SADDLE POINT SYSTEMS IN OPTIMIZATION

BY

HAFED AHMED MOHSEN SAEED

A Thesis Presented to the

DEANSHIP OF GRADUATE STUDIES

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS

DHAHRAN, SAUDI ARABIA

In Partial Fulfillment of the

Requirements for the Degree of

MASTER OF SCIENCE

In

MATHEMATICS

December, 2020

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS
DHAHRAN 31261, SAUDI ARABIA

DEANSHIP OF GRADUATE STUDIES

This thesis, written by **HAFED AHMED MOHSEN SAEED** under the direction of his thesis adviser and approved by his thesis committee, has been presented to and accepted by the Dean of Graduate Studies, in partial fulfillment of the requirements for the degree of **MASTER OF SCIENCE IN MATHEMATICS**.

Thesis Committee

8/2/2021

Dr. Mohammed Alshahrani (Adviser)

Dr.Faisal Fairag (Member)

Dr.Qamrul Hasan Ansari (Member)

Dr. Husain Salem Al-Attas
Department Chairman

Dr. Suliman Al-Homidan
Dean of Graduate Studies

23-2-2021
Date

Dedication

I dedicate my thesis work to my family. A special feeling of gratitude to my loving parents, whose words of encouragement and push for tenacity ring in my ears. I also wish to dedicate this work to my brothers and friends. This work would not be possible without their support. Special thank to my oldest brother Saleh, who never hesitated to support me during my study. I would also like to sincerely thank all teachers who taught me, as their encouragements and motivations have laid the foundation for this work. I also dedicate this work and give special thanks to my secondary school teacher, Hussain Naji Al-Mahd.

ACKNOWLEDGMENTS

Praise and thanks to Allah, Lord of the Worlds. I would like to express my deep gratitude to my advisor Dr. Mohammed Alshahrani, without his guidance, advice, instructions, and support, this work would have never been. I would like also to express my great appreciation to my thesis committee members, Dr. Faisal Fairag and Prof. Qamrul Hasan Ansar, for their energy, understanding and help throughout my work. My great thanks to all math department professors, academic staff, and employees, especially for the head of the department Dr. Husain Al-Attas for providing the best educational environment and support to all students in the department. Thanks to all my colleagues for their encouragement and support. Finally, I would like to thank all my family members including my dearest father and mother, my lovely wife, my little children, and my kind brothers and sisters for supporting me during the compilation of this work. Especial thank to my oldest brother, Saleh for his continuous help, support, and encouragement.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

THESIS ABSTRACT

NAME:	Hafed Ahmed Mohsen Saeed
TITLE OF STUDY:	Saddle Point Systems in Optimization
MAJOR FIELD:	Mathematics
DATE OF DEGREE:	December, 2020.

Saddle point systems appear in many applications such as optimization problems, discretizing Darcy and non-Darcy equations and solving partial differential equations. In our study, we consider the saddle point systems obtained from optimization problems. This study focuses on properties, particularly spectral properties, of the coefficient matrix and gives a view of the available numerical methods used to solve these systems. Also, in this woke, we establish a block triangular preconditioner matrix for solving saddle point systems that arise of primal-dual interior point algorithms in linear and quadratic programming. The preconditioner has the attractive property of improved eigenvalue clustering with increased ill-conditioning of the saddle point matrix. Therefore, the new preconditioner matrix with a new parameter γ give us well condition number whatever the value of the nullity (high or low).

ملخص الرسالة

الاسم: حافظ أحمد محسن سعيد

عنوان الدراسة: نظم نقطة السرج في التحسين

التخصص: الرياضيات

تاريخ الدرجة العلمية: ديسسمبر 2020

في السنوات الأخيرة ، تم تخصيص قدر كبير من العمل لهذه المشكلة لحل الأنظمة الخطية الكبيرة في شكل نقطة السرج. سبب ذلك الفائدة هي حقيقة أن مثل هذه المشاكل تنشأ في مجموعة متنوعة من التقنية والتطبيقات العلمية. على سبيل المثال ، مشكلة التحسين ،الشعبية المتزايدة باستمرار طرق العناصر المحدودة المختلطة في المجالات الهندسية مثل السوائل والصلبة كانت الميكانيكا مصدرًا رئيسيًا لأنظمة نقاط السرج. سبب آخر لهذا زيادة الاهتمام هو النجاح غير العادي لخوارزميات النقاط الداخلية في كل من التحسين الخطي وغير الخطي ، الأمر الذي يتطلب في جوهرهم حل سلسلة من الأنظمة في شكل نقطة سرج لذلك لقد قدمنا في هذه الاطروحة دراسة شاملة تبين كيف ظهرت نقطه السرج في مشاكل التحسين وكذلك دراسة الخصائص لذلك الانظمة ،وكذلك لقد قمنا بدراسة التهيئة المسبقة للانظمة الخطيه والتي تكون مهمه جدا لاختصار الحسابات في حل الانظمة الخطيه عالية الابعاد. لقد قدمنا بارامترات جديدة والتي تعمل على تحسين الحل والحصول على اقل عدد من الدورات وكذلك تحسين توزيع القيم الذاتية وقمنا بتطبيق امثلة عددية لايضاح كل ما تكلمنا عنه.

CHAPTER 1

INTRODUCTION

1.1 Motivation and Overview

Saddle point systems appear in many application such as: optimization problems, discrediting Stokes and Navier-Stokes problem, optimal control , electrical circuits and networks, image reconstruction problems, parameter identification problems, solving partial differential equations, finance, economics and discrediting Darcy and non-Darcy equations.

A saddle point problem is a linear system with the form:

$$\begin{bmatrix} G & B_1^T \\ B_2 & -C \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} f \\ g \end{bmatrix}, \tag{1.1}$$

where $G \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{m\times n}$, $C \in \mathbb{R}^{m\times m}$ with $n \geq m$. Problem (1.1) can be written as

$$\mathcal{A}z = b \tag{1.2}$$

with

$$A = \begin{bmatrix} G & B_1^T \\ B_2 & -C \end{bmatrix}, \quad z = \begin{bmatrix} u \\ v \end{bmatrix}, \quad b = \begin{bmatrix} f \\ g \end{bmatrix}$$

Problem (1.1) is called generalized saddle point system, in case $C = 0$, it is called saddle point system. The above system appears in many domains such as what we mentioned above [1].

If $B_1 = B_2 = B$ and $G = G^T, \quad C = C^T$ in (1.1) then the system is symmetric and becomes of the form:

$$\begin{bmatrix} G & B^T \\ B & -C \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} f \\ g \end{bmatrix} \tag{1.3}$$

Problem (1.3) appears in many domains. It emerges, for example, when constrained optimization problems are solved by interior point methods [1–3]. Also, problem (1.3) obtained from discretizing the Euler–Lagrange equations associated with the image deblurring problem as in [4].

If $B_1 = B_2 = B$ *and* $C = 0$ in equation (1.1), we have

$$\begin{bmatrix} G & B^T \\ B & 0 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} f \\ g \end{bmatrix} \tag{1.4}$$

Problem (1.4) arises in many optimization problems for instance it emerges as the the first-order optimality conditions for the next equality-constrained quadratic programming problem [5]

$$\begin{aligned} &\min \qquad f(u) = \tfrac{1}{2}u^T G u - f^T u \\ &\text{subject to} \\ &\qquad\qquad Bu = g \\ &\qquad\qquad u \in \mathbb{R}^n \end{aligned} \tag{1.5}$$

In this case the variable v represents the vector of Lagrange multipliers. Any solution (u_*, v_*) of (1.4) is a saddle point for the Lagrangian

$$L(u, v) = \frac{1}{2}u^T G u - f^T u + v^T(Bu - g)$$

hence the name 'saddle point problem' given to (1.4). Recall that a saddle point is a point $(u_*, v_*) \in \mathbb{R}^{n+m}$ that satisfies

$L(u_*, v) \leq L(u_*, v_*) \leq L(u, v_*)$ for any $u \in \mathbb{R}^n$ and $v \in \mathbb{R}^m$

or, equivalently,

$\underset{u}{min}\ \underset{v}{max}\ L(u, v) = L(u_*, v_*) = \underset{v}{max}\ \underset{u}{min}\ L(u, v)$ [1, 2].

Also, problem (1.4) arises in fluid dynamics (Stokes' problem) and constrained least squares problem [2]

If $C = 0$ in (1.1), we get the following problem

$$\begin{bmatrix} G & B_1^T \\ B_2 & 0 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} f \\ g \end{bmatrix} \tag{1.6}$$

System (1.6) appears in many applications, and numerical solution methods have been extensively explored for this type of problem [6].

We consider the saddle point systems of the form (1.3) and (1.4). These systems obtained from optimization problems. This work is focused on studying the properties of these system and numerical methods for solving them. These linear systems are large, indefinite and often have poor spectral properties. Due to these properties, numerical solutions of such systems have been a major challenge for solver developers because they require a huge amount of computer time and storage.

In the literature, one can identify two family of methods to solve Problem (1.3) and (1.4). The first are direct. The other are iterative. These include multigrid methods, stationary methods, and Krylov subspace methods [2], but the iterative methods that are nowadays applied to large-scale linear systems are mostly Krylov subspace methods. The converge of Krylov methods is very slow, because the condition number of such system is huge. To overcome this drawback preconditioning techniques are needed [7].

Preconditioning is a technique that converts the linear system (1.2) into another system

$$P^{-1}\mathcal{A}z = P^{-1}b. \tag{1.7}$$

The matrix P, called a preconditioner, is easy to invert and the preconditioned matrix $P^{-1}\mathcal{A}$ has a good clustering behavior of the eigenvalues.

In the literature, there are different types of preconditioners, such as

- Schur complement preconditioners.
- diagonal preconditioner.
- Hermitian and skew-Hermitian splitting (HSS) preconditioner.

- Triangular preconditioner.

1.2 Definite and indefinite linear systems

We consider the system (1.3) where G, B and C are as in (1.1). Then $\mathcal{A}$ is positive definite if $z^T\mathcal{A}z > 0$ for $z \neq 0$ or negative definite, i.e. $z^T\mathcal{A}z < 0$ for $z \neq 0$, other wise, the system is classified as indefinite.

If G and C are symmetric ($G^T = G$ and $C^T = C$), then A is symmetric ($\mathcal{A}^T = \mathcal{A}$).

1.3 Contributions

The thesis contributions are outline below:

- We introduce a new preconditioner

  P=$\begin{bmatrix} G + \frac{2k}{k+1} B^T W^{-1} B & \left(1 + \dfrac{2k}{k+1}\right) B^T \\ 0 & -W \end{bmatrix}$

  where $k \geq 1$.

- We introduce new parameters for $W = \gamma I_m$ for quadratic programming $\gamma = \frac{\|B\|_2^2}{\|G\|_2 max(G/c)}$ or $\gamma = \frac{\|B\|_1^2}{\|G\|_1 max(G/c)}$ (where c is constant) these parameters fit with the initial steps of interior point methods.

- We introduce new parameters for $W = \gamma I_m$ for quadratic programming $\gamma = \frac{\|B\|_2^2}{\|G\|_2 max(Q/c)}$ and $\gamma = \frac{1}{max(G/c)}$ (where c is constant) these parameters fit with all interior point methods steps.

1.4 Outline Of Thesis

The outline of this thesis is as follows. In Chapter 2, we introduce the saddle point systems that arise in optimization problems such as using interior point methods to solve linear, quadratic and nonlinear programming, and linear systems of saddle point type commonly arise when solving least squares problems. Chapter 3 discusses the properties of saddle point matrices (basic algebraic properties such as invertibility, the existence of their block factorizations, the expressions for their inverses and the analysis of their spectral properties such as eigenvalue localization). In Chapter 4, we present the fundamental definitions of preconditioning techniques, the literature review of preconditioning techniques, and show the main results analysis. Chapter 5, reports our numerical experiments results. Finally, in Chapter 6, we provide briefly the conclusion and the recommendations for future work.

CHAPTER 2

APPLICATIONS LEADING TO SADDLE POINT PROBLEMS IN OPTIMIZATION

As already mentioned, large-scale saddle point systems appear in many application such as : optimization problems, discreditizing Stokes and Navier-Stokes problem, electrical circuits and networks, image reconstruction problems, parameter identification problems, partial differential equations, finance, economics, optimal control and discreditizing Darcy and non-Darcy equations. In this study we consider saddle point systems obtained from optimization problems [1, 2, 8–14].

2.1 Saddle point systems from interior point methods

Interior point methods for linear, quadratic and nonlinear optimization in many details are different but they depend on the same linear algebra kernel. In the following we discuss the three cases

2.1.1 Linear Programming (LP) Problem

An LP is defined as maximizing or minimizing a linear function subject to linear constraints(equality or/and inequality). The standard form of LP problem is as follows,

$$\begin{aligned} \min \quad & c^T u \\ \text{subject to} \quad & \\ & Bu = g, \\ & u \geq 0 \end{aligned} \tag{2.1}$$

where $c,\ u \in \ \mathbb{R}^n, B \in \ \mathbb{R}^{m\times n},\ g \in \ \mathbb{R}^m, m < n$ (we assume that $m < n$, otherwise the system $Bu = g$ is infeasible or contains redundant rows. If $m = n$ the system $Bu = g$ has a unique solutions.)

A solution to the LP problem is a vector $u_* \in \ \mathbb{R}^n$ such that $c^T u_*$ is minimized and the constraints $Bu_* = g,\ \ u_* \in \ \mathbb{R}^n$ are each satisfied. Throughout this work, we assume that B has full row rank.

The Karush-Kuhn-Tucker (KKT) conditions

The Karush-Kuhn-Tucker (KKT) conditions can be summarised in the following theorem:

Theorem 2.1 *Consider the problem of minimizing a smooth objective function* $f : \mathbb{R}^n \to \mathbb{R}$ *with respect to a collection of constraints* $\{g_i(u) \geq 0 \vee g_i(u) = 0\}_{i=1}^m$ *where* $g_i : \mathbb{R}^n \to \mathbb{R}$ *is also a smooth function. Let* L *be the Lagrangian associated with this problem. Suppose* u^* *is a solution of this problem, and that the subset of constraints satisfying* $\{g_i(u^*) = 0\}$ *has the property that* $\{\nabla g_i(u^*)\}$ *is linearly independent. Then there is a vector of Lagrange multipliers* v^* *for which the following holds:*

- $\nabla_u(u^*, v^*) = 0$
- *All of the constraints are satisfied at* u^*.
- $v_i^* \geq 0$
- *For all* $1 \leq i \leq m, v_i^* g_i(u^*) = 0$.

Newton Search Directions

Suppose $F : \mathbb{R}^n \to \mathbb{R}^m$ is a continuously differentiable vector valued function, and write in general $F(y) = (F_1(y), F_2(y), ..., F_m(y))^T$. The Jacobian of $F(y)$, denoted $J(y)$, is the $m \times n$ matrix of first derivatives of F:

$J(y) = (\frac{\partial F_i(y)}{\partial(y_j)})_{i,j} \quad , 1 \leq i \leq m, 1 \leq j \leq n$. Since we are trying to solve the equation $F(y) = 0$. Then one iteration of Newton's given by

$J(y)p_k = -F(y)$

where p_k is the the Newton direction and $y = (u, v, s)$ is initial point. For the convergence of this method, (see [10]).

Notice that convexity of the problem ensures that KKT conditions are sufficient for a global minimum [15].

2.1.2 The primal approach

We associate a barrier function $D(u, \mu)$ defined as

$$D(u, \mu) = c^T u - \mu \sum_{j=1}^{n} \ln(u_j) \tag{2.2}$$

where $\mu > 0$ is real. Now we consider

$$\min_{u \in \mathbb{R}^n} D(u, \mu) \quad \textit{subject to} \quad Bu = g \tag{2.3}$$

where $\mu > 0$ is a barrier parameter.

From (2.2) $\ln(u_j)$ is concave, which implies $-\mu \ln(u_j)$ is convex. Therefore, the $D(u, \mu)$ is a positive sum of convex functions, which implies that $D(u, \mu)$ is convex then the KKT conditions are sufficient to get the optimal solution .

Consider the Lagrange function

$$L(u, v; \mu) = c^T u - \mu \sum_{j=1}^{n} \ln(u_j) - v^T (Bu - g) \tag{2.4}$$

where $v \in \mathbb{R}^m$ are the Lagrange multipliers for the $Bu = g$ and the KKT conditions are

$$\nabla_u L(u, v; \mu) = c - \mu U^{-1}e - B^T v = 0$$
$$\nabla_v L(u, v; \mu) = g - Bu \tag{2.5}$$

where $e = (1, 1, ..., 1)^T$ and $U = diag(u_1, u_2, ..., u_n)$

Introducing the vector $s = \mu U^{-1}e$. The KKT conditions in (2.5) can be rewritten in light of the definition of s to yield

$$Bu = g, \quad u > 0$$
$$B^T v + s = c \tag{2.6}$$
$$Us = \mu e \ \ or \ \ USe = \mu e$$

where $S = diag(s_1, s_2, ..., s_n)$

Since the objective function in (2.2) is convex, then the optimality (KKT) conditions are sufficient to get the optimal solution (see [16] p.244). The first set of equalities in (2.6) enforces primal feasibility, the second set enforces dual feasibility, and the last set is the complementarity condition perturbed by μ . Let $(u(\mu), v(\mu), s(\mu))$ be a solution to (2.6) for some $\mu > 0$. Then $u(\mu)$ is primal feasible. Furthermore, $B^T v(\mu) + s(\mu) = c$ and $s(\mu) = \mu U^{-1}e > 0$, which shows $(v(\mu), s(\mu))$ is dual feasible. In other words $(u(\mu), v(\mu), s(\mu))$ is a primal-dual feasible pair. Therefore, the duality gap can be computed as follows,

$$c^T u(\mu) - g^T v(\mu) = u^T(\mu)s(\mu) = e^T U(\mu)s(\mu) = e^T(\mu e) = \mu\, e^T e = \mu n \tag{2.7}$$

Any solution that satisfies (2.6), and is optimal solution to (2.3), defines a primal-dual feasible pair. The system (2.6) can be rewritten as $F(u,v,s) = \begin{bmatrix} Bu-g & B^Tv+s-c & USe-\mu e \end{bmatrix}^T = 0$ where $F: \mathbb{R}^{2n\times m} \rightarrow \mathbb{R}^{2n\times m}$. Solving the nonlinear equations (2.6) by Newton's method. That is solving the system

$$\begin{bmatrix} B & 0 & 0 \\ 0 & B^T & I \\ S & 0 & U \end{bmatrix} \begin{bmatrix} \Delta u \\ \Delta v \\ \Delta s \end{bmatrix} = \begin{bmatrix} \eta_p \\ \eta_d \\ \eta_\mu \end{bmatrix}, \tag{2.8}$$

where $\eta_p = g - Bu \quad \eta_d = c - B^Tv - s \quad \eta_\mu = \mu e - USe$ By elimination of $\Delta s = U^{-1}\eta_\mu - U^{-1}S\Delta u$ from the second equation we get the $(n\times m)\times(n\times m)$ symmetric indefinite augmented system of linear equations

$$\begin{bmatrix} \theta^{-1} & B^T \\ B & 0 \end{bmatrix} \begin{bmatrix} -\Delta u \\ \Delta v \end{bmatrix} = \begin{bmatrix} f \\ g \end{bmatrix}, \tag{2.9}$$

where $\theta^{-1} = U^{-1}S \;,\quad f = \eta_d - U^{-1}\eta_\mu \;,\quad g = -\eta_p$

where $\theta^{-1} \in \mathbb{R}^{n\times n}$ $(diagonal)$, $B \in \mathbb{R}^{m\times n}$, with $n \geq m$ and assuming B has full rank $(rank(B) = m)$. We first observe that the ill-conditioning in the saddle point system (2.9) is a consequence of the properties of the diagonal scaling matrix. From the complementarity condition for linear programs we know that, at the optimum, $u_js_j = 0,\;\; j \in 1,2,...,n$. The condition $u_js_j = 0$ is satisfied if at least one of the variables u_j and s_j is zero.

