Appl Math Optim 40:191–210 (1999)

Applied Mathematics & Optimization

© 1999 Springer-Verlag New York Inc.

A Dynamical Systems Analysis of Semidefinite Programming with Application to Quadratic Optimization with Pure Quadratic Equality Constraints*

R. J. Orsi,[1] R. E. Mahony,[2] and J. B. Moore[3]

[1] Department of Electrical and Electronic Engineering, University of Melbourne, Parkville, VIC 3052, Australia

[2] Heudiasyc - UTC UMR 6599, Centre de Recherche de Royallieu, BP 20529, 60205 Compiegne Cedex, France

[3] Department of Systems Engineering, RSISE, Australian National University, Canberra, ACT 0200, Australia

Abstract. This paper considers the problem of minimizing a quadratic cost subject to purely quadratic equality constraints. This problem is tackled by first relating it to a standard semidefinite programming problem. The approach taken leads to a dynamical systems analysis of semidefinite programming and the formulation of a gradient descent flow which can be used to solve semidefinite programming problems. Though the reformulation of the initial problem as a semidefinite programming problem does not in general lead directly to a solution of the original problem, the initial problem is solved by using a modified flow incorporating a penalty function.

Key Words. Dynamical systems, Semidefinite programming, Quadratic optimization, Quadratic equality constraints.

AMS Classification. 90C26, 90C30, 58F40.

* The authors wish to acknowledge the funding of the activities of the Cooperative Research Centre for Robust and Adaptive Systems by the Australian Commonwealth Government under the Cooperative Research Centre Program. The first author would also like to acknowledge the support of Telstra under the TRL Postgraduate Fellowship scheme.

1. Introduction

Constrained quadratic optimization problems form an important area of research and arise in many practical applications. Many of the problems that have been studied in this area fall into the category of linearly constrained, convex quadratic programming problems. A wealth of effective techniques are available for solving such problems, in particular we mention active set methods (Fletcher, 1987) and various interior-point methods (Faybusovich, 1991; Tits and Zhou, 1993; Nesterov and Nemirovskii, 1994). The situation is not as tractable when one allows nonlinear equality constraints, see for example Thng et al. (1996). Such constraints inherently lead to non-convex feasible sets which often consist of a number of disconnected components.

An important area of current research closely related to constrained quadratic optimization is that of semidefinite programming. Semidefinite programming is a convex optimization method that unifies a number of standard problems such as linear and quadratic programming and has a wide variety of applications from engineering to combinatorial optimization. Importantly, there exist many effective interior-point methods to solve semidefinite programming problems. These methods have polynomial worst-case complexity and perform well in practice (Vandenberghe and Boyd, 1996).

In this paper we consider the problem of minimizing a quadratic cost subject to purely quadratic equality constraints. Such problems are non-convex and their geometry is such that in many cases the resulting constraint set consists of the union of a number of disconnected subsets, each with their own local minima. To overcome the problem of multiple minima, we reformulate the problem in a novel manner. The reformulation involves the consideration of a sequence of linear optimization problems on the boundary of the positive definite matrices. Each of these problems is nested together in a manner that leads to a standard semidefinite programming problem on the interior of the positive definite matrices. The approach taken leads to the formulation of a gradient descent flow which can be used (in theory at least) to solve semidefinite programming problems. Though our reformulation of the initial problem as a semidefinite programming problem does not in general lead directly to a solution of the original problem, the initial problem is solved by using a modified flow incorporating a penalty function. The optimum of the semidefinite programming problem is used as the initial condition for this modified flow.

Our aims in this paper are twofold. We present both a method for minimizing a quadratic cost subject to quadratic equality constraints and we provide an analysis of semidefinite programming from the non-standard though very interesting viewpoint of dynamical systems. Though it is unlikely that the gradient flow developed will provide a practical approach to solving semidefinite programming problems, the analysis undertaken provides an interesting perspective into the geometry of such problems.

The paper is structured as follows. In Section 2 a quadratic optimization problem subject to pure quadratic equality constraints is introduced and then related through a number of steps into a semidefinite programming problem. In Section 3 the geometry of this problem is analyzed. A gradient flow to solve semidefinite programming problems is developed and analyzed in Section 4; Section 5 contains some further analysis. In Section 6 a modified version of the gradient flow incorporating a penalty function is introduced. In Section 7 various methods of solving the original quadratic optimization

problem based on the flow of Section 6 are discussed and a simulation example for one of the methods is presented. The paper ends with some concluding remarks.

2. Problem Formulation

In this section a quadratic optimization problem subject to purely quadratic equality constraints is presented. This problem is then reformulated as a sequence of linear optimization problems on the boundary of the positive definite matrices. Each of these problems is nested together in a manner that leads to a standard semidefinite programming problem.

Consider the quadratic optimization problem:

Problem 2.1. Given $A_0, A_1, \ldots, A_m \in \mathbb{R}^{n\times n}$ and $c_1, \ldots, c_m \in \mathbb{R}$,

$$\begin{aligned} &\text{minimize } \varphi(x) := x^T A_0 x \\ &\text{subject to} \quad x \in \mathbb{R}^n, && (2.1)\\ &\qquad\qquad \psi_i(x) := x^T A_i x = c_i, \qquad i = 1, \ldots, m. && (2.2)\end{aligned}$$

The feasible set, those points which satisfy the constraints (2.1), (2.2), will certainly not be convex, and in general will have a number of separate connected components. Indeed, without considerably more knowledge of the matrices $A_1, \ldots, A_m \in \mathbb{R}^{n\times n}$ and the scalars $c_1, \ldots, c_m \in \mathbb{R}$ it is unclear whether the feasible set is non-empty. To avoid dealing with a null problem of this form the following assumptions are made:

Assumption 2.2.

(i) The matrices $A_0, A_1, \ldots, A_m$ are symmetric.
(ii) The set of points satisfying the constraints (2.1), (2.2) is non-empty.
(iii) The matrices $A_1, \ldots, A_m$ are linearly independent.

The first of these assumptions can be made without loss of generality due to the symmetry of the functions φ and $\psi_1, \ldots, \psi_m$. The second assumption is for convenience while the third assumption ensures that the constraints (2.2) are non-redundant.

Remark 2.3. It is important not to require implicitly the structure of the feasible set to be known prior to the solution of Problem 2.1 being undertaken. Computing the set of feasible points is itself a difficult and time-consuming task.

