Journal of Global Optimization **14:** 105–115, 1999.

Computable Error Bounds for Semidefinite Programming

SIEN DENG and HUI HU
Department of Mathematical Sciences, Northern Illinois University, DeKalb, IL 60115, USA

(Received 24 April 1997; accepted in revised form 30 March 1998)

Abstract. We study computability and applicability of error bounds for a given semidefinite programming problem under the assumption that the recession function associated with the constraint system satisfies the Slater condition. Specifically, we give computable error bounds for the distances between feasible sets, optimal objective values, and optimal solution sets in terms of an upper bound for the condition number of a constraint system, a Lipschitz constant of the objective function, and the size of perturbation. Moreover, we are able to obtain an exact penalty function for semidefinite programming along with a lower bound for penalty parameters. We also apply the results to a class of statistical problems.

Key words: Error bounds, Patterned covariance matrices, Penalty functions, Semidefinite programming, Sensitivity analysis

1. Introduction

Given $n+1$ real $m \times m$ symmetric matrices $B_0, B_1, \ldots, B_n$, a general semidefinite programming problem (SDP) can be defined as

$$\text{(SDP)} \quad \text{minimize} \quad f(x) \qquad \text{subject to} \quad B(x) + B_0 \geqslant 0,$$

where $x \in \mathbb{R}^n$, f is a finite convex function, $B(x) = \sum_{i=1}^{n} (x)_i B_i$. The inequality $B(x)+B_0 \geqslant 0$ means that $B(x)+B_0$ is positive semidefinite [1, 2]. Since numerical methods for solving (SDP) in general can only provide approximate solutions, it is important to study a perturbed problem

$$\text{(P)} \quad \text{minimize} \quad f(x) \qquad \text{subject to} \quad B(x) + B_0 - P \geqslant 0,$$

where $P \in \mathbb{R}^{m\times m}$ is a symmetric matrix. For instance, if $P = \epsilon I$ with I the identity matrix and $\epsilon > 0$, then $B(x) + B_0 - P \geqslant 0$ is equivalent to $\lambda_{\min}[B(x) + B_0] \geqslant \epsilon$, where '$\lambda_{\min}$' indicates the smallest eigenvalue of a symmetric matrix, and a solution to this perturbed problem is a suboptimal solution to (SDP).

In this paper we focus on computability and applicability of error bounds for (SDP). We give error bounds for the distances between the feasible sets, the optimal values, and the optimal solution sets of (SDP) and (P) in terms of a Lipschitz constant of the objective function, the size of perturbation, and an upper bound for the condition number of the convex system that defines the feasible set of (SDP). Finally, we use the results to obtain an exact penalty function for (SDP) along with a lower bound for penalty parameters, and we apply the results to a class of statistical problems.

The constraint system for (SDP) is equivalent to a convex inequality system. Error bounds for convex inequality systems have received increasing attention in the recent literature due to their important roles in the sensitivity analysis of convex programming and the convergence analysis of iterative algorithms [3–6]. The paper of Lewis and Pang [6] gives an excellent survey on this active research area, and in which one can find many very interesting new results.

Throughout this paper, we use the following notation.

For any symmetric matrix P, let $S(P)$ be the feasible set of problem (P), and $\rho(x, S(P)) = \max\{-\lambda_{\min}[B(x) + B_0 - P], 0\}$. It is clear that $S(P) = \{x \mid -\lambda_{\min}[B(x) + B_0 - P] \leqslant 0\}$. Let $V(P)$, $X(P)$, and $X_\epsilon(P)$ denote the optimal value, optimal solution set, and ϵ-optimal solution set of (P) respectively, i.e., $V(P) = \inf\{f(x) \mid x \in S(P)\}$, $X(P) = \arg\min\{f(x) \mid x \in S(P)\}$, and $X_\epsilon(P) = \{x \in S(P) \mid f(x) \leqslant V(P) + \epsilon\}$, where $\epsilon > 0$. We will use the vector 2-norm and the matrix 2-norm for a vector and a matrix respectively, which are denoted by $\|\cdot\|$. The unit ball in $\mathbb{R}^n$ is denoted by $\mathbb{B}^n$, and the ball with radius γ is denoted by $\gamma\mathbb{B}^n$. For any nonempty closed convex set S, let $d(x, S) = \min\{\|x - y\| \mid y \in S\}$ denote the Euclidean distance from any x to S. Given nonempty closed convex sets S_1 and S_2 in $\mathbb{R}^n$, the Hausdorff distance between them is defined as

$$d_\infty(S_1, S_2) = \max\left\{\sup_{x\in S_1} d(x, S_2), \sup_{x\in S_2} d(x, S_1)\right\}.$$