The first step of the primal dual algorithm consists of solving (2.8),

$$
\begin{bmatrix} u_k \\ v_k \\ s_k \end{bmatrix} = \begin{bmatrix} u \\ v \\ s \end{bmatrix} + \alpha \begin{bmatrix} \Delta u \\ \Delta v \\ \Delta s \end{bmatrix} \tag{2.10}
$$

where u, v, s are the starting points and $\Delta u, \Delta v, \Delta s$ are the Newton's direction that obtained in (2.8) equation and then a new point is obtained for a suitable choice of α. For the choosing of α see([9])

2.1.3 A dual approach

The standard for the dual of (2.1):

$$
\begin{aligned}
\max \quad & g^T v \\
\text{subject to} \quad & \\
& B^T v + s = c, \\
& s \geq 0
\end{aligned} \tag{2.11}
$$

We can introduce the barrier term to the dual problem as follows.,

$$
\begin{aligned}
\max \quad & g^T v + \mu \textstyle\sum_{j=1}^{n} ln(s_j) \\
\text{subject to} \quad & \\
& B^T v + s = c, \\
& s > 0
\end{aligned} \tag{2.12}
$$

The Lagrange function is given by

$$L(u, v, s) = g^T v + \mu \sum_{j=1}^{n} ln(s_j) - u^T (B^T v + s - c). \tag{2.13}$$

where u denote the Lagrange multipliers

The optimality(KKT) conditions are,

$$\begin{aligned} \nabla_u L(u, v, s) &= c - s - B^T v = 0, \quad s > 0 \\ \nabla_v L(u, v, s) &= g - Bu = 0, \\ \nabla_s L(u, v, s) &= \mu S^{-1} e - u = 0. \end{aligned} \tag{2.14}$$

This equation is equivalent to

$$\begin{aligned} Bu &= g, \\ B^T v + s &= c, \quad s > 0 \\ Us &= \mu e \end{aligned} \tag{2.15}$$

The solution for the nonlinear (2.14) is exactly the same process what we did in (2.6).

2.1.4 The primal-dual approach

A set of the first optimal conditions for the barrier problem was obtained in both primal and dual cases. The combination of these two sets of optimal conditions gives

$$\begin{gathered} Bu = g, \qquad u > 0 \\ B^T v + s = c, \quad s > 0 \\ Us = \mu e \;\; or \;\; USe = \mu e \end{gathered} \tag{2.16}$$

These conditions are called the perturbed Karush-Kuhn-Tucker (KKT) conditions, because they are identical to the KKT conditions to original LP problem, except the complementary conditions have been perturbed by μ . Therefore, a solution to (2.16) for a sufficiently small μ is a good approximation to the optimal solution.

Solving the nonlinear equations (2.16) by Newton's method . That is solving the system

$$\begin{bmatrix} B & 0 & 0 \\ 0 & B^T & I \\ S & 0 & U \end{bmatrix} \begin{bmatrix} \Delta u \\ \Delta v \\ \Delta s \end{bmatrix} = \begin{bmatrix} \eta_p \\ \eta_d \\ \eta_\mu \end{bmatrix}, \tag{2.17}$$

where $\eta_p = g - Bu \quad \eta_d = c - B^T v - s \quad \eta_\mu = \mu e - USe$ By elimination of $\Delta s = U^{-1}\eta_\mu - U^{-1}S\Delta u$ from the second equation we get the symmetric indefinite augmented system of linear equations

$$\begin{bmatrix} \theta^{-1} & B^T \\ B & 0 \end{bmatrix} \begin{bmatrix} -\Delta u \\ \Delta v \end{bmatrix} = \begin{bmatrix} f \\ g \end{bmatrix}, \tag{2.18}$$

where

$\theta^{-1} = U^{-1}S$ (diagonal), $\quad f = \eta_d - U^{-1}\eta_\mu$, $\quad g = -\eta_p$

where $\theta^{-1} \in \mathbb{R}^{n\times\, n}$, $B \in \mathbb{R}^{m\times n}$, with $n \geq m$ and assuming $\mathrm{rank}(B) = m$. Exactly as (2.9).

Remark 1 *The conditions in system* (2.16) *are equivalent to central path equations and,* $\mu = \beta \frac{s^T u}{n}$ *where* β *is constant. This method give us better result of taking* $\mu = 0$ *even if* μ *is large value , and gradually reduce it toward* 0 *as the algorithm proceeds.*

Remark 2 *The primal-dual interior-point methods in the case of barrier function differ in one key point from the strict primal methods: the Newton direction is computed for* (v, s) *as well as for* u*, but the primal-dual in case* $\mu > 0$ *gives better result than primal-dual in the case* $\mu = 0$.

2.1.5 Quadratic programming

Consider the convex quadratic programming problem

$$\begin{aligned} \min \qquad & \tfrac{1}{2}u^T Q u + c^T u \\ \text{subject to} \qquad & \\ & Bu = g, \\ & u \geq 0 \end{aligned} \tag{2.19}$$

where $Q \in \mathbb{R}^{n\times n}$ is positive semidefinite matrix, $B \in \mathbb{R}^{m\times n}$ has full rank ($rank(B) = m$), $c,\ u \in \mathbb{R}^n$, $B \in \mathbb{R}^{m\times n}$ *and* $g \in \mathbb{R}^m$

As we did with linear programming above replacing inequality constraints with the logarithmic barriers to get

$$\begin{aligned} \min \quad & \tfrac{1}{2}u^T Q u + c^T u - \mu \textstyle\sum_{j=1}^{n} \ln(u_j) \\ \text{subject to} \quad & \\ & Bu = g, \end{aligned} \tag{2.20}$$

Where $\mu \geq 0$ is a barrier parameter. By The Lagrangian

$$L(u, v; \mu) = \frac{1}{2}u^T Q u + c^T u - \mu \sum_{j=1}^{n} \ln(u_j) - v^T(Bu - g)$$

where $v \in \mathbb{R}^m$ are the Lagrange multipliers and the KKT conditions are

$$\begin{aligned} \nabla_u L(u, v; \mu) = c + Qu - \mu U^{-1} e - B^T v = 0 \\ \nabla_v L(u, v, \mu) = g - Bu = 0 \end{aligned} \tag{2.21}$$

where $U = diag(u_1, u_2, ..., u_n)$

Setting $s = \mu U^{-1} e$, i.e $\quad USe = \mu e$

where $S = diag(s_1, s_2, ..., s_n)$ and $e = (1, 1, ..., 1)^T$.

The KKT conditions in (2.21) can be rewritten in light of the definition of s to yield

$$\begin{aligned} Bu = g \\ B^T v + s - Qu = c \\ USe = \mu e \\ (u, s) \geq 0 \end{aligned} \tag{2.22}$$

Solving the nonlinear equations (2.22) by Newton's method. That is solving the system

$$\begin{bmatrix} B & 0 & 0 \\ -Q & B^T & I \\ S & 0 & U \end{bmatrix} \begin{bmatrix} \Delta u \\ \Delta v \\ \Delta s \end{bmatrix} = \begin{bmatrix} \eta_p \\ \eta_d \\ \eta_\mu \end{bmatrix}, \tag{2.23}$$

where $\eta_p = g - Bu,\ \eta_d = c - B^T v - s + Qu,\ \eta_\mu = \mu e - USe$ and gradually decreasing the Barrier parameter μ. Eliminating $\Delta s = U^{-1}\eta_\mu - U^{-1}S\Delta u$ from the second equation we get the symmetric indefinite augmented system of linear equations

$$\begin{bmatrix} Q + \theta^{-1} & B^T \\ B & 0 \end{bmatrix} \begin{bmatrix} -\Delta u \\ \Delta v \end{bmatrix} = \begin{bmatrix} f \\ g \end{bmatrix}, \tag{2.24}$$

where $\theta^{-1} = U^{-1}S, \quad f = \eta_d - U^{-1}\eta_\mu\ , \quad g = -\eta_p$

where $(Q + \theta^{-1}) \in \mathbb{R}^{n\times\ n}, \quad B \in \mathbb{R}^{m\times n},$ with $n \geq m$ and assuming B has full rank $(rank(B) = m)$. As we saw with LP since the saddle point system (2.24) contains θ^{-1} then it is ill-conditioning for the same reasons. Solving the systems in (2.23) and (2.24) to obtain the Newton direction and for the solution do the same process as in (2.8) .

Remark 3 *If we replace the condition $u \geq 0$ by $u \in \mathbb{R}^n$ then using Lagrangian and* *the KKT condition we obtain the following matrix*

$$\begin{bmatrix} Q & B^T \\ B & 0 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} f \\ g \end{bmatrix}, \tag{2.25}$$

where $f = -c$ and u, v are the solution and the above system can be rewritten as the form in ([15] p.451).

Remark 4 *Strictly convex quadratic programming are those in which Q is positive definite.*

Remark 5 *In saddle point systems in* (2.19) *if the constraint $Bu = g$ is replaced by the inequality constraint then we get saddle point system as* (2.25) *but last block not equal zero for the details see ([15] p.480-483) but in this work we consider as in* (2.19).

2.1.6 Nonlinear convex programming

Consider the convex nonlinear optimization problem

$$\begin{aligned} &\min && f(u) \\ &\text{subject to} && \\ & && h(u) \leqslant 0, \end{aligned} \tag{2.26}$$

where $u \in \mathbb{R}^n$, and $f : \mathbb{R}^n \to \mathbb{R}$ and $h : \mathbb{R}^n \to \mathbb{R}^m$ with $h(u) = (h(u)_1, ..., h_m(u))^T$ are convex, twice differentiable. Having replaced inequality constraints with an equality $h(u) + x = 0$, where $x \in \mathbb{R}^m$ is a nonnegative slack variable, we can formulate the associated barrier problem:

$$\begin{aligned} &\min \qquad f(u) - \mu \textstyle\sum_{i=1}^{m} \ln(x_i) \\ &\text{subject to} \\ &\qquad\qquad h(u) + x = 0 \end{aligned} \tag{2.27}$$

where $\mu \geq 0$

The corresponding Lagrangian is

$L(u, v, x; \mu) = f(u) - \mu \sum_{i=1}^{m} \ln(x_i) + v^T(h(u) + x)$ Writing the KKT conditions for this problem

$$\begin{aligned} \nabla_u L(u, v, x; \mu) &= \nabla f(u) + \nabla h(u)^T v = 0 \\ \nabla_v L(u, v, x; \mu) &= h(u) + x = 0 \\ \nabla_x L(u, v, x; \mu) &= v - \mu X^{-1} e \end{aligned} \tag{2.28}$$

where $X = diag(x_1, x_2, ..., x_m)$, $e = (1, 1, ..., 1)^T$.

The KKT conditions (2.28) can be rewritten by the folowing form

$$\begin{gathered} \nabla f(u) + \nabla h(u)^T v = 0 \\ h(u) + x = 0 \\ V X e = \mu e \\ (v, x) \geq 0 \end{gathered} \tag{2.29}$$

where $V = diag(v_1, v_2, ..., v_m)$, $\nabla f(u)$ denotes the gradient and defined as $\nabla f(u) = (\frac{\partial f(u)}{\partial u_1}, ..., \frac{\partial f(u)}{\partial u_n})^T$, and the term $\nabla h(u)$ denotes here the matrix $\nabla h(u) \in \mathbb{R}^{m \times n}$ defined as $\nabla h(u) = (\nabla h_1(u), ..., \nabla h_m(u))^T$.

Solving the nonlinear equations (2.29) by Newton's method . That is solving the system

$$\begin{bmatrix} H(u,v) & B(u)^T & 0 \\ B(u) & 0 & I \\ 0 & X & V \end{bmatrix} \begin{bmatrix} \Delta u \\ \Delta v \\ \Delta x \end{bmatrix} = \begin{bmatrix} -\nabla f(u) - B(u)^T v \\ -h(u) - x \\ \mu e - VXe \end{bmatrix}, \tag{2.30}$$

where $B(u) = \nabla h(u) \in \mathbb{R}^{m\times n}$, $H(u,v) = \nabla^2 f(u) + \sum_{i=1}^{m} v_i \nabla^2 h_i(u) \in \mathbb{R}^{n\times n}$ and gradually decreasing the Barrier parameter μ

By elimination of $\Delta x = \mu V^{-1}e - Xe - XV^{-1}\Delta v$

from the second equation we get

$$\begin{bmatrix} H(u,v) & B(u)^T \\ B(u) & -XV^{-1} \end{bmatrix} \begin{bmatrix} \Delta u \\ \Delta v \end{bmatrix} = \begin{bmatrix} -\nabla f(u) - B(u)^T v \\ -h(u) - \mu V^{-1}e \end{bmatrix}, \tag{2.31}$$

The above saddle point system is symmetric and indefinite since f and g is convex, the matrix H is positive semi-definite and if f is strictly convex H is positive definite.[1, 2]

2.2 Least Squares with Equality Constraints

Linear systems of saddle point type commonly arise when solving least squares problems. Consider the following least squares problem with linear equality constraints:

$$\begin{aligned} &min \quad \tfrac{1}{2}\|Au-b\|_2^2 \\ &subject \quad to \qquad Bu = g \end{aligned} \tag{2.32}$$

where $A \in V^{q\times n}$, $u \in \mathbb{R}^n$, $b \in \mathbb{R}^q$, $B \in \mathbb{R}^{m\times n}$, $g \in \mathbb{R}^m$, $n > m, q \geq n$.

by Lagrangian

$$L(u,v) = \frac{1}{2}\|Au-b\|_2^2 + v^T(Bu-g). \tag{i}$$

where $v \in R^m$. Since

$$\frac{1}{2}\|Au-b\|_2^2 = \frac{1}{2}(Au-b)^T(Au-b) = \frac{1}{2}[u^TA^TAu - u^TA^Tb - b^TAu + b^Tb] \tag{ii}$$

from (i) and (ii)

$$L(u,v) = \frac{1}{2}[u^TA^TAu - u^TA^Tb - b^TAu + b^Tb] + v^T(Bu-g). \tag{iii}$$

by KKT conditions for (iii)

$$\begin{aligned} \nabla_u L(u,v) &= \tfrac{1}{2}[u^T(A^TA + A^TA) - b^TA - b^TA] + v^TB = 0 \\ &= u^TA^TA - b^TA + v^TB = 0 \\ &= A^TAu - A^Tb + B^Tv = 0 \\ \nabla_v L(u,v) &= (Bu-g)^T = Bu - g = 0 \end{aligned} \tag{2.33}$$

From the above systems

$$\begin{bmatrix} A^T A & B^T \\ B & 0 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} A^T b \\ g \end{bmatrix}, \tag{2.34}$$

rewrite the above system as

$$\begin{bmatrix} G & B^T \\ B & 0 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} f \\ g \end{bmatrix}, \tag{2.35}$$

where $G = A^T A \quad f = A^T b$

The above system called saddle point system which arise form solving optimization problem using least square method [17] .

Remark 6 *since* $u^T A^T A u = (Au)^T (Au) = \|A^T A\|_2^2 \geq 0$

therefore $G = A^T A$ *is positive semidefinite.*

CHAPTER 3

PROPERTIES OF SADDLE POINT MATRICES

This chapter is dedicated to the establishment of basic algebraic properties of saddle point matrix $\mathcal{A}$ such as the existence of different factorizations, invertibility, spectral properties and conditioning. In developing solution algorithms, knowledge of these properties is important.

3.1 Block factorizations and the Schur complement

The saddle point matrix $\mathcal{A}$ when G is nonsingular can be factorized in the three following block triangular factorizations:

$$\mathcal{A} = \begin{bmatrix} G & B^T \\ B & -C \end{bmatrix} = \begin{bmatrix} I & 0 \\ BG^{-1} & I \end{bmatrix} \begin{bmatrix} G & 0 \\ 0 & -C - BG^{-1}B^T \end{bmatrix} \begin{bmatrix} I & G^{-1}B^T \\ 0 & I \end{bmatrix} \tag{3.1}$$

$$= \begin{bmatrix} I & 0 \\ BG^{-1} & I \end{bmatrix} \begin{bmatrix} G & B^T \\ 0 & -C - BG^{-1}B^T \end{bmatrix} \tag{3.2}$$

$$= \begin{bmatrix} G & 0 \\ B & -C - BG^{-1}B^T \end{bmatrix} \begin{bmatrix} I & G^{-1}B^T \\ 0 & I \end{bmatrix} \tag{3.3}$$

where $S = -(C + BG^{-1}B^T)$ is the Schur complement of G in $\mathcal{A}$. In this work, we can't use these properties since G is singular, as we saw in chapter 2 then the Schur complement does not exit, but we can use augmented Lagrangian techniques to replace G by $M = G + B^TW^{-1}B$, where W^{-1} is $m \times m$ symmetric positive definite and M is nonsingular.

3.2 Solvability conditions

G is needed to be nonsingular, so $\mathcal{A}$ is invertable if only if S is. Unfortunately, we can say very little in general about the invertibility of S. Placing some restrictions on the matrices G, B and C are necessary. Of course, our saddle point system is symmetric whether G is singular or not. We begin with G symmetric positive definite in (1.4). The Schur complement in this case reduces to $S = -BG^{-1}B^T$, a symmetric negative semidefinite matrix. It is obvious that S, and thus $\mathcal{A}$, is invertible if and only if B has full row rank ($rank(B) = m$), hence S is symmetric negative definite. The discussion for saddle point system in (1.3) where G is symmetric positive definite, $C \neq 0$ is symmetric positive semidefinite and $S = -(C + BG^{-1}B^T)$ can be summarized in the following theorem

Theorem 3.1 *Assume G is symmetric positive definite and C is symmetric positive semidefinite. If null$(B^T) \cap$ null$(C) = 0$, then the saddle point matrix $\mathcal{A}$ is nonsingular. In particular, $\mathcal{A}$ is invertible if B has full rank.*

When the condition G positive definite is relaxed. Then $\mathcal{A}$ may be singular, even if B has full rank. See the example ([18], p.17). Now we have the following result for the case G is symmetric positive semidefinite.

Theorem 3.2 *Assume that G is symmetric positive semidefinite, B has full-row rank, and the last block $C = O$ in (3.1). Then null$(G) \cap$ null$(B) = \{0\}$ is a necessary and sufficient condition for the saddle point matrix $\mathcal{A}$ to be nonsingular.*

Proof. Let $z = [u^T \quad v^T]^T$ be such that $\mathcal{A}z = 0$. Hence, $Gu + B^T v = 0$ and $Bu = 0$, i.e., $u \in$ null(B). It follows that $u^T G u = -u^T B^T v = -(Bu)^T v = 0$. We prove that since G is symmetric positive semidefinite, $u^T G u = 0$ implies $Gu = 0$. (see [2], p.21). This gives $u \in$ null(G) then $u \in$ null$(G) \cap$ null$(B) = \{0\}$, implying $u = 0$. Also we have $v = 0$, since $B^T v = -Gu = 0$ and B^T has a full-column rank. Therefore, $z = 0$ and $\mathcal{A}$ is nonsingular.

Assume now that null$(G) \cap$ null$(B) \neq \{0\}$. Taking $u \neq 0$ such that $Bu = 0$ and $Gu = 0$, we have that

$$\mathcal{A} \begin{bmatrix} u \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{3.4}$$

This, however, implies that $\mathcal{A}$ is singular. ∎

Example 3.1 Consider the saddle point matrix :

$$\mathcal{A} = \left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 0 & 1 \\ \hline 0 & 1 & 0 \end{array}\right]$$

Since $G = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ is symmetric positive semidefinite and the off-diagonal block $B^T = \begin{bmatrix} 0 & 1 \end{bmatrix}^T$ satisfying $null(G) = span\{\begin{bmatrix} 0 & 1 \end{bmatrix}^T\}$, $null(B) = span\{\begin{bmatrix} 1 & 0 \end{bmatrix}^T\}$, and $null(G) \cap null(B) = \{\begin{bmatrix} 0 & 0 \end{bmatrix}^T\}$.

Then by theorem (3.2) $\mathcal{A}$ is nonsingular.

Example 3.2 Consider the saddle point matrix

$$\mathcal{A} = \left[\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 0 & 0 \\ \hline 1 & 0 & 0 \end{array}\right]$$

Since $G = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ is symmetric positive semidefinite and the off-diagonal block $B^T = \begin{bmatrix} 1 & 0 \end{bmatrix}$ satisfying $null(G) = span\{\begin{bmatrix} 0 & 1 \end{bmatrix}^T\}$, $null(B) = span\{\begin{bmatrix} 0 & 1 \end{bmatrix}^T\}$, and $null(G) \cap null(B) = span\{\begin{bmatrix} 0 & 1 \end{bmatrix}^T\}$.