The approach taken is to reformulate Problem 2.1 as a matrix optimization problem on the boundary of the positive definite matrices. Let tr denote the trace operator. Then, given any matrix $A \in \mathbb{R}^{n\times n}$ and any vector $x \in \mathbb{R}^n$, one has

$$x^T A x = \operatorname{tr}(x^T A x) = \operatorname{tr}(A x x^T) = \operatorname{tr}(AX),$$

where $X := xx^T$. The set of real $n \times n$ matrices that can be written in the form $X = xx^T$,

$x \neq 0$, is the set of symmetric, positive semidefinite matrices of rank 1. Let

$$S(1,n) = \{X \in \mathbb{R}^{n\times n} \mid X^T = X \geq 0,\ \operatorname{rank}(X) = 1\}$$

denote this set of matrices. Consider the set

$$\mathcal{M}_1 = \{X \in S(1,n) \mid \Psi_i(X) := \operatorname{tr}(A_i X) = c_i,\ i = 1,\ldots,m\}.$$

$\mathcal{M}_1$ is the set of all rank 1 matrices of the form xx^T where x is a feasible point for Problem 2.1. This leads to the following optimization problem:

Problem 2.4. Given $A_0, A_1, \ldots, A_m \in \mathbb{R}^{n\times n}$ and $c_1, \ldots, c_m \in \mathbb{R}$ satisfying Assumption 2.2,

$$\begin{aligned} &\text{minimize } \Phi(X) := \operatorname{tr}(A_0 X) \\ &\text{subject to} \quad X \in \mathcal{M}_1. \end{aligned}$$

Observe that in the new formulation both the cost and the explicit constraint functions, $\Psi_i(X)$, are linear in X. The nonlinearity in the problem is confined to the geometry of the set $S(1,n)$. Much is known about the geometry of $S(1,n)$. In particular, $S(1,n)$ can be thought of as a homogeneous orbit of the general linear group under congruence transformation (Helmke and Moore, 1994). The addition of linear constraints in the definition of $\mathcal{M}_1$ will generally divide the set into a number of separate connected components. However, the reformulation allows one to consider the generalized sets

$$S(r,n) = \{X \in \mathbb{R}^{n\times n} \mid X^T = X \geq 0,\ \operatorname{rank}(X) = r\} \tag{2.3}$$

and

$$\mathcal{M}_r = \{X \in S(r,n) \mid \operatorname{tr}(A_i X) = c_i,\ i = 1,\ldots,m\}. \tag{2.4}$$

This leads directly to the nested set of optimization problems:

Problem 2.5. Given $A_0, A_1, \ldots, A_m \in \mathbb{R}^{n\times n}$ and $c_1, \ldots, c_m \in \mathbb{R}$ satisfying Assumption 2.2 and r some integer $1 \leq r \leq n$,

$$\begin{aligned} &\text{minimize } \Phi(X) \\ &\text{subject to} \quad X \in \mathcal{M}_r. \end{aligned}$$

In fact each $\mathcal{M}_r$ suffers from the same difficulty as $\mathcal{M}_1$ with potentially several connected components. As the number r is increased the number of potential separate components reduces until $r = n$. In this final case then it is easily seen that $\mathcal{M}_n$ is simply the intersection of a set of affine constraints with the convex cone of positive definite matrices. Thus $\mathcal{M}_n$ consists of only a single connected component and by solving Problem 2.5 for $r = n$ one avoids the complication of local minima due to the geometry of the constraint sets. Unfortunately $\mathcal{M}_n$ is not a closed set and hence the problem could be ill posed.

To avoid this problem we consider the set

$$\mathcal{M} := \overline{\mathcal{M}_n} = \bigcup_{r=0,\ldots,n} \mathcal{M}_r, \tag{2.5}$$

where $\overline{\mathcal{M}_n}$ denotes the topological closure of $\mathcal{M}_n$. Note that $\mathcal{M}$ is a closed subset of $\mathbb{R}^{n\times n}$ with a single connected component. This leads to the well-posed optimization problem:

Problem 2.6. Given $A_0, A_1, \ldots, A_m \in \mathbb{R}^{n\times n}$ and $c_1, \ldots, c_m \in \mathbb{R}$ satisfying Assumption 2.2,

$$\begin{aligned} &\text{minimize } \Phi(X) \\ &\text{subject to} \quad X \in \mathcal{M}. \end{aligned}$$

Problem 2.6 is a standard semidefinite programming problem (Alizadeh, 1995). There exist many practical interior-point methods to solve such problems (see Vandenberghe and Boyd, 1996). While classical linear programming theory ensures that the optimum of Problem 2.6 always lies on the boundary of $\mathcal{M}$ (and hence that the optimum is rank degenerate), unfortunately, as will be discussed later in Section 5, the optimum will not in general be rank 1. In the next sections we introduce a gradient flow to solve Problem 2.6. Problem 2.4, and hence Problem 2.1, is then solved using a modified version of this flow incorporating a penalty function designed to penalize solutions of rank greater than 1. Simulations indicate that solution to Problem 2.4 lies close in terms of the cost $\operatorname{tr}(A_0 X)$ to the solution of Problem 2.6.

3. The Geometry of the Feasible Sets

In this section the geometry of the sets $\mathcal{M}_r$ is investigated. It is shown that, excluding a set of singular points of zero measure, each set $\mathcal{M}_r$ is a Riemannian manifold. Background material on differential geometry, Lie groups and related material used in this paper can be found in Boothby (1986) and Helmke and Moore (1994). The homogeneous space structure of $S(r, n)$ is discussed in Chapter 5 of Helmke and Moore (1994).

An advantage of dealing with semi-algebraic Lie groups and group actions (such as the general linear group and its group action on $S(r, n)$) is that the linearization of the group action can be used to provide an explicit algebraic representation of the geometric properties of the homogeneous spaces considered. Following the notation presented in Chapter 5 of Helmke and Moore (1994), denote the symmetric bracket of two matrices $A, B \in \mathbb{R}^{n\times n}$ by

$$\{A, B\} := AB + B^T A^T.$$

Theorem 3.1. *The set*

$$S(r, n) = \{X \in \mathbb{R}^{n\times n} \mid X^T = X \geq 0,\ \operatorname{rank}(X) = r\}$$

(*as previously defined, see* (2.3)) *is a smooth manifold whose tangent space at an element* $X \in S(r, n)$ *is the vector space*

$$T_X S(r, n) = \{\{\Delta, X\} \mid \Delta \in \mathbb{R}^{n\times n}\}.$$

Proof. Consider the map

$$\alpha\colon\ GL(n, \mathbb{R}) \times \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}, \qquad \alpha(Z, X) = ZXZ^T.$$

It is straightforward to show that

(i) If I is the identity matrix of $GL(n, \mathbb{R})$, then

$$\alpha(I, X) = X \qquad \text{for all} \quad X \in \mathbb{R}^{n\times n}.$$

(ii) If $A, B \in GL(\mathrm{n,R})$, then

$$\alpha(A, \alpha(B, X)) = \alpha(AB, X) \qquad \text{for all} \quad X \in \mathbb{R}^{n\times n}.$$

Hence α is a left group action. Let I_r denote the $r \times r$ identity matrix and let E_r be the $n \times n$ block matrix

$$E_r = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}.$$

Then

$$S(r, n) = \{ZE_rZ^T \mid Z \in GL(n, \mathbb{R})\}$$

and hence $S(r, n)$ is an orbit of the Lie group action α. As this group action is semi-algebraic, it follows that $S(r, n)$ is a smooth submanifold of $\mathbb{R}^{n\times n}$ (Gibson, 1979, p. 224).