When $S(P)$ is nonempty, the condition number of the system $-\lambda_{\min}[B(x) + B_0 - P] \leqslant 0$ is defined as [7, 8]

$$K(P) = \sup_{x\notin S(P)} \frac{d(x, S(P))}{-\lambda_{\min}[B(x) + B_0 - P]}.$$

When $K(P)$ is finite, it is the smallest positive scalar τ such that the following global error bound for $S(P)$ holds.

$$d(z, S(P)) \leqslant \tau\rho(x, S(P)), \quad \text{for all } z \in \mathbb{R}^n. \tag{1}$$

The above inequality bounds the distance between any point $z \in \mathbb{R}^n$ and $S(P)$ in terms of a constant multiple of the residual function $\rho(z, S(P))$. When $K(P)$ is 'small', for any z, $d(z, S(P))$ is small whenever $\rho(z, S(P))$ is small. The following example shows that $K(P)$ can be less than 1. When $B(x) + B_0 = 2x + 1$, it is easy to see that $K(P) = 1/2$.

2. Error bounds

We start with a general global error bound result for the following convex inequality system.

$$h(x) \leqslant 0, \quad x \in C, \tag{2}$$

where $h : \mathbb{R}^n \to \mathbb{R}$ is a finite convex function, and C is a nonempty closed convex set. Let $\tilde{S}$ be the set of solutions to the system (2), and suppose that $\tilde{S}$ is nonempty. The following proposition provides sufficient conditions under which $d(z, \tilde{S})$ can be bounded by the sum of *computable* constant multiples of $d(z, C)$ and $[h(z)]_+$, where $[a]_+ = \max\{a, 0\}$.

PROPOSITION 2.1. *Consider the system (2). Suppose that h is Lipschitz continuous on $\mathbb{R}^n$ with a Lipschitz constant l, and that there is a positive scalar γ such that*

$$d(z, \tilde{S}) \leqslant \gamma[h(z)]_+ \quad \textit{for all } z \in C. \tag{3}$$

Then

$$d(z, \tilde{S}) \leqslant (\gamma l + 1)d(z, C) + \gamma[h(z)]_+ \quad \textit{for all } z \in \mathbb{R}^n. \tag{4}$$

In particular, if there exists a unit vector $u \in C^\infty$ and a positive scalar τ such that $h^\infty(u) \leqslant -\tau^{-1}$, where h^∞ and C^∞ denote the recession function of h and the recession cone of C respectively, then (3) and (4) hold with $\gamma = \tau$.

Proof. For any $z \in \mathbb{R}^n$ but not in $\tilde{S}$, let z_C be the projection of z onto C, and $\bar{z}$ be the projection of z_C onto $\tilde{S}$. It is clear that

$$d(z, \tilde{S}) \leqslant \|z - \bar{z}\| \leqslant \|z - z_C\| + \|z_C - \bar{z}\| = d(z, C) + d(z_C, \tilde{S}).$$

Since $z_C \in C$, by (3), we have

$$d(z_C, \tilde{S}) \leqslant \gamma[h(z_C)]_+ \leqslant \gamma[h(z)]_+ + \gamma l\|z - z_C\| = \gamma[h(z)]_+ + \gamma l d(z, C),$$

where the second inequality follows from the fact that l can be used as a Lipschitz constant for $[h(\cdot)]_+$ on $\mathbb{R}^n$. Thus (4) holds. The last part follows from Theorem 2.3 in [9]. $\square$

REMARK . (a) Recently, Lewis and Pang have proved the existence of a global error bound of the kind (4) under weaker assumptions than those in Proposition 2.1, (Corollary 2(e) in [6]).

(b) A Lipschitz constant l for $h(x) = -\lambda_{\min}[B(x)+B_0]$ is $\sqrt{n}\max_{1\leqslant i\leqslant n}\{\|B_i\|_F\}$, where $\|\cdot\|_F$ is the Frobenious norm (Corollary A.5 in [10]).