Since theorem (3.2) is not satisfied then $\mathcal{A}$ is singular.

3.3 The inverse of a saddle point matrix

If G is nonsingular, then it is clear from equations (3.1),(3.2) and (3.3) that the matrix $\mathcal{A}$ is nonsingular if and only if $S = -(C + BG^{-1}B^T)$ is nonsingular, and we have the explicit expression for the inverse of A in the form

$$\mathcal{A}^{-1} = \begin{bmatrix} G & B^T \\ B & -C \end{bmatrix}^{-1} = \begin{bmatrix} I & G^{-1}B^T \\ 0 & I \end{bmatrix}^{-1} \begin{bmatrix} G^{-1} & 0 \\ 0 & S^{-1} \end{bmatrix} \begin{bmatrix} I & 0 \\ BG^{-1} & I \end{bmatrix}^{-1} \tag{3.5}$$

$$= \begin{bmatrix} G^{-1} + G^{-1}B^TS^{-1}BG^{-1} & -G^{-1}B^TS^{-1} \\ -S^{-1}BG^{-1} & S^{-1} \end{bmatrix} \tag{3.6}$$

In the above equation if C is nonsingular but G is singular, one can give an analogous expression if we assume that the matrix $G+B^TC^{-1}B$, the Schur complement of C in $\mathcal{A}$, is nonsingular. An interesting special case arises when G is symmetric positive definite, $C = 0$, $S = -BG^{-1}B^T$ is nonsingular, and $g = 0$. Then the explicit expression for $\mathcal{A}^{-1}$ shows that the solution (u_*, v_*) of (1.4)) is given by

$$\begin{bmatrix} u_* \\ v_* \end{bmatrix} = \begin{bmatrix} (I + G^{-1}B^TS^{-1}B)G^{-1}f \\ -S^{-1}BG^{-1}f \end{bmatrix} \tag{3.7}$$

An alternative expression can be given for the inverse of $\mathcal{A}$ when G is positive semidefinite, B has full rank, $C = 0$ and $\text{null}(G) \cap \text{null}(B) = 0$ which by theorem (3.2) imply that $\mathcal{A}$ nonsingular. Denote by $Z \in R^{n\times(n-m)}$ any matrix whose columns form a basis for null(B). It also follows from $\text{null}(G) \cap \text{null}(B) = 0$ that the matrix G is symmetric positive definite on null(B), i.e., $Z^TGZ > 0$ for all $Z \neq 0$ where Z^TGZ is a $n-m \times n-m$ nonsingular matrix and symmetric positive definite. Then, we can express the inverse of $\mathcal{A}$ as

$$\mathcal{A}^{-1} = \begin{bmatrix} G & B^T \\ B & 0 \end{bmatrix}^{-1} = \begin{bmatrix} X & (I-XG)B^T(BB^T)^{-1} \\ (BB^T)^{-1}B(I-GX) & -(BB^T)^{-1}B(G-GXG)B^T(BB^T)^{-1} \end{bmatrix} \tag{3.8}$$

where $X = Z(Z^T G Z^{-1})Z^T$

3.4 Spectral properties of saddle point matrices

To solve saddle point matrices by iterative methods we need to know their spectral properties. Now assume that G is symmetric positive definite and B^T has a full-column rank, and C (possibly zero) is symmetric positive semidefinite. Then from (3.1) we have

$$\begin{bmatrix} I & 0 \\ -BG^{-1} & I \end{bmatrix} \begin{bmatrix} G & B^T \\ B & -C \end{bmatrix} \begin{bmatrix} I & -G^{-1}B^T \\ 0 & I \end{bmatrix} = \begin{bmatrix} G & 0 \\ 0 & S \end{bmatrix} \tag{3.9}$$

where $S = -(C + BG^{-1}B^T)$ is symmetric negative definite, since $\mathcal{A}$ is congruent to the block diagonal matrix (3.9), it follows from that $\mathcal{A}$ is indefinite, with n positive and m negative eigenvalues [1, 2, 19].

This still true when G is symmetric positive semidefinite and the condition $\text{null}(G) \cap \text{null}(B) = \{0\}$ is satisfied. The proof as follows: using the QR factorization of B^T where $B^T \in \mathbb{R}^{n\times m}$

Let

$$B^T = \begin{bmatrix} Y & Z \end{bmatrix} \begin{bmatrix} R \\ 0 \end{bmatrix} \quad , \begin{bmatrix} Y & Z \end{bmatrix}^T \begin{bmatrix} Y & Z \end{bmatrix} = I \tag{3.10}$$

Where $Y \in \mathbb{R}^{n \times m}$ a matrix whose columns form an orthonormal basis of the $rang(B^T)$, and $Z \in \mathbb{R}^{n,n-m}$ is any matrix whose columns form an orthonormal basis of the null(B) so that $BZ = 0$ and the matrix $\begin{bmatrix} Z & Y \end{bmatrix}$ is orthogonal. Then the matrix is congruent to the 3-by-3 block matrix

$$\begin{bmatrix} \begin{bmatrix} Z & Y \end{bmatrix}^T & 0 \\ 0 & R^{-T} \end{bmatrix} \begin{bmatrix} G & B^T \\ B & 0 \end{bmatrix} \begin{bmatrix} \begin{bmatrix} Z & Y \end{bmatrix} & 0 \\ 0 & R^{-1} \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} Z & Y \end{bmatrix}^T G \begin{bmatrix} Z & Y \end{bmatrix} & \begin{bmatrix} Z & Y \end{bmatrix}^T B^T R^{-1} \\ R^{-T} B \begin{bmatrix} Z & Y \end{bmatrix} & 0 \end{bmatrix} =$$
$$\begin{bmatrix} Z^TGZ & Z^TGY & Z^TB^TR^{-1} \\ Y^TGZ & Y^TGY & Y^TB^TR^{-1} \\ R^{-T}BZ & R^{-T}BY & 0 \end{bmatrix} = \left[\begin{array}{c|cc} Z^TGZ & Z^TGY & 0 \\ \hline Y^TGZ & Y^TGY & I \\ 0 & I & 0 \end{array}\right] \tag{3.11}$$

since null$(G) \cap$ null$(B) = 0$ then $Z^TGZ \in \mathbb{R}^{n-m \times n-m}$ is SPD with $n - m$ positive eigenvalues.

The schur complement matrix of $\begin{bmatrix} Y^TGY & I \\ I & 0 \end{bmatrix}$ with respect to (3.11) is $Z^TGZ - \begin{bmatrix} Z^TGY & 0 \end{bmatrix} \begin{bmatrix} Y^TGY & I \\ I & 0 \end{bmatrix}^{-1} \begin{bmatrix} Y^TGZ \\ 0 \end{bmatrix} = Z^TGZ$ where $\begin{bmatrix} Y^TGY & I \\ I & 0 \end{bmatrix}^{-1} = \begin{bmatrix} 0 & I \\ I & -Y^TGY \end{bmatrix}$.

The spectrum of G is given by the schur complement Z^TGZ and $\begin{bmatrix} Y^TGY & I \\ I & 0 \end{bmatrix}$ as we saw above that Z^TGZ has $n-m$ positive eigenvalues and we need to find the eigenvalues of $\begin{bmatrix} Y^TGY & I \\ I & 0 \end{bmatrix}$ which has m positive eigenvalues and m negative eigenvalues as we will see in the following

Proposition 3.1 *Assume $G \in \mathbb{R}^{n\times n}$ be a symmetric positive semidefinite matrix with the eigenvalues $0 \le \lambda_n(G) \le ... \le \lambda_1(G)$, and assume $B \in \mathbb{R}^{n\times n}$ be equal to $B = I$. Then $A = \begin{bmatrix} G & I \\ I & 0 \end{bmatrix}$ has all nonzero eigenvalues and they are given as $\frac{1}{2}(\lambda_i(G) \pm \sqrt{\lambda_i^2(G)+4}$ i= 1,2,3 .....n*

Proof.

$$\begin{bmatrix} G & I \\ I & 0 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \lambda \begin{bmatrix} u \\ v \end{bmatrix} \tag{3.12}$$

from the above equation we follow:

$$Gu + v = \lambda u. \tag{a}$$

$$u = \lambda v. \tag{b}$$

where $\lambda \neq 0$ (if $\lambda = 0$ then from (a) and (b) $u = 0, v = 0$).

Substitute (b) in (a) to get

$\lambda Gv + v = \lambda^2 v$ for any non negative eigenvalues $\lambda_i(G) \ge 0$ of G then we have

$\lambda = \frac{1}{2}(\lambda_i(G) \pm \sqrt{\lambda_i^2(G)+4})$ for i= 1,2,3 .....n positive and eigenvalues. ∎

Remark 7 *(i) Assume* G *is nonsingular with eigenvalues* $0 < \lambda_n(G) \leq ... \leq \lambda_1(G)$. *Then it is clear that for all vectors* $y \in \mathbb{R}^n$, *we have* $\lambda_n(G)\|y\|^2 \leq y^T G y \leq \lambda_1(G)\|y\|^2$.

(ii) Assume B *has full-row rank with singular values* $0 < \sigma_m(B) \leq \cdots \leq \sigma_1(B)$. *Then we have for all* $z \in \mathbb{R}^m$ *and for all* $y \in$ *rang*$(B^T) = (null(B))^{\perp}$

$\sigma_m(B)\|z\| \leq \|B^T z\| \leq \sigma(B)\|z\| \quad , \sigma_m(B)y \leq \|By\| \leq \sigma_1(B)\|y\|$.

Proposition 3.2 *[1, 2] Assume* $G \in \mathbb{R}^{n\times n}$ *be symmetric positive definite with eigenvalues* $0 < \lambda_n(G) \leq ... \leq \lambda_1(G)$, *and let* $B \in \mathbb{R}^{n\times m}$ *be of full-row rank with singular values* $0 < \sigma_m(B) \leq ... \leq \sigma_1(B)$. *Let* $\sigma(\mathcal{A})$*denote the spectrum of* $\mathcal{A}$. *Then*

$\sigma(\mathcal{A}) \subset I^+ \cup I^-$

where

$I^+ = [\lambda_n(G), \frac{1}{2}(\lambda_1(G) + \sqrt{\lambda_1^2(G) + 4\sigma_1^2(B)}\ )]$

and

$I^- = [\frac{1}{2}(\lambda_n(G) - \sqrt{\lambda_n^2(G) + 4\sigma_1^2(B)}), \frac{1}{2}(\lambda_1(G) - \sqrt{\lambda_1^2(G) + 4\sigma_m^2(B)}\ )]$

Proof. For each eigenvalue $\lambda \in \sigma(\mathcal{A})$ and its corresponding eigenvector $\begin{bmatrix} u^T & v^T \end{bmatrix}^T \neq 0$

$$Gu + B^T v = \lambda \qquad \text{(a)}$$

$$Bu = \lambda v \qquad \text{(b)}$$

since $rank(B) = m$ and from (a) $B^T v = \lambda u - Gu$ if $u = 0$ the $v = 0$

So, $u \neq 0$

Let $\lambda > 0$ be a positive eigenvalues. Taking the inner product of (a) with u implies $u^TGu + u^TB^Tv = \lambda\|u\|^2$ (c). Taking the inner product of (b) with v imlies $v^TBu = \lambda\|v\|^2$ (d) from (c) and (d) we get $u^TGu+\lambda\|v\|^2 = \lambda\|u\|^2$ (e) from remark (7) we get $\lambda_n\|u\|^2 + \lambda\|v\|^2 \leq \lambda\|u\|^2$ this implies $(\lambda_n - \lambda)\|u\|^2 \leq -\lambda\|v\|^2 \leq 0$ which lead to the inequality $\lambda_n(G) \leq \lambda$. Now taking the inner product of (b) with Bu this gives $u^TB^Tv = \frac{1}{\lambda}\|Bu\|^2$ subtitiuting this into (c) we get $\lambda^2\|u\|^2 - \lambda u^TGu - \|Bu\|^2 = 0$ from remark(7) we can bounded the left half side from the the previous equation $(\lambda^2 - \lambda_1(G)\lambda - \sigma_1(B))\|u\|^2 \leq \lambda^2\|u\|^2 - \lambda u^TGu - \|Bu\|^2 = 0$ this gives the bound $\lambda \leq \frac{1}{2}(\lambda_1(G) + \sqrt{\lambda_1^2(G) + 4\sigma_1^2(B)})$

If $\lambda < 0$ be a negative eigenvalues and $(\lambda^2 - \lambda_n(G)\lambda - \sigma_1(B))\|u\|^2 \leq \lambda^2\|u\|^2 - \lambda u^TGu - \|Bu\|^2 = 0$ then $\lambda \geq \frac{1}{2}(\lambda_n(G) - \sqrt{\lambda_n^2(G) + 4\sigma_1^2(B)})$

At the end consider the orthogonal decomposition of $u = u|_{\text{null}(B)} + u|_{rang(B^T)}$ where $u|_{\text{null}(B)} \in \text{null}(B)$ and $u|_{rang(B^T)} \in rang(B)$ substituting u into (b) it follows that $v = \frac{1}{\lambda}Bu|_{rang(B^T)}$. Taiking the inner product of (a) with $u|_{rang(B^T)}$ to get $u|_{reng(B^T)}^TGu + u|_{reng(B^T)}^TB^Tv = \lambda u|_{rang(B^T)}^Tu$ substituting u and v into this equation we get

$u|_{rang(B^T)}^TGu|_{\text{null}(B)} = -u|_{rang(B^T)}^TGu|_{rang(B^T)} - \frac{1}{\lambda}\|Bu|_{rang(B^T)}\| + \lambda\|u|_{rang(B^T)}\|^2$ bounding this equation from below we got $u|_{rang(B^T)}^TGu|_{null(B)} \geq (\lambda - \lambda_1(G) - \frac{\sigma_m^2(B)}{\lambda})\|u|_{rang(B^T)}\|^2$. To obtain the upper bound take the inner product of (a) with $u|_{\text{null}(B)}$ we got $u|_{rang(B^T)}^TGu|_{\text{null}(B)} \leq (\lambda - \lambda_n(G))u|_{\text{null}(B)}\|^2 \leq 0$ so $\lambda^2 - \lambda_1(G)\lambda - \sigma_m^2\|u|_{rang(B^T)}\|^2 \geq 0$ if $u|_{rang(B^T)}^T = 0$ then $v = 0$ and $Gu|_{\text{null}(B)} = \lambda u|_{\text{null}(B)}$ with $\lambda < 0$ which can not be the case. Therefore we have $\lambda^2 - \lambda_1(G)\lambda - \sigma_m^2\|u|_{rang(B^T)}\|^2 \geq 0$ and $\lambda \leq \frac{1}{2}(\lambda_1(G) - \sqrt{\lambda_1^2(G) + 4\sigma_m^2(B)})$ ∎

Proposition 3.3 *([2]) Assume $G \in \mathbb{R}^{n\times n}$ be symmetric positive semidefinite with eigenvalues $0 = \lambda_n(G) \leq ... \leq \lambda_1(G)$ satisfying also $y^T Gy \geq \lambda_0(G)\|y\|^2$ for some $\lambda_0(G) > 0$ and for all $y \in$ null(B). Let $B \in \mathbb{R}^{m\times n}$ be of a full-row rank with singular values $0 < \sigma_m(B) \leq ... \leq \sigma_1(B)$.Let $\sigma(A)$ denote the spectrum of A. Then*

$\sigma(A) \subset I^+ \cup I^-$

where

$I^+ = [\lambda_0(G), \frac{1}{2}(\lambda_1(G) + \sqrt{\lambda_1^2(G) + 4\sigma_1^2(B)}\ )]$

and

$I^- = [\frac{1}{2}(\lambda_n(G) - \sqrt{\lambda_n^2(G) + 4\sigma_1^2(B)}), \frac{1}{2}(\lambda_1(G) - \sqrt{\lambda_1^2(G) + 4\sigma_m^2(B)}\ )]$

Remark 8 *Everything done in this chapter was when B has full rank. For the rank-deficient (see [1, 2])*

3.5 Conditioning issues

In practice several saddle point systems can be conditioned very poorly, therefore, when applying and developing solution algorithms attention must be paid during the development. When $G = G^T$ is a positive definite, B has a full rank, and $C = O$. In this case $\mathcal{A}$ is symmetric and the condition number of $\mathcal{A}$ $K(\mathcal{A}) = \frac{\lambda_n}{\lambda}$ where λ_n is the maxmium and λ is the minimum eigenvalues of $\mathcal{A}$. From Proposition 3.2 one can notice that the condition number of $\mathcal{A}$ grows unboundedly as either $\lambda_1 = \lambda_{min}$ or $\sigma_1 = \sigma_{min}(B)$ goes to zero where $\sigma_{min}(B)$ is the minimum singular value of B and λ_{min} is the minimum eigenvalue of G .

3.6 Krylov Subspaces Methods

Let

$$\mathcal{A}z = b, \tag{3.13}$$

be a linear system

where $\mathcal{A}$ is a nonsingular matrix and b a given vector. Assume that z_0 is an initial guess for the solution z of (3.13) and starting with the initial residual $r_0 = b - \mathcal{A}z_0$, the sequence of nested spaces K_k and L_k (constraint spaces) for $k = 0, 1, ...$ are build and the sequence of approximate solutions z_k are computed such that they satisfy the Petrov-Galerkin condition

$$z_k \in \ z_0 + K_k, r_k = b - \mathcal{A}z_0 \perp L_k \tag{3.14}$$

It is clear that the Krylov subspaces form a nested sequence that ends with dimension $d \equiv dim K_{n+m}(\mathcal{A}, r_0) \leq n + m$, i.e., $K_1(\mathcal{A}, r_0) \subset \cdots \subset K_d(A, r_0) = \cdots = K_{n+m}(\mathcal{A}, r_0)$ [1]. In particular, for each $d \geq k$, the Krylov subspace $K_k(\mathcal{A}, r_0)$ has dimension k. There are three iterative methods to solve large-scale linear systems such as multigrid methods, stationary methods, and Krylov subspace methods [2] but the Krylov subspace methods are the iterative methods that are nowadays applied to large-scale linear systems. This class of methods corresponds to $K_k = K_k(\mathcal{A}, r_0)$, where $K_k(\mathcal{A}, r_0)$ denotes the k-th Krylov subspace associated with $\mathcal{A}$ and r_0 defined as $K_k(\mathcal{A}, r_0) = span(r_0, \mathcal{A}r_0, ..., A^{k-1}r_0)$. It follows then from (3.14) that for the error $z_* - z_k$ and for the residual r_k, we have $z_* - z_k = p_k(\mathcal{A})(z_* - z_0)$ and $r_k = p_k(A)r_0$, respectively, where p_k stands for some polynomial of degree at most k satisfying $p_k(0) = 1$. The whole class of such polynomials will be denoted as p_k. The convergence of Krylov methods is very slowly, because the condition number of such system is huge to overcome this drawback preconditioning technique is needed.

Preconditioning is a technique that uses to fast the convergence of Krylov methods. Precisely, it is to convert the linear system $\mathcal{A}z = b$ into another system $P^{-1}\mathcal{A}z = P^{-1}b$ where P is the preconditioner matrix such that it is easy to invert and the preconditioned $P^{-1}\mathcal{A}$ has a good clustering behavior of its eigenvalues. The different versions of Krylov subspace methods arise from different choices of the subspace L_k and from the ways in which the system is preconditioned. Two broad choices for L_k give rise to the best-known techniques:

- $L_k = K_k(\mathcal{A}, r_0)$ FOM method for nonsymmetric systems and Conjugate Gradient Method (CG) method for symmetric positive definite systems.

- $L_k = \mathcal{A}K_k(\mathcal{A}, r_0)$ Generalized Minimal Residual Method (GMRES) method for nonsymmetric systems and) Minimal Residual Method (MINRES) method for symmetric indefinite systems Biorthogonalization methods

- $L_k = K_k(\mathcal{A}^T, r_0)$ (QMR method for nonsymmetric systems, Bi-CG method for nonsymmetric systems)

Remark 9 *For Preconditioned Krylov subspace methods we use :*

(i) Conjugate Gradient Method (CG) for symmetric positive-definite system and symmetric positive-definite preconditioner.