For any $X \in S(r, n)$, consider the map

$$\alpha_X\colon\ GL(n, \mathbb{R}) \to S(r, n), \qquad Z \mapsto \alpha(Z, X) = ZXZ^T.$$

As α is a smooth action of a Lie group on a smooth manifold and the orbit $S(r, n)$ is a smooth submanifold, it follows that the map α_X is a submersion (Gibson, 1979, p. 74). Hence, $D\alpha_X|_I$, the differential of α_X evaluated at $Z = I$, is a linear map from $T_I GL(n, \mathbb{R})$ onto $T_X S(r, n)$ and

$$D\alpha_X|_I\,(\Delta) = \Delta X + X\Delta^T.$$

Noting that $T_I GL(n, \mathbb{R}) = \mathbb{R}^{n\times n}$ and that X is arbitrary completes the proof. □

Consider the map

$$F\colon\ S(r, n) \to \mathbb{R}^m, \qquad X \mapsto (\mathrm{tr}(A_1X)\ \cdots\ \mathrm{tr}(A_mX))^T.$$

The set $\mathcal{M}_r$ (see (2.4)) is a fibre of this map given by $\mathcal{M}_r = F^{-1}(c_1, \ldots, c_m)$. The derivative of F in direction $\{\Delta, X\} \in T_X S(r, n)$ is[1]

$$\begin{aligned} DF|_X(\{\Delta, X\}) &= (\mathrm{tr}(A_1\{\Delta, X\}) \ \cdots \ \mathrm{tr}(A_m\{\Delta, X\}))^T \\ &= 2(\mathrm{tr}(A_1 \Delta X) \ \cdots \ \mathrm{tr}(A_m \Delta X))^T \\ &= 2(\mathrm{vec}(A_1) \ \cdots \ \mathrm{vec}(A_m))^T (X \otimes I)\, \mathrm{vec}(\Delta). \end{aligned}$$

The Fibre Theorem (Helmke and Moore, 1994, p. 346) implies that $\mathcal{M}_r = F^{-1}(c_1, \ldots, c_m)$ is a smooth submanifold of $S(r, n)$ if the derivative of F is full rank at every point in the fibre. That is, if

$$(\mathrm{vec}(A_1) \ \cdots \ \mathrm{vec}(A_m))^T (X \otimes I) \tag{3.1}$$

is full rank for all $X \in \mathcal{M}_r$. In addition, at every point X where the derivative of F is full rank, $\mathcal{M}_r$ is locally a manifold and the tangent space of $\mathcal{M}_r$ at such points is

$$\begin{aligned} T_X \mathcal{M}_r &= \ker DF|_X \\ &= \{\{\Delta, X\} \mid \Delta \in \mathbb{R}^{n \times n},\ \mathrm{tr}(A_i\{\Delta, X\}) = 0,\ i = 1, \ldots, m\}. \end{aligned}$$

Definition 3.2. Any point $X \in \mathcal{M}_r$ for which (3.1) is not full rank is termed a *singular point*.

Unfortunately in practice it is difficult to know in advance when singular points may arise. It follows from Sard's theorem (Hirsch, 1976, p. 69) that the set of points $(c_1, \ldots, c_m) \in \mathbb{R}^m$ for which $\mathcal{M}_r$ is not a manifold has measure zero in $\mathbb{R}^m$. Consequently, for an arbitrary choice of matrices $A_1, \ldots, A_m$ and scalars $c_1, \ldots, c_m$, it is unlikely that $\mathcal{M}_r$ will contain singularities.

---

[1] Let A and B be real $n \times n$ matrices. The *vec* of the matrix $A \in \mathbb{R}^{n \times n}$ is the n^2 length column vector $\mathrm{vec}(A) := [A(:, 1); \ldots; A(:, m)]$. It is easily verified that

$$\mathrm{tr}(AB) = (\mathrm{vec}(A^T))^T \ \mathrm{vec}(B).$$

Let A_{ij} denote the ijth entry of the matrix A. The *Kronecker product* of the matrices A and B is defined by

$$A \otimes B = \begin{pmatrix} A_{11}B & \cdots & A_{1n}B \\ \vdots & & \vdots \\ A_{n1}B & \cdots & A_{nn}B \end{pmatrix} \in \mathbb{R}^{n^2 \times n^2}.$$

Some readily verified identities involving the vec operation and the Kronecker product are (Helmke and Moore, 1994, p. 314)

$$\mathrm{vec}(AB) = (I \otimes A) \ \mathrm{vec}(B) = (B^T \otimes I) \ \mathrm{vec}(A)$$

and

$$(A \otimes B)^T = (A^T \otimes B^T).$$

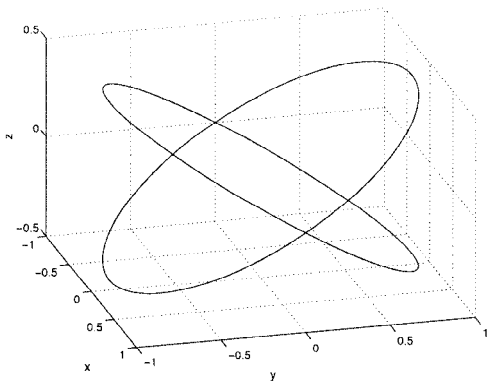

Figure 3.1. A constraint set that is not a manifold.

The following is an example of the geometry of a set $\mathcal{M}_r$ that contains singular points. For $r = 1$, $n = 3$ and $m = 2$ constraints, let $\mathcal{M}_1$ be defined by $A_1 = I$, $A_2 = \operatorname{diag}\left(1, \frac{1}{4}, 4\right)$ and $c_1 = c_2 = 1$. It is easily verified that the matrix $(\operatorname{vec}(A_1)\ \operatorname{vec}(A_2))^T (X \otimes I)$ is rank degenerate at $X = \operatorname{diag}(1, 0, 0) \in \mathcal{M}_1$. In local coordinates any $X \in \mathcal{M}_1$ can be represented by $x \in \mathbb{R}^n$ satisfying $X = xx^T$. The fact that the set is not a manifold is clearly demonstrated in Figure 3.1 which is a plot of the constraint set in local coordinates. The plot shows that the points $(\pm 1, 0, 0)$ are degenerate and hence that the constraint set does not form a manifold.

Before proceeding we make the following definition:

Definition 3.3. $$W := (\operatorname{vec}(A_1)\ \cdots\ \operatorname{vec}(A_m)) \in \mathbb{R}^{n^2 \times m}.$$

Assumption 2.2 implies W has rank m. Using the definition of W a singular point now becomes a point $X \in \mathcal{M}_r$ for which $W^T(X \otimes I)$ is not full rank.

Lemma 3.4. *Given $A_0, A_1, \ldots, A_m \in \mathbb{R}^{n \times n}$ and $c_1, \ldots, c_m \in \mathbb{R}$ satisfying Assumption* 2.2, *and W as in Definition* 3.3, *for each $r = 1, \ldots, n$, the set*

$$\mathcal{N}_r = \{X \in \mathcal{M}_r \mid W^T(X \otimes I) \text{ is full rank}\}$$

is a smooth submanifold of $S(r, n)$ that differs from the set $\mathcal{M}_r$ by at most a set of measure zero. The tangent space of $\mathcal{N}_r$ at a point X can be represented by the vector space

$$T_X\mathcal{N}_r = \{\{\Delta, X\} \mid \Delta \in \mathbb{R}^{n \times n},\ \operatorname{tr}(A_i\{\Delta, X\}) = 0,\ i = 1, \ldots, m\}.$$

Proof. The fact that $\mathcal{N}_r$ differs from $\mathcal{M}_r$ by at most a set of measure zero is a conse-

quence of the fact that $\mathcal{M}_r - \mathcal{N}_r$ is a proper algebraic subset, defined by

$$\det\left(W^T (X \otimes I)^2 W\right) = 0,$$

of the semi-algebraic set $\mathcal{M}_r$. The remainder of the theorem follows from discussion above using the Fibre Theorem (Helmke and Moore, 1994, p. 346). □

The manifolds $\mathcal{N}_r$, $r = 1, \ldots, n$, are submanifolds of the homogeneous spaces $S(r, n)$. It is possible to give $S(r, n)$ a Riemannian structure derived from the normal metric on the general linear group (Helmke and Moore, 1994, Chapter 5). This metric is known as the normal metric on $S(r, n)$. A key property of the metric used is that the algebraic structure of the metric is equivalent for each of the manifolds $S(r, n)$, $r = 1, \ldots, n$. The manifolds $\mathcal{N}_r$ inherit this Riemannian structure as submanifolds of $S(r, n)$. The explicit form of the normal metric is given in the following discussion.