Motivated by a class of practical problems in statistics [11], we make the following assumption on the constraint system. Under this assumption, Proposition 2.1 can be used to derive useful error bounds for (SDP).

ASSUMPTION 1. There exists a unit vector $\hat{x} \in \mathbb{R}^n$ such that $\lambda_{\min}[B(\hat{x})] = \tau^{-1} > 0$.

$-\lambda_{\min}[B(\cdot)]$ is positively homogeneous and convex [10]. Assumption 1 implies that the recession function $-\lambda_{\min}[B(\cdot)]$ of $-\lambda_{\min}[B(\cdot) + B_0]$ satisfies the Slater condition. Thus $S(P)$ is unbounded for any symmetric matrix P. For patterned covariance matrix problems [11], one can easily find an $\hat{x}$ satisfying Assumption 1 (see examples in Section 3.2). For a general semidefinite programming problem, one can find an $\hat{x}$ satisfying Assumption 1 by minimizing $-\lambda_{\min}[B(\cdot)]$, if such an $\hat{x}$ exits.

A consequence of Proposition 2.1 is the following corollary which can also be derived from Theorem 2 in [12].

COROLLARY 2.1. *Suppose that Assumption 1 is satisfied. Then for any symmetric matrix P,*

$$d\,(x, S(P)) \leqslant \tau\rho(x, S(P)) \quad \text{for all } x \in \mathbb{R}^n.$$

In view of Corollary 2.1, we immediately have an upper bound for the Hausdorff distance between $S(P)$ and $S(Q)$ in terms of τ and the distance between P and Q.

THEOREM 2.1. *If Assumption 1 is satisfied, then for any symmetric matrices P and Q,*

$$d_\infty\,(S(P), S(Q)) \leqslant \tau\|P - Q\|.$$

Proof. It suffices to show that $d(x, S(P)) \leqslant \tau\|P - Q\|$ for any $x \in S(Q)$. From Corollary 2.1,

$$d(x, S(P)) \leqslant \tau\rho(x, S(P)) = \tau \max\{-\lambda_{\min}[B(x) + B_0 - P], 0\}.$$

As $-\lambda_{\min}[\cdot]$ is positively homogeneous and convex, and $\lambda_{\min}[B(x) + B_0 - Q] \geqslant 0$, we have

$$\begin{aligned} -\lambda_{\min}[B(x) + B_0 - P] &\leqslant -\lambda_{\min}[B(x) + B_0 - Q] - \lambda_{\min}[Q - P] \\ &\leqslant -\lambda_{\min}[Q - P] = \lambda_{\max}[P - Q] \\ &\leqslant \|P - Q\|. \end{aligned}$$

It follows that $d(x, S(P)) \leqslant \tau\|P - Q\|$ for any $x \in S(Q)$ and this completes the proof. □

A consequence of Theorem 2.1 on the optimal value of (P) is the following proposition. Note that a finite convex function is Lipschitz continuous on any compact set in $\mathbb{R}^n$ [13, Theorem 10.4].

PROPOSITION 2.2. *If Assumption 1 is satisfied, then for any P, Q with $\|P - Q\| \leqslant \tau^{-1}$ and any bounded set S such that $X(P) \cap S \neq \emptyset$ and $X(Q) \cap S \neq \emptyset$, we have*

$$|V(P) - V(Q)| \leqslant \tau L_{S+\mathbb{B}^n} \|P - Q\|,$$

where $L_{S+\mathbb{B}^n}$ is a Lipschitz constant of f on $S + \mathbb{B}^n$.

Proof. Let $x' \in X(P) \cap S$, $x'' \in X(Q) \cap S$, and $x^* \in S(Q)$ such that

$$\|x' - x^*\| = d(x', S(Q)). \tag{5}$$

By Theorem 2.1, $\|x' - x^*\| \leqslant \tau \|P - Q\| \leqslant 1$ and thus $x^* \in S + \mathbb{B}^n$. It follows that

$$\begin{aligned} V(Q) = f(x'') \leqslant f(x^*) &= f(x') + (f(x^*) - f(x')) \\ &\leqslant V(P) + \tau L_{S+\mathbb{B}^n} \|P - Q\|. \end{aligned}$$

Similarly, $V(P) \leqslant V(Q) + \tau L_{S+\mathbb{B}^n} \|P - Q\|$. The result follows. □

An important class of problems in semidefinite programming are problems with linear objective functions. Now we study this case and assume that $f(x) = c^T x$. We first state the following lemma on the compactness of $X(P)$. Its proof can be easily derived from Theorem 27.1 in [13].