(ii) Minimal Residual Method (MINRES) for symmetric indefinite system and symmetric positive-definite preconditioner.

(iii) Generalized Minimal Residual Method (GMRES) for Symmetric indefinite system and nonsymmetric indefinite preconditioner.

CHAPTER 4

PRECONDITIONERS FOR SADDLE POINT SYSTEMS

4.1 Preconditioning techniques

Preconditioning is a technique that converts the linear system $\mathcal{A}z = b$ into another system $P^{-1}\mathcal{A}z = P^{-1}b$. The matrix P, called a preconditioner, is easy to invert and the preconditioned matrix $P^{-1}\mathcal{A}z$ has a good clustering behavior of the eigenvalues.

4.2 Literature Review

The systems obtained in optimization problems are huge and ill-conditioned. These properties complicate developing an efficient numerical algorithms. We know that using direct methods for solving such systems (1.3) and (1.4) requires $O(n^3)$ and hence they are not applicable. For these systems, iterative methods like Krylov subspace methods are applicable. However, their convergence is too slow because they are sensitive to the condition number. Hence the idea of preconditioning is needed to accelerate the convergence of the Krylov subspace methods. In the literature, many preconditioners in [1, 20] are developed for a saddle point problem. Preconditioning is a technique that uses a preconditioning matrix P to convert a linear system of the form $\mathcal{A}z = b$ into another system to improve the spectral properties of the system matrix. A preconditioner is a matrix P such that this matrix is easy to invert and the preconditioned matrix $P^{-1}\mathcal{A}$ has a good clustering behavior of the eigenvalues. Because rapid convergence is often associated with a clustered spectrum of $P^{-1}\mathcal{A}$. In the Preconditioning technique, we solve $P^{-1}\mathcal{A}z = P^{-1}b$ instead of solving the original one $\mathcal{A}z = b$ because the new system $P^{-1}\mathcal{A}z = P^{-1}b$ will converge rapidly when we use a suitable preconditioner. To apply the preconditioner matrix P within a Krylov subspace method, we need to compute a matrix times a vector of the form $x = Pr$ at each iteration. Hence, evaluating this product must be cheap. In the literature, several preconditioners [1] are developed for a special linear system such as block preconditioners and constraint preconditioners. For diagonal preconditioner we refer to Silvester and Wathen [21] & [22]. For the block triangular preconditioners we refer to Bramble and Pasciak [23] and also [21, 24–27] and the references therein. For the Constrain preconditioners, for example Axelsson and Neytcheva [28]. For the augmented Lagrangian preconditioners, Fortin and Glowinski [29]. Another preconditioners based on the Hermitian/skew-Hermitian splitting are studied in [30], [31], [30] and [32]. Other preconditioners are used for solving the systems (1.3) and/or (1.4). Sue Dolla et. al. [33] showed how new methods for the solution of saddle point system of the form (1.3) that arises in optimization problems which can be derived from the Bramble–Pasciak conjugate gradient method. Yvan Notay [34] use a matrix of symmetric saddle point system of the linear systems using the iterative solution method. He addressed several solution techniques which depends on the knowledge of a preconditioner. Also, there are many methods used to solve system of the form (1.4) Jennifer Pestana and Tyrone Rees [35] depend on incomplete versions of a particular null-space factorization to develop preconditioning. The study compared the equivalent Schur complement based preconditioners with their performance. Tyrone Rees Jennifer Scott [36] addressed the fitness of exploiting null-space factorizations to derive sparse direct methods. They attend numerical results for both academic and practical problems. Susanne Bradley [37] studied the iterative solution of symmetric saddle point problems with a rank deficient leading block. Two preconditioners were developed by the study such that, under certain assumptions on the rank structure of the system, yield a preconditioned matrix with a constant number of eigenvalues. Also, there are different methods for solving system of the form (1.6). Ron Estrin and Chen Greif [38] introduced a new way of saddle-point minimum residual methods. He used a minimum residual or quasi-minimum method to solve saddle-point problem.

In our study, we propose a suitable preconditioner and iterative method for solving (1.3) and/or (1.4).

4.3 Literature Review of Preconditioning Techniques

We adopt the general notation for saddle point systems that arise of using interior point to solve linear and quadratic problems and for saddle point systems that arise when solving least squares problems as:

$$\mathcal{A} = \begin{bmatrix} G & B^T \\ B & 0 \end{bmatrix} \tag{4.1}$$

where $G \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{m\times n}$ has full rank (rank$B = m$) and $n \geq m$.

to represent the saddle point matrices (2.9), (2.17), (2.24) and (2.35).

Remark 10 *As we saw in chapter 2 that G is symmetric positive semidefinite and assuming its nullity is s and B has full row rank. Where the assumption that $\mathcal{A}$ is nonsingular implies that $null(G) \cap null(B) = 0$, which we use in our analysis later.*

We start our discussion when G is nonsingular, there are two ideal block diagonal Schur complement preconditioners :

$$P_1^+ = \begin{bmatrix} G & 0 \\ 0 & BG^{-1}B^T \end{bmatrix} \textit{ and } P_1^- = \begin{bmatrix} G & 0 \\ 0 & -BG^{-1}B^T \end{bmatrix} \tag{4.2}$$

where $\pm BG^{-1}B^T$ is the Schur complement of G in $\mathcal{A}$. Considering (4.1) with (4.2) it has been shown that the preconditioned matrices

$$P^{+}{}_{1}^{-1}\mathcal{A} = \begin{bmatrix} I & G^{-1}B^T \\ (BG^{-1}B^T)^{-1}B & 0 \end{bmatrix} \; and \;\; P^{-}{}_{1}^{-1}\mathcal{A} = \begin{bmatrix} I & G^{-1}B^T \\ -(BG^{-1}B^T)^{-1}B & 0 \end{bmatrix} \tag{4.3}$$

are diagonalizable. $P^{+}{}_{1}^{-1}\mathcal{A}$ has only three distinct eigenvalues 1, $\frac{(1\pm\sqrt{5})}{2}$ which implies that preconditioned MINRES is expected to converge (in the absent of roundoff errors) within three iterations ([2, 3]). $P^{-}{}_{1}^{-1}\mathcal{A}$ has only three distinct eigenvalues 1, $\frac{(1\pm i\sqrt{3})}{2}$ implies that preconditioned GMRES is expected to converge (in the absent of roundoff errors) within three iterations also, there are two ideal block triangular preconditioners

$$P_2^{+} = \begin{bmatrix} G & 0 \\ B & BG^{-1}B^T \end{bmatrix} \; and \;\; P_2^{-} = \begin{bmatrix} G & 0 \\ B & -BG^{-1}B^T \end{bmatrix} \tag{4.4}$$

Considering (4.1) with (4.4) it has been shown that the preconditioned matrices

$$P^{+}{}_{2}^{-1}\mathcal{A} = \begin{bmatrix} I & G^{-1}B^T \\ 0 & -I \end{bmatrix} \; and \;\; P^{-}{}_{2}^{-1}\mathcal{A} = \begin{bmatrix} I & G^{-1}B^T \\ 0 & I \end{bmatrix} \tag{4.5}$$

have only two distinct eigenvalues $-1, 1$ and one distinct eigenvalue 1 respectively which implies that preconditioned GMERS is expected to converge within two and one iterations respectively (in the absent of roundoff errors), but the first one is diagonalizable, whereas the others not. [1–3, 39–42]. However, when G is singular(positive semidefinite) or ill-conditioned which is our studying focus as we mentioned in chapter 2 . Of course the Schur complement does not exist when G is singular in (4.1) because

G can not be inverted. So dealing with the systems by augmentation, i.e. by replacing G with $G + B^T W^{-1} B$ is one possible way, where $W \in \mathbb{R}^{m\times m}$ is a symmetric positive definite weight matrix.

The following block triangular preconditioner was considered

$$P_1 = \begin{bmatrix} G + B^T W^{-1} B & kB^T \\ 0 & W \end{bmatrix} \tag{4.6}$$

where W is $m \times m$ symmetric positive definite weight matrix and k is scalar.

The authors in [43] showed that $P_1^{-1}\mathcal{A}$ has eigenvalues $\lambda = 1$ with multiplicity $n-m$ and $\lambda = \frac{-k\pm\sqrt{4+k^2}}{2}$, each with multiplicity s. The remaining $2(m-s)$ eigenvalues satisfy the relation

$$\lambda = \frac{-(k\mu - 1) \pm \sqrt{(k\mu - 1)^2 + 4\mu(1 + \mu)}}{2(1 + \mu)},$$

where μ are the $2(m - s)$ positive generalized eigenvalues of $\mu G u = B^T W^{-1} B u$. In case $m = s$, only three distinct eigenvalues were obtained $\lambda = 1$ with multiplicity $n - m$ and $\lambda = \frac{-k\pm\sqrt{4+k^2}}{2}$, each with multiplicity m. If $k = 0$ (see [43, 44]), $P_1^{-1}\mathcal{A}$ has eigenvalues $\lambda = 1$ and $\lambda = -1$ with multiplicity n and s respectively, the remaining eigenvalues lie in the interval $(-1, 0)$ and satisfy the relation $\lambda = -\frac{\mu}{\mu+1}$ where μ are the $m-s$ positive generalized eigenvalues of $\mu G u = B^T W^{-1} B u$. Furthermore, in case $m = s$, only two distinct eigenvalues $\lambda = 1$ and $\lambda = -1$ with multiplicity n and m respectively. The authors in [43] showed if $k = -1$, $P_1^{-1}\mathcal{A}$ has eigenvalues $\lambda = 1$ with multiplicity $n - m$ and $\lambda = \frac{1\pm\sqrt{5}}{2}$, each with multiplicity s, the remaining eigenvalues

$2(m-s)$ lie in the interval $(\frac{1-\sqrt{5}}{2},0)\cup(1,\frac{1+\sqrt{5}}{2})$. Huang et al.[40] established the following preconditioner

$$P_2 = \begin{bmatrix} G+B^TW^{-1}B & (1-k)B^T \\ 0 & kW \end{bmatrix} \tag{4.7}$$

where $k \neq 0$ is a parameter, and

$$P_3 = \begin{bmatrix} G+kB^TW^{-1}B & kB^T \\ 0 & \frac{1-k}{k}W \end{bmatrix} \tag{4.8}$$

where $1 \neq k > 0$ is a parameter.

The authors in [40] showed that the preconditioned matrix $P_2^{-1}\mathcal{A}$ has the eigenvalues $\lambda_1 = 1$ with multiplicity n and $\lambda_2 = \frac{-1}{k}$ with multiplicity s. The remaining $m-s$ eigenvalues lie in the interval $(0,\frac{-1}{k})$ when $k<0$ or $(\frac{-1}{k},0)$ when $k>0$. The authors discussed the case $k=-1$, and showed that the matrix $P_2^{-1}\mathcal{A}$ has the eigenvalues $\lambda = 1$ with multiplicity $n+s$ and the remaining $m-s$ eigenvalues lie in the interval $(0,1)$. The preconditioner $P_3^{-1}\mathcal{A}$ has the eigenvalues $\lambda_1 = 1$ with multiplicity n and $\lambda_2 = \frac{1}{k-1}$ with multiplicity s. The remaining $m-s$ eigenvalues lie in the interval $(0,\frac{1}{k-1})$ when $k>1$ or $(\frac{1}{k-1},0)$ when $k<1$. The authors discussed the case $k=2$, and showed that the matrix $P_3^{-1}\mathcal{A}$ has the eigenvalues $\lambda = 1$ with multiplicity $n+s$ and the remaining $m-s$ eigenvalues lie in the interval $(0,1)$. Furthermore, the authors conclude that the proposed preconditioners are more efficient than those of Tim Rees and Chen Grief in [43]. Cui-Xia Li et al.[45] introduced two augmentation block triangular preconditioners

$$P_4 = \begin{bmatrix} G + kB^TW^{-1}B & (1+k)B^T \\ 0 & -W \end{bmatrix} \text{ and } P_5 = \begin{bmatrix} G + kB^TW^{-1}B & (1-k)B^T \\ 0 & W \end{bmatrix} \tag{4.9}$$

where k is assumed to be positive.

The authors showed that $P_4^{-1}\mathcal{A}$ has two eigenvalues $\lambda = 1$ and $\lambda = \frac{1}{k}$ with multiplicity n and s, respectively. The remaining $m - s$ eigenvalues satisfy the relation $\lambda = \frac{\mu}{k\mu+1}$ where μ are the $(m - s)$ positive generalized eigenvalues of $\mu G u = B^TW^{-1}Bu$. The authors also, discussed the case $k = 1$, and showed that the matrix $P_4^{-1}\mathcal{A}$ has eigenvalues $\lambda = 1$ with multiplicity $n + s$ and the remaining $m - s$ eigenvalues satisfy the relation $\lambda = \frac{\mu}{\mu+1}$. If null$(G) = m$ then $P_4^{-1}\mathcal{A}$ has one eigenvalue $\lambda = 1$ with multiplicity $n + m$. For the matrix $P_5^{-1}\mathcal{A}$ the discussion is similar to $P_4^{-1}\mathcal{A}$. Yuping Zeng, Chenliang Li citezeng2011new introduce the following preconditioner involving two parameters

$$P_{\eta,k} = \begin{bmatrix} G + \eta B^TW^{-1}B & (1-\eta k)B^T \\ 0 & kW \end{bmatrix} \tag{4.10}$$

where $\eta > 0$ and $k \neq 0$.

The authors showed that the preconditioned matrix $P_{\eta,k}^{-1}\mathcal{A}$ has the eigenvalues $\lambda_1 = 1$ with multiplicity n and $\lambda_2 = \frac{-1}{\eta k}$ with multiplicity s. The remaining $m{-}s$ eigenvalues lie in the interval $(0, \frac{-1}{\eta k})$ when $k < 0$ or $(\frac{-1}{\eta k}, 0)$ when $k > 0$. The authors discussed the case $\frac{1}{\eta k} = -1$, and showed that the matrix $P_{\eta,k}^{-1}\mathcal{A}$ has the eigenvalues

$\lambda = 1$ with multiplicity $n+s$ and the remaining $m{-}s$ eigenvalues lie in the interval $(0,1)$. Zhi-Hao Cao [46] considered a set of augmentation block preconditioners:

$$P_{j,k} = \begin{bmatrix} G + B^T W^{-1} B & jB^T \\ 0 & kW \end{bmatrix} \tag{4.11}$$

where $k, j \neq 0$.

The author showed that the preconditioned matrix $P_{j,k}^{-1}\mathcal{A}$ has the eigenvalues $\lambda = 1$ with multiplicity $n-m$ while the other eigenvalues discussed in the following cases

Case 1

When $k \neq \frac{-j^2}{4}$. Then $\lambda = \frac{-j \pm \sqrt{j^2+4k}}{2k}$, each with multiplicity s

If $k \neq 1-j$. The remaining of $2(m{-}s)$ eigenvalues satisfy the relation $\lambda = \frac{k - j\mu \pm \sqrt{(k-j\mu)^2 + 4k\mu(\mu+1)}}{2k(\mu+1)}$ where μ are the $m{-}s$ finite nonzero eigenvalues of the generalized eigenvalue problem $\mu G u = B^T W^{-1} B u$. If $k = 1-j$,then 1 is eigenvalue with multiplicity $m-s$ ($\lambda = 1$ with multiplicity $n - m + 2s + m - s$) which is$n+s$, the remaining $(m{-}s)$ eigenvalues satisfy the relation $\frac{-\mu}{(j-1)(\mu+1)}$

Case 2

When $k = \frac{-j^2}{4}$. Then $\frac{2}{j}$ is an eigenvalue of multiplicity $2s$

If $j \neq 2$, then the remaining $2(m{-}s)$ eigenvalues satisfy the relation $\lambda = \frac{4\mu + j \pm \sqrt{(j+4\mu)^2 - 16\mu(\mu+1)}}{2j(\mu+1)}$

If $j = 2$, then $\lambda = 1$ with multiplicity $n+s$. The remaining $(m{-}s)$ eigenvalues satisfy the relation $\lambda = \frac{\mu}{\mu+1}$ where μ as mentioned above. Then he discussed when $nullity(G) = m$ and got better results of all previous cases. Li-Tao Zhang [47] discussed the preconditioner matrix (4.10), but the author assumed $\eta \neq 0$ instead of

$\eta > 0$. The author showed that the preconditioned matrix $P_{\eta,k}^{-1}A$ at $\neq -1$ has the eigenvalues $\lambda_1 = 1$ with multiplicity n and $\lambda_2 = \frac{-1}{\eta k}$ with multiplicity s. The remaining $m{-}s$ eigenvalues lie in the interval $(0, \frac{-1}{\eta k})$ when $\eta k < 0$ or $(\frac{-1}{\eta k}, 0)$ when $\eta k > 0$ and discussed the case $m = s$ and got $\lambda_1 = 1$ with multiplicity n and $\lambda_2 = \frac{-1}{\eta k}$ with multiplicity m . The author also, discussed the case $\eta k = -1$ and showed that the matrix $P_{\eta,k}^{-1}A$ has the eigenvalues $\lambda = 1$ with multiplicity $n + s$ and the remaining $m{-}s$ eigenvalues lie in the interval $(0, 1)$ and the author discussed the case $m = s$ and got $\lambda_1 = 1$ with multiplicity n and $\lambda_2 = \frac{-1}{\eta k}$ with multiplicity m. He conclude that When $\eta k = -1$, then preconditioner $P_{\eta,k}A$ is the optimal in the augmentation block preconditioner set $P_{\eta,k}^{-1}A$ where $\eta, k \neq 0$ real. If $\eta k = -1$ (which is the optimal parameter) $\lambda = 1$ with multiplicity $n + m$ at the end he conclude at the optimalm parameter of $P_{\eta,k}$ $(\eta k = -1)$ and $P_{j,k}$ (j=2,k=-1) have the same spectral clustering. Qingbing Liu [48] discussed two preconditioners

$$P_6 = \begin{bmatrix} G + B^T W^{-1} B & jB^T \\ 0 & -W \end{bmatrix} \quad and \quad P_7 = \begin{bmatrix} G + B^T W^{-1} B & jB^T \\ 0 & (1-j)W \end{bmatrix} \tag{4.12}$$