The proof of Theorem 3.1 indicates that the tangent space $T_X S(r, n)$ can be considered as the image of the surjective linear map

$$D\alpha_X|_I : \mathbb{R}^{n\times n} \to T_X S(r, n), \qquad \Delta \mapsto \{\Delta, X\}.$$

The kernel of $D\alpha_X|_I$ is

$$K = \ker D\alpha_X|_I = \{\Delta \in \mathbb{R}^{n\times n} \mid \{\Delta, X\} = 0\}.$$

With respect to the standard inner product on $\mathbb{R}^{n\times n}$,

$$\langle A, B\rangle = \operatorname{tr}(A^T B),$$

the orthogonal complement of $\ker D\alpha_X|_I$ is

$$K^\perp = \{Z \in \mathbb{R}^{n\times n} \mid \operatorname{tr}(Z^T \Delta) = 0 \ \ \forall\, \Delta \in K\}.$$

This leads to the following orthogonal decomposition of $\mathbb{R}^{n\times n}$,

$$\mathbb{R}^{n\times n} = K \oplus K^\perp.$$

Hence, every element $\Delta \in \mathbb{R}^{n\times n}$ has a unique decomposition

$$\Delta = \Delta_X + \Delta^X, \tag{3.2}$$

where $\Delta_X \in K$ and $\Delta^X \in K^\perp$.

The map $D\alpha_X|_I$ is surjective with kernel K and hence induces an isomorphism of $K^\perp \subset \mathbb{R}^{n\times n}$ onto $T_X S(r, n)$. Thus defining an inner product on $T_X S(r, n)$ is equivalent to defining an inner product on $K^\perp$. For $\{\Delta_1, X\}, \{\Delta_2, X\} \in T_X S(r, n)$, set

$$\langle\langle\{\Delta_1, X\}, \{\Delta_2, X\}\rangle\rangle := 2\operatorname{tr}((\Delta_1^X)^T \Delta_2^X), \tag{3.3}$$

where Δ_1^X and Δ_2^X are defined by (3.2). The factor of 2 is added purely for convenience. This defines a positive definite, inner product on $T_X S(r, n)$. Since all the constructions are

algebraic it is easily verified that the construction depends smoothly on X and generates a Riemannian metric on $S(r, n)$ (Helmke and Moore, 1994). This metric is referred to as the normal Riemannian metric on $S(r, n)$.

Finally, as $\mathcal{N}_r$ is a submanifold of $S(r, n)$, the restriction of this metric to $\mathcal{N}_r$ is a Riemannian metric on $\mathcal{N}_r$.

4. Gradient Flow

In this section Problem 2.5 is considered and a gradient descent flow of the cost Φ on the smooth manifolds $\mathcal{N}_r$ introduced. Existence and uniqueness of solutions of the flow are established along with some convergence properties.

Theorem 4.1. *Given $A_0, A_1, \ldots, A_m \in \mathbb{R}^{n\times n}$ and $c_1, \ldots, c_m \in \mathbb{R}$ satisfying Assumption* 2.2, *let $X \in \mathcal{N}_r$ for some r, $1 \le r \le n$. Then there is a unique solution $(d_1, \ldots, d_m)^T$ to the linear equation*

$$\begin{pmatrix} \operatorname{tr}(A_1A_1XX) & \cdots & \operatorname{tr}(A_1A_mXX) \\ \vdots & & \vdots \\ \operatorname{tr}(A_mA_1XX) & \cdots & \operatorname{tr}(A_mA_mXX) \end{pmatrix} \begin{pmatrix} d_1 \\ \vdots \\ d_m \end{pmatrix} = - \begin{pmatrix} \operatorname{tr}(A_1A_0XX) \\ \vdots \\ \operatorname{tr}(A_mA_0XX) \end{pmatrix} \tag{4.1}$$

and the gradient of $\Phi(X) = \operatorname{tr}(A_0X)$ with respect to the normal Riemannian metric (3.3) *is given by*

$$\begin{aligned} \operatorname{grad}\Phi(X) &:= \{A_0X + d_1A_1X + \cdots + d_mA_mX, X\} \\ &= A_0XX + XXA_0 + \sum_{i=1}^{m} d_i(A_iXX + XXA_i). \end{aligned} \tag{4.2}$$

Proof. For a unique solution to (4.1) to exist it is sufficient to show that

$$D(X) = \begin{pmatrix} \operatorname{tr}(A_1A_1XX) & \cdots & \operatorname{tr}(A_1A_mXX) \\ \vdots & & \vdots \\ \operatorname{tr}(A_mA_1XX) & \cdots & \operatorname{tr}(A_mA_mXX) \end{pmatrix}$$

is full rank. Observing that

$$\operatorname{tr}(A_iA_jXX) = \operatorname{tr}((A_iX)^TA_jX) = (\operatorname{vec}(A_iX))^T \operatorname{vec}(A_jX),$$

it can be verified that

$$D(X) = W^T(X \otimes I)(X \otimes I)W.$$

Recall that $W^T(X \otimes I)$ is full rank for all $X \in \mathcal{N}_r$ and hence $D(X)$ is full rank.

The gradient of $\Phi\colon \mathcal{N}_r \to \mathbb{R}$ with respect to the normal Riemannian metric is the unique vector field grad Φ which satisfies the following conditions:

(i) $\operatorname{grad}\Phi(X) \in T_X\mathcal{N}_r$ for all $X \in \mathcal{N}_r$.
(ii) $D\Phi|_X(\{\Delta, X\}) = \langle\langle \operatorname{grad}\Phi(X), \{\Delta, X\}\rangle\rangle$ for all $\{\Delta, X\} \in T_X\mathcal{N}_r$.

The first of these conditions implies that, for all $X \in \mathcal{N}_r$,

$$\operatorname{grad}\Phi(X) = \{\Omega, X\}$$

for some $\Omega \in \mathbb{R}^{n\times n}$ which possibly depends on X. In addition $\operatorname{grad}\Phi(X)$ must also satisfy

$$\operatorname{tr}(A_i \ \operatorname{grad}\Phi(X)) = 0 \qquad \text{for} \quad i = 1, \ldots, m. \tag{4.3}$$

Consider

$$\Omega = A_0X + d_1A_1X + \cdots + d_mA_mX, \tag{4.4}$$

where $d_1, \ldots, d_m$ are given by (4.1). With Ω defined by (4.4) it is straightforward to show that $\operatorname{grad}\Phi(X) = \{\Omega, X\}$ satisfies (4.3) and hence that $\operatorname{grad}\Phi(X) = \{\Omega, X\}$ satisfies condition (i).

