambda \mathrm{Diag}(\mathbf{A}\mathbf{D}\mathbf{A}^T + B_{\mathcal{J}_1}^T B_{\mathcal{J}_1}) + \epsilon I. \tag{4.47}$$

Clearly, to use the preconditioner above, we need the explicit form of V, which may lead to memory difficulty when n is large. Thus when n is too large for V to be stored explicitly, we just use the following simple diagonal preconditioner

$$M' = \lambda \mathrm{Diag}(\mathbf{A}\mathbf{A}^T + B^T B) + \epsilon I. \tag{4.48}$$

Chapter 5

Numerical results of ADM and SNDPPA for matrix norm approximation problems

In this Chapter, we first employ the alternating direction method (ADM) to solve the MNA problem and mainly report the numerical performance of the SNDPPA and the ADM we have implemented to solve four different types of problem. All the codes are written in MATLAB 7.6 and run an Intel 2.10GHz PC with 4GB memory.

We use R_p, R_d and gap to denote respectively the primal infeasibility, dual infeasibility and primal-dual relative gap, namely

$$R_p = \frac{\|[\mathcal{A}^* y + X - A_0; \Pi_{Q^*}(b - By)]\|}{1 + \|[A_0; b]\|}, \ \ R_d = \frac{\|\mathcal{A}Z + B^T w\|}{1 + \|[A; B^T]\|},$$

and

$$\text{gap} = \frac{|\text{pobj} - \text{dobj}|}{1 + |\text{pobj}| + |\text{dobj}|},$$

where pobj and dobj are the primal and dual objective values, respectively.

In our experiments, we start the ADM from the point $(X, y, z, Z, w) = (0, 0, 0, 0, 0)$ and is stopped when

$$\max\{R_p, R_d\} < 10^{-6} \tag{5.1}$$

or the maximum number of iterations exceeds 2000. Furthermore, the penalty parameter β in the ADM is adjusted according to the following rule dynamically. Starting from the initial value of 10, we adjust β at every fifth step as follows:

$$\beta_{k+1} = \begin{cases} \min(10^3, 2\beta_k), & \text{if } R_p^k / R_d^k < 0.1, \\ \max(10^{-2}, 0.5\beta_k), & \text{if } R_p^k / R_d^k > 10, \\ \beta_k, & \text{otherwise.} \end{cases} \tag{5.2}$$

For the SNDPPA, we first use the proposed ADM to generate a good starting point by running ADM for at most 50 iterations. The ADM is stopped when $\max\{R_p, R_d\} < 5 \cdot 10^{-3}$. The SNDPPA is stopped when the condition

$$\max\{R_p, R_d\} < 10^{-6}$$

is met. For each outer iteration, we cap the number of Newton iterations for solving an inner subproblem to 40. In solving the linear system associated with Newton direction, the maximal number of PCG steps is set as 500. As the parameter λ plays a critical role in the convergence speed of a PPA-based algorithm, we need to tune it carefully. In our implementation, the parameter λ is initialized as 10 and updated according to the following rule:

$$\lambda_{k+1} = \begin{cases} 3\lambda_k, & R_p^{k+1}/R_p^k > 0.5 \text{ and } R_p^{k+1} > 10^{-4}, \\ 2\lambda_k, & R_p^{k+1}/R_p^k > 0.5 \text{ and } R_p^{k+1} < 10^{-4}, \\ \lambda_k, & \text{otherwise}, \end{cases} \tag{5.3}$$

where λ_k is the kth value of the penalty parameter.

5.1 Random matrix norm approximation

We first consider randomly generated matrix norm approximation problems with/without constraints. In the experiments, the matrices $A_0, A_1, \ldots, A_p$ are generated independently from the multivariate uniform distribution on $[0, 1]^{m\times n}$.

In Table 5.1, we report the numerical performance of the SNDPPA and ADM for solving different random matrix approximation instances without constraints. The number of outer iterations (iter), primal infeasibility (R_p), dual infeasibility (R_d), primal objective value (pobj), relative gap (gap), and the CPU time (time) taken are listed in the table. To better understand the performance of the SNDPPA, we also report the number of Newton systems solved (itersub) and the average number PCG steps (pcg) taken to solve each of the systems.

$p \mid m \mid n$	Algo.	it (itersub | pcg)	pobj | gap	$R_p \mid R_d$	time
300 | 300 | 300	PPA	14(15|3.9)	9.44515934 0| 2.8-6	4.4-7| 3.2-8	22.8
	ADM	300	9.44520938 0| 4.2-6	9.7-7| 2.5-7	73.8
500 | 500 | 500	PPA	17(18|3.9)	1.22905150 1| 3.7-6	4.3-7| 2.3-8	117.3
	ADM	619	1.22905586 1| 1.9-5	6.8-7| 9.9-7	672.8
100 | 100 | 3000	PPA	16(18|3.8)	1.83807818 1| 8.3-6	9.4-7| 4.1-8	31.8
	ADM	821	1.83807914 1| 6.5-6	9.9-7| 1.7-7	268.6
100 | 100 | 5000	PPA	16(17|3.8)	2.31039070 1| 5.6-6	9.3-7| 4.0-8	56.3
	ADM	443	2.31040515 1| 3.2-6	9.9-7| 9.7-7	282.3
100 | 100 | 10000	PPA	18(19|3.8)	3.16771120 1| 2.8-6	5.4-7| 1.1-7	127.0
	ADM	740	3.16774836 1| 7.4-6	9.5-7| 9.9-7	1096.7

$p \mid m \mid n$	Algo.	it (itersub | pcg)	pobj | gap	$R_p \mid R_d$	time
100 | 100 | 20000	PPA	16(17|3.8)	4.37704442 1| 1.2-7	7.3-7| 2.7-9	546.1
	ADM	654	4.37704413 1| 9.9-6	4.6-7| 9.9-7	5774.7

Table 5.1: Results for unconstrained random matrix norm approximation problems.

As can been observed in Table 5.1, both the ADM and SNDPPA are able to solve the unconstrained random matrix approximation problems to relatively high accuracy. The SNDPPA substantially outperforms the ADM in terms of the CPU time taken to solve the problems. For example, the ADM takes about 1.5 hours to solve the last instance while our SNDPPA solves it in 10 minutes and with a better accuracy. It is worth noting that for the instances with $(p, m) = (100, 100)$, the CPU time taken by each iteration of the SNDPPA and ADM increases almost linearly with n. But for a solver (say the algorithm in [95]) that attempts to solve (1.1) via the SDP reformulation (1.14), the cost per iteration would grow at least quadratically in n. This observation is consistent with the fact mentioned in the previous section that our SNDPPA is capable of exploiting the flat rectangular structure of the matrices involved.

Next, we test the SNDPPA on the MNA problems with constraints. A simple example is to find a convex combination of given matrices $A_0, A_1, \ldots, A_p$ having the minimal spectral norm, i.e.,

$$\begin{aligned} \min \quad & \|A_0 - \mathcal{A}^* y\|_2 \\ \text{s.t.} \quad & \sum_{i=1}^{p} y_i = 1, \quad y \geq 0. \end{aligned} \tag{5.4}$$

In what follows, we investigate the performance of the SNDPPA and ADM applied to (6.59) where the matrices $A_1, \ldots, A_p$ are randomly generated as before.

$p \mid m \mid n$	Algo.	it (itersub | pcg)	pobj | gap	$R_p \mid R_d$	time
300 | 300 | 300	PPA	17(32|13.0)	9.59409291 0| 2.7-5	8.0-7| 3.2-8	90.2
	ADM	2000	9.59309978 0| 3.8-6	4.2-6| 6.8-6	474.8
500 | 500 | 500	PPA	19(36|14.7)	1.24537638 1| 5.1-5	7.3-7| 6.0-7	447.7
	ADM	2000	1.24556407 1| 1.2-4	1.1-5| 1.1-5	8822.5

$p \mid m \mid n$	Algo.	it (itersub | pcg)	pobj | gap	$R_p \mid R_d$	time
100 | 100 | 3000	PPA	21(27|4.7)	1.83873310 1| 3.5-6	5.2-7| 2.4-8	48.4
	ADM	2000	1.83863304 1| 1.6-4	7.1-6| 6.7-6	678.3
100 | 100 | 5000	PPA	18(24|5.2)	2.31091411 1| 3.5-6	4.2-7| 2.8-8	78.8
	ADM	2000	2.31077010 1| 1.0-4	8.3-6| 2.8-6	1325.5
100 | 100 | 10000	PPA	19(25|4.7)	3.16803831 1| 6.2-6	8.7-7| 1.6-8	169.6
	ADM	2000	3.16798808 1| 1.0-5	3.0-6| 9.9-7	2957.5
100 | 100 | 20000	PPA	21(25|3.9)	4.37736720 1| 2.2-5	7.8-7| 9.3-7	965.3
	ADM	2000	4.37716525 1| 6.2-5	4.5-6| 2.1-6	17563.7

Table 5.2: Results for the matrix norm approximation problem (6.59).

Table 5.2 lists the numerical results obtained by the SNDPPA and ADM. For this collection of problems, we can easily see the superiority of the SNDPPA over the first order algorithm ADM. While our SNDPPA solves all the tested instances to the accuracy of 10^{-6} within 36 semismooth Newton iterations, the ADM fails to achieve the required accuracy even after 2000 iterations. For the problem with $(m, n, p) = (100, 20000, 100)$, the ADM obtains a solution with the the accuracy in the order of $5 * 10^{-6}$ after running for 4.5 hours while our SNDPPA is able to solve the problem in about 15 mins. As one may deduce from the results in Table 2, the ADM may encounter both computational and accuracy difficulties when constraints are imposed on y.

5.2 Chebyshev polynomials of matrices

In this subsection, we apply the proposed SNDPPA to compute the Chebyshev polynomials of a given matrix A. Since the power basis $I, A, \ldots, A^t$ is usually highly ill conditioned, in [83] the authors suggested replacing this basis by a better-conditioned alternative $Q_1, Q_2, \ldots, Q_{t+1}$ and consider the resulting problem

$$\min_{y \in \Re^t} \|Q_{t+1} - \sum_{i=1}^{t} y_i Q_i\|_2. \tag{5.5}$$

From the solution of (5.5), one can easily compute the coefficients of the Chebyshev polynomials via Theorem 2 in [83]. In our experiments, the test examples are taken from Section 6 in [83] and $Q_1, Q_2, \ldots, Q_{t+1}$ is the orthogonal basis corresponding to the power basis of A.

problem	Algo.	$N \mid t$	it (itersub | pcg)	pobj | gap	$R_p \mid R_d$	time
Rand	PPA	500 | 50	16(18|7.7)	2.19977200-1| 3.5-7	3.4-7| 1.8-7	46.3
		1000 | 100	14(15|8.7)	1.84595015-1| 1.4-7	2.5-7| 1.8-7	306.1
	ADM	500 | 50	354	2.19977287-1| 6.2-8	3.0-7| 8.7-7	145.0
		1000 | 100	661	1.84594835-1| 3.1-7	1.1-7| 9.1-7	2490.3
Randtri	PPA	500 | 50	6(9|10.9)	4.14987173-1| 8.5-8	4.1-8| 2.7-7	27.4
		1000 | 100	9(12|13.9)	3.56518026-1| 4.8-7	3.0-7| 6.6-7	279.9
	ADM	500 | 50	663	4.14987166-1| 1.4-6	7.1-8| 8.8-7	252.9
		1000 | 100	786	3.56517734-1| 1.1-6	7.3-8| 9.1-7	2030.6
Diag	PPA	500 | 50	16(44|9.3)	7.20404744-2| 5.8-8	1.4-9| 9.3-8	51.7
		1000 | 100	15(27|9.1)	4.85090599-2| 1.8-7	3.1-8| 1.1-7	264.4
	ADM	500 | 50	2000	7.20772064-2| 3.0-4	3.0-5| 3.6-4	308.6
		1000 | 100	396	4.85093725-2| 2.7-7	4.6-7| 8.7-7	557.5
Bidiag	PPA	500 | 50	11(37|19.1)	1.90877134-1| 6.0-7	2.0-7| 4.1-7	85.6
		1000 | 100	18(80|38.9)	1.38036105-1| 4.6-7	2.0-7| 2.8-7	2212.3
	ADM	500 | 50	1482	1.90877146-1| 2.2-7	2.0-7| 8.8-7	364.1
		1000 | 100	2000	1.38036596-1| 5.4-7	1.0-6| 9.1-7	4122.5
Ellipse	PPA	500 | 50	9(14|3.8)	5.51257423-2| 5.5-10	1.7-7| 1.7-9	14.2
		1000 | 100	11(20|4.4)	3.90141290-2| 6.7-11	4.7-7| 4.6-10	188.0
	ADM	500 | 50	269	5.51257424-2| 4.4-7	4.7-7| 8.7-7	62.2
		1000 | 100	370	3.90141292-2| 3.6-7	4.8-7| 5.5-7	771.0
Grcar	PPA	500 | 50	17(42|8.6)	7.19041068-2| 6.5-8	3.8-7| 3.2-8	57.3
		1000 | 100	11(25|8.4)	5.07326772-2| 2.5-7	4.0-7| 4.0-7	313.0
	ADM	500 | 50	2000	7.19051700-2| 2.2-6	2.4-6| 7.3-7	547.6
		1000 | 100	865	5.07326793-2| 9.0-7	2.6-7| 9.1-7	2056.3
Lemniscate1	PPA	500 | 50	2(2|1.0)	4.71404521-2| 4.2-15	4.1-10| 6.0-15	3.3
		1000 | 100	2(2|0.5)	3.33333333-2| 1.0-14	2.5-10| 5.4-15	24.4
	ADM	500 | 50	17	4.71404521-2| 8.0-6	4.2-7| 4.5-14	4.7
		1000 | 100	17	3.33333333-2| 5.6-6	2.9-7| 1.2-13	45.8
Lemniscate2	PPA	500 | 50	17(78|9.2)	8.09231049-2| 8.7-8	2.7-7| 3.0-7	116.5
		1000 | 100	20(76|18.3)	3.33334398-2| 8.9-7	3.6-7| 1.6-7	1231.1
	ADM	500 | 50	1104	8.09229960-2| 3.1-8	2.1-7| 8.7-7	315.6
		1000 | 100	1154	3.33337454-2| 1.8-6	4.9-7| 8.6-7	2534.3

problem	Algo.	$N \mid t$	it (itersub | pcg)	pobj | gap	$R_p \mid R_d$	time
Wilkinson	PPA	500 | 50	14(26|6.0)	2.02888114-1| 3.5-8	1.9-7| 6.9-8	27.8
		1000 | 100	13(29|7.6)	1.92544545-1| 5.8-7	1.5-7| 2.7-7	276.9
	ADM	500 | 50	859	2.02888114-1| 1.2-6	5.0-7| 8.5-7	194.0
		1000 | 100	1723	1.92544547-1| 9.2-8	5.8-8| 9.1-7	3487.7
Chebyshev	PPA	500 | 50	10(15|8.7)	2.24960671-1| 2.8-6	1.5-7| 7.2-7	23.3
		1000 | 100	12(18|11.7)	2.06618546-1| 8.8-7	3.8-7| 7.8-7	300.4
	ADM	500 | 50	788	2.24960549-1| 1.3-6	3.4-7| 8.7-7	209.3
		1000 | 100	2000	2.06629894-1| 2.7-4	4.8-6| 2.4-5	4030.0

Table 5.3: Chebyshev polynomials of matrices.