The preconditioner matrix P_7 discussed above by letting $k = 1{-}j$ in (4.11) and showed that the preconditioned matrix $P_6^{-1}A$ has the eigenvalues $\lambda = 1$ with multiplicity $n - m$ and $\lambda = \frac{j \pm \sqrt{j^2-4}}{2}$, each with multiplicity s. The remaining $2(m{-}s)$ eigenvalues satisfy the relation $\lambda = \frac{(1+j\mu) \pm \sqrt{(1+j\mu)^2 - 4\mu(\mu+1)}}{2(\mu+1)}$ where μ as mentioned in all the above preconditioners. The author considered the case $j = 1$ and showed that $\lambda = 1$ with multiplicity $n - m$ and $\lambda = \frac{1 \pm i\sqrt{3}}{2}$, each with multiplicity s. The remaining $2(m{-}s)$

eigenvalues satisfy the relation $\lambda = \frac{1}{2} \pm \frac{\sqrt{-3\mu^2-2\mu+1}}{2(\mu+1)}$. Since the $\lambda = \frac{j \pm \sqrt{j^2-4}}{2}$ with multiplicity $2s$ are unbounded as $j \to \infty$, he concluded that j should be of moderate size and they conclude that $j = 2$ P_6, P_7 are the same and they compare their studying with Tim Rees and Chen Grief [43] and got better results. Yi-Fen Ke ect. [49] have introduced augmentation block triangular preconditioners

$$P_{t,j,k} = \begin{bmatrix} G + tB^TW^{-1}B & jB^T \\ 0 & kW \end{bmatrix} \tag{4.13}$$

and showed that $P^{-1}_{t,j,k}A$ for any t, j, k ($tk \neq 0$) the preconditioned matrix $P^{-1}_{t,j,k}A$ has eigenvalue $\lambda = 1$ of multiplicity $n{-}m$: when $j^2 + 4tk \neq 0$ and $tk + j = 1$ (that is, $kt \neq -1$ and$j \neq 2$), the matrix $P^{-1}_{t,1-tk,k}A$ has eigenvalue$\lambda = 1$ of multiplicity n $(n - m + s + m - s)$, $\lambda = \frac{-1}{kt}$ with multiplicity s and the remaining $(m{-}s)$ eigenvalues satisfy the relation $\lambda = \frac{-\mu}{k(t\mu+1)}$. When $j^2 + 4tk = 0$ and $j = 2$(that is, $tk = -1$), the matrix $P^{-1}_{t,2,k}A$ has eigenvalue $\lambda = 1$ of multiplicity $n + s$ and the remaining $(m{-}s)$ eigenvalues satisfy the relation $\lambda = \frac{\mu}{k-\mu}$, very strong spectral clustering and he showed when $null(G) = m$, the preconditioned matrix $P^{-1}_{t,j,k}A$ has exactly one eigenvalue $\lambda = 1$ of multiplicity $n{+}m$. Therefore, they concluded tha $j = 2$ and $tk = -1$ are the optimal parameters, and he introduced The preconditioner matrix

$$P_{t,2,\frac{-1}{t}} = \begin{bmatrix} G + tB^TW^{-1}B & jB^T \\ 0 & \frac{-1}{t}W \end{bmatrix} \tag{4.14}$$

when $j = 2$ is the strong cluster for any $t \neq 0$. Litao Zhang [50] introduced a new preconditioner matrix :

$$P_8 = \begin{bmatrix} G + kB^T W^{-1} B & (1-k)B^T \\ 0 & W \end{bmatrix} \tag{4.15}$$

and discussed the same as the discussion in P_5 in (4.9), but he assumed that $k \neq 0$. Jun He, Ting-Zhu Huang [51] introduced different augmentation preconditioner from all the previously preconditioners as the following

$$P_9 = \begin{bmatrix} G + B^T W^{-1} B & B^T \\ B & -W \end{bmatrix} \tag{4.16}$$

and showed that the preconditioned matrices $P_9^{-1} A$ has the eigenvalues $\lambda = 1$ with multiplicity $n - m$ and $\lambda = \frac{1 \pm i}{2} (i = \sqrt{-1})$, each with multiplicity s. The remaining $2(m-s)$ eigenvalues satisfy the relation $\lambda = \frac{1}{2} \pm \frac{1}{2}\sqrt{\frac{1-2\mu}{1+2\mu}}$. Cui-Xia Li et al.[52] introduced a new preconditioner matrix

$$P_{10} = \begin{bmatrix} G + B^T W^{-1} B & (1+\sqrt{2})B^T \\ (1-\sqrt{2})B & -2W \end{bmatrix} \tag{4.17}$$

to improved (4.16), and showed that the preconditioned matrix $P_{10}^{-1} A$ has the eigenvalues $\lambda = 1$ with multiplicity $n - m + 2s$. The remaining $2(m-s)$ eigenvalues satisfy the relation $\lambda = \frac{\mu}{\mu+2}$ which we can write it in anther form the preconditioned matrices $P^{-1} A$ has the eigenvalues $\lambda = 1$ with multiplicity $n + s$. The remaining $(m-s)$ eigenvalues satisfy the relation $\lambda = \frac{\mu}{\mu+2}$ and we can change the number 2 by any constant

t with difference results.

4.4 Main Results

In this section, we consider the following preconditioner

$$P = \begin{bmatrix} G + \frac{2k}{k+1} B^T W^{-1} B & \left(1 + \dfrac{2k}{k+1}\right) B^T \\ 0 & -W \end{bmatrix} \tag{4.18}$$

where $k \geq 1$.

Theorem 4.1 *Assume that $\mathcal{A}$ is nonsingular and its* (1.4) *block G is symmetric positive semidefinite with nullity s. Then the preconditioned matrix $P^{-1}\mathcal{A}$ has two distinct eigenvalues which are given by*

$$\lambda = 1 \quad \textit{and} \quad \lambda = \frac{k+1}{2k} \tag{4.19}$$

with algebraic multiplicity n and s, respectively. The remaining $m - s$ eigenvalues satisfy the relation

$$\lambda = \frac{\mu}{\left(\left(\frac{2k}{k+1}\right)\mu + 1\right)} \tag{4.20}$$

where μ is a positive generalized eigenvalue of

$$\mu G u = B^T W^{-1} B u \tag{4.21}$$

Let $\{x_i\}_{i=1}^{n-m}$ be a basis of the null space of B, let $\{y_i\}_{i=1}^{s}$ be a basis of the null space of G, *and* $\{z_i\}_{i=1}^{m-s}$ *be a set of linearly independent vectors that complete* $null(G) \cup null(B)$ *to a basis of* R^n. *Then the* $n-m$ *vectors* $[x_i^T, 0^T]^T$ $(i = 1, 2, ..., n-m)$*, the* s *vectors* $[y_i^T, -(W^{-1}By_i)^T]^T (i = 1, ..., s)$*, and the* $m-s$ *vectors* $[z_i^T, -(W^{-1}Bz_i)^T]^T$ $(i = 1, 2, ..., m-s)$ *are linearly independent eigenvectors associated with* $\lambda = 1$ *and the* s *vectors* $[y_i^T, -\frac{2k}{k+1}(W^{-1}By_i)^T]^T (i = 1, ..., s)$ *are linearly independent eigenvectors associated with* $\lambda = \frac{(k+1)}{2k}$.

Proof. Let λ be an eigenvalue of $P^{-1}\mathcal{A}$ with eigenvector $[u^T, v^T]^T$. So, we have

$$P^{-1}\mathcal{A} \begin{bmatrix} u \\ v \end{bmatrix} = \lambda \begin{bmatrix} u \\ v \end{bmatrix} \tag{4.22}$$

Moreover, equation (4.22) satisfies the generalized eigenvalue problem

$$\begin{bmatrix} G & B^T \\ B & 0 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \lambda \begin{bmatrix} G + \frac{2k}{k+1} B^T W^{-1} B & \left(1 + \dfrac{2k}{k+1}\right) B^T \\ 0 & -W \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} \tag{4.23}$$

or

$$Gu + B^T v = \lambda \left(G + \frac{2k}{k+1} B^T W^{-1} B\right) u + \lambda \left(1 + \frac{2k}{k+1}\right) B^T v \tag{4.24}$$

$$Bu = -\lambda W v \tag{4.25}$$

Because $\mathcal{A}$ is nonsingular, it follows that $\lambda \neq 0$. Substituting for $v = \frac{-1}{\lambda} W^{-1} Bu$ from (4.25) in (4.24) we obtain

$$(\lambda^2 - \lambda) Gu + \left[\left(\frac{2k}{k+1}\right) \lambda^2 - \left(1 + \frac{2k}{k+1}\right) \lambda + 1\right] B^T W^{-1} Bu = 0 \tag{4.26}$$

which we can write as

$$(\lambda - 1)\left[\lambda G u + \left(\lambda\left(\frac{2k}{k+1}\right) - 1\right) B^T W^{-1} B u\right] = 0 \tag{4.27}$$

If $\lambda = 1$, then equation (4.27) is satisfied for any nonzero vector $u \in R^n$, and from equation (4.25) $(u, -W^{-1}Bu)$ is an eigenvector of $P^{-1}\mathcal{A}$. To show that $\lambda = 1$ has algebraic multiplicity n, let $u \in \text{null}(G)$ to get, from equation (4.27)

$$(\lambda - 1)\left(\lambda\left(\frac{2k}{k+1}\right) - 1\right) B^T W^{-1} B u = 0 \tag{4.28}$$

since $B^T W^{-1} B$ is symmetric positive definite then $B^T W^{-1} B > 0$ which implies that $\lambda = \frac{k+1}{2k}$ or $\lambda = 1$. The eigenvalue $\lambda = \frac{k+1}{2k}$ has $(u, -\frac{2k}{k+1} W^{-1} B u)$ as an eigenvector. The eigenvalue $\lambda = 1$ has $(u, -W^{-1}Bu)$ as an eigenvector. This shows that both eigenvalues have algebraic multiplicity s.

If $u \in \text{null}(B)$, then (4.27) gives

$$(\lambda - 1)\lambda G u = 0. \tag{4.29}$$

The nonzero vector $u \notin \text{null}(G)$ because $\text{rank} B = m$ and $\text{null}(G) \cap \text{null}(B) = 0$. Therefore $Gu \neq 0$ and $(\lambda - 1)\lambda = 0$. It follows that $(u, -W^{-1}Bu)$ is an eigenvector associated with $\lambda = 1$. This shows that the eigenvalue $\lambda = 1$ has algebraic multiplicity $n - m$.

Next, we consider the remaining $2(m - s)$ eigenvalues; so we assume $\lambda \neq \frac{k+1}{2k}$. From equation (4.26), we find that by letting $r = \frac{2k}{k+1}$

$$(\lambda^2 - \lambda)Gu + \left[r\lambda^2 - (1+r)\,\lambda + 1\right] B^TW^{-1}Bu = 0 \tag{4.30}$$

from the above equation

$$\frac{\lambda - \lambda^2}{\left[r\lambda^2 - (1+r)\,\lambda + 1\right]} = B^TW^{-1}Bu. \tag{4.31}$$

Since $u \notin \text{null}(G)$ and $u \notin \text{null}(B)$, then $u^TB^TW^{-1}Bu > 0$ and therefore $u^TGu > 0$. Hence $\frac{\lambda-\lambda^2}{[r\lambda^2-(1+r)\lambda+1]} > 0$. Let

$$\frac{\lambda - \lambda^2}{\left[r\lambda^2 - (1+r)\,\lambda + 1\right]} = \mu \quad \text{for some } \mu > 0.$$

Solving for λ,

$$\lambda = \frac{(1+\mu(1+r)) \pm \sqrt{(1+\mu(1+r))^2 - 4\mu(1+\mu r)}}{2(1+\mu r)} \tag{4.32}$$

then $\lambda_1 = 1$ and $\lambda_2 = \frac{\mu}{(1+r\mu)}$, each with multiplicity $m - s$

We therefore get $\lambda \in (0, \frac{k+1}{2k})$ as $\mu \to 0, \infty$. Let $\{x_i\}_{i=1}^{n-m}$ be a basis of the null space of B, and let $\{y_i\}_{i=1}^{s}$ be a basis of the null space of G. Because $\text{null}(B) \cap \text{null}(G) = \{0\}$, the vectors $\{x_i\}_{i=1}^{n-m}$ and $\{y_i\}_{i=1}^{s}$ are linearly independent and together span the subspace $\text{null}(B) \cup \text{null}(G)$. Let the vector $\{z_i\}_{i=1}^{m-s}$ complete $\text{null}(G) \cup \text{null}(B)$ to a basis of R^n. It follows the vectors $[x_i^T, 0^T]^T$ $(i = 1, 2, ...n-m)$, the vectors $[y_i^T, -(W^{-1}By_i)^T]^T(i = 1, ..., s)$, and the vectors $[z_i^T, -(W^{-1}Bz_i)^T]^T$

$(i = 1, 2, ..., m - s)$, are linearly independent eigenvectors associated to $\lambda = 1$ and the vectors $[{y_i}^T, -\frac{(2k)}{k+1}(W^{-1}By_i)^T]$ are linearly independent eigenvectors associated with $\lambda = \frac{(k+1)}{2k}$. ∎

Corollary 1 *Let $k = 1$, then the matrix $P^{-1}\mathcal{A}$ has one eigenvalue which is given by $\lambda = 1$ with algebraic multiplicity $n + s$. The remaining $m - s$ eigenvalues satisfy the relation $\lambda = \frac{\mu}{\mu+1}$ where μ are some positive generalized eigenvalue of $\mu Gu = B^T W^{-1} Bv$.*

Corollary 2 *If $null(G) = m$, then the matrix $P^{-1}\mathcal{A}$ has two eigenvalues which are given by $\lambda = 1$ with algebraic multiplicity n and $\lambda = \frac{k+1}{2k}$ with algebraic multiplicity s.*

Corollary 3 *If $null(G) = m$ and $k = 1$, then the matrix $P^{-1}\mathcal{A}$ has one eigenvalues which is given by $\lambda = 1$ with algebraic multiplicity $n + m$.*

CHAPTER 5

NUMERICAL EXPERIMENTS

In this chapter, we report numerical experiments to illustrate the performance of $P^{-1}\mathcal{A}$ with our preconditioner matrix P (4.18) and $W = \gamma I_m$. We compare $P^{-1}\mathcal{A}$ at new positive parameter $\gamma_1 = \frac{\|B\|_2^2}{\|G\|_2 max(G/c)}$ or $\gamma_2 = \frac{\|B\|_1^2}{\|G\|_1 max(G/c)}$ (where c is constant) with $P^{-1}\mathcal{A}$ at $\gamma_3 = \frac{\|B\|_2^2}{\|G\|_2}$ or $\gamma_4 = \frac{\|B\|_1^2}{\|G\|_1}$ as discussed in the literature.

We consider the following version of the CUTEst (constrained and unconstrained testing environment for numerical optimization.) matrix CVXQP1 (quadratic programming example)

$$\mathcal{A} = \begin{bmatrix} G & B^T \\ B & 0 \end{bmatrix}, \tag{5.1}$$

where $G \in \mathbb{R}^{1000\times1000}$ is symmetric positive semidefinite and $B \in \mathbb{R}^{500\times1000}$ has full rank.

This system is obtained in solving the quadratic programming problem by interior point method, (2.19). From the matrix $\mathcal{A}$ in (5.1) we construct the following four saddle point-type matrices

$$\mathcal{A}_i = \begin{bmatrix} G_i & B^T \\ B & 0 \end{bmatrix}, \quad i = 1, 2, 3, 4 \tag{5.2}$$

where G_i is constructed from G in two ways, the first case is constructed by making its first $i \times m/4,\ i = 1, 2, 3, 4$ columns and rows equal to zero entries. Note that G_i is symmetric positive semidefinite real and its nullity is $i \times m/4$. Since our new preconditioner matrix P works with any nullity with the new positive parameters $\gamma_1 = \frac{\|B\|_2^2}{\|G\|_2 max(G/c)}$ and $\gamma_2 = \frac{\|B\|_1^2}{\|G\|_1 max(G/c)}$ we can construct the second case for G_i from G by making its first $i,\ i = 1, ..., m$ columns and rows equal to zero entries. Note that G_i is symmetric positive semidefinite real and its nullity is i.

We use GMRES, BICGSTAB and MINRES methods to solve the following four saddle point-type systems:

$$\mathcal{A}_i z = b_i, \quad i = 1, 2, 3, 4 \tag{5.3}$$

We also take $i = 1$ and $i = 10$ from the second case where the right-hand side b_i in the two cases is taken such that the solution z is all ones. The stopping criterion is $1e^{-12}$ for the relative residual norm $\frac{\|\mathcal{A}z-b\|}{\|b\|} \leq 10^{-12}$, where the initial guess is zero, which is the default of GMRES, BICGSTAB and MINRES.

In our experiment, we choose the value of c as the following:

1. At the parameter ($k \geq 1$), if G_i has the nullity m then, c can be taken between 1 and 100 with γ_1 (in our experiment we set $c = 10$).

2. At the optimal parameter ($k = 1$) and the nullity of G is close to m, then it is

recommended to take γ_1 or γ_2 and c between 10 and 40.

3. At the optimal parameter ($k = 1$) and G is almost less than 5% of m, then it is recommended to take γ_1 and c between 1 and 5 (when i is less than 167 we take $c = 1$ in our work).

4. At $k > 1$ do the same case 2 and case 3. However, when i is less than 200 we take $c = 1.5$.

In Table 5.1, we use $\gamma_4 = \frac{\|B\|_1^2}{\|G\|_1}$ in our preconditioner matrix P for $G1, G2, G3, G4$ with restart 10.

k	$Nullity$	$It_{GMRES(10)(P^{-1}\mathcal{A})}$	$Res_{GMRES(10)(P^{-1}\mathcal{A})}$	$It_{BICGSTAB(P^{-1}\mathcal{A})}$	$Res_{BICGSTAB(P^{-1}\mathcal{A})}$
1	500	1(4)	1.8×10^{-17}	2.5	9.6×10^{-13}
1	375	289(10)	9.9×10^{-13}	301	9.6×10^{-13}
1	250	1236(6)	1×10^{-12}	non conv. to this tol.	—
1	125	non conv. to this tol.	—	460.5	9.7×10^{-13}
100	500	1(4)	2.7×10^{-17}	3	2×10^{-13}
100	375	254(10)	1×10^{-12}	619.5	9.8×10^{-13}
100	250	981(10)	1×10^{-12}	non conv. to this tol.	—
100	125	1231(3)	1×10^{-12}	491.5	7.3×10^{-13}

Table 5.1: Iteration numbers of GMRES(10) and BICGSTAB of preconditioned matrix $P^{-1}\mathcal{A}$ with γ_4 at $k = 1, 100$ for $G1, G2, G3, G4$

In Table 5.2, we use $\gamma_1 = \frac{\|B\|_2^2}{\|G\|_2 max(G/c)}$ in our preconditioner matrix P for $G1, G2, G3, G4$ with restart 10.

k	$Nullity$	$It_{GMRES(10)(P^{-1}A)}$	$Res_{GMRES(10)(P^{-1}\mathcal{A})}$	$It_{BICGSTAB(P^{-1}A)}$	$Res_{BICGSTAB(P^{-1}\mathcal{A})}$	c
1	500	1(4)	6.7×10^{-15}	4	2.6×10^{-14}	10
1	375	2(4)	4.8×10^{-13}	8	4.1×10^{-13}	4
1	250	2(4)	9.3×10^{-13}	8.5	1.9×10^{-13}	2
1	125	2(4)	7.3×10^{-13}	8.5	9.3×10^{-14}	1
100	500	1(4)	1.7×10^{-14}	4	3×10^{-14}	10
100	375	1(10)	1.5×10^{-13}	8.5	1.4×10^{-13}	4
100	250	1(10)	8.5×10^{-14}	8.5	9×10^{-13}	2
100	125	2(5)	1.5×10^{-13}	9.5	4.7×10^{-13}	2

Table 5.2: Iteration numbers of GMRES(10) and BICGSTAB of preconditioned matrix $P^{-1}\mathcal{A}$ with γ_1 at $k = 1, 100$ for $G1, G2, G3, G4$

Comparing Table 5.1 and Table 5.2, whenever $k = 1$ or $k = 100$, we notice that the number of iterations in Table 5.2 is better than the number of iterations in Table 5.1 except when the value of nullity of G is full nullity (nullity(G)=500) . However, the value at the nullity $(G) = 500$ shown in Table 5.2 is still very close to the value at the nullity(G)=500 shown in table 5.1. Therefore, our work gives good results for any nullity.

Remark 11 *We apply $\gamma_4 = \frac{\|B\|_1^2}{\|G\|_1}$ that discussed in the literature with tolerance $e - 12$ to our preconditioner matrix P for $i = 1$ and $i = 10$ at $k = 1, 100$ but no solution (convergence) found in this case.*

In Table 5.3, we use $\gamma_1 = \frac{\|B\|_2^2}{\|G\|_2 max(G/c)}$ in our preconditioner matrix P for $i = 1$ and $i = 10$ at $k = 1, 100$.

k	$Nullity$	$It_{GMRES(10)(P^{-1}\mathcal{A})}$	$Res_{GMRES(10)(P^{-1}\mathcal{A})}$	$It_{BICGSTAB(P^{-1}\mathcal{A})}$	$Res_{BICGSTAB(P^{-1}\mathcal{A})}$	c
1	i=1	2(7)	9.5×10^{-14}	9	2.3×10^{-13}	1
1	i=10	2(7)	9.6×10^{-14}	8.5	9×10^{-13}	1
100	i=1	2(5)	6.3×10^{-13}	10	8.6×10^{-14}	1.5
100	i=10	2(5)	4×10^{-13}	9	7.4×10^{-13}	1.5

Table 5.3: Iteration numbers of GMRES(10) and BICGSTAB of preconditioned matrix $P^{-1}\mathcal{A}$ with γ_1 at $k = 1, 100$ for $i = 1$ and $i = 10$

Comparing Remark (11) with Table 5.3, there is a convergence whenever $k = 1$ or $k = 100$, but when using Remark (11) for preconditioner matrix P when $i = 1$ and $i = 10$ at $k = 1, 100$, no convergence was found.