The derivative of Φ at X is

$$D\Phi|_X(\{\Delta, X\}) = \operatorname{tr}(A_0\{\Delta, X\}).$$

Condition (ii) requires

$$\begin{aligned}\operatorname{tr}(A_0\{\Delta, X\}) &= \langle\langle \operatorname{grad}\Phi(X), \{\Delta, X\}\rangle\rangle \\ &= \langle\langle \{\Omega, X\}, \{\Delta, X\}\rangle\rangle \\ &= 2\operatorname{tr}((\Omega^X)^T\Delta^X)\end{aligned}$$

for all $\{\Delta, X\} \in T_X\mathcal{N}_r$.

We now show that $\Omega = \Omega^X$. Let $\Lambda \in K$. Then

$$\begin{aligned}\operatorname{tr}(\Omega^T\Lambda) &= \operatorname{tr}((XA_0 + d_1XA_1 + \cdots + d_mXA_m)\Lambda) \\ &= \operatorname{tr}(\Lambda XA_0 + d_1\Lambda XA_1 + \cdots + d_m\Lambda XA_m) \\ &= \tfrac{1}{2}\operatorname{tr}((\Lambda X + X\Lambda^T)A_0 + d_1(\Lambda X + X\Lambda^T)A_1 + \cdots + d_m(\Lambda X + X\Lambda^T)A_m) \\ &= \tfrac{1}{2}\operatorname{tr}(\{\Lambda, X\}A_0 + d_1\{\Lambda, X\}A_1 + \cdots + d_m\{\Lambda, X\}A_m) \\ &= 0 \qquad \text{as} \quad \Lambda \in K\end{aligned}$$

and hence $\Omega \in K^\perp$ and $\Omega = \Omega^X$. This implies $\operatorname{tr}((\Omega^X)^T\Delta^X) = \operatorname{tr}(\Omega^T\Delta)$. Finally we show that $2\operatorname{tr}(\Omega^T\Delta) = \operatorname{tr}(A_0\{\Delta, X\})$ for all $\{\Delta, X\} \in T_X\mathcal{N}_r$. Let $\{\Delta, X\} \in T_X\mathcal{N}_r$. Then

$$\begin{aligned}2\operatorname{tr}(\Omega^T\Delta) &= 2\operatorname{tr}((XA_0 + d_1XA_1 + \cdots + d_mXA_m)\Delta) \\ &= \operatorname{tr}(A_0(\Delta X + X\Delta^T) + d_1A_1(\Delta X + X\Delta^T) + \cdots + d_mA_m(\Delta X + X\Delta^T)) \\ &= \operatorname{tr}(A_0\{\Delta, X\} + d_1A_1\{\Delta, X\} + \cdots + d_mA_m\{\Delta, X\}) \\ &= \operatorname{tr}(A_0\{\Delta, X\}) \qquad \text{as} \quad \operatorname{tr}(A_i\{\Delta, X\}) = 0 \quad \text{for} \quad i = 1, \ldots, m.\end{aligned}$$

This completes the proof. □

An important aspect of this construction is that the algebraic representation of the gradient, (4.2), as a function from $\mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$ is independent of the rank of X. Thus, except at singular points, it is possible to consider the algebraic equation

$$\dot{X} = -\left\{ A_0 X + \sum_{i=1}^{m} d_i A_i X, X \right\} \tag{4.5}$$

as a differential equation on $\mathcal{M}$ (see (2.5)). There are several advantages in this interpretation of the problem. In particular, the fact that the flow will be defined on a closed (and, in most cases of interest, compact) set. However, before proceeding with the analysis it is necessary to consider how to deal with singular points should they occur.

Observe that, whenever the matrices $A_1, \ldots, A_m$ satisfy the linear independence requirement of Assumption 2.2, if $X > 0$ the matrix given by (3.1) is always full rank.[2] Thus, if a singular point does occur it will always occur on the boundary of the positive definite cone.

Suppose X is a non-singular point that approaches a singular point X_s via some continuous path. Consider the coefficients $d_1, \ldots, d_m$ defined by (4.1). These d_i's are in fact simply projection coefficients that ensure $\operatorname{tr}(A_i \operatorname{grad} \Phi(X)) = 0$, $i = 1, \ldots, m$. Hence the functions $d_1, \ldots, d_m$ must be continuous. Now, even though the matrix $D(X)$ becomes singular as $X \to X_s$, the gradient direction $-\operatorname{grad} \Phi(X)$ remains bounded and has a well-defined limit. Consequently, along any solution $X(t)$ of the gradient flow (4.5) for which there exists a time $t_s > 0$ such that $X(t_s) = X_s$, the continuous limit of the gradient

$$\operatorname{grad} \Phi(X_s) = \lim_{t \to t_s} \operatorname{grad} \Phi(X(t)) \tag{4.6}$$

exists. Since the solution up to this point is unique, the extension of the gradient field in this manner is unique for a given initial condition. Of course if a different initial condition is chosen, the gradient extension may be different. Considering a single initial condition and applying classical existence and uniqueness theory of ordinary differential equations, it follows that a solution of (4.5), extended via (4.6), must continue to exist beyond time t_s. The convention of choosing the gradient extension as mentioned above ensures that the solution obtained in this way is unique and (due to continuity) will continue to satisfy the constraints that preserve the solution in $\mathcal{M}_r$. Since the cost is analytic, the solution will pass through the singular surface instantaneously at time t_s and then continue evolving in $\mathcal{N}_r$.

Recall the definition of $\mathcal{M}$ (see (2.5)). The set $\mathcal{M}$ is a closed subset of the symmetric positive semidefinite matrices. While it can be readily verified that $\mathcal{M}$ is convex, $\mathcal{M}$ may not be bounded and hence there may exist trajectories $X(t)$ which are unbounded. Such solutions correspond to the case where the infimum of the cost $\Phi(X) = \operatorname{tr}(A_0 X)$ over $\mathcal{M}$ is negative infinity. This possibility is of little interest and its analysis is beyond the scope of the present paper. In what follows we assume that the cost Φ is bounded below on $\mathcal{M}$. Indeed, we assume the slightly stronger condition, which considerably simplifies

[2] If $A, B \in \mathbb{R}^{n\times n}$ and $\lambda_1, \ldots, \lambda_n$ and $\mu_1, \ldots, \mu_n$ are the eigenvalues of A and B, respectively, the eigenvalues of $A \otimes B$ are $\lambda_i \mu_j$ for all i, j. Hence if X is invertible, so is $X \otimes I$.

the analysis, that the cost Φ has compact sublevel sets, i.e., that, for each $c \in \mathbb{R}$, the set $\{X \in \mathcal{M} \mid \Phi(X) \leq c\}$ is a (possibly empty) compact subset of $\mathcal{M}$.

Remark 4.2. A common condition encountered in practice is that one or more of the constraint matrices $A_1, \ldots, A_m$ is positive definite. In this case it can be shown that $\mathcal{M}$ is compact and hence that Φ has compact sublevel sets.