Table 5.3 shows that for most of the test instances, both the SNDPPA and ADM are capable of obtaining solutions with an accuracy of less than 10^{-6}. However, for hard examples such as Bidiag(1000), the ADM fails to solve it within 2000 iterations while the SNDPPA succeeds in achieving the required accuracy for all the instances. This illustrates that our SNDPPA performs much more stably than the ADM. Moreover, the SNDPPA is much more efficient than the ADM in terms of computing time. Specifically, the former is about 5 to 10 times faster than the latter. This is not surprising since for most instances, the SNDPPA takes less than 30 semismooth Newton iterations to generate a highly accurate solution and the average number of PCG steps needed to solve each of the Newton systems is less than 15.

5.3 FMMC/FDLA Problem

In this subsection, we investigate the numerical performance of the two algorithms for solving the fastest Markov mixing chain (FMMC) and fastest distributed linear averaging (FDLA) problems. We first generate a family of graphs, all with 1000 vertices as follows. First we generate a symmetric matrix $\Re \in \Re^{1000 \in 1000}$, whose entries R_{ij} for $i \leq j$, are independent and uniformly distributed on the interval $[0, 1]$.

$p \mid n$	Algo.	it (itersub | pcg)	pobj | gap	$R_p \mid R_d$	time
3960| 1000	PPA	4(4|2.0)	1.00000000 0| 2.1-15	2.0-8| 0.0-16	62.6
FDLA	ADM	34	1.00000000 0| 1.7-6	7.4-7| 2.8-8	72.0

$p \mid n$	Algo.	it (itersub | pcg)	pobj | gap	$R_p \mid R_d$	time
3960| 1000	PPA	7(7|3.4)	1.00000000 0| 1.5-14	5.2-9| 0.0-16	95.4
FMMC	ADM	2000	1.00000000 0| 4.3-3	1.7-6| 1.6-5	4154.8
8988| 1000	PPA	12(42|35.1)	4.49231089-1| 3.9-4	4.3-7| 6.2-7	838.4
FDLA	ADM	2000	4.49650982-1| 5.8-4	6.9-5| 2.0-5	6829.1
8988| 1000	PPA	16(49|65.7)	4.56475004-1| 3.2-6	7.6-7| 8.7-8	2780.8
FMMC	ADM	2000	4.56682293-1| 3.4-4	5.5-6| 1.7-5	7262.0
13882| 1000	PPA	12(34|44.3)	3.49261397-1| 3.1-4	4.0-7| 4.3-7	1416.1
FDLA	ADM	2000	3.49460210-1| 3.9-4	3.8-5| 1.9-5	7288.6
13882| 1000	PPA	19(44|53.8)	3.51488679-1| 3.1-6	3.7-7| 4.6-8	2161.3
FMMC	ADM	2000	3.53454384-1| 1.8-3	1.9-4| 6.1-5	7391.8
19032| 1000	PPA	10(22|31.8)	2.83598169-1| 1.8-5	4.4-7| 1.1-7	767.1
FDLA	ADM	2000	2.84265573-1| 2.8-3	5.4-5| 6.0-5	7726.7
19032| 1000	PPA	14(42|46.7)	2.84032758-1| 2.2-6	5.8-7| 1.5-8	1991.3
FMMC	ADM	2000	2.84198686-1| 1.2-3	9.9-6| 3.0-5	7599.6
24094| 1000	PPA	11(18|31.6)	2.45400905-1| 8.7-5	5.0-7| 1.6-7	640.4
FDLA	ADM	2000	2.45475940-1| 6.3-4	4.1-6| 2.6-5	7141.7
24094| 1000	PPA	16(36|83.2)	2.45692830-1| 1.9-7	3.7-7| 9.9-9	2892.0
FMMC	ADM	2000	2.85592249-1| 1.4-1	3.5-4| 2.7-4	7494.2
29170| 1000	PPA	14(15|15.0)	2.17515735-1| 3.2-6	9.3-7| 6.2-8	354.0
FDLA	ADM	2000	2.17595850-1| 4.1-4	1.5-5| 2.0-5	7599.6
29170| 1000	PPA	18(36|54.7)	2.17715035-1| 4.1-7	9.4-7| 3.4-9	2116.8
FMMC	ADM	2000	2.17985919-1| 9.2-4	3.4-5| 2.9-5	8705.7

Table 5.4: Performance of the SNDPPA and ADM for FMMC/FDLA problems on random connected graphs.

As we can see clearly, our SNDPPA outperforms the ADM both in the sense of accuracy and CPU time. While our SNDPPA is able to achieve the required accuracy of less than 10^{-6} for all the test examples, the adm only succeed to achieve this accuracy for the first instance. Moreover, the relative gap obtained by the ADM is of the order 10^{-3} to 10^{-3}, which is substantially lower than that obtained by the SNDPPA.

The following tested graph instances are taken from the University of Florida sparse matrix collection [18] but some are slightly modified to make them connected. The data set is available at: http://www2.research.att.com/∼gyifanhu/GALLERY/GRAPHS/search.html.

problem	$p \mid n$	Algo.	it (itersub | pcg)	pobj | gap	$R_p \mid R_d$	time
FDLA–Cage	2562| 366	PPA	4(6|1.7)	4.58554209-1| 9.2-6	6.8-7| 3.0-9	5.5
		ADM	2000	4.75216058-1| 3.4-1	1.1-3| 1.3-3	231.8
FMMC–Cage	2562| 366	PPA	5(6|1.7)	4.58545022-1| 5.7-7	4.1-8| 7.6-10	4.4
		ADM	2000	5.87384195-1| 4.1-1	7.0-5| 1.5-3	198.0
FDLA–Erdos981	1381| 485	PPA	6(6|2.7)	1.00000000 0| 2.3-14	0.0-8| 0.0-16	6.6
		ADM	20	1.00000000 0| 3.7-5	7.1-7| 3.0-8	5.5
FMMC–Erdos981	1381| 485	PPA	7(9|4.0)	1.00000000 0| 1.1-7	0.0-9| 4.5-9	10.2
		ADM	2000	1.00001117 0| 1.3-3	3.9-5| 1.8-5	777.8
FDLA–NotreDame_yeast	2203| 2114	PPA	11(15|4.1)	1.00000000 0| 4.6-14	0.0-9| 0.0-16	1395.1
		ADM	97	1.00000000 0| 3.5-13	5.8-7| 1.3-8	2562.1
FMMC–NotreDame_yeast	2203| 2114	PPA	8(8|3.0)	1.00000000 0| 2.0-12	0.0-8| 0.0-16	741.1
		ADM	2000	1.00002029 0| 6.7-3	7.5-6| 2.4-5	49781.3
FDLA–G46	9990| 1000	PPA	10(32|31.1)	4.17341777-1| 6.0-6	4.2-7| 2.8-7	590.0
		ADM	429	4.17346539-1| 9.4-5	7.2-7| 9.9-7	971.6
FMMC–G46	9990| 1000	PPA	12(30|38.5)	4.19935830-1| 1.9-6	8.4-7| 4.4-8	629.5
		ADM	2000	4.21142429-1| 9.8-4	1.2-4| 6.2-5	4228.1
FDLA–G15	4661| 800	PPA	15(39|36.4)	7.31904357-1| 7.2-5	9.4-7| 4.3-7	484.3
		ADM	1122	7.31899758-1| 4.1-4	3.7-7| 9.9-7	1345.1
FMMC–G15	4661| 800	PPA	14(72|92.0)	7.85245424-1| 8.5-5	6.4-7| 3.8-7	1520.4
		ADM	2000	7.86626977-1| 1.1-3	4.6-6| 3.6-5	2418.5
FDLA–G43	9990| 1000	PPA	11(51|39.7)	4.21305708-1| 1.5-5	4.7-7| 1.1-7	1279.3
		ADM	490	4.21308585-1| 1.6-4	5.3-7| 9.9-7	1216.9
FMMC–G43	9990| 1000	PPA	16(49|59.7)	4.25983919-1| 4.2-7	4.6-7| 2.6-8	1637.9
		ADM	2000	4.26209610-1| 8.5-4	6.3-6| 3.3-5	4808.9
FDLA–G54	5916| 1000	PPA	13(55|51.3)	7.32246398-1| 1.9-4	5.8-7| 8.0-7	1877.5
		ADM	2000	7.33611791-1| 2.0-4	7.8-5| 2.9-5	5402.2
FMMC–G54	5916| 1000	PPA	15(71|75.6)	7.86520590-1| 2.5-5	7.1-7| 1.5-7	2847.4
		ADM	2000	7.88923019-1| 1.7-3	4.2-6| 3.5-5	5227.8
FDLA–G3	19176| 800	PPA	11(16|31.8)	2.40597734-1| 3.2-4	9.3-7| 9.6-7	208.7
		ADM	2000	2.41026286-1| 3.9-4	1.1-4| 5.3-6	2407.0
FMMC–G3	19176| 800	PPA	16(24|43.7)	2.40914608-1| 2.4-7	5.4-7| 1.0-8	388.9
		ADM	2000	2.41009134-1| 8.7-4	5.4-6| 3.0-5	2445.3

Table 5.5: Performance of the SNDPPA and ADM for FMMC/FDLA problems on connected graphs.

Table 5.5 shows that our SNDPPA is able to achieve the required accuracy of less than 10^{-6} for all the test examples. However, by comparing the results for FMMC/FDLA with those for random matrix approximation and Chebyshev polynomial problems, we see that the SNDPPA is slower for the former cases. This behavior is understandable because for FMMC/FDLA problems, the average PCG steps taken to compute the Newton directions and the total number of semismooth Newton iterations are significantly larger. For the FMMC problem, it is also not surprising that the ADM fails to obtain solutions with the desired accuracy after 2000 iterations for all the instances. In fact, the ADM can only obtain an approximate solution with the accuracy in the order 10^{-4} to 10^{-5} for about 50% of the instances and the relative gap in the order 10^{-3} to 10^{-4} for most computed examples.

Chapter 6

A squared smoothing Newton method

6.1 Introduction

In this section, we briefly review the squared smoothing Newton method for the nonsmooth equations $F(x) = 0$, where $F : \Re^n \to \Re^n$ is a locally Lipschitz continuous function. Then, we apply this method to solving the equivalent nonsmooth equation reformulation of the matrix norm approximation problems.

The feature of smoothing methods is to construct a smoothing approximation function $G_u : \Re_{++} \times \Re^n \to \Re^n$ of F such that for $\varepsilon > 0$, $G_u(\varepsilon, \cdot)$ is continuously differentiable on $\Re^n$ and for any $x \in \Re^n$,

$$\|G_u(\varepsilon, x) - F(x)\| \to 0, \text{ as } \varepsilon \downarrow 0$$

and then to find a solution of $F(x) = 0$ by solving a sequential smooth equations for a positive sequence $\{\varepsilon^k\}$, $k = 0, 1, 2, \dots,$

$$G_u(\varepsilon^k, x) = 0.$$

With the kth iterate (ε^k, x^k), a natural idea to generate x^{k+1} is via

$$x^{k+1} = x^k - t_k[(G_u)_x{}'(\varepsilon^k, x^k)]^{-1}F(x^k), \tag{6.1}$$

where $\varepsilon^k > 0$, $(G_u)_x{}'(\varepsilon^k, x^k)$ is the derivative of G with respect to x at (ε^k, x^k) and $t_k > 0$ is the stepsize. The smoothing Newton method (6.1) has attracted much attention from lots of researchers, see [15, 65, 66] and references therein. Under certain conditions depending strongly on the Jacobian consistency property, they proved that each accumulation point is a solution of $F(x) = 0$. In [68], the authors proposed a class of squared smoothing Newton method to solve the nonsmooth equation $F(x) = 0$ and established its convergence without the Jacobian consistency property condition. Define the operator $E : \Re \times \Re^n \to \Re^{n+1}$ by

$$E_u(\varepsilon, x) := \begin{bmatrix} \varepsilon \\ G_u(\varepsilon, x) \end{bmatrix}$$

for any $(\varepsilon, x) \in \Re \times \Re^n$. Let ϕ_u be the merit function associated with E_u, i.e.,

$$\phi_u(\varepsilon, x) := \|E_u(\varepsilon, x)\|^2 .$$

Choose $r \in (0, 1)$. Let

$$\psi_u(\varepsilon, x) := r \min\{1, \phi_u(\varepsilon, x)\}.$$

The the smoothing Newton method [68] can be briefly descried as follows.

Algorithm 6.1 (A squared smoothing Newton method)

Step 0. Choose $\hat{\varepsilon} \in (0, +\infty)$ such that $\delta := r\hat{\varepsilon} < 1$. Select constants $\rho \in (0, 1)$ and $\sigma \in (0, 1/2)$. Let $\varepsilon^0 := \hat{\varepsilon}$ and $x^0 \in \Re^n$ be an arbitrary point. $k := 0$.

Step 1. If $E_u(\varepsilon^k, x^k) = 0$ then stop. Otherwise, compute $\beta_k := \psi_u(\varepsilon^k, x^k)$.

Step 2. Solving the following equation

$$E_u(\varepsilon^k, x^k) + E_u'(\varepsilon^k, x^k) \begin{bmatrix} \Delta\varepsilon^k \\ \Delta x^k \end{bmatrix} = \begin{bmatrix} \beta_k \hat{\varepsilon} \\ 0 \end{bmatrix}, \tag{6.2}$$

where $\Delta\varepsilon^k := -\varepsilon^k + \beta_k \hat{\varepsilon}$.