In Table 5.4, we use $\gamma_4 = \frac{\|B\|_1^2}{\|G\|_1}$ in our preconditioner matrix P for $G1, G2, G3, G4$ with restart 500.

k	$Nullity$	$It_{GMRES(500)(P^{-1}\mathcal{A})}$	$Res_{GMRES(500)(P^{-1}\mathcal{A})}$	$It_{BICGSTAB(P^{-1}\mathcal{A})}$	$Res_{BICGSTAB(P^{-1}\mathcal{A})}$
1	500	1(4)	1.8 $\times 10^{-17}$	2.5	9.6 $\times 10^{-13}$
1	375	1(154)	6.1 $\times 10^{-13}$	301	9.6 $\times 10^{-13}$
1	250	1(293)	5.9 $\times 10^{-13}$	non conv. to this tol.	—
1	125	1(436)	3.5 $\times 10^{-14}$	460.5	9.7 $\times 10^{-13}$
100	500	1(4)	2.7 $\times 10^{-17}$	3	2 $\times 10^{-13}$
100	375	1(152)	2.6 $\times 10^{-16}$	619.5	9.8 $\times 10^{-13}$
100	250	1(289)	1.1 $\times 10^{-13}$	non conv. to this tol.	—
100	125	1(423)	6.7 $\times 10^{-13}$	491.5	7.3 $\times 10^{-13}$

Table 5.4: Iteration numbers of GMRES(500) and BICGSTAB of preconditioned matrix $P^{-1}\mathcal{A}$ with γ_4 at $k = 1, 100$ for $G1, G2, G3, G4$

In Table 5.5 we use $\gamma_1 = \frac{\|B\|_2^2}{\|G\|_2 max(G/c)}$ in our preconditioner matrix P for $G1, G2, G3, G4$ with restart 500.

k	$Nullity$	$It_{GMRES(500)(P^{-1}\mathcal{A})}$	$Res_{GMRES(500)(P^{-1}\mathcal{A})}$	$It_{BICGSTAB(P^{-1})\mathcal{A}}$	$Res_{BICGSTAB(P^{-1}\mathcal{A})}$	c
1	500	1(4)	6.7 $\times 10^{-15}$	4	2.6 $\times 10^{-14}$	10
1	375	1(11)	3.6 $\times 10^{-13}$	8	4.1 $\times 10^{-13}$	4
1	250	1(11)	2.3 $\times 10^{-13}$	8.5	1.9 $\times 10^{-13}$	2
1	125	1(13)	9.5 $\times 10^{-14}$	10	2 $\times 10^{-13}$	2
100	500	1(4)	1.7 $\times 10^{-14}$	4	3 $\times 10^{-14}$	10
100	375	1(10)	1.5 $\times 10^{-13}$	8.5	1.4 $\times 10^{-13}$	4
100	250	1(10)	8.5 $\times 10^{-14}$	8.5	9 $\times 10^{-13}$	2
100	125	1(11)	3.2 $\times 10^{-13}$	9.5	4.7 $\times 10^{-13}$	2

Table 5.5: Iteration numbers of GMRES(500) and BICGSTAB of preconditioned matrix $P^{-1}\mathcal{A}$ with γ_1 at $k = 1, 100$ for$G1, G2, G3, G4$

Comparing Table 5.4 and Table 5.5, whenever $k = 1$ or $k = 100$, we notice that the number of iterations in Table 5.5 is better than the number of iterations in Table 5.4 except when the value of nullity of G is full nullity (nullity(G)=500). However, the value at nullity(G)=500 shown in Table 5.5 is still very close to the value at nullity(G)=500 shown in table 5.1. Therefore, our work gives good results for any nullity.

In Table 5.6, we use $\gamma_4 = \frac{\|B\|_1^2}{\|G\|_1}$ in our preconditioner matrix P for $i = 1$ and $i = 10$ at $k = 1, 100$ with restart 500.

k	$Nullity$	$It_{GMRES(500)(P^{-1}\mathcal{A})}$	$Res_{GMRES(500)(P^{-1}\mathcal{A})}$	$It_{BICGSTAB(P^{-1}\mathcal{A})}$	$Res_{BICGSTAB(P^{-1}\mathcal{A})}$
1	i=1	2(63)	9.9 $\times 10^{-13}$	508.5	8.8 $\times 10^{-13}$
1	i=10	2(83)	1 $\times 10^{-12}$	465.5	9.6 $\times 10^{-13}$
100	i=1	1(488)	9.5 $\times 10^{-13}$	523.5	9.5 $\times 10^{-13}$
100	i=10	1(495)	6.4 $\times 10^{-13}$	non conv. to this tol.	—

Table 5.6: Iteration numbers of GMRES(500) and BICGSTAB of preconditioned matrix $P^{-1}A$ with γ_4 at $k = 1, 100$ for $i = 1$ and $i = 10$

In Table 5.7, we use $\gamma_1 = \frac{\|B\|_2^2}{\|G\|_2 max(G/c)}$ in our preconditioner matrix P for $i = 1$ and $i = 10$ at $k = 1, 100$ with restart 500.

k	$Nullity$	$It_{GMRES(500)(P^{-1}\mathcal{A})}$	$Res_{GMRES(500)(P^{-1}\mathcal{A})}$	$It_{BICGSTAB(P^{-1}\mathcal{A})}$	$Res_{BICGSTAB(P^{-1}\mathcal{A})}$	c
1	i=1	1(14)	3.4 $\times 10^{-14}$	10.5	1.1 $\times 10^{-13}$	2
1	i=10	1(14)	3.1 $\times 10^{-14}$	10.5	3.1 $\times 10^{-13}$	2
100	i=1	1(13)	3.7 $\times 10^{-13}$	11	5.9 $\times 10^{-13}$	3
100	i=10	1(13)	2.1 $\times 10^{-13}$	11	2.3 $\times 10^{-13}$	3

Table 5.7: Iteration numbers of GMRES(500) and BICGSTAB of preconditioned matrix $P^{-1}\mathcal{A}$ with γ_1 at $k = 1, 100$ for $i = 1$ and $i = 10$

Comparing Table 5.6 and Table 5.7, whenever $k = 1$ or $k = 100$, we notice that the number of iterations in Table 5.7 is better than the number of iterations in Table 5.6.

In Table 5.8 and Table 5.9 we use BICGSTAB method with $\gamma_1 = \frac{\|B\|_2^2}{\|G\|_2 max(G/c)}$ in our preconditioner matrix P with $c = 1$ for any nullity at $k = 1$ and $c = 1.5$ for any $k > 1$ for any nullity.

k	$Nullity$	$It_{BICGSTAB(P^{-1}\mathcal{A})}$	$Res_{BICGSTAB(P^{-1}\mathcal{A})}$	c
1	500	4	2.6 $\times 10^{-14}$	10
1	490	4	4.4 $\times 10^{-15}$	1
1	430	6	9.7 $\times 10^{-14}$	1
1	350	6.5	1.2 $\times 10^{-13}$	1
1	300	7	2 $\times 10^{-13}$	1
1	260	7.5	7.4 $\times 10^{-14}$	1
1	150	8.5	7.2 $\times 10^{-14}$	1
1	70	8.5	3 $\times 10^{-13}$	1
1	30	9	9.5 $\times 10^{-14}$	1
1	7	9	1.9 $\times 10^{-13}$	1
1	1	9	2.3 $\times 10^{-13}$	1
1	0	9	2.4 $\times 10^{-13}$	1

Table 5.8: Iteration numbers of BICGSTAB of preconditioned matrix $P^{-1}\mathcal{A}$ with γ_1 at $k = 1$ for different random nullity.

k	Nullity	$It_{BICGSTAB(P^{-1}\mathcal{A})}$	$Res_{BICGSTAB(P^{-1}\mathcal{A})}$	c	$It_{BICGSTAB(P^{-1}\mathcal{A})}$	$Res_{BICGSTAB(P^{-1}\mathcal{A})}$	c
100	500	5	8.9 $\times 10^{-13}$	1	5	3.8 $\times 10^{-14}$	1.5
100	490	5	2.1 $\times 10^{-13}$	1	5	6.2 $\times 10^{-15}$	1.5
1000	430	6.5	3.2 $\times 10^{-13}$	1	6.5	4.1 $\times 10^{-13}$	1.5
2000	350	7.5	1.1 $\times 10^{-13}$	1	7.5	1.1 $\times 10^{-13}$	1.5
5000	300	7.5	6 $\times 10^{-13}$	1	8	4.9 $\times 10^{-14}$	1.5
10000	260	7.5	4.7 $\times 10^{-13}$	1	8	6.5 $\times 10^{-14}$	1.5
10000	150	8	1.2 $\times 10^{-13}$	1	8.5	3.4 $\times 10^{-13}$	1.5
20000	70	8.5	3.7 $\times 10^{-13}$	1	9	9.7 $\times 10^{-13}$	1.5
90000	30	9	3.2 $\times 10^{-14}$	1	10	6.5 $\times 10^{-14}$	1.5
100000	7	8	3.7 $\times 10^{-13}$	1	10	3 $\times 10^{-14}$	1.5
100000	1	8.5	8.2 $\times 10^{-13}$	1	9.5	8.8 $\times 10^{-13}$	1.5
1000000	0	8.5	5.6 $\times 10^{-13}$	1	9.5	6.1 $\times 10^{-13}$	1.5

Table 5.9: Iteration numbers of BICGSTAB of preconditioned matrix $P^{-1}\mathcal{A}$ with γ_1 for different random k and different random nullity

In Table 5.10, we use γ_4 and γ_1 in preconditioner matrix P_0 for different nullity. Now we apply our new γ_1 in P_0

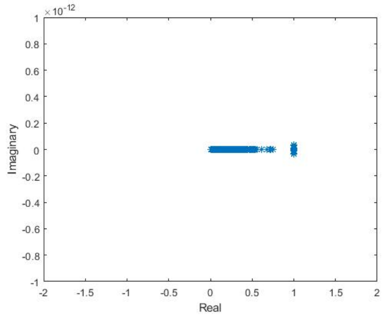
$$P_0 = \begin{bmatrix} G + B^T W^{-1} B & 0 \\ 0 & W \end{bmatrix} \tag{5.4}$$

to find the number of iteration of $P_0^{-1}\mathcal{A}$ and compare it with $P_0^{-1}\mathcal{A}$ at γ_4 as discussed in the literature.

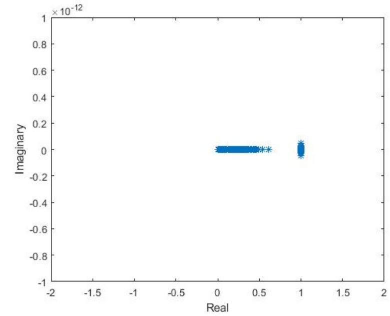
Nullity	$It_{MINERS(P_0^{-1}\mathcal{A})}$ at γ_4	$Res_{MINERS(P_0^{-1}\mathcal{A})}$ at γ_4	$It_{MINERS(P_0^{-1}\mathcal{A})}$ at γ_1	$Res_{MINERS(P_0^{-1}\mathcal{A})}$ at γ_1
500	2	2.7 $\times 10^{-13}$	5	7.6 $\times 10^{-14}$
500	2	2.7 $\times 10^{-13}$	4 at c=10	4.1 $\times 10^{-14}$
375	220	9.7 $\times 10^{-13}$	7	2 $\times 10^{-14}$
250	403	8.1 $\times 10^{-13}$	8	9.7 $\times 10^{-16}$
125	528	6.6 $\times 10^{-13}$	8	9.9 $\times 10^{-15}$
10	715	8 $\times 10^{-13}$	8	3.3 $\times 10^{-14}$
1	729	6.8 $\times 10^{-13}$	8	4.7 $\times 10^{-14}$
0	722	9 $\times 10^{-13}$	8	4.8 $\times 10^{-14}$

Table 5.10: Iteration numbers of MINERS of preconditioned matrix $P_0^{-1}\mathcal{A}$ at γ_4 and γ_1 (c=.01)

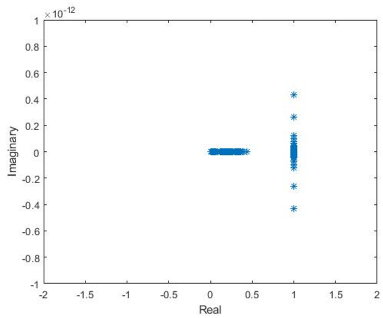
In Figure 5.1 and Figure 5.2, we use $\gamma_4 = \frac{\|B\|_1^2}{\|G\|_1}$ in our preconditioner matrix P for $G1, G2, G3, G4$ to plot the distribution eigenvalues of $P^{-1}\mathcal{A}$ at $k = 1$ and $k = 100$, respectively.



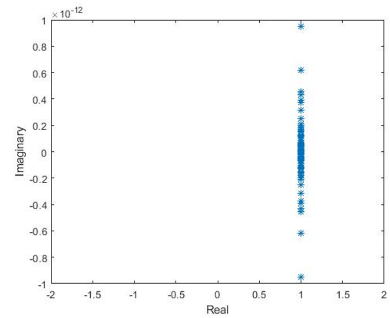
(a) Nullity=m/4



(b) Nullity=m/2

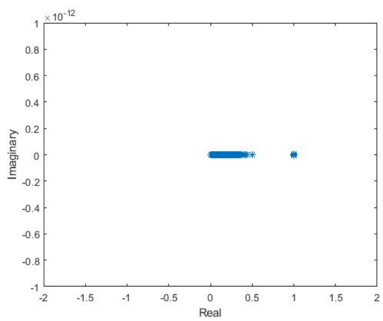


(c) Nullity=3m/4

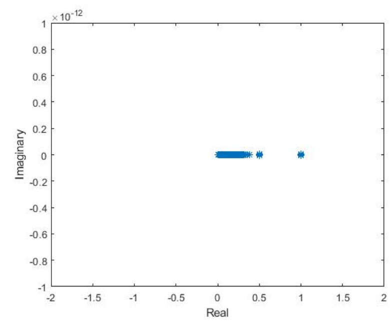


(d) Nullity=m

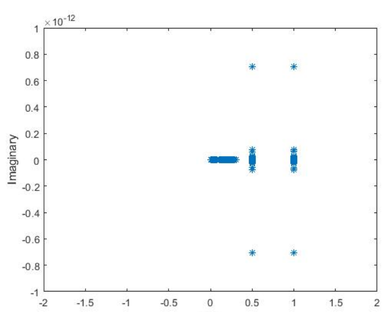
Figure 5.1: The eigenvalue distribution of preconditioned systems $P^{-1}\mathcal{A}$ with $k = 1$ and γ_4



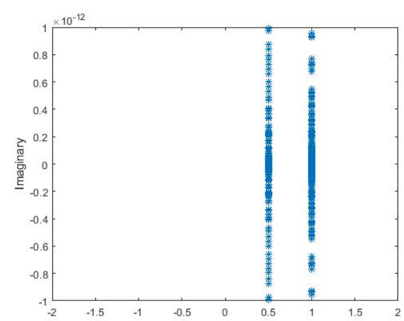
(a) Nullity=m/4



(b) Nullity=m/2



(c) Nullity=3m/4

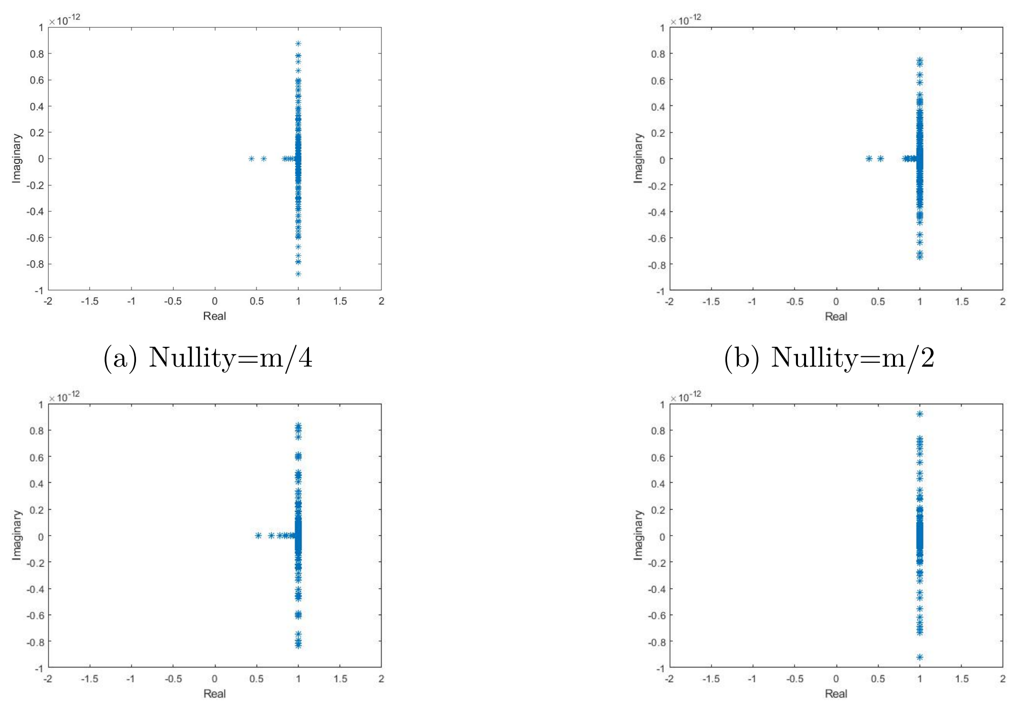


(d) Nullity=m

Figure 5.2: The eigenvalue distribution of preconditioned systems $P^{-1}A$ with $k = 100$ and γ_4.

Comparing a, b, c and d in Figure 5.1 and Figure 5.2 with a, b, c and d in Figure 5.3 and Figure 5.4, respectively, whenever $k = 1$ or $k = 100$ we notice that the distribution and cluster of the eigenvalues in Figure 5.3 and Figure 5.4 are better than the distribution and cluster of the eigenvalues in Figure 5.1 and Figure 5.2 except when the value of nullity of G at d is full nullity. However, the distribution and cluster of the eigenvalues d shown in Figure 5.3 and Figure 5.4 is still very close to the distribution and cluster of the eigenvalues d shown in Figure 5.1 and Figure 5.2. Therefore, our work gives good distribution and cluster of the eigenvalues for any nullity.

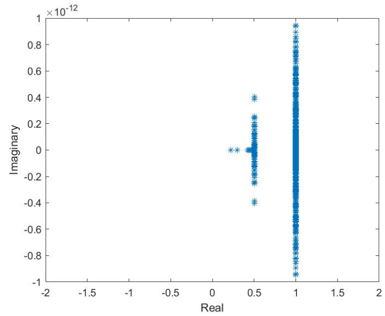
In Figure 5.3 and Figure 5.4, we use $\gamma_1 = \frac{\|B\|_2^2}{\|G\|_2 max(G/c)}$ in our preconditioner matrix P for $G1, G2, G3, G4$ to plot the distribution eigenvalues of $P^{-1}\mathcal{A}$ at $k = 1$ and $k = 100$, respectively.



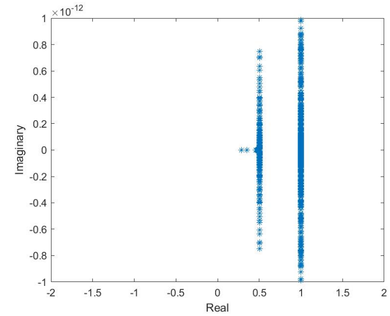
(a) Nullity=m/4 (b) Nullity=m/2

(c) Nullity=3m/4 (d) Nullity=m

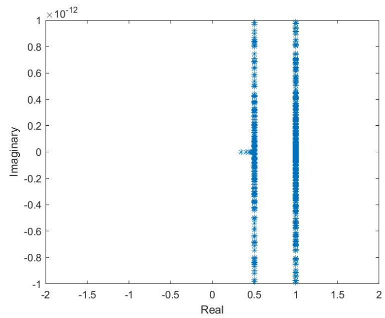
Figure 5.3: The eigenvalue distribution of preconditioned systems $P^{-1}\mathcal{A}$ with $k = 1$ and γ_1.