Theorem 4.3. *Given $A_0, A_1, \ldots, A_m \in \mathbb{R}^{n\times n}$ and $c_1, \ldots, c_m \in \mathbb{R}$ satisfying Assumption* 2.2, *let $X(0) = X_0 \in \mathcal{M}$ be any non-singular point* (*see Definition* 3.2). *Assume Φ has compact sublevel sets. Then the solution $X(t)$ to*

$$\begin{aligned}\dot{X} &= -\operatorname{grad}\Phi(X)\\ &= -A_0XX - XXA_0 - \sum_{i=1}^{m} d_i(A_iXX + XXA_i),\end{aligned} \tag{4.7}$$

extended to all points in $\mathcal{M}$ via the extension (4.6), *satisfies*:

(i) *The solution $X(t)$ exists and is unique for all time $t \geq 0$ and remains in $\mathcal{M}$.*
(ii) *The rank of the solution* $\operatorname{rank}(X(t))$ *remains constant for all time.*
(iii) *The equilibria of* (4.7) *are characterized by those points $X \in \mathcal{M}$ such that*

$$(A_0 + d_1A_1 + \cdots + d_mA_m)X = 0. \tag{4.8}$$

(iv) *The cost $\Phi(X(t))$ is a monotonically decreasing function of time. The solutions $X(t)$ converge to a connected component of the set of equilibria given by* (4.8).

Proof. Existence and uniqueness of the solution of (4.7) over a small time around a point $X(t)$ is guaranteed by classical ODE theory and the discussion of the extension of the flow onto singular points. (Note that singular initial conditions, for which the extension (4.6) is unclear, are excluded from consideration.) The cost Φ having compact sublevel sets implies that the solution remains bounded and it follows that $X(t)$ exists and is unique for all time $t \geq 0$.

The preservation of the rank of $X(t)$ for all time is a direct consequence of the local existence results and the fact that locally $X(t)$ remains in $\mathcal{M}_r$.

To verify the characterization of the critical points consider a critical point $X \in \mathcal{M}$. Then

$$\operatorname{grad}\Phi(X) = \{\mathcal{A}X, X\} = 0, \tag{4.9}$$

where $\mathcal{A} := A_0 + d_1A_1 + \cdots + d_mA_m$. Expanding (4.9) and multiplying by $\mathcal{A}$ on the left implies

$$\mathcal{A}\mathcal{A}XX + \mathcal{A}XX\mathcal{A} = 0.$$

Taking the trace of the above equation and using properties of the trace operator implies that

$$\operatorname{tr}\left((\mathcal{A}X)^T(\mathcal{A}X)\right) = 0$$

and hence that $\mathcal{A}X = 0$. Substituting for $\mathcal{A}$ in this equation recovers the characterization in the theorem statement. Sufficiency of the condition is easily verified.

Part (iv) is a consequence of the gradient nature of the flow on all but a set of measure zero in $\mathcal{M}$. The convergence result follows by considering $\Phi(X)$ as a Lyapunov function for the flow. □

5. Further Analysis

In this section a phase portrait analysis of the flow (4.7) is developed.

In the generic case that Φ is not constant over $\mathcal{M}$, any critical point of Problem 2.6 will occur on the boundary of the positive definite cone and hence will be rank degenerate. This follows in a straightforward manner from the cost and constraint functions being linear.

Remark 5.1. The situation that the cost functional Φ is constant over $\mathcal{M}$ will occur if one has a degenerate situation such as A_0 being a linear combination of the constraint matrices $A_1, \ldots, A_m$.

In addition, any minima of Problem 2.6 are global minima and the set of all such minima form a convex subset of $\mathcal{M}$ (Luenberger, 1969). In fact, generically, there is a unique solution to Problem 2.6 (Alizadeh et al., 1996).

Classical dynamical systems theory now ensures that the attractive basin of the set of global minima is almost all of the set $\mathcal{M}$. Indeed, the authors believe that the attractive sets of the non-minimal critical points are confined to the boundary of the positive definite cone, however, we have no satisfactory proof for this claim.

It is of interest to study more carefully the characteristics of the critical points. Consider (4.8). In the case where there is a single constraint, the critical point condition becomes a standard generalized eigenvalue problem:

$$A_0 X = -d_1 A_1 X.$$

Furthermore, when A_1 is positive definite, the one constraint case can be solved analytically. Suppose $A_1 > 0$ and recall the original problem statement, Problem 2.1, with $m = 1$ constraints: *minimize* $x^T A_0 x$ *subject to* $x^T A_1 x = c_1$. As $A_1 > 0$, there exists an invertible, symmetric square root of A_1, denoted $A_1^{1/2}$. Define $y := c_1^{-1/2} A_1^{1/2} x$ and $A_0' := c_1 A_1^{-1/2} A_0 A_1^{-1/2}$. Then the optimization problem can be rewritten as *minimize* $y^T A_0' y$ *subject to* $y^T y = 1$. The solution to this problem is well known: it is the unit length eigenvector corresponding to the minimal eigenvalue of A_0'. However, this is just the same as claiming that x is the solution corresponding to the smallest eigenvalue of the generalized eigenvalue problem for the matrix pair A_0, A_1.

Denoting the minimal generalized eigenvalue of A_0, A_1 by λ, one has the following conditions for $X = xx^T$ to be a global minimum of Problem 2.6:

(i) $(A_0 - \lambda A_1)x = 0$,

(ii) $x^T A_1 x = c_1$,
(iii) $A_0 - \lambda A_1 \geq 0$.

Equations (i) and (iii) are solved by choosing x as the eigenvector corresponding to the minimal generalized eigenvalue of A_0, A_1. Equation (ii) can be satisfied by scaling x.

In fact, in the semidefinite programming literature it is shown that this characterization of the global minima of the single constraint case generalizes to conditions for a global minima in the multi-constraint case.

Theorem 5.2. *$X \in \mathcal{M}$ is an optimal solution to Problem* 2.6 *if and only if*

(i) $(A_0 + d_1 A_1 + \cdots + d_m A_m)X = 0$,
(ii) $\mathrm{tr}(A_i X) = c_i,\ i = 1, \ldots, m$,
(iii) $A_0 + d_1 A_1 + \cdots + d_m A_m \geq 0$.

Proof. See Alizadeh et al. (1996). □

The final issue that needs to be considered is the question of whether a minimum of Problem 2.6 on the full set $\mathcal{M}$ relates to a minima of Problem 2.1, the original problem on $\mathbb{R}^n$. Unfortunately, a minimum of Problem 2.6 will not always be rank 1 and hence one is not always able to solve Problem 2.1 directly from the solution of Problem 2.6.

The next theorem gives upper and lower bounds on the rank of X. The result is taken from Alizadeh et al (1996). Before proceeding we introduce some notation. Let

$$n^{\overline{2}} := n(n+1)/2.$$

For $h \geq 0$, let $\lfloor h \rfloor$ denote the largest integer less than or equal to h. Define

$$\sqrt[\overline{2}]{k} = \lfloor h \rfloor \qquad \text{where } h \text{ is the positive real root of} \quad h^{\overline{2}} = k.$$

Theorem 5.3. *If X is an minima of Problem* 2.6, *then, generically,*

$$n - \sqrt[\overline{2}]{n^{\overline{2}} - m} \leq \mathrm{rank}(X) \leq \sqrt[\overline{2}]{m}.$$

Proof. See Alizadeh et al. (1996). □

An example of the sorts of bounds produced by Theorem 5.3 are given in Table 5.1 (Alizadeh et al., 1996). Bounds are given for $n = 20$ and various values of m.