Step 3. Let l_k be the smallest nonnegative integer l satisfying

$$\phi_u(\varepsilon^k + \rho^{l_k}\Delta\varepsilon^k,\, x^k + \rho^l \Delta x^k) \leq [1 - 2\sigma(1-\sigma)\rho^l]\phi_u(\varepsilon^k,\, x^k). \tag{6.3}$$

Define

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

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

Let $\mathcal{K}$ be the epigraph cone of the spectral norm, i.e., $\mathcal{K} = \{(t, X) \mid t \geq \|X\|_2\}$. Introducing an auxiliary scalar $t \in \Re$, we can rewrite the matrix norm approximation problem (1.1) as

$$\begin{array}{ll} \min & t \\ \text{s.t.} & \mathcal{A}^* y + X = A_0, \\ & By \in b + Q,\ (t, X) \in \mathcal{K}. \end{array} \tag{6.4}$$

Assume the strong duality holds for the problem (6.4) and its dual and there exists at least one saddle point. Then solving the MNA problem is equivalent to the following KKT system:

$$\begin{cases} (t, X) = \Pi_{\mathcal{K}}(t-1, X+Z) \\ \mathcal{A}Z + B^T w = 0, \\ \mathcal{A}^* y + X = A_0, \\ B_1 y = b_1, \\ w_2 = \Pi_{\Re_+^{n_2}}(w_2 - B_2 y + b_2), \end{cases} \tag{6.5}$$

where Z and w are Lagrangian multipliers; $w^T = [w_1^T \; w_2^T]$ with $w_1 \in \Re^{n_1}$ and $w_2 \in \Re^{n_2}$. Write $W^T = [t \; X^T \; y^T \; Z^T \; w^T]$ and let

$$T(\varepsilon, W) = \begin{bmatrix} (t, X) - G(\varepsilon, t-1, X+Z) \\ \mathcal{A}Z + B^T w \\ \mathcal{A}^* y + X - \mathcal{A}_0 \\ B_1 y - b_1 \\ w_2 - H_u(\varepsilon, w_2 - B_2 y + b_2) \end{bmatrix}. \tag{6.6}$$

Redefine the operator $E : \Re \times \Re^{m\times n} \times \Re^{m\times n} \times \Re^n \times \Re^p$ by

$$E_u(\varepsilon, W) := \begin{bmatrix} \varepsilon \\ T(\varepsilon, W) \end{bmatrix}.$$

Let ϕ_u be the merit function associated with E_u, i.e.,

$$\phi_u(\varepsilon, W) := \|E_u(\varepsilon, W)\|^2 .$$

Choose $r \in (0, 1)$. Let

$$\psi_u(\varepsilon, W) := r \min\{1, \phi_u(\varepsilon, W)\}.$$

Thus the smoothing Newton algorithm 6.1 can be directly applied to solve the KKT system (6.5), which byproduct solves the MNA problem (1.1).

Remark 6.1. Up to now, a lot of variants of smoothing Newton methods have been proposed to solve nonsmooth equations. Most of them can be easily extended to solve the problem (1.1). For instance, the inexact smoothing Newton method developed in [31].

6.2 The Newton systems

For given $(\varepsilon, t, X) \in \Re_{++} \times \Re \times \Re^{m\times n}$, let $\mathcal{J}_1$ and $\mathcal{J}_2$ denote the partial derivatives of $G(\cdot,\cdot,\cdot)$ with respect to the last two variables and the first argument at (ε, t, X), respectively. We first claim that the linear system $\mathcal{J}_1[t_1,\, Y] = [t_2,\, X]$ has a unique solution. Suppose X admits the following singular value decomposition:

$$X = U[\mathrm{Diag}(\sigma)\, 0]V^T,$$

where $\sigma = [\sigma_1, \sigma_2, \ldots, \sigma_m]^T$ and $U \in \Re^{m\times m}$ and $V \in \Re^{n\times}$ are orthogonal matrices. Let $\tilde{Y} = U^T Y V$ and $\tilde{X} = U^T X V$. Let $y(\cdot,\cdot,\cdot), s(\cdot,\cdot,\cdot)$ and $H(\cdot,\cdot,\cdot)$ be defined by (2.22). Denote by h' and g' the partial Jacobian of $s(\cdot,\cdot,\cdot)$ and $y(\cdot,\cdot,\cdot)$ with respect to the last two variables at the point (ε, t, σ), respectively. For $\sigma_i \neq \sigma_j$, we write

$$g^-_{ij} = \frac{y_i - y_j}{\sigma_i - \sigma_j},$$

and

$$g^+_{ij} = \frac{y_i + y_j}{\sigma_i + \sigma_j}$$

for $\sigma_i\sigma_j \neq 0$. Let $u_1 > u_2 > \ldots > u_r > u_{r+1} = 0$ be the distinct singular values of X. Define $a_k := \{i \,|\, \sigma_i = u_k\}$. By Proposition 2.4 (ii), it holds that

$$\begin{cases} t_2 = h'\big[t_1; \mathrm{diag}(\tilde{Y})\big], \\ \tilde{X} = \Big[\Omega \circ S(\tilde{Y}_1) + \Gamma \circ T(\tilde{Y}_1) + \mathrm{Diag}(g'\big[t_1; \mathrm{diag}(\tilde{Y})\big]),\ \mathcal{F} \circ \tilde{Y}_2\Big], \end{cases} \tag{6.7}$$

where $\tilde{Y} = [\tilde{Y}_1\ \tilde{Y}_2]$ with $\tilde{Y}_1 \in \Re^{m\times m}$ and $\tilde{Y}_2 \in \Re^{m\times(n-m)}$; Ω, Γ and $\mathcal{F}$ are defined by (2.4)-(2.7). We consider the following five cases.

Case 1: $i \in a_{k_1}, j \in a_{k_2}$, where $1 \leq k_1 \neq k_2 \leq r$. In this case, it follows from (6.7) that

$$\begin{cases} \tilde{X}_{ij} = \frac{1}{2} g^-_{ij}(\tilde{Y}_{ij} + \tilde{Y}_{ji}) + \frac{1}{2} g^+_{ij}(\tilde{Y}_{ij} - \tilde{Y}_{ji}) \\ \tilde{X}_{ji} = \frac{1}{2} g^-_{ij}(\tilde{Y}_{ij} + \tilde{Y}_{ji}) + \frac{1}{2} g^+_{ij}(\tilde{Y}_{ji} - \tilde{Y}_{ij}) \end{cases}. \tag{6.8}$$

Then solving the linear equation directly yields that

$$\tilde{Y}_{ij} = \frac{(g^-_{ij} + g^+_{ij})\tilde{X}_{ij} - (g^-_{ij} - g^+_{ij})\tilde{X}_{ji}}{2 g^-_{ij} g^+_{ij}} \tag{6.9}$$

Case 2: $i \in a_k, j \in a_k$, where $i \neq j$ and $1 \leq k \leq r$. Let

$$\eta_k = g'_{i(i+1)} - g'_{i(j+1)}$$

and

$$g_{ij}^{+} = \frac{y_i + y_j}{\sigma_i + \sigma_j} = \frac{y_i}{\sigma_i}.$$

By equation (6.7), it follows that

$$\begin{cases} \tilde{X}_{ij} = \frac{1}{2}\eta_k(\tilde{Y}_{ij} + \tilde{Y}_{ji}) + \frac{1}{2}g_{ij}^{+}(\tilde{Y}_{ij} - \tilde{Y}_{ji}), \\ \tilde{X}_{ji} = \frac{1}{2}\eta_k(\tilde{Y}_{ij} + \tilde{Y}_{ji}) + \frac{1}{2}g_{ij}^{+}(\tilde{Y}_{ji} - \tilde{Y}_{ij}). \end{cases} \tag{6.10}$$

Then solving directly the linear equations above yield that

$$\tilde{Y}_{ij} = \frac{(\eta_k + g_{ij}^{+})\tilde{X}_{ij} - (\eta_k - g_{ij}^{+})\tilde{X}_{ji}}{2\eta_k g_{ij}^{+}}. \tag{6.11}$$

Case 3: $i \in a_k, j \in a_{r+1} \cup \{m+1, m+2, \ldots, n\}$ or $i \in a_{r+1}, j \in a_k$, where $1 \leq k \leq r$. In this case, we know from (6.7) that

$$\tilde{X}_{ij} = \frac{y_i}{\sigma_i}\tilde{Y}_{ij}, \tag{6.12}$$

which implies

$$\tilde{Y}_{ij} = \frac{\sigma_i \tilde{X}_{ij}}{y_i}. \tag{6.13}$$

Case 4: $i \in a_{r+1}, j \in a_{r+1} \cup \{m+1, m+2, \ldots, n\}$. Similarly, it follows easily from (6.7) that

$$\tilde{Y}_{ij} = \frac{\tilde{X}_{ij}}{g'_{i(i+1)}}. \tag{6.14}$$

Case 5: $i = j \in a_k$, where $1 \leq k \leq r$. In this case, by (6.7), it is easy to establish that

$$\begin{pmatrix} t_2 \\ \tilde{X}_{11} \\ \tilde{X}_{22} \\ \vdots \\ \tilde{X}_{ll} \end{pmatrix} = \begin{bmatrix} g' \\ h' \end{bmatrix}_{dd} \begin{pmatrix} t_1 \\ \tilde{Y}_{11} \\ \tilde{Y}_{22} \\ \vdots \\ \tilde{Y}_{ll} \end{pmatrix}, \tag{6.15}$$

where $d = \{1, 2, \ldots, l+1\}$ and $l = |a_1| + |a_2| + \ldots + |a_r|$. By Proposition 2.11, we know the symmetric matrix $\begin{bmatrix} g' \\ h' \end{bmatrix}_{dd}$ is positive definite and thus nonsingular. Hence, it holds that

$$\begin{pmatrix} t_1 \\ \tilde{Y}_{11} \\ \tilde{Y}_{22} \\ \vdots \\ \tilde{Y}_{ll} \end{pmatrix} = \begin{bmatrix} g' \\ h' \end{bmatrix}_{dd}^{-1} \begin{pmatrix} t_2 \\ \tilde{X}_{11} \\ \tilde{X}_{22} \\ \vdots \\ \tilde{X}_{ll} \end{pmatrix}. \tag{6.16}$$

Remark 6.2. By similar analysis, for given $(t_2,\ X) \in \Re \times \Re^{m\times n}$, it is not difficult to show the following system

$$\begin{bmatrix} t_1 \\ 0 \end{bmatrix} - \mathcal{J}_1 \begin{bmatrix} t_1 \\ Y \end{bmatrix} = \begin{bmatrix} t_2 \\ X \end{bmatrix}$$

also has a unique solution. To simplify the later discussion, we write this solution as

$$(t_1, Y) = \begin{bmatrix} \mathcal{P}_1 \\ \mathcal{P}_2 \end{bmatrix} (t_2, X).$$

With the above preparation, we are ready to prove the nonsingularity of the Newton system (6.2) in Algorithm 6.1.

Proposition 6.1. *Suppose $A_1, \cdots, A_p$ are linearly independent and B has full row rank. Then for any $W = (\varepsilon, t, X, y, Z, w) \in \Re \times \Re^{m\times n} \times \Re^{m\times n} \times \Re^q \times \Re^p$ with $\varepsilon \neq 0$, $E_u'(W)$ is nonsingular.*

Proof. Suppose there exists $(\Delta\varepsilon, \Delta t, \Delta X, \Delta y, \Delta Z, \Delta w) \in \Re \times \Re^{m\times n} \times \Re^{m\times n} \times \Re^q \times \Re^p$ such that

$$E_u'(\varepsilon, W)(\Delta\varepsilon, \Delta t, \Delta X, \Delta y, \Delta Z, \Delta w) = 0\,,$$

i.e.,

$$\begin{cases} \Delta\varepsilon = 0, \\ (\Delta t, \Delta X) = \mathcal{J}_1(\Delta t, \Delta X + \Delta Z) + \mathcal{J}_2 \Delta\varepsilon \\ \mathcal{A}\Delta Z + B^T \Delta w = 0 \\ \mathcal{A}^* \Delta y + \Delta X = 0 \\ B_1 \Delta y = 0 \\ \Delta w_2 = \mathcal{J}_3(\Delta w_2 - B_2 \Delta y) + \mathcal{J}_4 \Delta\varepsilon \end{cases}, \tag{6.17}$$

where $\mathcal{J}_3$ and $\mathcal{J}_4$ are the partial Jacobian of $H_u(\cdot,\cdot)$ with respect to the first and second variable at the point $(\varepsilon, w_2 - B_2 y + b_2)$. Since $\Delta\varepsilon = 0$, it follows from the second equality in (6.17) that

$$(0, \Delta Z) = (\mathcal{J}^{-1} - \mathcal{I})(\Delta t, \Delta X).$$

By the discussion on $\mathcal{J}_1^{-1}$ above and Proposition 2.11 and 2.12, it can be verify easily that $\langle \Delta X,\ \Delta Z\rangle \geq 0$; furthermore, $\Delta X = \Delta Z = 0$ if and only if ΔX is orthogonal to

ΔZ. Using the last equality of (6.17), we know there exists a diagonal matrix M with nonnegative diagonal entries such that $\Delta w_2 = -M(B_2\Delta y)$. Hence, it holds

$$\begin{aligned}\langle \Delta X\,, \Delta Z\rangle &= -\langle \mathcal{A}^*\Delta y,\, \Delta Z\rangle \\ &= -\langle \Delta y,\, \mathcal{A}\Delta Z\rangle \\ &= \langle \Delta y,\, B_1^T\Delta w_1 + B_2^T\Delta w_2\rangle \\ &= \langle B_1\Delta y,\, w_1\rangle + \langle B_2\Delta y,\, \Delta w_2\rangle \\ &= -\langle B_1\Delta y,\, MB_1\Delta y\rangle \\ &\leq 0, \end{aligned} \tag{6.18}$$

which implies the orthogonality of ΔX and ΔZ. This shows $\Delta X = \Delta Z = 0$ and therefore $\Delta t = 0$. Moreover since $A_1, A_2, \cdots, A_p$ are linearly independent and B has full row rank, we have $\Delta y = 0$ and $\Delta w = 0$. This means the linear system (6.17) has only zero solution and the proposition follows. □

Note that the Newton system (6.2) is equivalent to

$$\begin{cases} \Delta\varepsilon^k = -\varepsilon^k + \beta\hat{\varepsilon} \\ (\Delta t^k, \Delta X^k) - \mathcal{J}_1(\Delta t^k, \Delta X^k + \Delta Z^k) = \delta_1 \\ \mathcal{A}\Delta Z^k + B^T\Delta w^k = \delta_2 \\ \mathcal{A}^*\Delta y^k + \Delta X^k = \delta_3 \\ B_1\Delta y^k = \delta_4 \\ \Delta w_2^k - \mathcal{J}_3(\Delta w_2^k - B_2\Delta y^k) = \delta_5 \end{cases}, \tag{6.19}$$

where $\delta_1 = \mathcal{J}_2\Delta\varepsilon^k + G(\varepsilon^k, t^k - 1, X^k + Z^k) - (t^k, X^k), \delta_2 = -\mathcal{A}Z^k - B^Tw^k$, $\delta_3 = -\mathcal{A}^*y^k - X^k$, $\delta_4 = -B_1y^k$ and $\delta_5 = \mathcal{J}_4\Delta\varepsilon^k + H_u(\varepsilon^k, w_2^k - B_2y^k) - w_2^k$. By direct computation applied to (6.19) yields

$$\begin{cases} \Delta X^k = \delta_3 - \mathcal{A}^*\Delta y_k, \\ \begin{bmatrix} \Delta t^k \\ \Delta Z^k \end{bmatrix} = \begin{bmatrix} \mathcal{P}_1 \\ \mathcal{P}_2 \end{bmatrix} \left[(\mathcal{J}_1 - I)\begin{bmatrix} 0 \\ \Delta X^k \end{bmatrix} + \delta_1\right] \\ \Delta w_2^k = (I - \mathcal{J}_3)^{-1}\left[-\mathcal{J}_3B_2\Delta y^k + \delta_5\right] \\ \mathcal{A}\Delta Z^k + B^T\Delta w^k = \delta_2 \\ B_1\Delta y^k = \delta_4 \end{cases}. \tag{6.20}$$

Define the linear operator $\chi_1(\cdot)$ by

$$\chi_2(X) := \mathcal{P}_2(\mathcal{J}_1 - I)\begin{bmatrix} 0 \\ X \end{bmatrix}, \quad \forall\, X \in \Re^{m\times n}.$$

Substituting the first three equalities into the last two equalities, we can easily deduce that

$$\begin{cases} [\mathcal{A}\chi_2\mathcal{A}^* + B_2^T(I-\mathcal{J}_3)^{-1}\mathcal{J}_3 B_2]\Delta y^k - B_1^T\Delta w_1^k = \delta_6, \\ B_1\Delta y^k = \delta_4, \end{cases} \tag{6.21}$$

where

$$\delta_6 = \mathcal{A}(\chi_2\delta_3 + \mathcal{P}_2\delta_1) + B_2^T(I-\mathcal{J}_3)^{-1}\delta_5 - \delta_2.$$

Now, we are ready to describe clearly the procedure for solving the Newton system in the smoothing Newton method.