LEMMA 2.1. *If $f(x) = c^T x > 0$ whenever $B(x) \geqslant 0$, then for any symmetric matrix P, the solution set $X(P)$ of* (P) *is nonempty and bounded.*

Since a linear function $c^T x$ has a uniform Lipschitz constant $\|c\|$ for all bounded sets, by Proposition 2.2 and Lemma 2.1, we have the following theorem.

THEOREM 2.2. *If Assumption 1 is satisfied, $\|P - Q\| \leqslant \tau^{-1}$, and $f(x) = c^T x$ satisfies the conditions stated in Lemma 2.1, then*

(a) *for any symmetric matrices P and Q, $|V(P) - V(Q)| \leqslant \tau \|c\| \|P - Q\|$;*

(b) *for any symmetric matrices P and Q with $\|P - Q\| \leqslant \epsilon(4\tau\|c\|)^{-1}$, any given $\epsilon > 0$, and $0 < \eta < \epsilon/2$, $\sup_{x \in X_\eta(Q)} d(x, X_\epsilon(P)) \leqslant \tau \|P - Q\|$.*

Proof. Part (a) follows from Proposition 2.2 and Lemma 2.1. To prove (b), let $x \in X_\eta(Q)$. From part (a), we have

$$c^T x \leqslant V(Q) + \eta \leqslant V(P) + \tau \|c\| \|P - Q\| + \eta. \tag{6}$$

Choose $x^* \in S(P)$ such that $\|x - x^*\| = d(x, S(P))$. By Theorem 2.1,

$$\|x - x^*\| \leqslant d_\infty(S(Q), S(P)) \leqslant \tau \|P - Q\|.$$

It follows from (6) that

$$\begin{aligned} c^T x^* &\leqslant c^T x + \|c\| \|x^* - x\| \\ &\leqslant V(P) + \eta + 2\tau \|c\| \|P - Q\| \\ &\leqslant V(P) + \epsilon, \end{aligned}$$

which means that $x^* \in X_\epsilon(P)$. Thus $d(x, X_\epsilon(P)) \leqslant \tau\|P - Q\|$ for all $x \in X_\eta(Q)$. □

In the rest of this section, we make the following assumption on the objective function.

ASSUMPTION 2. f is continuously differentiable on $\mathbb{R}^n$, and there exist constants $\alpha > 0$ and $\beta > 1$ such that for all $x',\ x'' \in \mathbb{R}^n$

$$f(x') - f(x'') \geqslant \nabla f(x'')^T(x' - x'') + \alpha\|x' - x''\|^\beta.$$

For example, a strictly convex quadratic function satisfies Assumption 2 with $\beta = 2$ and $\alpha = \lambda_{\min}[\nabla^2 f(x)/2]$. It is not difficult to verify that if f satisfies Assumption 2, then f has the following properties.

(i) f is strictly convex on $\mathbb{R}^n$.
(ii) For any nonempty compact convex set S, f is Lipschitz continuous on S with the Lipschitz constant $L_S = \max_{x\in S}\|\nabla f(x)\|$.
(iii) For any real number γ, the set $\{x | f(x) \leqslant \gamma\}$ is closed and bounded.
(iv) For any nonempty closed convex set S, $\arg\min_{x\in S} f(x)$ is a singleton.