Remark 5.4. From Theorem 5.3 it follows that, generically, if $n \geq 2$, then the solution to Problem 2.6 is rank 1 if $m \leq 2$.

Table 5.1. Generic bounds on $\text{rank}(X)$ for $n = 20$.

m	Bounds on $r = \text{rank}(X)$
10	$1 \le r \le 4$
20	$1 \le r \le 5$
30	$2 \le r \le 7$
40	$3 \le r \le 8$
50	$3 \le r \le 9$

6. Gradient Flow with Penalty Function

In this section we consider a modified version of the flow that incorporates a penalty function designed to encourage the solution of the flow to converge to a rank 1 matrix.

Consider the cost function

$$\Omega(X) = ||X||_F^2 - ||X||_2^2, \tag{6.1}$$

where $||X||_F = \text{tr}(X^2)^{1/2}$ is the Frobenius norm of X and $||X||_2 = \max_{||v||=1} ||Xv||$ is the 2-norm of X. If $0 \le \lambda_1 \le \cdots \le \lambda_n$ are the eigenvalues of X, then $\Omega(X) = \sum_{i=1}^{n-1} \lambda_i^2$ and can be thought of as a measure of how close the matrix X is to being rank 1. If $\Omega(X) \ge 0$ is small, then X is close to being rank 1 and indeed $\Omega(X) = 0$ if and only if X is rank 1 or $X = 0$. (Note that generally $X = 0$ will not be a member of $\mathcal{M}$.)

Remark 6.1. The cost Ω is non-convex on $\mathcal{M}$. All functions which are continuous on $\mathcal{M}$ and are minimized only on the set of rank 1 matrices will be non-convex.

Consider the following optimization problem:

Problem 6.2. Given $A_0, A_1, \ldots, A_m \in \mathbb{R}^{n\times n}$ and $c_1, \ldots, c_m \in \mathbb{R}$ satisfying Assumption 2.2, and $\varepsilon > 0$,

$$\begin{aligned} &\text{minimize } \Theta(X) := \Phi(X) + \varepsilon \log(\Omega(X)) \\ &\text{subject to} \quad X \in \mathcal{M}, \end{aligned}$$

where $\Phi(X) = \text{tr}(A_0 X)$, the cost of Problem 2.6, and $\Omega(X)$ is defined by (6.1).

The term $\varepsilon \log(\Omega(X))$ can be thought of as a penalty function that penalizes solutions of rank larger than 1. Let X_{opt} denote an optimal solution of Problem 2.6 and let X_ε denote a solution of Problem 6.2 for a given $\varepsilon > 0$. By varying ε one can trade off how close $\Phi(X_\varepsilon)$ is to $\Phi(X_{\text{opt}})$ versus how close X_ε is to being rank 1. (Note that $\Phi(X_{\text{opt}}) \le \Phi(X_\varepsilon)$ for all $\varepsilon > 0$.)

In order to solve Problem 6.2 we develop a gradient descent flow in the same way a gradient flow was developed to solve Problem 2.6. We now proceed to do this. Consider again (6.1) which can be rewritten as

$$\Omega(X) = \text{tr}(X^2) - v^T X^2 v,$$

where $v = v(X)$ is the unit length eigenvector corresponding to the maximum eigenvalue of X. In the derivation that follows, we are required to take the derivative of v with respect to X. Though v will generally not be differentiable everywhere, v is differentiable almost everywhere and this is sufficient for our purposes. The derivative of $\Omega(X)$ in direction $\{\Delta, X\} \in T_X\mathcal{M}$ is

$$D\Theta|_X(\{\Delta, X\}) = \operatorname{tr}(A_0\{\Delta, X\}) + \varepsilon \frac{D\Omega|_X(\{\Delta, X\})}{\Omega(X)},$$

where

$$\begin{aligned} D\Omega|_X(\{\Delta, X\}) &= 2\operatorname{tr}(X\{\Delta, X\}) - 2v^T X\{\Delta, X\}v - 2v^T X^2\ Dv|_X(\{\Delta, X\}) \\ &= 2\operatorname{tr}(X\{\Delta, X\}) - 2v^T X\{\Delta, X\}v - 2\lambda_{\max}^2(X)v^T\ Dv|_X(\{\Delta, X\}). \end{aligned}$$

As $v^T v = 1$, it follows that $v^T\ Dv|_X(\{\Delta, X\}) = 0$. Hence

$$D\Omega|_X(\{\Delta, X\}) = 2\operatorname{tr}((X - vv^T X)\{\Delta, X\})$$

and

$$D\Theta|_X(\{\Delta, X\}) = \operatorname{tr}\left(\left(A_0 + \frac{2\varepsilon(X - vv^T X)}{\Omega(X)}\right)\{\Delta, X\}\right).$$

This implies that the gradient flow that solves Problem 6.2 is the same as the one that solves Problem 2.6, see (4.7), if one replaces A_0 with

$$A_0^\varepsilon := A_0 + \frac{2\varepsilon(X - vv^T X)}{\Omega(X)}.$$

Note that A_0^ε is a function of X. Explicitly, the flow is

$$\dot{X} = -\left\{A_0^\varepsilon X + \sum_{i=1}^m d_i^\varepsilon A_i X, X\right\}, \tag{6.2}$$

where $d_1^\varepsilon, \ldots, d_m^\varepsilon$ satisfy

$$\begin{pmatrix} \operatorname{tr}(A_1 A_1 X X) & \cdots & \operatorname{tr}(A_1 A_m X X) \\ \vdots & & \vdots \\ \operatorname{tr}(A_m A_1 X X) & \cdots & \operatorname{tr}(A_m A_m X X) \end{pmatrix} \begin{pmatrix} d_1^\varepsilon \\ \vdots \\ d_m^\varepsilon \end{pmatrix} = -\begin{pmatrix} \operatorname{tr}(A_1 A_0^\varepsilon X X) \\ \vdots \\ \operatorname{tr}(A_m A_0^\varepsilon X X) \end{pmatrix}.$$

7. Solution Methods

If Problem 2.6 has a rank 1 optimal solution, a standard semidefinite programming algorithm can be used to find the optimum solution of Problem 2.6 efficiently and hence solve Problem 2.1. In this section we discuss how the modified flow (6.2) can be used to solve Problem 2.1 in the case that Problem 2.6 does not have a rank 1 solution.

Before proceeding we consider the problem of initial conditions. Before a semidefinite program or a gradient flow can be used, a matrix $X = X^T > 0$ satisfying $\mathrm{tr}(A_i X) = c_i$, $i = 1, \ldots, m$, must be found. This is a standard problem in semidefinite programming and many methods of solving this problem exist, see, for example, pp. 86–88 of Vandenberghe and Boyd (1996).