Algorithm 6.2: Solving Newton system (6.2)

Step 1. Computer $\Delta\varepsilon^k = -\varepsilon^k + \beta\hat{\varepsilon}$.

Step 2. Compute Δy^k and Δw_2^k from (6.21) (in fact this is the Schur complement equation of the original Newton system).

Step 3. Computer $\Delta X^k, \Delta t^k, \Delta Z^k$ and Δw_1^k from the first three equation in (6.20).

6.3 Convergence analysis

In this section, we establish the suplinear convergence of the smoothing Newton method under the constraint nondegeneracy conditions. Let $(\bar{t}, \bar{X}, \bar{y})$ be a feasible solution of (6.4). Then the primal constraint nondegeneracy of problem (6.4) at $(\bar{t}, \bar{X}, \bar{y})$ is

$$\begin{bmatrix} \mathcal{A}^* & \mathcal{I} \\ B & 0 \\ 0 & \mathcal{I} \end{bmatrix} \begin{bmatrix} \Re^d \\ \Re \times \Re^{m\times n} \end{bmatrix} + \begin{bmatrix} \{0\}^{m\times n} \\ \mathrm{lin}(T_Q(B\bar{y}-b)) \\ \mathrm{lin}(T_{\mathcal{K}}(\bar{t},\bar{X})) \end{bmatrix} = \begin{bmatrix} \Re^{m\times n} \\ \Re^s \\ \Re\times\Re^{m\times n} \end{bmatrix}, \tag{6.22}$$

or equivalently

$$\begin{bmatrix} \mathcal{A}^* \\ B \end{bmatrix} \Re^d + \begin{bmatrix} \hat{\mathcal{K}} \\ \mathrm{lin}(T_Q(B\bar{y}-b)) \end{bmatrix} = \begin{bmatrix} \Re^{m\times n} \\ \Re^s \end{bmatrix}, \tag{6.23}$$

where $\hat{\mathcal{K}}$ is the projection of $\mathrm{lin}(T_{\mathcal{K}}(\bar{t},\bar{X}))$ onto $\Re^{m\times n}$. Denote by $\mathcal{B}$ the unit nuclear ball. It is not difficult to see the dual constraint nondegeneracy associated with (6.4) at

$(\bar{Z}, \bar{w})$ is of the form

$$\begin{bmatrix} \mathcal{A} & B^T \\ \mathcal{I} & 0 \\ 0 & \mathcal{I} \end{bmatrix} \begin{bmatrix} \Re^{m\times n} \\ \Re^s \end{bmatrix} + \begin{bmatrix} \{0\}^d \\ \mathrm{lin}(T_{\mathcal{B}}(\bar{Z})) \\ \mathrm{lin}(T_{Q^*}(\bar{w})) \end{bmatrix} = \begin{bmatrix} \Re^d \\ \Re^{m\times n} \\ \Re^s \end{bmatrix}, \tag{6.24}$$

or equivalently

$$\mathcal{A}\mathrm{lin}(T_{\mathcal{B}}(\bar{Z})) + B^T \mathrm{lin}(T_{Q^*}(\bar{w})) = \Re^d. \tag{6.25}$$

Let $\tau_1, \tau_{21}, \tau_{22}$ and τ_3 denotes, respectively, the index sets identified by the equality constraints, the strongly active inequality constraints, the weakly active inequality constraints and the nonactive constraints. Let $\tau_2 = \tau_{21} \cup \tau_{22}$.

Proposition 6.2. *Let* $(\bar{t}, \bar{X}, \bar{y}, \bar{Z}, \bar{w})$ *be a solution of the smoothing system* (6.5)*. Then the following statements hold.*

(i) *There exist two orthogonal matrices* $U \in \Re^{m\times m}$ *and* $V \in \Re^{n\times n}$ *such that*

$$\bar{X} = U[\mathrm{Diag}(\sigma(\bar{X}))\ 0]V^T \quad \text{and} \quad \bar{Z} = U[\mathrm{Diag}(\sigma(\bar{Z}))\ 0]V^T$$

with $\sigma(\bar{X})$ *and* $\sigma(\bar{Z})$ *arranged in non-increasing order. Moreover, if the multiplicity of the largest singular value of* $\bar{X}$ *is* α *and* $\bar{X} \neq 0$*, then*

$$\sigma_1(\bar{Z}) \geq \sigma_2(\bar{Z}) \ldots \geq \sigma_\alpha(\bar{Z}) \geq 0 = \sigma_{\alpha+1}(\bar{Z}) = \ldots = \sigma_m(\bar{Z})$$

with $\sigma_1(\bar{Z}) + \sigma_2(\bar{Z}) + \ldots + \sigma_\alpha(\bar{Z}) = 1$.

(ii) *Suppose* $\bar{X} \neq 0$*. Then the primal constraint nondegeneracy holds at* $(\bar{t}, \bar{X}, \bar{y})$ *if and only if*

$$\begin{cases} \mathcal{A}X + B^T_{\tau_1\cup\tau_2} w = 0 \\ XV^{(2)} = 0, (U^{(2)})^T X = 0 \\ (U^{(1)})^T X V^{(1)} \in \mathcal{S}^\alpha \\ \mathrm{Tr}[(U^{(1)})^T X V^{(1)}] = 0 \end{cases} \Longrightarrow \begin{cases} X = 0 \\ w = 0 \end{cases}, \tag{6.26}$$

where $U^{(1)} = [U_1, U_2, \ldots, U_\alpha]$ *and* $V^{(1)} = [V_1, V_2, \ldots, V_\alpha]$.

(iii) *Suppose* $\bar{X} = 0$*. Then the primal constraint nondegeneracy holds at* $(0, 0, \bar{y})$ *if and only if*

$$\mathcal{A}X + B^T_{\tau_1\cup\tau_2} w = 0 \Longrightarrow \begin{cases} X = 0 \\ w = 0 \end{cases}. \tag{6.27}$$

(iv) *If* $\bar{X} \neq 0$*, then* $\|\bar{Z}\|_* = 1$ *and the dual constraint nondegeneracy holds at* $(\bar{Z}, \bar{w})$ *if and only if, for any* $k \in \Re$*,*

$$\begin{cases} B_{\tau_1 \cup \tau_{21}} y = 0 \\ (U^1)^T(\mathcal{A}^* y)V^1 = kI \\ (U^1)^T(\mathcal{A}^* y)V^2 = 0 \\ (U^2)^T(\mathcal{A}^* y)V^1 = 0 \end{cases} \Longrightarrow y = 0, \tag{6.28}$$

where $U^1 = [U_1, U_2, \ldots, U_\beta]$ *and* $V^1 = [V_1, V_2, \ldots, V_\beta]$ *and* $\beta \leq \alpha$ *is the number of the nonzero singular values of* $\bar{Z}$*.*

(v) *If* $\|\bar{Z}\|_* < 1$*, the dual constraint nondegeneracy holds at* $(\bar{Z}, \bar{w})$ *if and only if*

$$\begin{cases} B_{\tau_1 \cup \tau_{21}} y = 0 \\ \mathcal{A}^* y = 0 \end{cases} \Longrightarrow y = 0. \tag{6.29}$$

Proof. (i) From the first equality of (6.5) and the properties of projection, we know

$$\mathcal{K} \ni \begin{bmatrix} \bar{t} \\ \bar{X} \end{bmatrix} \perp \begin{bmatrix} 1 \\ -\bar{Z} \end{bmatrix} \in \mathcal{K}^*, \tag{6.30}$$

which implies

$$\bar{t} = \langle \bar{X}, \bar{Z} \rangle \quad \text{and} \quad \|\bar{Z}\|_* \leq 1.$$

By the von Neumann's trace inequality applied to $\bar{Y}$ and $\bar{Z}$, it holds that

$$\langle \bar{X}, \bar{Z} \rangle \leq \langle \sigma(\bar{X}), \sigma(\bar{Z}) \rangle \leq \sigma_1(\bar{X}). \tag{6.31}$$

Since $(\bar{t}, \bar{X}) \in \mathcal{K}$, then $\bar{t} \geq \sigma_1(\bar{X})$ and hence the inequality above becomes equality. Then part (i) of this proposition follows.

(ii) Since by direct computation $\text{lin}(T_Q(B\bar{y} - b))$ can be derived as follows:

$$\text{lin}(T_Q(B\bar{y} - b)) = \begin{bmatrix} \{0\}^{n_1 + |\tau_2|} \\ \Re^{|\tau_3|} \end{bmatrix},$$

the primal constraint nondegeneracy condition (6.23) reduces to

$$\begin{bmatrix} \mathcal{A}^* \\ B_{\tau_1 \cup \tau_2} \end{bmatrix} \Re^d + \begin{bmatrix} \hat{\mathcal{K}} \\ 0 \end{bmatrix} = \begin{bmatrix} \Re^{m \times n} \\ \Re^{n_1 + |\tau_2|} \end{bmatrix}, \tag{6.32}$$

which is equivalent to

$$\left\{\begin{array}{l}\mathcal{A}X + B^T_{\tau_1\cup\tau_2}w = 0\\ X \in \hat{\mathcal{K}}^{\perp}\end{array}\right. \Longrightarrow \left\{\begin{array}{l}X = 0\\ w = 0\end{array}\right. . \tag{6.33}$$

Since $(\bar{t}, \bar{X})$ is on the boundary of $\mathcal{K}$, by Clark's classical results on the characterization of the tangent cone [16, Proposition 2.3.6 and Theorem 2.4.9], we deduce that

$$\begin{aligned}T_{\mathcal{K}}(\bar{t}, \bar{X}) &= \{(s,d) \in \Re \times \Re^{m\times n} \,|\, \sigma_1'(\bar{X}; d) \leq s\}\\ &= \{(s,d) \in \Re \times \Re^{m\times n} \,|\, \lambda_1[\frac{(U^{(1)})^T dV^{(1)} + (V^{(1)})^T d^T U^{(1)}}{2}] \leq s\}.\end{aligned}$$

Then the lineality space of $\mathcal{K}$ can be written

$$\text{lin}T_{\mathcal{K}}(\bar{t}, \bar{X}) = \{(s,d) \in \Re \times \Re^{m\times n} \,|\, (U^{(1)})^T dV^{(1)} + (V^{(1)})^T d^T U^{(1)} = 2sI\}$$

and its projection $\hat{\mathcal{K}}$ is

$$\hat{\mathcal{K}}(\bar{t}, \bar{X}) = \{d \in \Re^{m\times n} \,|\, \exists s \in \Re, (U^{(1)})^T dV^{(1)} + (V^{(1)})^T d^T U^{(1)} = sI\}.$$

By this equality and a quick computation, it can be shown that the orthogonal complement of $\hat{\mathcal{K}}$ is given by

$$\hat{\mathcal{K}}^{\perp} = \{X \in \Re^{m\times n} |\, XV^{(2)} = 0,\ (U^{(2)})^T X = 0,\ (U^{(1)})^T XV^{(1)} \in \mathcal{S}^{\alpha},\ \text{Tr}((U^{(1)})^T XV^{(1)}) = 0\}.$$

Combining with (6.33), we know the conclusion of part (ii) holds.

(iii) Since $\bar{X} = 0$, it follows from [16, Proposition 2.3.6 and Theorem 2.4.9] that

$$\begin{aligned}T_{\mathcal{K}}(\bar{t}, \bar{X}) &= \{(s,d) \in \Re \times \Re^{m\times n} \,|\, \sigma_1'(\bar{X}; d) \leq s\}\\ &= \{(s,d) \in \Re \times \Re^{m\times n} \,|\, \|d\|_2 \leq s\},\end{aligned}$$

which implies $\text{lin}T_{\mathcal{K}}(\bar{t}, \bar{X}) = \{0\} \times \{0\}^{m\times n}$ and $\hat{\mathcal{K}} = \{0\}^{m\times n}$. Hence the primal constraint nondegeneracy condition (6.23) reduces to

$$\left[\begin{array}{c}\mathcal{A}^*\\ B_{\tau_1\cup\tau_2}\end{array}\right]\Re^d = \left[\begin{array}{c}\Re^{m\times n}\\ \Re^{n_1+|\tau_2|}\end{array}\right], \tag{6.34}$$

which is equivalent to

$$\mathcal{A}X + B^T_{\tau_1\cup\tau_2}w = 0 \Longrightarrow \left\{\begin{array}{l}X = 0\\ w = 0\end{array}\right. . \tag{6.35}$$

(iv) Direct computation establishes that

$$\operatorname{lin}(T_{Q^*}(\bar{w})) = \begin{bmatrix} \Re^{n_1+|\tau_{21}|} \\ \{0\}^{|\tau_{22}|+|\tau_3|} \end{bmatrix}$$

and then the dual constraints nondegeneracy condition (6.25) is reduced to

$$\mathcal{A}\operatorname{lin}(T_{\mathcal{B}}(\bar{Z})) + B^T_{\tau_1\cup\tau_{21}}\Re^{n_1+|\tau_{21}|} = \Re^d, \tag{6.36}$$

which is equivalent to

$$\begin{cases} B^T_{\tau_1\cup\tau_{21}}y = 0 \\ \mathcal{A}^* y \in \operatorname{lin}(T_{\mathcal{B}}(\bar{Z}))^{\perp} \end{cases} \Longrightarrow y = 0. \tag{6.37}$$

By part (ii) of this proposition we know $\bar{Z}$ is on the boundary of $\mathcal{B}$. Invoking [16, Proposition 2.3.6 and Theorem 2.4.9] and [71, Theorem 2.7], one can establish

$$\operatorname{lin}(T_{\mathcal{B}}(\bar{Z})) = \{Y \in \Re^{m\times n} |\, HV^1 \in (U^1)^{\perp},\, (U^2)^T HV^2 = 0\}.$$

A simple calculation shows

$$\operatorname{lin}(T_{\mathcal{B}}(\bar{Z}))^{\perp} = \{Y \in \Re^{m\times n} |\, \exists k \in \Re, (U^1)^T Y V^1 = kI_\beta, (U^1)^T Y V^2 = 0, (U^2)^T Y V^1 = 0\},$$

which, together with (6.37), completes the proof of this part.