LEMMA 2.2. *For any $z \in \mathbb{R}^n$, the set $\{x | f(x) \leqslant f(z) + 1\}$ is contained in $z + \mu\mathbb{B}^n$, where*

$$\mu = \max\left\{\left(2\alpha^{-1}\right)^{1/\beta}, \left(2\alpha^{-1}\|\nabla f(z)\|\right)^{1/(\beta-1)}\right\}. \tag{7}$$

Proof. If there is an $\tilde{x}$ in $\{x | f(x) \leqslant f(z) + 1\}$, but not in $z + \mu\mathbb{B}^n$, then by Assumption 2,

$$1 \geqslant f(\tilde{x}) - f(z) \geqslant \nabla f(z)^T(\tilde{x} - z) + \alpha\|\tilde{x} - z\|^\beta. \tag{8}$$

We may assume that $\nabla f(z) \neq 0$, for otherwise, we would have

$$1 \geqslant f(\tilde{x}) - f(z) \geqslant \alpha\|\tilde{x} - z\|^\beta \geqslant \alpha\left(2\alpha^{-1}\right)^{(1/\beta)\beta} = 2. \tag{9}$$

It follows from (7) and (8) that

$$\begin{aligned}\frac{1}{\|\tilde{x} - z\|^\beta} &\geqslant \alpha - \frac{\|\nabla f(z)\|\|\tilde{x} - z\|}{\|\tilde{x} - z\|^\beta} = \alpha - \frac{\|\nabla f(z)\|}{\|\tilde{x} - z\|^{\beta-1}} \\ &> \alpha - \frac{\|\nabla f(z)\|}{\mu^{\beta-1}} \geqslant \alpha - \frac{\|\nabla f(z)\|}{2\alpha^{-1}\|\nabla f(z)\|} \\ &= \alpha - \alpha/2 = \alpha/2.\end{aligned}$$

Hence, $\|\tilde{x} - z\| < \left(2\alpha^{-1}\right)^{1/\beta} \leqslant \mu$, which contradicts the hypothesis that $\tilde{x} \notin z + \mu\mathbb{B}^n$. □

THEOREM 2.3. *Let z be a feasible solution of* (SDP)*, and $L_{\bar{\mu}}$ be a Lipschitz constant of f on $z + (\mu + 1)\mathbb{B}^n$, where μ is defined by* (7) *in Lemma 2.2. If Assumption 1 is satisfied and* $\max\{\|P\|, \|Q\|\} \leqslant 1/2 \min\left\{\tau^{-1}, \tau^{-1}L_{\bar{\mu}}^{-1}\right\}$*, then*

(a) $|V(P) - V(Q)| \leqslant \tau L_{\bar{\mu}}\|P - Q\|$;

(b) $\|X(P) - X(Q)\| \leqslant k\|P - Q\|^{1/\beta}$*, where* $k = (2\tau L_{\bar{\mu}}\alpha^{-1})^{1/\beta} + \tau(\min\{\tau^{-1}, \tau^{-1}L_{\bar{\mu}}^{-1}\})^{1-1/\beta}$.

Proof. Choose $x'' \in S(P)$ such that $\|z - x''\| = d(z, S(P))$. By Theorem 2.1,

$$\|z - x''\| \leqslant d_\infty(S(0), S(P)) \leqslant \tau\|P\|. \tag{10}$$

As $\tau\|P\| \leqslant 1/2$ and f is Lipschitz continuous on $z + (\mu + 1)\mathbb{B}^n$ with a Lipschitz constant $L_{\bar{\mu}}$, we have that $x'' \in z + \mathbb{B}^n \subset z + (\mu + 1)\mathbb{B}^n$ and

$$\begin{aligned} f(X(P)) &\leqslant f(x'') = f(z) + (f(x'') - f(z)) \\ &\leqslant f(z) + \tau L_{\bar{\mu}}\|P\| \leqslant f(z) + 1. \end{aligned}$$

From Lemma 2.2,

$$X(P) \in z + \mu\mathbb{B}^n. \tag{11}$$

Similarly, $X(Q) \in z + \mu\mathbb{B}^n$. Part (a) thus follows from Proposition 2.2.

It remains to prove part (b). Choose $x' \in S(Q)$ such that $\|X(P) - x'\| = d(X(P), S(Q))$. From Theorem 2.1 and the assumption on P and Q,

$$\|X(P) - x'\| \leqslant d_\infty(S(P), \quad S(Q)) \leqslant \tau\|P - Q\| \leqslant 1. \tag{12}$$