Let X_{opt} denote an optimum solution of Problem 2.6. One method of solving Problem 2.1 is to find X_{opt} using semidefinite programming techniques and then to use X_{opt} as an initial condition for the flow (6.2). Starting from X_{opt} is an intuitively appealing idea and this method has been found to work well in practice. Let $X_{\mathrm{rank}\ 1}$ denote the limit of the flow (6.2). By choosing ε sufficiently small, one can ensure that, to any desired computational accuracy, $X_{\mathrm{rank}\ 1}$ is indeed rank 1. Note that $\Phi(X_{\mathrm{opt}})$ provides a lower bound on $\Phi(X_{\mathrm{rank}\ 1})$. Indeed, if the difference $|\Phi(X_{\mathrm{opt}}) - \Phi(X_{\mathrm{rank}\ 1})|$ is small, then one can be confident that a good solution to Problem 2.1 has been obtained. For many applications, obtaining a feasible solution which is guaranteed to be within a known small margin of the actual optimal cost provides a good practical solution. The authors know of no other "tractable" numerical method that yields this information for this problem.

Simulation Example. The following is a typical simulation example with $n = 20$ and $m = 10$. We denote the initial feasible point by X_0, the optimal solution of Problem 2.6 by X_{opt} and the limit of the flow by $X_{\mathrm{rank}\ 1}$. In this particular simulation the optimal solution was found to be rank 2. The costs and eigenvalue spectrums of X_0, X_{opt} and $X_{\mathrm{rank}\ 1}$ are displayed in Table 7.1 and Figure 7.1, respectively. It is interesting to note that the flow does not appear to change the eigenvalue spectrum of X_{opt} except to reduce those eigenvalues that were associated with additional rank.

The method described above is only one way of using the flow (6.2). Another method would be the following. Starting with an initial feasible solution, one could start the flow with $\varepsilon \approx 0$. Once the flow is close to converging for this value of ε (convergence could be monitored by checking how close the norm of the gradient of the flow is to zero), ε could be increased by a certain small amount and the flow allowed to evolve further. Again, once the flow is close to converging, ε could be increased further and the whole process repeated until a solution was obtained with the $n - 1$ smallest eigenvalues of X sufficiently small. In this manner the penalty function can be thought of as a sort of pseudobarrier function. This method has not been tried in practice.

An area for possible future research could be to try to develop a more efficient method of solving the modified gradient flow. Using a numerical ODE solver to solve this flow is computationally quite expensive. Ideally, one would like to develop an explicit

Table 7.1. A cost comparison.

X	Cost $\Phi(X)$
X_0	1.6000
X_{opt}	−3.8126
$X_{\mathrm{rank}\ 1}$	−3.1908

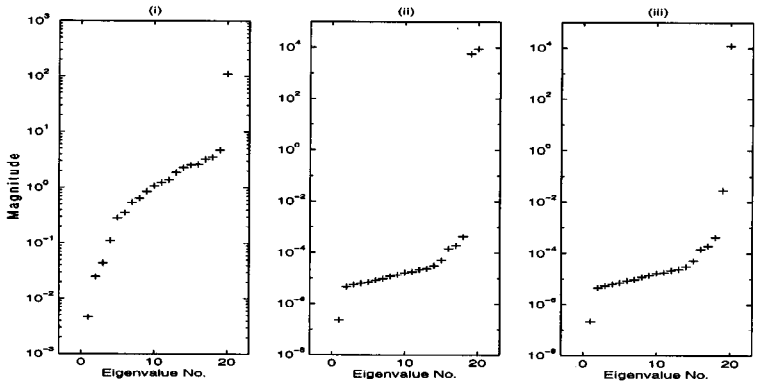

Figure 7.1. The eigenvalue spectrums of (i) X_0, (ii) X_{opt} and (iii) $X_{\text{rank } 1}$.

numerical scheme for solving the flow. For the present, consider again the flow

$$\dot{X} = -\{\mathcal{A}X, X\}, \tag{7.1}$$

where $\mathcal{A} = A_0 + d_1 A_1 + \cdots + d_m A_m$. Exploiting the homogeneous structure of $S(r, n)$ (see Section 2), the matrix X can be written as $X = SX_0S^T$ where $S \in GL(n, \mathbb{R})$ and $X_0 \in \mathbb{R}^{n\times n}$ satisfies $X_0 = X_0^T \geq 0$. Substituting $X = SX_0S^T$ into (7.1) produces

$$\{\dot{S}, X_0S^T\} = -\{\mathcal{A}SX_0S^TS, X_0S^T\}.$$

Instead of a flow on X, one can now consider the following flow:

$$\dot{S} = -\mathcal{A}SX_0S^TS$$

on $S(t) \in GL(n, \mathbb{R})$. This S flow can be thought of as a sort of square root version of the original flow and one would expect it to be numerically better conditioned. In fact, it has a number of other numerical advantages over the old flow, for while in theory the solution of (7.1), $X(t)$, will always be positive semidefinite, numerical inaccuracies may possibly lead to $X(t)$ leaving the positive semidefinite cone. (Negative eigenvalues of $X(t)$ often lead to unstable numerical behaviour.) Conversely, the S flow guarantees that $X(t)$ will remain positive semidefinite. Furthermore, if the solution method encounters numerical problems due to $S(t)$ (and hence $X(t)$) becoming singular, it could be re-initialized. The new value of X_0 could be set to the current value of $X(t) = S(t)X_0S(t)^T$ and $S(t)$ could be re-initialized to the identity matrix.

8. Conclusion

In this paper we have provided a dynamical systems analysis of semidefinite programming. We have also developed methods of minimizing a quadratic cost subject to purely

quadratic constraints based on a gradient descent flow incorporating a penalty function. One of these methods was simulated and found to work well in practice. Despite the encouraging results, at present it is not known whether the methods developed will always find the optimal solution. Further analysis is required in this area.

References

1. Alizadeh, F. (1995). Interior point methods in semidefinite programming with application to combinatorial optimization, SIAM Journal on Optimization 5:13–51.
2. Alizadeh, F., Haeberly, J. P., and Overton, M. (1996). Complementarity and nondegeneracy in semidefinite programming, Mathematical Programming (Series B). To appear.
3. Boothby, W. (1986). An Introduction to Differentiable Manifolds and Riemannian Geometry, 2nd edn, Academic Press, San Diego, CA.
4. Faybusovich, L. E. (1991). Dynamical systems which solve optimization problems with linear constraints, IMA Journal of Information and Control 8:135–149.
5. Fletcher, R. (1987). Practical Methods of Optimization, 2nd edn, Wiley, Chichester.
6. Gibson, C. G. (1979). Singular Points of Smooth Mappings, Pitman, London.
7. Helmke, U., and Moore, J. B. (1994). Optimization and Dynamical Systems, Springer-Verlag, London.
8. Hirsch, M. W. (1976). Differential Topology, Springer-Verlag, New York.
9. Luenberger, D. (1969). Optimization by Vector Space Methods, Wiley, New York.
10. Nesterov, Y., and Nemirovskii, A. (1994). Interior-Point Polynomial Algorithms in Convex Programming, SIAM, Philadelphia, PA.
11. Thng, I., Cantoni, A., and Leung, Y. (1996). Analytical solutions to the optimization of a quadratic cost function subject to linear and quadratic equality constraints, Applied Mathematics & Optimization 34(2):161–182.
12. Tits, A. L., and Zhou, J. L. (1993). A simple, quadratically convergent interior point algorithm for linear programming and convex quadratic programming, in W. W. Hager, D. W. Hearn and P. M. Pardalos (eds), Large Scale Optimization: State of the Art, Kluwer, Dordrecht.
13. Vandenberghe, L., and Boyd, S. (1996). Semidefinite programming, SIAM Review 38(1):49–95.

Accepted 10 *March* 1998