(v) Since $\bar{Z} = 0$, by direct computation, it holds that

$$T_{\mathcal{B}}(\bar{Z}) = \Re^{m\times n} \text{ and } \operatorname{lin}(T_{\mathcal{B}}(\bar{Z})) = \Re^{m\times n}.$$

Substituting the equality above into (6.37), we can easily obtain the conclusion. □

Lemma 6.3. *Let $H(\cdot,\cdot,\cdot)$ be the smoothing function of $\Pi_\infty(\cdot,\cdot)$ defined by* (2.22) *and $\bar{t} \in \Re$ and $\bar{x} \in \Re^{m\times n}$ be given. Then for any $(t,x) \in \Re\times\Re^m$ and $V \in \partial H(0,\bar{\varepsilon},\bar{x})$, the following inequality holds:*

$$\langle V(0,t,x),\, (t,x)\rangle \geq \|V(0,t,x)\|^2. \tag{6.38}$$

Proof. Denote $\mathcal{M} = \{0\}\times\Re\times\Re^m$ whose Lebesgue measure is 0 and let

$$\partial_{\mathcal{M}} H(0,\bar{t},\bar{x}) := \left\{ \lim_{k\to\infty} H'(\varepsilon^k, t^k, x^k) : (\varepsilon^k, t^k, x^k) \to (0,\bar{t},\bar{x}),\, \varepsilon^k \neq 0 \right\}.$$

From Proposition 2.11 part (i), for any $(\varepsilon_k, t_k, x_k) \in \Re\backslash\{0\}\times\Re\times\Re^m$, it follows that

$$\langle H'(\varepsilon^k,t^k,x^k)(0,t,x),\, (t,x)\rangle \geq \|H'(\varepsilon^k,t^k,x^k)(0,t,x)\|^2. \tag{6.39}$$

By taking limits for $k \to +\infty$ in (6.39), we know the inequality (6.38) is valid for any $V \in \partial_{\mathcal{M}} H(0, \bar{t}, \bar{x})$. Let $V \in \partial H(0, \bar{t}, \bar{x})$. Since the generalized Jacobian is blind to sets of zero measure, then there exists a positive integer $m > 0$, $V_i \in \partial_{\mathcal{M}} H(0, \bar{t}, \bar{x})$, $\lambda_i \geq 0,\ i = 1, 2, \ldots, m$, such that $\sum_{i=1}^m \lambda_i = 1$ and $V = \sum_{i=1}^m \lambda_i V_i$. Therefore,

$$\begin{aligned} \langle V(0,t,x),\, (t,x) \rangle &= \textstyle\sum_{i=1}^m \lambda_i \langle V_i\,(0,t,x),\, (t,x) \rangle \\ &\geq \sum_{i=1}^m \lambda_i \| V_i\,(0,t,x) \|^2 \\ &\geq \left\| \sum_{i=1}^m \lambda_i V_i\,(0,t,x) \right\|^2 \\ &= \| V(0,t,x) \|^2, \end{aligned} \tag{6.40}$$

which completes the proof. □

Lemma 6.4. *Let* $(\bar{t}, \bar{X}, \bar{y}, \bar{Z}, \bar{w})$ *be a solution of the smoothing system* (6.5)*. Suppose* $\bar{t} > 0$*. For any* $V \in \partial_{\mathcal{M}} H(0, \bar{t} - 1, \sigma(\bar{X}) + \sigma(\bar{Z}))$ *and* $(t, x) \in \Re \times \Re^m$*, it holds that*

$$V(0,t,x) = \begin{bmatrix} f & f[1^\beta \ l \ 0^{m-\alpha}] \\ f[1^\beta \ l \ 0^{m-\alpha}]^T & \mathrm{Diag}([0^\beta \ 1 - l \ 1^{m-\alpha}]) + f[1^\beta \ l \ 0^{m-\alpha}]^T[1^\beta \ l \ 0^{m-\alpha}] \end{bmatrix} \begin{bmatrix} t \\ x \end{bmatrix},$$

where $f \in (0,1)$ *and* $l \in [0,1]^{\alpha - \beta}$.

Proof. Let V be an element in $\partial_{\mathcal{M}} H(0, \bar{t} - 1, \sigma(\bar{X}) + \sigma(\bar{Z}))$. Then there exists a sequence $\{(\varepsilon_k, t_k, x_k)\} \in \Re \backslash \{0\} \times \Re \times \Re^m$ with $(\varepsilon_k, t_k, x_k) \to (0, \sigma(\bar{X}) + \sigma(\bar{Z}), \bar{t} - 1)$ such that $V = \lim_{k \to \infty} H'(\varepsilon_k, t_k, x_k)$. For any given $(0, t, x)$, it holds that

$$H'(\varepsilon_k, t_k, x_k)(0,t,x) = \frac{\partial H(\varepsilon_k, t_k, x_k)}{(t,x)}(t,x). \tag{6.41}$$

On the other hand, by a direct computation applied to (2.24), we have

$$\frac{\partial H(\varepsilon_k, t_k, x_k)}{(t,x)} = \begin{bmatrix} \alpha_k & -\alpha_k \Big(\dfrac{\varepsilon_k^2 a_k}{1 + \varepsilon_k^2 b_k} \Big)^T \\ -\alpha \dfrac{\varepsilon_k^2 a_k}{1 + \varepsilon_k^2 b_k} & \mathrm{Diag}(1 + \varepsilon_k^2 b_k)^{-1} + \alpha_k \dfrac{\varepsilon_k^2 a_k}{1 + \varepsilon_k^2 b_k} \Big(\dfrac{\varepsilon_k^2 a_k}{1 + \varepsilon_k^2 b_k} \Big)^T \end{bmatrix},$$

where, for each k and $i = 1, 2, \ldots, m$

$$(a_k)_i = \frac{1}{((y_k)_i + s_k)^2} - \frac{1}{((y_k)_i - s_k)^2}, \ (b_k)_i = \frac{1}{((y_k)_i + s_k)^2} + \frac{1}{((y_k)_i - s_k)^2}$$

and

$$\alpha_k = \frac{1}{1 + \varepsilon_k^2 \sum_{i=1}^{m} (b_k)_i - \sum_{i=1}^{m} \frac{\varepsilon_k^4 (a_k)_i^2}{1 + \varepsilon_k^2 (b_k)_i}}.$$

Since $(\varepsilon_k, t_k, x_k) \to (0, \bar{t} - 1, \sigma(\bar{X}) + \sigma(\bar{Z}))$, by Proposition 2.11 and 6.2, we can easily deduce that

$$\lim_{k\to\infty} y_k = \sigma(\bar{X}), \ \lim_{k\to\infty} s(\varepsilon_k, t_k, x_k) = \sigma_1(\bar{X}).$$

Using the above equalities and (6.6), it is easy to see

$$\lim_{k\to\infty} \frac{\varepsilon_k^2}{((y_k)_i - s_k)^2} = 0, \quad \alpha + 1 \le i \le m$$

and

$$\lim_{k\to\infty} \frac{\varepsilon_k^2}{((y_k)_i - s_k)^2} = \infty, \quad 1 \le i \le \beta.$$

Since $\varepsilon_k^2 a_k/(1 + \varepsilon_k^2 b_k)$ is bounded, by taking a subsequence if necessary, there exists a vector $l \in [0, 1]^{\alpha-\beta}$ such that

$$\lim_{k\to\infty} \varepsilon_k^2 a_k/(1 + \varepsilon^2 b_k) = [1^\beta \ l \ 0^{m-\alpha}]^T,$$

which implies

$$\lim_{k\to\infty} \frac{1}{1 + \varepsilon_k^2 b_k} = [0^\beta \ 1 - l \ 1^{m-\alpha}]^T, \ f := \lim_{k\to\infty} \alpha_k = 1/(1 + \beta + \sum_{i=1}^{\beta-\alpha} l_i) \in (0, 1).$$

This, together with (6.41), completes the proof.

□

Proposition 6.5. *Let $\Omega_1, \Omega_2 \in [0, 1]^{m\times m}$ be two symmetric matrices. Then for any $A \in \Re^{m\times m}$,*

$$\langle A, \, \Omega_1 \circ S(A) + \Omega_2 \circ T(A) \rangle \ge \|\Omega_1 \circ S(A) + \Omega_2 \circ T(A)\|_F^2.$$

Proof. By simple manipulation, it can be seen immediately that

$$\begin{aligned} & \langle A, \, \Omega_1 \circ S(A) + \Omega_2 \circ T(A) \rangle \\ = \ & \langle S(A), \, \Omega_1 \circ S(A) \rangle + \langle T(A), \, \Omega_2 \circ T(A) \rangle \\ \ge \ & \|\Omega_1 \circ S(A)\|_F^2 + \|\Omega_2 \circ T(A)\|_F^2 \\ \ge \ & \|\Omega_1 \circ S(A) + \Omega_2 \circ T(A)\|_F^2. \end{aligned}$$

The proof is completed. □

Proposition 6.6. *Let* $(\bar{t}, \bar{X}, \bar{y}, \bar{Z}, \bar{w})$ *be a solution of* (6.5)*. If the primal constraint nondegeneracy condition* (6.23) *holds at* $(\bar{t}, \bar{X}, \bar{y})$ *and the dual constraint nondegeneracy conditions* (6.25) *holds at* $(\bar{Z}, \bar{w})$*, then any element in* $\partial E(0, \bar{t}, \bar{X}, \bar{y}, \bar{Z}, \bar{w})$ *is nonsingular.*

Proof. For any $\mathcal{W} \in \partial E(0, \bar{t}, \bar{X}, \bar{y}, \bar{Z}, \bar{w})$, let $(\Delta\varepsilon, \Delta t,\ \Delta X, \Delta y, \Delta Z, \Delta w)$ be a solution of the linear system:

$$\mathcal{W}\,[\Delta\varepsilon, \Delta t, \Delta X, \Delta y, \Delta, \lambda, \Delta u] = 0.$$

Therefore, there exist a $\Xi_1 \in \partial G(0, \bar{t} - 1, \bar{X} + \bar{Z})$ and a $\Xi_2 \in \partial H_u(0, \bar{u}_2 - B\bar{y} + b)$ such that

$$\begin{cases} \Delta\varepsilon = 0, \\ (\Delta t, \Delta X) = \Xi_1(\Delta\varepsilon, \Delta t, \Delta X + \Delta Z) \\ \mathcal{A}\Delta Z + B^T \Delta w = 0 \\ \mathcal{A}^* \Delta y + \Delta X = 0 \\ B_1 \Delta y = 0 \\ \Delta w_2 = \Xi_2(\Delta\varepsilon, \Delta w_2 - B_2 \Delta y) \end{cases} . \tag{6.42}$$

By the equation above, we know that:

(i) if $i \in \tau_{21}$, thus $(H_u)_i(\cdot,\cdot)$ is continuously differentiable at $(0, (\bar{w}_2 - B_2\bar{y} + b)_i)$ with $(0, 1)$ as its derivative. This implies $B_{\tau_{21}}\Delta y = 0$.

(ii) if $i \in \tau_{22}$, then $(\Delta w_2)_i = \xi_i(\Delta w_2 - B_2\Delta y)_i$, where $0 \le \xi_i \le 1$. This shows

$$\langle (\Delta w)_{\tau_{22}},\ (B_2\Delta y)_{\tau_{22}} \rangle \le 0.$$

(iii) if $i \in \tau_3$, then $(B_2\bar{y} - b)_i > 0$ and $(\bar{w}_2)_i = 0$. $(H_u)_i(\varepsilon, x)$ is continuously differentiable with the derivative $(0, 0)$, and hence $\Delta w_{\tau_3} = 0$.

In summary, for any $i \in \tau_2 \cup \tau_3$, it holds that $\langle \Delta w_2,\ B_2\Delta y \rangle \ge 0$. and then

$$\langle \Delta X,\ \Delta Z \rangle = -\langle \Delta y,\ (B_2)^T \Delta w \rangle \le 0. \tag{6.43}$$

We next proceed to prove this proposition by considering the following two cases.

Case 1: $\bar{t} = 0$ and then $\bar{X} = 0$. Since the primal constraint nondegeneracy condition holds at $(\bar{t}, \bar{X}, \bar{y})$, we know $\bar{Z} = 0$ and $\bar{w} = 0$. Let $(\varepsilon^k, t^k, Z^k)$ be a sequence converging to $(0, -1, 0, 0))$ with $\varepsilon^k \neq 0$. Denote by σ_k the singular value of Z^k and write

$(s_k, y_k) = H(\varepsilon^k, t^k, \sigma_k)$. By the definition of $(s_k, (y_k)_1)$, we know that

$$\begin{cases} (y_k)_1 - (\sigma_k)_1 + \varepsilon_k^2 \dfrac{2(y_k)_1}{s_k^2 - (y_k)_1^2} = 0, \\ s_k - t_k - \varepsilon_k^2 \displaystyle\sum_{i=1}^{m} \dfrac{2s_k}{s_k^2 - (y_k)_i^2} = 0, \; i = 1, 2, \ldots, m \end{cases}. \tag{6.44}$$

Since $(y_k)_1$ is the largest element in y_k, we know

$$\liminf_{k\to\infty} \varepsilon_k^2 \frac{2s_k}{s_k^2 - (y_k)_1^2} \geq \frac{1}{m},$$

which, together with the first equality of (6.44), implies that

$$\lim_{k\to\infty} \frac{(y_k)_1}{s_k} = 0$$

and then, for $i = 1, 2, \ldots, m$, it holds

$$\lim_{k\to\infty} \frac{(y_k)_i}{s_k} = 0, \; \lim_{k\to\infty} \frac{s_k - (y_k)_i}{s_k + (y_k)_i} = 1.$$

By taking a subsequence if necessary, we are able to show, for $i = 2, 3, \ldots, m$,

$$\lim_{k\to\infty} \frac{2s_k\varepsilon_k^2}{s_k^2 - (y_k)_i^2} = \frac{1}{m},$$

which implies that

$$\lim_{k\to\infty} \frac{\varepsilon_k}{s_k - (y_k)_i} = \lim_{k\to\infty} \frac{\varepsilon_k}{s_k + (y_k)_i} = \infty.$$

By simple algebraic manipulation, we know

$$\alpha_k \to 0, \; \frac{1}{1 + \varepsilon_k^2 b_k} \to 0$$

and

$$\frac{(y_k)_i - (y_k)_j}{(\sigma_k)_i - (\sigma_k)_j} \to 0, \; \frac{(y_k)_i + (y_k)_j}{(\sigma_k)_i + (\sigma_k)_j} \to 0.$$

This shows $\Xi_1 = 0$ and, then $\Delta t = 0$ and $\Delta X = -\mathcal{A}^*\Delta y = 0$. On the other hand, by the analysis above, it holds that

$$\begin{cases} \mathcal{A}\Delta Z + B_{\tau_1\cup\tau_2}^T y = 0 \\ B_{\tau_1\cup\tau_{21}} y = 0 \end{cases} \tag{6.45}$$

Then one can invoke part (iii) and (v) in Proposition 6.2 and deduce that

$$\Delta Z = 0 \quad \text{and} \quad \Delta w = 0, \; \Delta y = 0.$$

This proves the proposition for this case.