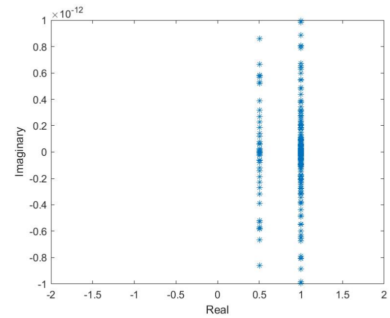
(a) Nullity=m/4



(b) Nullity=m/2



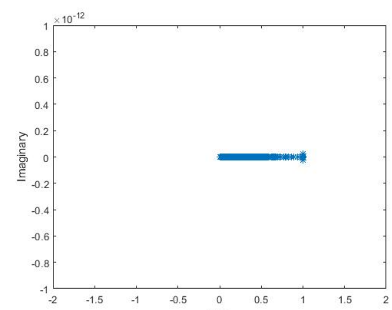
(c) Nullity=3m/4



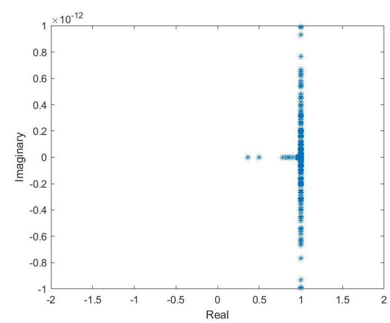
(d) Nullity=m

Figure 5.4: The eigenvalue distribution of preconditioned systems $P^{-1}\mathcal{A}$ with $k = 100$ and γ_1.

In Figure 5.5, we use $\gamma_4 = \frac{\|B\|_1^2}{\|G\|_1}$ and $\gamma_1 = \frac{\|B\|_2^2}{\|G\|_2 max(G/c))}$ in our preconditioner matrix P for $nullity = 1$ and $k = 1$, respectively to plot the distribution eigenvalues of $P^{-1}\mathcal{A}$.



(a) Nullity=1 and $k = 1$ at γ_4

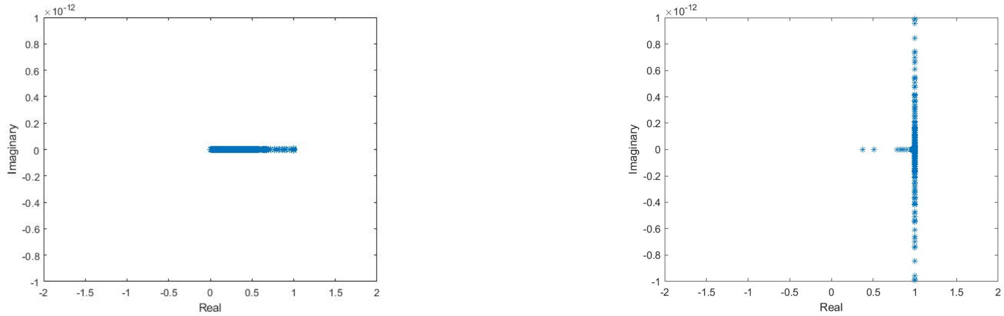


(b) Nullity=1 and $k = 1$ at γ_1

Figure 5.5: The eigenvalue distribution of preconditioned systems $P^{-1}\mathcal{A}$ with $k = 1$ and $i = 1$ at γ_4 and γ_1, respectively.

In Figure 5.6, we use $\gamma_4 = \frac{\|B\|_1^2}{\|G\|_1}$ and $\gamma_1 = \frac{\|B\|_2^2}{\|G\|_2 max(G/c)}$ in our preconditioner matrix

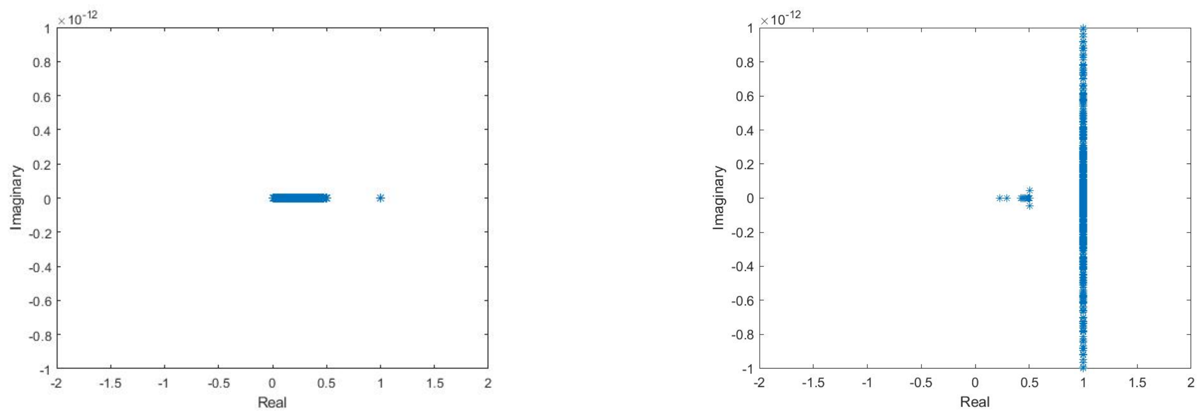
P for $nullity = 10$ and $k = 1$, respectively to plot the distribution eigenvalues of $P^{-1}\mathcal{A}$.



(a) Nullity=10 and $k = 1$ at γ_4 (b) Nullity=10 and $k = 1$ at γ_1

Figure 5.6: The eigenvalue distribution of preconditioned systems $P^{-1}\mathcal{A}$ with $k = 1$ and $i = 10$ at γ_4 and γ_1, respectively.

In Figure 5.7, we use $\gamma_4 = \frac{\|B\|_1^2}{\|G\|_1}$ and $\gamma_1 = \frac{\|B\|_2^2}{\|G\|_2 max(G/c)}$ in our preconditioner matrix P for nullity $k = 1$ and $k = 100$, respectively to plot the distribution eigenvalues of $P^{-1}\mathcal{A}$.



(a) Nullity=1, $k = 100$ at γ_4 (b) Nullity=10, $k = 100$ at γ_1

Figure 5.7: The eigenvalue distribution of preconditioned systems $P^{-1}\mathcal{A}$ with $k = 100$ and $i = 1$ at γ_4 and γ_1, respectively

From Figure 5.5, Figure 5.6 and Figure 5.7 that the distribution and cluster of the eigenvalues at $\gamma_1 = \frac{\|B\|_2^2}{\|G\|_2 max(G/c)}$ is better than $\gamma_4 = \frac{\|B\|_1^2}{\|G\|_1}$, whatever the value of k and i.

We can notice that in all the previous cases we take $\mathcal{A}$ where the initial point $u_0 > 0$ is given and s_0 obtained from $s_0 = c+Qu-B^T v > 0$. $v_0 \in R^m$ is also, taken as random vector that make $s_0 > 0$ and all the previous tables and figures are taken at the first iteration. It is noticed that in some saddle point systems at the beginning of the solution of interior point point methods that the value γ_3 and γ_4 are large which makes the condition number of $P^{-1}\mathcal{A}$ is larger than the condition number of $\mathcal{A}$. We introduce new parameters γ_1 and γ_2 which overcome the drawbacks of using γ_3 and γ_4 and give good condition number and good cluster behavior for eigenvalues of $P^{-1}\mathcal{A}$. However, there are some problems with using γ_1 and γ_2 when the solution approach the optimal value in interior point method. Therefor, we introduce $\gamma_5 = \frac{1}{max(G/c)}$ which works with all the iterations of interior point method. Also, we replete G by Q in $max(\frac{Q}{c})$ in γ_1 and γ_2 to become $\gamma_1 = \frac{\|B\|_2^2}{\|G\|_2 max(Q/c)}$ and $\gamma_2 = \frac{\|B\|_1^2}{\|G\|_1 max(Q/c)}$ and this modification make them work with all iterations in interior point method.

In Table 5.11 we use BICGSTAB method with γ_5 and γ_4 respectively in our preconditioner matrix P with $c = .01$, $tol = 10^{-06}$ and start point u_0 and s_0 where j is the number of interior point method iteration.

j	$It_{BICGSTAB(P^{-1}\mathcal{A})}$ at γ_5	Time	$Nullity(G)$	$It_{BICGSTAB(P^{-1}\mathcal{A})}$ at γ_4	Time	$Nullity(G)$
1	6.5	2.5416	0	150	33.5430	0
2	4.5	1.8149	0	206.5	45.7184	0
3	2	1.3275	0	62.5	15.3329	0
4	4	1.7099	0	123.5	28.1549	0
5	2	1.2617	0	50	12.7219	0
6	2	1.2728	0	24	6.1942	0
7	1.5	1.1759	0	28	7.0523	0
8	1.5	1.1777	0	11.5	3.3315	0
9	1.5	1.1635	0	12.5	3.6717	0
10	1.5	1.1762	0	5	1.8797	0
11	1.5	1.1435	0	2	1.3607	4
12	1.5	1.2013	10	1.5	1.1405	10
13	1.5	1.2008	10	1	1.0027	10
14	1.5	1.1560	12	1	0.9932	12
15	1.5	1.1376	12	1	0.9877	12
16	1.5	1.1544	13	1	1.0683	13

Table 5.11: Iteration numbers and time of BICGSTAB of preconditioned matrix $P^{-1}\mathcal{A}$ at γ_5 and γ_4 respectively with $c = .01$, $k = 1$ and start point u_0 and s_0.

In Table 5.12 we use MINERS method with γ_5 and γ_4 respectively in the preconditioner matrix P_0 with $c = .01$, $tol = 10^{-06}$ and start point u_0 and s_0 where j is the number of interior point method iteration.

.

j	$It_{MINERS(P_0^{-1}\mathcal{A})}$ at γ_5	Time	$Nullity(G)$	$It_{MINERS(P_0^{-1}\mathcal{A})}$ at γ_4	Time	$Nullity(G)$
1	11	2.4569	0	396	28.3714	0
2	8	1.4527	0	270	19.2489	0
3	4	1.3476	0	91	7.3024	0
4	7	1.5472	0	229	16.2154	0
5	4	1.2259	0	47	4.3188	0
6	4	1.1563	0	85	7.0031	0
7	3	1.1771	0	48	4.2704	0
8	3	1.1508	0	32	3.0822	0
9	3	1.11152	0	25	2.6134	0
10	3	1.1302	0	7	1.4183	0
11	3	1.0759	10	4	1.2869	0
12	4	1.1341	10	3	1.0803	10
13	4	1.1894	12	2	1.1284	11
14	4	1.3183	12	2	0.9751	12
15	4	1.1397	12	2	0.9805	12
16	4	1.3925	13	2	0.9810	15
17	—-	—	—	2	1.0054	15

Table 5.12: Iteration numbers and time of MINERS of preconditioned matrix $P_0^{-1}\mathcal{A}$ at γ_5 and γ_4 respectively with $c = .01$ and start point u_0 and s_0

In Table 5.13 we use BICGSTAB method with γ_5 and γ_4 respectively in our preconditioner matrix P with $c = .01$, $tol = 10^{-06}$ and start point u_0 and $\frac{s_0}{2}$ where j is the number of interior point method iteration.

j	$It_{BICGSTAB(P^{-1}\mathcal{A})}$ at γ_5	Time	$Nullity(G)$	$It_{BICGSTAB(P^{-1}\mathcal{A})}$ at γ_4	Time	$Nullity(G)$
1	6.5	2.6178	0	115	26.7769	0
2	6	2.3249	0	124	28.1630	0
3	2.5	1.5154	0	88.5	20.9376	0
4	4.5	1.9500	0	138.5	32.1706	0
5	3	1.6595	0	112	26.4020	0
6	2.5	1.5451	0	105.5	24.6976	0
7	2	1.5687	0	42	10.3280	0
8	1.5	1.2483	0	13	3.7856	0
9	1.5	1.2638	0	11	3.3543	0
10	1	1.1838	0	4	1.7884	0
11	1	1.122 8	0	2.5	1.3444	0
12	1	1.1512	10	1.5	1.3637	10
13	1	1.1838	10	1	1.0060	10
14	1	1.2231	1	1	1.0270	12
15	1	1.1582	12	—	—	—

Table 5.13: Iteration numbers and time of BICGSTAB of preconditioned matrix $P^{-1}\mathcal{A}$ at γ_5 and γ_4 respectively with $c = .01$, $k = 1$ and start point u_0 and $\frac{s_0}{2}$.

In Table 5.14 we use MINERS method with γ_5 and γ_4 respectively in the preconditioner matrix P_0 with $c = .01$, $tol = 10^{-06}$ and start point u_0 and $\frac{s_0}{2}$ where j is the number of interior point method iteration.

j	$It_{MINERS(P_0^{-1}\mathcal{A})}$ at γ_5	Time	$Nullity(G)$	$It_{MINERS(P_0^{-1}\mathcal{A})}$ at γ_4	Time	$Nullity(G)$
1	10	2.0755	0	370	26.8552	0
2	10	1.6612	0	312	22.4239	0
3	5	1.3158	0	98	7.9203	0
4	7	1.4364	0	310	23.0764	0
5	6	1.2971	0	215	16.5441	0
6	5	1.2546	0	148	1.4082	0
7	4	1.1979	0	67	5.6200	0
8	3	1.1313	0	38	3.6168	0
9	3	1.0863	0	19	2.3427	0
10	2	1.0164	0	9	1.5774	0
11	2	1.0299	0	5	1.2131	0
12	2	1.0037	10	3	1.1134	10
13	2	1.0443	10	2	1.0262	10
14	3	1.1335	12	2	1.0508	12
15	3	1.0811	12	2	1.1795	12

Table 5.14: Iteration numbers and time of MINERS of preconditioned matrix $P_0^{-1}\mathcal{A}$ at γ_5 and γ_4 respectively with $c = .01$ and start point u_0 and $\frac{s_0}{2}$.

In Table 5.15 we use BICGSTAB method with γ_6 and γ_4 respectively in our preconditioner matrix P with $c = .01$, $tol = 10^{-06}$ and start point u_0 and s_0 where j is the number of interior point method iteration.

j	$It_{BICGSTAB(P^{-1}\mathcal{A})}$ at γ_6	Time	$Nullity(G)$	$It_{BICGSTAB(P^{-1}\mathcal{A})}$ at γ_4	Time	$Nullity(G)$
1	6.5	3.5009	0	150	36.9578	0
2	4	2.4951	0	206.5	50.1548	0
3	2	2.3890	0	62.5	16.2623	0
4	3.5	2.3745	0	123.5	30.4663	0
5	2	1.9907	0	50	13.9289	0
6	2	1.9566	0	24	6.6771	0
7	1.5	1.8042	0.	28	7.7575	0
8	1.5	2.2438	0	11.5	3.3.7104	0
9	1.5	2.2593	0	12.5	4.0371	0
10	1.5	1.8060	0	5	2.2369	0
11	1.5	1.8104	0	2	1.4411	4
12	1.5	1.8028	10	1.5	1.2988	10
13	1.5	1.8035	10	1	1.1701	10
14	1.5	1.8420	12	1	1.1408	12
15	1.5	1.8548	13	1	1.1683	12
16	1.5	1.8413	13	1	1.1700	13

Table 5.15: Iteration numbers and time of BICGSTAB of preconditioned matrix $P^{-1}\mathcal{A}$ at γ_6 and γ_4 respectively with $c = 1$, $k = 1$ and start point u_0 and s_0

In Table 5.16 we use MINERS method with γ_6 and γ_4 respectively in the preconditioner matrix P_0 with $c = 1$, $tol = 10^{-06}$ and start point u_0 and s_0 where j is the number of interior point method iteration.

.

j	$It_{MINERS(P_0^{-1}\mathcal{A})}$ at γ_6	Time	$Nullity(G)$	$It_{MINERS(P_0^{-1}\mathcal{A})}$ at γ_4	Time	$Nullity(G)$
1	3	1.4269	0	396	28.2646	0
2	4	1.3517	0	270	19.8509	0
3	4	1.2986	0	91	7.3303	0
4	4	1.2715	0	229	16.4352	0
5	3	1.3199	0	47	4.1907	0
6	4	1.2708	0	85	6.8095	0
7	3	1.1687	0	48	4.2661	0
8	3	1.1102	0	32	3.0494	0
9	3	1.1425	0	25	2.6029	0
10	3	1.1209	0	7	1.3856	0
11	3	1.1176	0	4	1.1108	0
12	3	1.0845	10	3	1.0545	10
13	3	1.1452	10	2	1.1023	10
14	5	1.1628	12	2	0.9567	11
15	5	1.3354	12	2	0.9750	12
16	4	1.2352	13	2	0.9877	12
17	—-	—	—	2	0.9745	15

Table 5.16: Iteration numbers and time of MINERS of preconditioned matrix $P_0^{-1}\mathcal{A}$ at γ_6 and γ_4 respectively with $c = 1$ and start point u_0 and s_0.

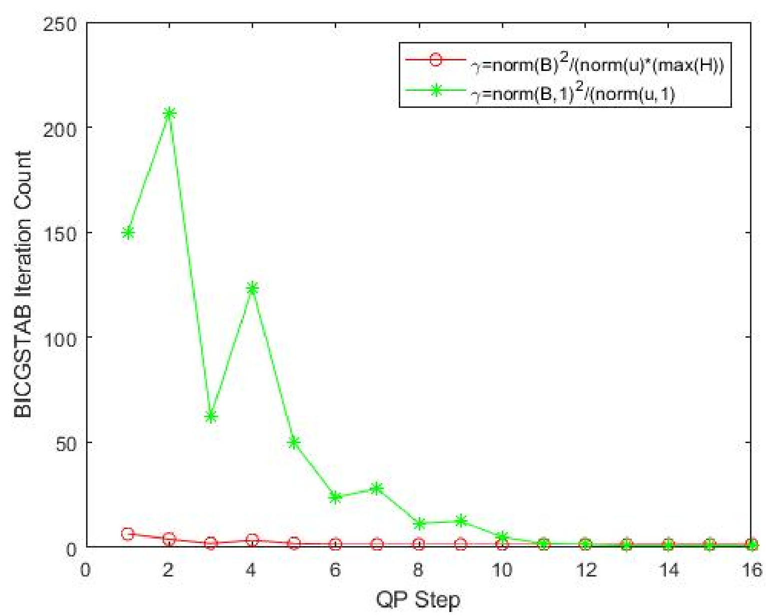


Figure 5.8: BICGSTAB iteration count for "CVXQP1",n=1000,m=500 preconditioner P iterations with γ_6 are represented by 'o',preconditioner P iterations with γ_4 are represented by '*'. The preconditioner P iterations with γ_6 is consistently better and approaches theoretical convergence.

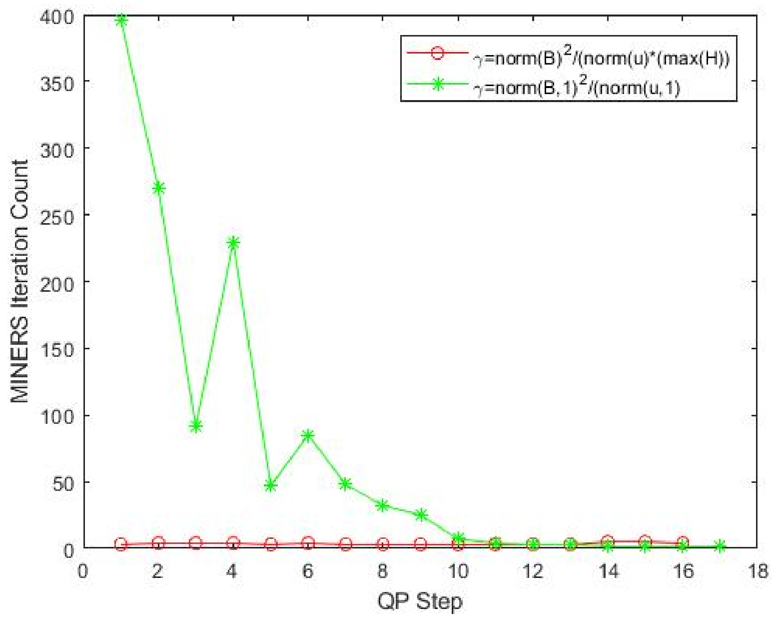


Figure 5.9: MINERS iteration count for "CVXQP1",n=1000,m=500 preconditioner P iterations with γ_6 are represented by 'o',preconditioner P iterations with γ_4 are represented by '*'. The preconditioner P iterations with γ_6 is consistently better and approaches theoretical convergence.