Case 2: $\bar{t} > 0$. In this case, we break the proof into two steps.
Step 1: In this step, we prove $\Delta y = 0$ and $\Delta X = 0$. Denote $\mathcal{N} = \{0\} \times \Re \times \Re^{m\times n}$ whose Lebesgue measure is 0 and let

$$\partial_{\mathcal{N}} G(0, \bar{t}-1, \bar{Z}+\bar{\lambda}) := \left\{ \lim_{k\to\infty} G'(\varepsilon^k, t^k, Z^k) : (\varepsilon^k, t^k, Z^k) \to (0, \bar{t}-1, \bar{Z}+\bar{\lambda}),\, \varepsilon^k \neq 0 \right\}.$$

Since the generalized Jacobian is blind to sets of zero measure, we obtain that

$$\partial G(0, \bar{t}-1, \bar{Z}+\bar{\lambda}) = \mathrm{conv}\partial_{\mathcal{N}} G(0, \bar{t}-1, \sigma(\bar{Z}+\bar{\lambda})).$$

Let $\Delta\tilde{X} = U^T X V$ and $\Delta\tilde{Z} = U^T \Delta Z V$. Let $u_1 > u_2 > \ldots > u_r > u_{r+1} = 0$ be the distinct singular values of $X + Z$. Redefine $a_k := \{i \,|\, \sigma_i = u_k\}$. We consider 10 subcases.

Subcase 1: $i \in a_{k_1}, j \in a_{k_2}, k_1 \neq k_2$, where $1 \leq i \leq \beta$ and $1 \leq j \leq \alpha$. Since

$$\bar{\Omega}_{ij} := \lim_{k\to\infty} \Omega^k_{ij} = 0$$

and

$$\bar{\Gamma}_{ij} := \lim_{k\to\infty} \Gamma^k_{ij} = \frac{2\sigma_1(X)}{2\sigma_1(X) + \sigma_i(Z) + \sigma_j(Z)} \in (0,1).$$

By Proposition 2.4, it follows easily that

$$\begin{cases} \Delta\tilde{X}_{ij} = \dfrac{\bar{\Gamma}_{ij}}{2}\,[\Delta\tilde{X}_{ij} + \Delta\tilde{Z}_{ij} - \Delta\tilde{X}_{ji} - \Delta\tilde{Z}_{ji}], \\ \Delta\tilde{X}_{ji} = \dfrac{\bar{\Gamma}_{ij}}{2}\,[\Delta\tilde{X}_{ji} + \Delta\tilde{Z}_{ji} - \Delta\tilde{X}_{ij} - \Delta\tilde{Z}_{ij}]. \end{cases} \tag{6.46}$$

By solving the linear system above, we obtain that

$$\Delta\tilde{X}_{ij} = -\Delta\tilde{X}_{ji} = \frac{\bar{\Gamma}_{ij}}{2(1-\bar{\Gamma}_{ij})}[\Delta\tilde{Z}_{ij} - \Delta\tilde{Z}_{ji}]\,.$$

which, by direct computation, yields that

$$\langle \Delta\tilde{X}_{ij},\, \Delta\tilde{Z}_{ij}\rangle + \langle \Delta\tilde{X}_{ji},\, \Delta\tilde{Z}_{ji}\rangle = \frac{2(1-\bar{\Gamma}_{ij})}{\bar{\Gamma}_{ij}}(\Delta\tilde{X}_{ij})^2\,.$$

Subcase 2: $1 \leq i \leq \beta$ and $\alpha + 1 \leq j \leq m$. In this subcase,

$$\bar{\Omega}_{ij} := \lim_{k\to\infty} \Omega^k_{ij} = \frac{\sigma_1(X) - \sigma_j(X)}{\sigma_1(X) + \sigma_i(Z) - \sigma_j(X)} \in (0,1),$$

and

$$\bar{\Gamma}_{ij} := \lim_{k\to\infty} \Gamma_{ij}^k = \frac{\sigma_1(X) + \sigma_j(X)}{\sigma_1(X) + \sigma_i(Z) + \sigma_j(Z)} \in (0,1).$$

By Proposition 2.4, we can easily deduce that

$$\begin{cases} \Delta\tilde{X}_{ij} = \dfrac{\bar{\Omega}_{ij}}{2}\,[\Delta\tilde{X}_{ij} + \Delta\tilde{Z}_{ij} + \Delta\tilde{X}_{ji} + \Delta\tilde{Z}_{ji}] + \dfrac{\bar{\Gamma}_{ij}}{2}\,[\Delta\tilde{X}_{ij} + \Delta\tilde{Z}_{ij} - \Delta\tilde{X}_{ji} - \Delta\tilde{Z}_{ji}], \\ \Delta\tilde{X}_{ji} = \dfrac{\bar{\Omega}_{ij}}{2}\,[\Delta\tilde{X}_{ji} + \Delta\tilde{Z}_{ji} + \Delta\tilde{X}_{ij} + \Delta\tilde{Z}_{ij}] + \dfrac{\bar{\Gamma}_{ij}}{2}\,[\Delta\tilde{X}_{ji} + \Delta\tilde{Z}_{ji} - \Delta\tilde{X}_{ij} - \Delta\tilde{Z}_{ij}], \end{cases}$$

which by simple algebraic computation shows that

$$\begin{cases} \Delta\tilde{Z}_{ij} = \dfrac{1}{2}(\epsilon_1 + \epsilon_2)\Delta\tilde{X}_{ij} + \dfrac{1}{2}(\epsilon_1 - \epsilon_2)\tilde{\Delta X}_{ji} \\ \Delta\tilde{Z}_{ji} = \dfrac{1}{2}(\epsilon_1 + \epsilon_2)\Delta\tilde{X}_{ji} + \dfrac{1}{2}(\epsilon_1 - \epsilon_2)\tilde{\Delta X}_{ij} \end{cases}, \tag{6.47}$$

where

$$\epsilon_1 = \frac{\sigma_i(Z)}{\sigma_1(X) - \sigma_j(X)},\ \epsilon_2 = \frac{\sigma_i(Z)}{\sigma_1(X) + \sigma_j(X)}.$$

From (6.47), taking into account $\epsilon_1, \epsilon_2 > 0$, we can write

$$\langle \Delta\tilde{X}_{ij},\, \Delta\tilde{Z}_{ij}\rangle + \langle \Delta\tilde{X}_{ji},\, \Delta\tilde{Z}_{ji}\rangle = \frac{\epsilon_1}{2}(\Delta\tilde{X}_{ij} + \Delta\tilde{X}_{ji})^2 + \frac{\epsilon_2}{2}(\Delta\tilde{X}_{ij} - \Delta\tilde{X}_{ji})^2$$

Subcase 3: $1 \le i \le \beta$ and $m+1 \le j \le n$. In this subcase,

$$\bar{\mathcal{F}}_{ij} := \lim_{k\to\infty} \mathcal{F}_{ij}^k = \frac{\sigma_1(X)}{\sigma_1(X) + \sigma_i(Z)},$$

and it therefore holds

$$\Delta\tilde{X}_{ij} = \bar{\mathcal{F}}_{ij}(\Delta\tilde{X}_{ij} + \Delta\tilde{Z}_{ij}).$$

This implies

$$\langle \Delta\tilde{X}_{ij},\, \Delta\tilde{Z}_{ij}\rangle = \frac{\sigma_i(Z)}{\sigma_1(X)}(\Delta\tilde{X}_{ij})^2.$$

Subcase 4: $\alpha+1 \le i \le m$ and $\beta+1 \le j \le \alpha$; or $\alpha+1 \le i \le r$ and $r+1 \le j \le m$;
or $i \in a_{k_1}$ and $j \in a_{k_2}$, where $k_1 \ne k_2$ and $\alpha+1 \le i,j \le r$. In these subcases, by direct computation, we can easily deduce that

$$\bar{\Omega}_{ij} := \lim_{k\to\infty} \Omega_{ij}^k = 1$$

and

$$\bar{\Gamma}_{ij} := \lim_{k\to\infty} \Gamma_{ij}^k = 1.$$

Hence, by Proposition 2.4 (ii), it holds that

$$\begin{cases} \Delta\tilde{X}_{ij} = \Delta\tilde{X}_{ij} + \Delta\tilde{Z}_{ij} \\ \Delta\tilde{X}_{ji} = \Delta\tilde{X}_{ji} + \Delta\tilde{Z}_{ji} \end{cases}, \tag{6.48}$$

which implies $\Delta\tilde{Z}_{ij} = \Delta\tilde{Z}_{ji} = 0$.

Subcase 5: $\beta + 1 \leq i \leq r$ and $m + 1 \leq j \leq n$. Since $\bar{\mathcal{F}}_{ij} := \lim_{k\to\infty} \mathcal{F}_{ij}^k = 1$, we know from Proposition 2.4 (ii) that

$$\Delta\tilde{X}_{ij} = \Delta\tilde{X}_{ij} + \Delta\tilde{Z}_{ij},$$

which means $\Delta\tilde{Z}_{ij} = 0$.

Subcase 6: $i, j \in a_k$, where $1 \leq i, j \leq \beta$ and $i \neq j$. By an easily manipulation applied to (2.23), it holds that

$$\frac{(\sigma_k)_i - (\sigma_k)_j}{(y_k)_i - (y_k)_j} - 1 = \frac{\varepsilon_k^2}{(s_k - (y_k)_i)(s_k - (y_k)_j)} + \frac{\varepsilon_k^2}{(s_k + (y_k)_i)(s_k + (y_k)_j)}, \tag{6.49}$$

which implies

$$\lim_{k\to\infty} \frac{(y_k)_i - (y_k)_j}{(\sigma_k)_i - (\sigma_k)_j} = 0.$$

Combining the equality above with lemma 6.4 and Proposition 2.4, we know

$$\begin{cases} \Delta\tilde{X}_{ij} = \dfrac{\sigma_1(X)}{2\sigma_1(X) + \sigma_i(Z) + \sigma_j(Z)}(\Delta\tilde{X}_{ij} + \Delta\tilde{Z}_{ij} - \Delta\tilde{X}_{ji} - \Delta\tilde{Z}_{ji}) \\ \Delta\tilde{X}_{ji} = \dfrac{\sigma_1(X)}{2\sigma_1(X) + \sigma_i(Z) + \sigma_j(Z)}(\Delta\tilde{X}_{ji} + \Delta\tilde{Z}_{ji} - \Delta\tilde{X}_{ij} - \Delta\tilde{Z}_{ij}) \end{cases}, \tag{6.50}$$

which, by simple algebraic computation, shows that

$$\Delta\tilde{X}_{ij} = -\Delta\tilde{X}_{ji} = \frac{\sigma_1(X)}{\sigma_i(X) + \sigma_j(X)}(\Delta\tilde{Z}_{ij} - \Delta\tilde{Z}_{ji}).$$

It therefore holds that

$$\langle \Delta\tilde{X}_{ij}, \Delta\tilde{Z}_{ij} \rangle + \langle \Delta\tilde{X}_{ji}, \Delta\tilde{Z}_{ji} \rangle = \frac{\sigma_i(X) + \sigma_j(X)}{\sigma_1(X)}(\Delta\tilde{X}_{ji})^2.$$

Subcase 7: $i, j \in a_k$, where $\alpha + 1 \leq i, j \leq r$. From the subcase 6, we have

$$\lim_{k\to\infty} \frac{(y_k)_i - (y_k)_j}{(\sigma_k)_i - (\sigma_k)_j} = 1.$$

Therefore, it is easy to check that

$$\begin{aligned}\Delta\tilde{X}_{a_ka_k} &= \sum_{i=1}^{v}\mu_i Q_i\left\{{Q_i}^T\frac{\Delta\tilde{X}_{a_ka_k}+\Delta\tilde{Z}_{a_ka_k}+(\Delta\tilde{X}_{a_ka_k}+\Delta\tilde{Z}_{a_ka_k})^T}{2}Q_i\right\}{Q_i}^T\\ &\quad+\frac{1}{2}\left[\Delta\tilde{X}_{a_ka_k}+\Delta\tilde{Z}_{a_ka_k}-(\Delta\tilde{X}_{a_ka_k}+\Delta\tilde{Z}_{a_ka_k})^T\right],\\ &=\Delta\tilde{X}_{a_ka_k}+\Delta\tilde{Z}_{a_ka_k},\end{aligned}$$

which implies that $\Delta\tilde{Z}_{a_ka_k}=0$.

Subcase 8: $r+1\le i,j\le m$. For $i\neq j$, by simple manipulation applied to equation (2.23) , we know the following equality holds

$$\frac{(\sigma_k)_i+(\sigma_k)_J}{(y_k)_i+(y_k)_J}-1=\frac{\varepsilon^2}{(s_k-(y_k)_i)(s_k+(y_k)_j)}+\frac{\varepsilon^2}{(s_k-(y_k)_j)(s_k+(y_k)_i)},$$

which together with (6.49) implies

$$\bar{\Omega}_{ij}:=\lim_{k\to\infty}\Omega_{ij}^k=1,\ \bar{\Gamma}_{ij}:=\lim_{k\to\infty}\Gamma_{ij}^k=1.$$

Then, similar as subcase 7, we can deduce easily that $\Delta\tilde{Z}_{a_ka_k}=0$.

Subcase 9: $r+1\le i\le m$ and $m+1\le j\le n$. Using simple algebraic computation, we know

$$\frac{(\sigma_k)_i}{(y_k)_i}-1=\frac{2\varepsilon_k^2}{(s_k)^2-(y_k)_i^2}$$

and it then holds

$$\lim_{k\to\infty}\frac{(y_k)_i}{(\sigma_k)_i}=1.$$

Similarly, we have $\Delta\tilde{Z}_{ij}=0$ for any i,j in this subcase.

Subcase 10: $\beta+1\le i\neq j\le\alpha$; or $1\le i=j\le\beta$. For notational simplicity, we use I_1,I_2 and I to denote the following indexes:

$$I_1:=\{1,2\ldots,\beta\},\ I_2:=\{\beta+1,\beta+2,\ldots,\alpha\},\ I:=\{1,2,\ldots,\alpha+1\}.$$

Since

$$\begin{bmatrix}\Delta t\\ \Delta X\end{bmatrix}=\Xi_1(\Delta\varepsilon,\Delta t,\Delta X+\Delta Z),\tag{6.51}$$

it follows from lemma 6.4 that

$$\begin{bmatrix} \Delta t \\ \mathrm{diag}(\Delta \tilde{X})_{I_1} \\ \Delta \tilde{X}_{I_2 I_2} \end{bmatrix} = \sum_{i=1}^{v} u_i \begin{bmatrix} 0 \\ 0 \\ Q_i(\bar{\Omega}^i \circ S(Q_i^T(\Delta \tilde{X} + \Delta \tilde{Z})Q_i) + \bar{\Gamma}^i \circ T(Q_i^T \Delta(\tilde{X} + \tilde{Z})Q_i)Q_i^T) \end{bmatrix}$$

$$+ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & Q_i \end{bmatrix} \mathrm{Diag}\left[\vartheta^i_{I\cup\{\alpha+2\}, I\cup\{\alpha+2\}} \begin{bmatrix} 0 \\ \Delta t \\ \mathrm{diag}(\Delta \tilde{X} + \Delta \tilde{Z})_{I_1} \\ \mathrm{diag}(Q_i^T(\Delta \tilde{X} + \Delta \tilde{Z})_{I_2 I_2} Q_i) \end{bmatrix} \right] \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & Q_i^T \end{bmatrix}, \tag{6.52}$$

where $Q_i \in \Re^{(\alpha-\beta)\times(\alpha-\beta)}$ is orthogonal matrix, $u_i > 0$, $\sum_{i=1}^{v} u_i = 1$ and $\vartheta_i \in \partial H(0, \bar{t} - 1, \sigma(\bar{X} + \bar{Z}))$ for $i = 1, 2, \ldots, v$. Then by Lemma 6.3 and Lemma 6.5, using a simple manipulation, we have

$$\langle \mathrm{diag}(\Delta \tilde{X})_{I_1}, \mathrm{diag}(\Delta \tilde{Z})_{I_1} \rangle + \langle \Delta \tilde{X}_{I_2 I_2}, \Delta \tilde{Z}_{I_2 I_2} \rangle \geq 0.$$

On the other hand, by the structure of ϑ_i and equation (6.52) , we can also deduce that $\mathrm{diag}(\Delta \tilde{X})_{I_1}$ is a quantity vector.

After checking all the subcases above and noting (6.43), we can see clearly that Δy and ΔX satisfy

$$\begin{cases} B_{\tau_1 \cup \tau_{21}} \Delta y = 0 \\ (U^1)^T \Delta X V^1 = kI \\ (U^1)^T \Delta X V^2 = 0 \\ (U^2)^T \Delta X V^1 = 0 \end{cases} \tag{6.53}$$

Using the fact $\Delta X = -\mathcal{A}^* \Delta y$ and part (iii) of Proposition (6.2), we know $\Delta y = 0$ and $\Delta X = 0$.

Step2: Consider $i, j \in a_k$, where $\beta + 1 \leq i \neq j \leq \alpha$. By the construction of generalized Jacobian, there exists a η_{ij} depending on $S(\Delta \tilde{Z}_{a_k a_k})$, such that

$$\begin{cases} 0 = \eta_{ij} + \dfrac{\Delta \tilde{Z}_{ij} - \Delta \tilde{Z}_{ji}}{2} \\ 0 = \eta_{ij} + \dfrac{\Delta \tilde{Z}_{ji} - \Delta \tilde{Z}_{ij}}{2} \end{cases}. \tag{6.54}$$

Hence, $\Delta \tilde{Z}_{ji} = \Delta \tilde{Z}_{ij}$. In order to use the primal constraint nondegeneracy of (1.1), we need to prove $\mathrm{Tr}[(U^{(1)})^T \Delta Z V^{(1)}] = 0$. Since

$$\begin{bmatrix} \Delta t \\ 0 \end{bmatrix} = \Xi_1(\Delta \varepsilon, \Delta t, \Delta Z), \tag{6.55}$$

it follows from Lemma 6.4 that

$$\begin{cases} \Delta t = \sum_{i=1}^{v} u_i f_i(\Delta t + \mathrm{Tr}(\Delta \tilde{Z}_{I_1 I_1})) + \sum_{i=1}^{v} u_i f_i l^i \mathrm{diag}[Q_i^T \Delta \tilde{Z}_{I_2 I_2} Q_i], \\ 0 = \sum_{i=1}^{v} u_i f_i(\Delta t + \mathrm{Tr}(\Delta \tilde{Z}_{I_1 I_1})) + \sum_{i=1}^{v} u_i f_i l^i \mathrm{diag}[Q_i^T \Delta \tilde{Z}_{I_2 I_2} Q_i]. \end{cases} \tag{6.56}$$

Observing the first two equalities above, we know

$$\Delta t = 0 \quad \text{and} \quad \sum_{i=1}^{v} u_i f_i (\mathrm{Tr}(\Delta \tilde{Z}_{I_1 I_1}) + l^i \mathrm{diag}[Q_i^T \Delta \tilde{Z}_{I_2 I_2} Q_i]) = 0. \tag{6.57}$$

Hence, it holds that

$$\begin{aligned} 0 = \sum_{i=1}^{v} & u_i Q_i \big(\mathrm{Diag}[f_i (l^i)^T \mathrm{Tr}(\Delta \tilde{Z}_{I_1 I_1}) + (\mathrm{Diag}(1 - (l^i)^T) + f_i (l^i)^T l^i) \mathrm{diag}(Q_i^T \Delta \tilde{Z}_{I_2 I_2} Q_i)]\big) Q_i^T \\ & + u_i Q_i \big(\bar{\Omega}_i \circ S(Q_i^T \Delta \tilde{Z}_{I_2 I_2} Q_i)\big) Q_i^T. \end{aligned} \tag{6.58}$$

Using Lemma 6.3 and Lemma 6.5, we are able to show

$$\begin{cases} \mathrm{Tr}(\Delta \tilde{Z}_{I_1 I_1}) + l^i \mathrm{diag}[Q_i^T \Delta \tilde{Z}_{I_2 I_2} Q_i] = 0 \\ \mathrm{Tr}(\Delta \tilde{Z}_{I_2 I_2}) = l^i \mathrm{diag}[Q_i^T \Delta \tilde{Z}_{I_2 I_2} Q_i] \end{cases}, \forall\, i \in \{j \,|\, u_j > 0, j = 1, 2, \ldots, v\},$$

which implies $\mathrm{Tr}(\Delta \tilde{Z})_{1:\alpha,1:\alpha} = 0$. Then one can invoke the part (ii) in Proposition 6.2 and establish that

$$\Delta Z = 0 \quad \text{and} \quad \Delta w = 0.$$

This completes the proof. □

By combining Proposition 2.11, Proposition 6.1 and Proposition 6.6, we immediately obtain the convergence of smoothing Newton method.

Theorem 6.7. *Suppose $A_1, \cdots, A_p$ are linearly independent and B has full row rank. Then an infinite sequence of $\{\varepsilon_k, W^k\}$ is generated by Algorithm 6.1 and each accumulation point $(0, \bar{W})$ of $\{\varepsilon_k, W^k\}$ is a solution of $E_u(\varepsilon, W) = 0$. Let $\bar{W} = (\bar{t}, \bar{X}, \bar{y}, \bar{Z}, \bar{w}) \in \Re \times \Re^{m\times n} \times \Re^p \times \Re^{m\times n} \times \Re^{n_1+n_2}$. If the primal constraint nondegeneracy condition* (6.23) *holds at $(\bar{t}, \bar{X}, \bar{y})$ and the dual constraint nondegeneracy condition* (6.25) *holds at $(\bar{Z}, \bar{w})$, then the whole sequence $\{\varepsilon_k, W^k\}$ converges to $(0, \bar{W})$ in the order of $1 + \gamma$ where $\gamma > 0$ is some rational number, i.e.,*

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

Furthermore, if $\bar{t} > 0$, in particular $A_0, A_1, \ldots, A_p$ are linearly independent, then the the convergence rate becomes quadratic.

6.4 Preliminary numerical results

In this section, we consider the unconstrained randomly generated matrix norm approximation problems

$$\min \quad \|A_0 - \mathcal{A}^* y\|_2 \tag{6.59}$$

In the experiments, the matrices $A_0, A_1, \ldots, A_p$ are generated independently from the multivariate uniform distribution on $[0,1]^{m\times n}$. For the squared smoothing Newton method, we generate a warm starting point by running ADM for at most 50 iterations. The ADM is stopped when $\max\{R_p, R_d\} < 5 \cdot 10^{-3}$. The squared smoothing Newton method is stopped when the condition $\phi(\varepsilon^k, W^k) < 10^{-5}$ is met.

$m \mid n \mid p$	ϕ	it	pobj | gap	$R_p \mid R_d$	time
5 | 800 | 100	9.7-7	5	8.39614503 0| 1.2-6	7.6-7| 1.5-7	0.4
10| 800 | 100	4.3-6	7	8.53608711 0| 2.0-6	7.5-7| 1.4-8	0.3
20| 800 | 100	5.7-7	8	8.91663710 0| 8.6-7	1.9-6| 5.0-7	2.1
30| 800 | 100	1.6-6	7	9.43725030 0| 3.3-6	5.4-7| 6.4-8	2.3
50| 800 | 100	3.3-6	11	9.84953792 0| 3.5-6	6.6-7| 5.3-8	4.1
80| 800 | 100	4.9-7	8	1.05945024 1| 3.9-6	7.7-7| 3.3-7	7.8
100| 800 | 100	5.2-6	7	1.07990485 1| 4.6-6	5.7-6| 8.6-6	30.6
200| 800 | 100	3.8-6	15	1.20325115 1| 4.3-6	6.1-7| 8.7-7	49.4
400| 800 | 100	7.8-8	12	1.37761589 1| 4.4-6	4.2-7| 6.1-8	85.1
600| 800 | 100	4.1-6	15	1.49803951 1| 8.4-6	9.7-7| 2.3-7	204.7
800| 800 | 100	3.5-7	13	1.60569712 1| 5.4-7	7.6-7| 5.3-8	656.1

Table 6.1: Results for unconstrained random matrix norm approximation problems.

Table 6.1 shows the smoothing Newton methods work well for all the tested examples, especially for the case $m << n$. It is able to obtain solutions with relative high accuracy in a few iterations. However, with the increased scale of m, the consumed time grows up quickly. Then, as one can expect from the Interior point methods, the smoothing Newton method would encounter the high computational cost issues, which limits its application to large scale problems.

Chapter 7

Conclusion remark

In this thesis, we designed efficient algorithms for solving the matrix norm approximation problems. We first proposed a first order alternating direction method (ADM) to solve this problem. At each iteration, the subproblem involved can either be solved by a fast algorithm or it has a closed form solution, which makes the ADM easily implementable.

To obtain solutions of high accuracy, we also proposed a semismooth Newton-CG dual proximal point algorithm (SNDPPA) to solve large scale matrix norm approximation problems. In each iteration, the dual PPA solves its subproblem by a semismooth Newton method and the Newton direction is computed inexactly by a PCG solver. Theoretical results to guarantee the global convergence and local superlinear convergence of the dual PPA are established based on the classical analysis of proximal point algorithms. Capitalizing on the recent advances on spectral operator and related perturbation analysis, we also characterize the nonsingularity of the semismooth Newton systems. The latter property is an important condition for the fast convergence of the semismooth Newton method. Extensive numerical experiments on problems arising from different areas are conducted to evaluate the performance of the SNDPPA against the ADM. The numerical results show that the SNDPPA is very efficient and robust, and it substantially outperforms the ADM.

Motivated by the great success of Interior point methods for the second order cone programming, we also designed a squared smoothing Newton method for the MNA problem in which the matrix is of much more columns than rows. Suplinear convergence of this method is also established under the primal and dual constraint nondegenerate conditions for the MNA problems and its dual at the primal dual optimal solution pairs. We conduct preliminary numerical experiments to investigate the performance of the smoothing Newton method. When the problem scale is small or moderate, the numerical results reveals that our smoothing Newton method is robust and efficient for the matrix norm approximation problems which is of the flat structure.

Bibliography

[1] Aldous, D., 1983. Random walks on finite groups and rapidly mixing markov chains. Séminaire de Probabilités XVII 1981/82 , 243–297.

[2] Andersen, M., Dahl, J., Vandenberghe, L., 2010. Implementation of nonsymmetric interior-point methods for linear optimization over sparse matrix cones. Mathematical Programming Computation 2, 167–201.

[3] Billera, L., Diaconis, P., 2001. A geometric interpretation of the metropolis-hastings algorithm. Statistical Science , 335–339.

[4] Bochnak, J., Coste, M., Roy, M., 1998. Real algebraic geometry. volume 36. Springer Verlag.

[5] Bolte, J., Daniilidis, A., Lewis, A., 2009. Tame functions are semismooth. Mathematical Programming 117, 5–19.

[6] Bonnans, J.F., Shapiro, A., 2000. Perturbation analysis of optimization problems. Springer Series in Operations Research, Springer-Verlag, New York.

[7] Boyd, S., Diaconis, P., Parrilo, P., Xiao, L., 2009. Fastest mixing Markov chain on graphs with symmetries. SIAM J. Optim. 20, 792–819.

[8] Boyd, S., Diaconis, P., Sun, J., Xiao, L., 2006. Fastest mixing Markov chain on a path. Amer. Math. Monthly 113, 70–74.

[9] Boyd, S., Diaconis, P., Xiao, L., 2004. Fastest mixing Markov chain on a graph. SIAM Rev. 46, 667–689.

[10] Brucker, P., 1984. An O(n) algorithm for quadratic knapsack problems. Operations Research Letters 3, 163–166.

[11] Cai, X., Gu, G., He, B., Yuan, X., 2011. A relaxed customized proximal point algorithm for separable convex programming. Technical Report. Technical report, Department of Mathematics, Hong Kong Baptist University, China.

[12] Calamai, P., Moré, J., 1987. Quasi-newton updates with bounds. SIAM journal on numerical analysis , 1434–1441.

[13] Chan, Z., Sun, D., 2009. Constraint nondegeneracy, strong regularity, and nonsingularity in semidefinite programming. SIAM Journal on optimization 19.

[14] Chen, C., He, B., Yuan, X., 2012. Matrix completion via an alternating direction method. IMA Journal of Numerical Analysis 32, 227–245.

[15] Chen, X., Qi, L., Sun, D., 1998. Global and superlinear convergence of the smoothing newton method and its application to general box constrained variational inequalities. Mathematics of Computation 67, 519–540.

[16] Clarke, F.H., 1990. Optimization and nonsmooth analysis. volume 5 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA. second edition.

[17] Cosares, S., Hochbaum, D., 1994. Strongly polynomial algorithms for the quadratic transportation problem with a fixed number of sources. Mathematics of operations research , 94–111.

[18] Davis, T., Hu, Y., 2009. University of florida sparse matrix collection .