In Tables 5.17, 5.18 and 5.19 we use BICGSTAB method at γ_6 and γ_4 respectively with some preconditioners matrices that done in the literature at $c = 1$, $tol = 10^{-06}$ and difference k.

j	$It_{BICGSTAB(P^{-1}\mathcal{A})}$ at γ_6	Time	$Nullity(G)$	$It_{BICGSTAB(P^{-1}\mathcal{A})}$ at γ_4	Time	$Nullity(G)$
1	15	5.3661	0	stop without conv.	—	—
2	10	4.0756	0	stop without conv.	—	—
3	5.5	2.9261	0	579	136.2355	0
4	9.5	3.8164	0	stop without conv.	—	—
5	4	2.3729	0	565	131.780	0
6	5.5	2.9767	0	144.5	34.1069	0
7	5.5	3.7269	0	193.5	45.3543	0
8	3	3.0304	0	73	17.6165	0
9	3	2.9347	0	13.5	4.0265	0
10	4	2.9780	0	10	3.0802	0
11	4	2.6697	0	6	2.2708	0
12	4	3.4023	10	3.5	1.6317	10
13	4	3.2255	10	3	1.5438	10
14	2.5	2.8212	15	2.5	1.4273	12
15	2.5	2.1700	12	2.5	1.4116	12
16	2.5	2.0767	13	2.5	1.4059	13

Table 5.17: Iteration numbers and time of BICGSTAB of preconditioned matrix $P_1^{-1}\mathcal{A}$ at γ_6 and γ_4 respectively with $c = 1$, $k = 2$ and start point u_0 and s_0.

j	$It_{BICGSTAB(P^{-1}\mathcal{A})}$ at γ_6	Time	$Nullity(G)$	$It_{BICGSTAB(P^{-1}\mathcal{A})}$ at γ_4	Time	$Nullity(G)$
1	5	2.8313	0	231	56.1978	0
2	3.5	2.2095	0	79	22.5667	0
3	2	1.8470	0	83.5	11.6092	0
4	3	2.0551	0	128.5	32.9967	0
5	2	1.8925	0	23.5	6.1621	0
6	2	1.7839	0	35.5	9.0220	0
7	1.5	1.6989	0	16	4.6034	0
8	1.5	1.6835	0	12.5	3.7525	0
9	1.5	1.7234	0	12.5	3.7258	0
10	1.5	1.7325	0	10.5	3.7457	0
11	1.5	1.7121	0	2.5	1.6809	0
12	1.5	1.6954	10	1.5	1.2969	6
13	1.5	1.6486	10	1.5	1.1737	10
14	1.5	1.6706	12	1	1.0675	11
15	1.5	1.6786	12	1	1.0690	12
16	1.5	1.6505	13	1	1.0665	12

Table 5.18: Iteration numbers and time of BICGSTAB of preconditioned matrix $P_3^{-1}\mathcal{A}$ at γ_6 and γ_4 respectively with $c = 1$, $k = 2$ and start point u_0 and s_0.

j	$It_{BICGSTAB(P^{-1}\mathcal{A})}$ at γ_6	Time	$Nullity(G)$	$It_{BICGSTAB(P^{-1}\mathcal{A})}$ at γ_4	Time	$Nullity(G)$
1	8.5	4.0315	0	185	44.7071	0
2	5	2.8112	0	306.5	71.4652	0
3	2	2.3809	0	82.5	19.7976	0
4	4	2.3695	0	180.5	42.3379	0
5	2	1.8691	0	62	15.2820	0
6	2	2.2070	0	34.5	9.1848	0
7	2	3.1611	0	26.5	6.8895	0
8	1.5	2.2923	0	24.5	6.5375	0
9	1.5	1.8842	0	6	2.2257	0
10	1.5	2.1527	0	3	1.5721	0
11	1.5	2.4856	0	2.5	1.4051	0
12	1.5	2.0094	10	1.5	1.1811	10
13	1.5	1.9305	10	1	1.0287	10
14	1.5	1.9225	12	1	1.0354	12
15	1.5	1.8862	12	1	1.0456	12
16	1.5	1.9773	13	1	1.0482	13

Table 5.19: Iteration numbers and time of BICGSTAB of preconditioned matrix $P_{10}^{-1}\mathcal{A}$ at γ_6 and γ_4 respectively with $c = 1$, and start point u_0 and s_0.

We consider another example of the CUTEst matrix MOSARQP2 (quadratic programming example) where $n = 2500$, $m = 700$, G is symmetric positive definite and $\text{rank}(B) = 700$. We use this example for any $u \in \mathbb{R}^n$ as in (2.25).

k	$Nullity$	$It_{BICGSTAB(P^{-1}\mathcal{A})}$ at γ_6	$Res_{BICGSTAB(P^{-1}\mathcal{A})}$	c	$It_{BICGSTAB(P^{-1}\mathcal{A})}$ at γ_4	$\text{Res}_{BICGSTAB(P^{-1}\mathcal{A})}$
1	700	1.5	1×10^{-13}	1	2	3.1×10^{-14}
1	525	4.5	8.8×10^{-14}	0.001	60.53	9.3×10^{-13}
1	350	5	8.8×10^{-14}	0.001	100.5	7.3×10^{-13}
1	175	5	9.1×10^{-13}	0.001	129.5	5.4×10^{-13}
1	10	5.5	8.4×10^{-14}	0.001	132.5	7.9×10^{-13}
1	1	5.5	8.3×10^{-14}	0.001	130	8.1×10^{-13}
1	0	5.5	8.5×10^{-14}	0.001	129.5	8.6×10^{-13}

Table 5.20: Iteration numbers of BICGSTAB of preconditioned matrix $P^{-1}\mathcal{A}$ with γ_6 and γ_4 respectively at $k = 1$ and $tol = 10^{-12}$ for saddle point in (4.25).

k	$Nullity$	$It_{BICGSTAB(P^{-1}\mathcal{A})}$ at γ_6	$Res_{BICGSTAB(P^{-1}\mathcal{A})}$	c	$It_{BICGSTAB(P^{-1}\mathcal{A})}$ at γ_4	$\text{Res}_{BICGSTAB(P^{-1}\mathcal{A})}$
100	700	1.5	2.2×10^{-13}	0.001	2	8.2×10^{-15}
100	525	5	1.3×10^{-13}	0.001	54	8.4×10^{-13}
100	350	6	8.9×10^{-14}	0.001	88.5	8.6×10^{-13}
100	175	6	9.5×10^{-14}	0.001	115.5	6.5×10^{-13}
100	10	6	8.7×10^{-14}	0.001	120	5.6×10^{-13}
100	1	6	8.7×10^{-14}	0.001	125	6.3×10^{-13}

Table 5.21: Iteration numbers of BICGSTAB of preconditioned matrix $P^{-1}\mathcal{A}$ with γ_6 and γ_4 respectively at $k = 100$ and $tol = 10^{-12}$ for saddle point in (2.25).

CHAPTER 6

CONCLUSION AND FUTURE WORK

6.1 Conclusion

In this work, we explain how the saddle point systems arise from optimization problems and we study the proprieties of these saddle point systems. The knowledge of these proprieties help to solve the saddle point systems. As we have seen in literature, each study had preconditioner matrices with optimal parameters, and those parameters are optimal when the nullity of $null(G) = m$, otherwise, those parameters are not necessary to be optimal. For example, in P_4, if the nullity of G is small, then the GMRES method, when k is large, produces better results than taking the optimal parameter $k = 1$ (as shown in [45]). We introduce a block triangular preconditioner matrix with new γ for saddle point systems whose coefficient has singular blocks (1.4) or very ill-conditions. The preconditioner has the attractive property of improved eigenvalues clustering when using the optimal parameter in practice. Also, with the new γ, the eigenvalues clustering are improved for any nullity and any parameter k. Numerical experiments further confirm the effectiveness of our preconditioner. we notice that in all cases we take $\mathcal{A}$ where the initial point $u_0 > 0$ is given and s_0 obtained from $s_0 = c + Qu - B^T v > 0$. $v_0 \in R^m$ is also, taken as random vector that make $s_0 > 0$. It is noticed that in some saddle point systems at the beginning of the solution of interior point point methods that the value γ_3 and γ_4 are large which makes the condition number of $P^{-1}\mathcal{A}$ is larger than the condition number of $\mathcal{A}$. We introduce new parameters γ_1 and γ_2 which overcome the drawbacks of using γ_3 and γ_4 and give good condition number and good cluster behavior for eigenvalues of $P^{-1}\mathcal{A}$. However, there are some problems with using γ_1 and γ_2 when the solution approach the optimal value in interior point method. Therefor, we introduce $\gamma_5 = \frac{1}{max(G/c)}$ which works with all the iterations of interior point method. Also, we replace G by Q in $max(\frac{Q}{c})$ in γ_1 and γ_2 to become $\gamma_1 = \frac{\|B\|_2^2}{\|G\|_2 max(Q/c)}$ and $\gamma_2 = \frac{\|B\|_1^2}{\|G\|_1 max(Q/c)}$ and this modification make them work with all iterations in interior point method.

6.2 Future Work

As a future work, we suggest to consider the saddle point system of the form :

$$\mathcal{A} = \begin{bmatrix} G & B^T \\ B & -C \end{bmatrix}$$

and study some preconditioners matrices for this form in three cases :

1. When G is symmetric positive definite and C is symmetric positive semidefinite.

2. When G is symmetric positive semidefinite and C is symmetric symmetric positive definite.
3. When G is symmetric positive semidefinite and C is symmetric positive semidefinite.

REFERENCES

[1] M. Benzi, G. H. Golub, and J. Liesen, "Numerical solution of saddle point problems," *Acta numerica*, vol. 14, pp. 1–137, 2005.

[2] M. Rozložník, *Saddle-point problems and their iterative solution.* Springer, 2018.

[3] M. F. Murphy, G. H. Golub, and A. J. Wathen, "A note on preconditioning for indefinite linear systems," *SIAM Journal on Scientific Computing*, vol. 21, no. 6, pp. 1969–1972, 2000.

[4] K. Chen, F. Fairag, and A. Al-Mahdi, "Preconditioning techniques for an image deblurring problem," *Numerical Linear Algebra with Applications*, vol. 23, no. 3, pp. 570–584, 2016.

[5] S. Lungten, "Solution methods for indefinite systems of linear equations in saddle-point form," 2016.

[6] R. Estrin and C. Greif, "Spmr: A family of saddle-point minimum residual solvers," *SIAM Journal on Scientific Computing*, vol. 40, no. 3, pp. A1884–A1914, 2018.

[7] M. Benzi and A. J. Wathen, "Some preconditioning techniques for saddle point

problems," in *Model order reduction: theory, research aspects and applications.* Springer, 2008, pp. 195–211.

[8] L. Bergamaschi, J. Gondzio, and G. Zilli, "Preconditioning indefinite systems in interior point methods for optimization," *Computational Optimization and Applications*, vol. 28, no. 2, pp. 149–171, 2004.

[9] M. Glavic, "Interior point methods: A survey, short survey of applications to power systems, and research opportunities," University of Liege, Tech. Rep., 2004.

[10] R. Robere, "Interior point methods and linear programming," *University of Toronto*, 2012.

[11] G. Al-Jeiroudi, J. Gondzio, and J. Hall, "Preconditioning indefinite systems in interior point methods for large scale linear optimisation," *Optimisation Methods and Software*, vol. 23, no. 3, pp. 345–363, 2008.

[12] F. A. Potra, "Interior-point methods florian a. potra and stephen j. wright," 2000.

[13] Y. Cui, K. Morikuni, T. Tsuchiya, and K. Hayami, "Implementation of interior-point methods for lp based on krylov subspace iterative solvers with inner-iteration preconditioning," *Computational Optimization and Applications*, vol. 74, no. 1, pp. 143–176, 2019.

[14] J.-S. Chai and K.-C. Toh, "Preconditioning and iterative solution of symmetric indefinite linear systems arising from interior point methods for linear programming," *Computational Optimization and Applications*, vol. 36, no. 2-3, pp. 221–247, 2007.

[15] S. Wright and J. Nocedal, "Numerical optimization," *Springer Science*, vol. 35, no. 67-68, p. 7, 1999.

[16] S. Boyd and L. Vandenberghe, *Convex optimization.* Cambridge university press, 2004.

[17] D. James, "Conjugate gradient methods for constrained least squares problems," AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OH, Tech. Rep., 1990.

[18] M. A. S. Salman, "Preconditioning techniques for saddle point problems," Ph.D. dissertation, King Fahd University of Petroleum and Minerals (Saudi Arabia), 2011.

[19] R. A. Horn and C. R. Johnson, *Matrix analysis.* Cambridge university press, 2012.

[20] F. Fairag, M. Alshahrani, and H. Tawfiq, "Bramble–pasciak-type conjugate gradient method for darcy's equations," *SIAM Journal on Matrix Analysis and Applications*, vol. 37, no. 1, pp. 469–489, 2016. [Online]. Available: https://doi.org/10.1137/151003143

[21] H. C. Elman, D. J. Silvester, and A. J. Wathen, *Finite elements and fast iterative solvers: with applications in incompressible fluid dynamics.* Numerical Mathematics and Scie, 2014.

[22] A. Wathen and D. Silvester, "Fast iterative solution of stabilised stokes systems.

part i: Using simple diagonal preconditioners," *SIAM Journal on Numerical Analysis*, vol. 30, no. 3, pp. 630–649, 1993.

[23] J. H. Bramble and J. E. Pasciak, "A preconditioning technique for indefinite systems resulting from mixed approximations of elliptic problems," *Mathematics of Computation*, vol. 50, no. 181, pp. 1–17, 1988.

[24] Z.-H. Cao, "Positive stable block triangular preconditioners for symmetric saddle point problems," *Applied numerical mathematics*, vol. 57, no. 8, pp. 899–910, 2007.

[25] A. Klawonn, "Block-triangular preconditioners for saddle point problems with a penalty term," *SIAM Journal on Scientific Computing*, vol. 19, no. 1, pp. 172–184, 1998.

[26] J. Pestana, "On the eigenvalues and eigenvectors of block triangular preconditioned block matrices," *SIAM Journal on Matrix Analysis and Applications*, vol. 35, no. 2, pp. 517–525, 2014.

[27] V. Simoncini, "Block triangular preconditioners for symmetric saddle-point problems," *Applied Numerical Mathematics*, vol. 49, no. 1, pp. 63–80, 2004.

[28] O. Axelsson and M. Neytcheva, "Preconditioning methods for linear systems arising in constrained optimization problems," *Numerical linear algebra with applications*, vol. 10, no. 1-2, pp. 3–31, 2003.

[29] M. Fortin and R. Glowinski, *Augmented Lagrangian methods: applications to the numerical solution of boundary-value problems.* Elsevier, 2000, vol. 15.

[30] M. Benzi and M. K. Ng, "Preconditioned iterative methods for weighted toeplitz least squares problems," *SIAM Journal on Matrix Analysis and Applications*, vol. 27, no. 4, pp. 1106–1124, 2006.

[31] Z.-Z. Bai, G. H. Golub, and C.-K. Li, "Optimal parameter in hermitian and skew-hermitian splitting method for certain two-by-two block matrices," *SIAM Journal on Scientific Computing*, vol. 28, no. 2, pp. 583–603, 2006.

[32] M. K. Ng and J. Pan, "Weighted toeplitz regularized least squares computation for image restoration," *SIAM Journal on Scientific Computing*, vol. 36, no. 1, pp. B94–B121, 2014.

[33] H. S. Dollar, N. I. Gould, M. Stoll, and A. J. Wathen, "Preconditioning saddle-point systems with applications in optimization," *SIAM Journal on Scientific Computing*, vol. 32, no. 1, pp. 249–270, 2010.

[34] Y. Notay, "Convergence of some iterative methods for symmetric saddle point linear systems," *SIAM Journal on Matrix Analysis and Applications*, vol. 40, no. 1, pp. 122–146, 2019.

[35] J. Pestana and T. Rees, "Null-space preconditioners for saddle point systems," *SIAM Journal on Matrix Analysis and Applications*, vol. 37, no. 3, pp. 1103–1128, 2016.

[36] T. Rees and J. Scott, "A comparative study of null-space factorizations for sparse symmetric saddle point systems," *Numerical Linear Algebra with Applications*, vol. 25, no. 1, p. e2103, 2018.

[37] S. Bradley, "Ideal preconditioners for saddle point systems with a rank-deficient leading block," *arXiv preprint arXiv:1807.08590*, 2018.

[38] R. Kobayashi, "Studies on methods for verifying the accuracy of numerical solutions of symmetric saddle point linear systems," Ph.D. dissertation, Waseda University, 2018.

[39] Z.-H. Cao, "Augmentation block preconditioners for saddle point-type matrices with singular (1, 1) blocks," *Numerical Linear Algebra with Applications*, vol. 15, no. 6, pp. 515–533, 2008.

[40] T. Huang, G. Cheng, and L. Li, "New block triangular preconditioners for saddle point linear systems with highly singular (1, 1) blocks," *Mathematical Problems in Engineering*, vol. 2009, 2009.

[41] Z.-H. Cao, "A note on block diagonal and constraint preconditioners for non-symmetric indefinite linear systems," *International Journal of Computer Mathematics*, vol. 83, no. 4, pp. 383–395, 2006.

[42] E. de Sturler and J. Liesen, "Block-diagonal and constraint preconditioners for nonsymmetric indefinite linear systems. part i: Theory," *SIAM Journal on Scientific Computing*, vol. 26, no. 5, pp. 1598–1619, 2005.

[43] T. Rees and C. Greif, "A preconditioner for linear systems arising from interior point optimization methods," *SIAM Journal on Scientific Computing*, vol. 29, no. 5, pp. 1992–2007, 2007.

[44] C. Greif and D. Schötzau, "Preconditioners for saddle point linear systems with highly singular (1, 1) blocks," *ETNA, Special Volume on Saddle Point Problems*, vol. 22, pp. 114–121, 2006.

[45] C.-X. Li and S.-L. Wu, "Two block triangular preconditioners for asymmetric saddle point problems," *Applied Mathematics and Computation*, vol. 269, pp. 456–463, 2015.

[46] Z.-H. Cao, "A note on spectrum analysis of augmentation block preconditioned generalized saddle point matrices," *Journal of Computational and Applied Mathematics*, vol. 238, pp. 109–115, 2013.

[47] L.-T. Zhang, "A new preconditioner for generalized saddle point matrices with highly singular (1, 1) blocks," *International Journal of Computer Mathematics*, vol. 91, no. 9, pp. 2091–2101, 2014.

[48] Q. Liu, "New preconditioners for nonsymmetric saddle point systems with singular block," *ISRN Computational Mathematics*, vol. 2013, 2013.

[49] Y.-F. Ke and C.-F. Ma, "Spectrum analysis of a more general augmentation block preconditioner for generalized saddle point matrices," *BIT Numerical Mathematics*, vol. 56, no. 2, pp. 489–500, 2016.

[50] L. Zhang, Y. Zhou, X. Zuo, C. Li, and Y. Li, "A note on block preconditioner for generalized saddle point matrices with highly singular (1, 1) block," *JOURNAL OF APPLIED ANALYSIS AND COMPUTATION*, vol. 9, no. 3, pp. 916–927, 2019.

[51] J. He and T.-Z. Huang, “Two augmentation preconditioners for nonsymmetric and indefinite saddle point linear systems with singular (1, 1) blocks,” *Computers & Mathematics with Applications*, vol. 62, no. 1, pp. 87–92, 2011.

[52] C.-X. Li and S.-L. Wu, “An augmentation preconditioner for asymmetric saddle point problems with singular (1, 1) blocks.”

VITAE

- Name: Hafed Ahmed Mohsen Saeed
- Nationality: Yemen
- Date of Birth: 25/12/1984
- Email: *hafed2006@gmail.com*
- Permenant Address: Department of Mathematics, Ibb University, Ibb, Yemen.
- Zip Code: 00967

- B.S.c Mathematics, College of Faculty Education, Ibb University, Republic of Yemen,
- Honors and Awards: [2009] Awarded the prize of excellence by the president of the Republic of Yemen.
- Research Interests:
  1. Preconditioning Techniques for Saddle point system in optimization.
  2. Saddle-Point Problems and Their Iterative Solution.
  3. Differential Equations.