[19] Diaconis, P., Stroock, D., 1991. Geometric bounds for eigenvalues of markov chains. The Annals of Applied Probability 1, 36–61.

[20] Ding, C., 2012. An introduction to a class of matrix optimization problem. Ph.D. thesis. National University of Singapore. Singapore.

[21] Ding, C., Sun, D., Toh, K.C., 2010. An introduction to a class of matrix cone programming. Available on http://www.math.nus.edu.sg/ mattohkc/publist.html .

[22] Douglas, J., Rachford, H., 1956. On the numerical solution of heat conduction problems in two and three space variables. Trans. Amer. Math. Soc 82, 421–439.

[23] Eckstein, J., 1994. Some saddle-function splitting methods for convex programming. Optimization Methods and Software 4, 75–83.

[24] Eckstein, J., Bertsekas, D., 1992. On the douglas—rachford splitting method and the proximal point algorithm for maximal monotone operators. Mathematical Programming 55, 293–318.

[25] Erdmann, K., Schroll, S., 2011. Chebyshev polynomials on symmetric matrices. Linear Algebra and its Applications .

[26] Faber, V., Liesen, J., Tichỳ, P., . On best approximation by polynomials of matrices .

[27] Facchinei, F., Pang, J.S., 2003. Finite-dimensional variational inequalities and complementarity problems. Vol. II. Springer Series in Operations Research, Springer-Verlag, New York.

[28] Fazel, M., Pong, T., Sun, D., Tseng, P., 2012. Hankel matrix rank minimization with applications in system identification and realization. Preprint, National University of Singapore. Singapore.

[29] Gabay, D., 1983. Chapter ix applications of the method of multipliers to variational inequalities. Studies in mathematics and its applications 15, 299–331.

[30] Gabay, D., Mercier, B., 1976. A dual algorithm for the solution of nonlinear variational problems via finite element approximations. Computational Mathematics with Applications 2, 17–40.

[31] Gao, Y., Sun, D., 2009. Calibrating least squares semidefinite programming with equality and inequality constraints. SIAM Journal on Matrix Analysis and Applications 31, 1432.

[32] Glowinski, R., 1984. Numerical methods for nonlinear variational problems. Springer Verlag.

[33] Glowinski, R., Le Tallec, P., 1989. Augmented Lagrangian and operator-splitting methods in nonlinear mechanics. volume 9. Society for Industrial Mathematics.

[34] Glowinski, R., Marrocco, A., 1975. Sur l'approximation, par éléments finis d'ordre un, et la résolution, par pénalisation-dualité, d'une classe de problèmes de Dirichlet non linéaires. Rev. Française Automat. Informat. Recherche Opérationnelle 9, 41–76.

[35] Gol'shtein, E., Tret'yakov, N., 1979. Modified lagrangians in convex programming and their generalizations. Point-to-Set Maps and Mathematical Programming , 86–97.

[36] Greenbaum, A., Trefethen, L.N., 1994. GMRES/CR and Arnoldi/Lanczos as matrix approximation problems. SIAM J. Sci. Comput. 15, 359–368. Iterative methods in numerical linear algebra (Copper Mountain Resort, CO, 1992).

[37] Ha, C.D., 1990. A generalization of the proximal point algorithm. SIAM J. Control Optim. 28, 503–512.

[38] Hastings, W., 1970. Monte carlo sampling methods using markov chains and their applications. Biometrika 57, 97–109.

[39] He, B., Liao, L., Han, D., Yang, H., 2002. A new inexact alternating directions method for monotone variational inequalities. Mathematical Programming 92, 103–118.

[40] He, B., Xu, M., Yuan, X., 2011. Solving large-scale least squares semidefinite programming by alternating direction methods. SIAM Journal on Matrix Analysis and Applications 32, 136.

[41] Held, M., Wolfe, P., Crowder, H.P., 1974. Validation of subgradient optimization. Math. Programming 6, 62–88.

[42] Helgason, R., Kennington, J., Lall, H., 1980. A polynomially bounded algorithm for a singly constrained quadratic program. Mathematical Programming 18, 338–343.

[43] Hestenes, M., 1969. Multiplier and gradient methods. Journal of optimization theory and applications 4, 303–320.

[44] Hille, E., 2005. Analytic function theory. volume 2. American Mathematical Society (RI).

[45] Hiriart-Urruty, J.B., Lemaréchal, C., 1993a. Convex analysis and minimization algorithms. I. volume 305 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin. Fundamentals.

[46] Hiriart-Urruty, J.B., Lemaréchal, C., 1993b. Convex analysis and minimization algorithms. II. volume 306 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin. Advanced theory and bundle methods.

[47] Huang, Z., Qi, L., Sun, D., 2004. Sub-quadratic convergence of a smoothing newton algorithm for the p 0–and monotone lcp. Mathematical programming 99, 423–441.

[48] Ioffe, A., 2009. An invitation to tame optimization. SIAM Journal on Optimization 19, 1894.

[49] Jiang, K., 2012. Algorithms for large scale nuclear norm minimization and convex quadractic semidefinite programming problems. Ph.D. thesis. National University of Singapore. Singapore.

[50] Kahale, N., 1997. A semidefinite bound for mixing rates of markov chains. Random Structures & Algorithms 11, 299–313.

[51] Lemaréchal, C., Sagastizábal, C., 1997. Practical aspects of the Moreau-Yosida regularization: theoretical preliminaries. SIAM J. Optim. 7, 367–385.

[52] Lewis, A.S., Sendov, H.S., 2005. Nonsmooth analysis of singular values. II. Applications. Set-Valued Anal. 13, 243–264.

[53] Liang, W., Baer, R., Saravanan, C., Shao, Y., Bell, A., Head-Gordon, M., 2004. Fast methods for resumming matrix polynomials and chebyshev matrix polynomials. Journal of Computational Physics 194, 575–587.

[54] Lin, Z.C., Chen, M.M., Wu, L.Q., Y., M., 2009. The augmented Lagrange multiplier method for exact recovery of corrupted low-rank Matrices. Preprint, UIUC.

[55] Lions, P., Mercier, B., 1979. Splitting algorithms for the sum of two nonlinear operators. SIAM Journal on Numerical Analysis , 964–979.

[56] Liu, Y., Sun, D., Toh, K., 2009. An implementable proximal point algorithmic framework for nuclear norm minimization. Mathematical Programming , 1–38.

[57] Lu, Z., 2008. Smooth optimization approach for sparse covariance selection. SIAM Journal on Optimization 19, 1807–1827.

[58] Martinet, B., 1970. Brève communication. régularisation d'inéquations variationnelles par approximations successives. Revue Française d' Informatique et de Recherche Opérationelle 4, 154–158.

[59] Meng, F., Sun, D., Zhao, G., 2005. Semismoothness of solutions to generalized equations and the Moreau-Yosida regularization. Math. Program. 104, 561–581.

[60] Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, A., Teller, E., 1953. Equation of state calculations by fast computing machines. The journal of chemical physics 21, 1087.

[61] Mifflin, R., 1977. Semismooth and semiconvex functions in constrained optimization. SIAM Journal on Control and Optimization 15, 959.

[62] Moreau, J.J., 1965. Proximite et dualite dans un espace hilbertien. Bulletin de la Societe Mathematique de France 93.

[63] von Neumann, J., 1962. Some matrix-inequalities and metrization of matric-space, in: Collected Works. Pergamon, Oxford. volume IV, pp. 205–218.

[64] Powell, M., Authority, U.K.A.E., 1967. " A method for non-linear constraints in minimization problems". UKAEA.

[65] Qi, L., 1995. Trust region algorithms for solving nonsmooth equations. SIAM Journal on Optimization 5, 219.

[66] Qi, L., Chen, X., 1995. A globally convergent successive approximation method for severely nonsmooth equations. SIAM Journal on Control and Optimization 33, 402–418.

[67] Qi, L., Sun, D., 2002. Smoothing functions and smoothing newton method for complementarity and variational inequality problems. Journal of Optimization Theory and Applications 113, 121–147.

[68] Qi, L., Sun, D., Zhou, G., 2000. A new look at smoothing newton methods for nonlinear complementarity problems and box constrained variational inequalities. Mathematical Programming 87, 1–35.

[69] Qi, L.Q., Sun, J., 1993. A nonsmooth version of Newton's method. Math. Programming 58, 353–367.

[70] Robinson, S.M., 1984. Local structure of feasible sets in nonlinear programming. II. Nondegeneracy. Math. Programming Stud. , 217–230Mathematical programming at Oberwolfach, II (Oberwolfach, 1983).

[71] Rockafellar, R.T., 1970. Convex analysis. Princeton Mathematical Series, No. 28, Princeton University Press, Princeton, N.J.

[72] Rockafellar, R.T., 1976a. Augmented Lagrangians and applications of the proximal point algorithm in convex programming. Math. Oper. Res. 1, 97–116.

[73] Rockafellar, R.T., 1976b. Monotone operators and the proximal point algorithm. SIAM J. Control Optimization 14, 877–898.

[74] Shen, Y., Wen, Z., Zhang, Y., 2011. Augmented lagrangian alternating direction method for matrix separation based on low-rank factorization. TR11-02, Rice University .

[75] Sinclair, A., 1992. Improved bounds for mixing rates of markov chains and multicommodity flow. Combinatorics, probability and Computing 1, 351–370.

[76] Sturm, J.F., 1999. Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. Optim. Methods Softw. 11/12, 625–653. Interior point methods.

[77] Sun, D., 2006. The strong second-order sufficient condition and constraint nondegeneracy in nonlinear semidefinite programming and their implications. Mathematics of Operations Research , 761–776.

[78] Sun, D., Sun, J., Zhang, L., 2008. The rate of convergence of the augmented lagrangian method for nonlinear semidefinite programming. Mathematical Programming 114, 349–391.

[79] Sun, J., Sun, D., Qi, L., 2004. A squared smoothing newton method for nonsmooth matrix equations and its applications in semidefinite optimization problems. SIAM Journal on Optimization 14, 783–806.

[80] Tang, P., 1988. A fast alglorithm for linear complexchebyshev approximat ions. Mathematics of Computation 51.

[81] Tao, M., Yuan, X., 2011. Recovering low-rank and sparse components of matrices from incomplete and noisy observations. SIAM Journal on Optimization 21, 57.

[82] Toh, K.C., Todd, M.J., Tütüncü, R.H., 1999. SDPT3—a MATLAB software package for semidefinite programming, version 1.3. Optim. Methods Softw. 11/12, 545–581. Interior point methods.

[83] Toh, K.C., Trefethen, L.N., 1999. The Chebyshev polynomials of a matrix. SIAM J. Matrix Anal. Appl. 20, 400–419.

[84] Wang, C., Sun, D., Toh, K.C., 2010. Solving log-determinant optimization problems by a Newton-CG primal proximal point algorithm. SIAM J. Optim. 20, 2994–3013.

[85] Warga, J., 1981. Fat homeomorphisms and unbounded derivate containers. J. Math. Anal. Appl. 81, 545–560.

[86] Watson, G.A., 1992. Characterization of the subdifferential of some matrix norms. Linear Algebra Appl. 170, 33–45.

[87] Wu, B., Ding, C., Sun, D.F., Toh, K.C., 2011. On the Moreau-Yosida regularization of the vector k-norm related functions. Preprint, National University of Singapore. Singapore.

[88] Xiao, L., Boyd, S., 2004. Fast linear iterations for distributed averaging. Systems Control Lett. 53, 65–78.

[89] Yang, J., Zhang, Y., 2011. Alternating direction algorithms for ℓ_1-problems in compressive sensing. SIAM J. Sci. Comput. 33, 250–278.

[90] Yang, J., Zhang, Y., Yin, W., 2010. A fast alternating direction method for tvl1-l2 signal reconstruction from partial fourier data. Selected Topics in Signal Processing, IEEE Journal of 4, 288–297.

[91] Ye, C., Yuan, X., 2007. A descent method for structured monotone variational inequalities. Optimisation Methods and Software 22, 329–338.

[92] Yosida, K., 1964. Functional analysis. Springer Verlag, Berlin.

[93] Yuan, X., 2011. Alternating direction method for covariance selection models. Journal of Scientific Computing , 1–13.

[94] Zhang, X., Burger, M., Osher, S., 2011. A unified primal-dual algorithm framework based on bregman iteration. Journal of Scientific Computing 46, 20–46.

[95] Zhao, X.Y., Sun, D., Toh, K.C., 2010. A Newton-CG augmented Lagrangian method for semidefinite programming. SIAM J. Optim. 20, 1737–1765.

[96] Zygmunt, M., 2002. Matrix chebyshev polynomials and continued fractions. Linear algebra and its applications 340, 155–168.

Acknowledgements

In the last five years, there have been many people, who have encouraged, supported and helped me through the journey to pursuit my Ph.D. degree. Without them, the work would never be done.

First of all, I would like to show my most respect and express my deepest gratitude to my supervisor Prof. He Bingsheng at Nanjing university. I thank him for accepting me as his Ph.D. student, and leading me into the amazing optimization world. Prof. He is very familiar with optimization field and always full of enthusiasm in doing research. His great creativity and stringent academic attitude impress me significantly. It is his patient guidance, endless support and continuous encouragement which provide me a strong impetus to growing up in academy. His wisdom and attitude will always be a guide to me.

My deepest thanks also go to my co-supervisors Prof. Sun Defeng and Prof. Toh Kim-Chuan at National University of Singapore. Prof. Sun is always of great enthusiasm about discussing issues on optimization and willing to share fresh ideas. He put great effort on offering me an extensive theoretical training, including convex and nonsmooth analysis, sensitivity analysis, polynomial optimization, copositive programming and many others, which has greatly facilitated my research progress. I also would like to thank Prof. Toh for offering me the great opportunity to visit the optimization group at NUS. During my staying in Singapore, Prof. Toh provided me his generous support and meticulous care in my daily life and encouragement in my academic research. He provided me an extensive training on numerical computation, especially in matrix optimization computation, which has enriched my skills in algorithms and implementation.

I would like to express my gratitude to Prof. Liu Yongjin, Dr. Li Min and Dr. Yuan Xiaoming for the co-authorship of several papers and for their invaluable suggestions. My thanks also go to the members of optimization group at Nanjing university and National university of Singapore for sharing their knowledge with me, including Dr. Gu Guoyong, Dr. Shen Yuan, Dr. Tao Min, Dr. Yang Junfeng, Dr. Zhang Wenxing, Cai Xingju, Wang Xiangfeng and You Yanfei at NJU; Dr. Ding Chao, Dr. Jiang Kaifeng, Du Mengyu, Gong Zheng , Li Xudong, Miao Weimin, Shi DongJian and Wu Bin at

NUS. I should give special thanks to Shen Yuan and Zhang Wenxing for their help in revising my thesis.

On the personal side, I would like to thank my parents, for their unconditional love and support all through my life. Last but not least, I am also greatly indebted to my wife for her understanding and patience throughout the years of my research.