From (11) and (12),

$$\|x' - z\| \leqslant \|x' - X(P)\| + \|X(P) - z\| \leqslant 1 + \mu, \tag{13}$$

which implies that $x' \in z + (\mu + 1)\mathbb{B}^n$. Hence,

$$\begin{aligned} \alpha\|x' - X(Q)\|^\beta &\leqslant f(x') - f(X(Q)) && (14) \\ &= \big(f(x') - f(X(P))\big) + (V(P) - V(Q)) \\ &\leqslant 2\tau L_{\bar{\mu}}\|P - Q\|, && (15) \end{aligned}$$

where (14) is from Assumption 2 and the optimality of $X(Q)$, and (15) is from (12), (13) and part (a). Therefore,

$$\begin{aligned} \|X(Q) - X(P)\| &\leqslant \|x' - X(Q)\| + \|x' - X(P)\| \\ &\leqslant (2\tau L_{\bar{\mu}}\alpha^{-1}\|P - Q\|)^{1/\beta} + \tau\|P - Q\| && (16) \\ &= k\|P - Q\|^{1/\beta}, \end{aligned}$$

where (16) is from (15) and (12). $\square$

3. Applications

3.1. EXACT PENALTY FUNCTIONS

An important class of solution methods in optimization are exact penalty function methods, which convert a given constrained nonlinear programming problem into an equivalent unconstrained problem or into a problem with simple constraints. But, if a lower bound of penalty parameters is unknown, then one must solve a sequence of unconstrained problems with increasing penalty parameters and this can cause ill-conditioning and numerical difficulties. Here we use the results from Section 2 to obtain an exact penalty function and a lower bound of penalty parameters for (SDP). We consider the following unconstrained penalty problem.

$$(\mathcal{Q}) \quad \text{minimize } \; h(x) = f(x) + \beta \max\{-\lambda_{\min}[B(x) + B_0], 0\} \quad \text{over } \mathbb{R}^n,$$

where $\beta > 0$ is a penalty parameter. In view of Corollary 2.1, it is not difficult to prove the following theorem, which gives a computable lower bound for β.

THEOREM 3.1. *Suppose that f is convex and Lipschitz continuous on $\mathbb{R}^n$ with a Lipschitz constant L (e.g., an affine function or a convex piecewise linear function). If Assumption 1 is satisfied and $\beta > \tau L$, then a vector is an optimal solution of (SDP) if and only if it is an optimal solution of $(\mathcal{Q})$.*

REMARK 3.1. Under different assumptions [14], Warga shows that there exist constants c and N such that the constrained problem: minimize $f(x)$, subject to $g(x) = 0$, $x \in \tilde{C}$ can be reduced to an unconstrained problem: minimize $f(x) + c|g(x)|^{1/N}$ over $x \in \tilde{C}$, where $\tilde{C}$ must be compact (see [14] for details). Due to the special structure of our problem, we are able to obtain a *computable* penalty parameter $c > \tau L$ along with $N = 1$.

For example, consider the problem of minimizing $f(x) = \max\{c_1^T x + b_1, c_2^T x + b_2\}$ subject to $\sum_{i=1}^{2}(x)_i B_i + B_0 \geqslant 0$, where $c_1 = (3, 1)$, $c_2 = (1, -2)$, $b_1 = 2$, $b_2 = 0$,

$$B_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad B_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad B_0 = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}.$$

With the choice of $\hat{x} = (1, 0)$, we have $\lambda_{\min}[B(\hat{x})] = 1 = \tau^{-1}$. Since $L = \sqrt{10}$, for the unconstrained problem $(\mathcal{Q})$, we can choose a penalty parameter $\beta > \sqrt{10}$.

3.2. PATTERNED COVARIANCE MATRICES

Patterned covariance matrices arise naturally from statistical models in the physical, biological, psychological, and social sciences [11]. For example, in a model for meteorological measurements, the covariance matrix exhibits the pattern

$$\Sigma_1 = \begin{pmatrix} a_0 & a_1 & a_2 & a_3 & \cdots & a_2 & a_1 \\ a_1 & a_0 & a_1 & a_2 & \cdots & a_3 & a_2 \\ a_2 & a_1 & a_0 & a_1 & \cdots & a_4 & a_3 \\ \vdots & \vdots & & \ddots & & \vdots & \vdots \\ a_2 & a_3 & a_4 & a_5 & \cdots & a_0 & a_1 \\ a_1 & a_2 & a_3 & a_4 & \cdots & a_1 & a_0 \end{pmatrix}$$

which is a circulant. In models for demography and migration, the covariance matrices exhibit the patterns

$$\Sigma_2 = \begin{pmatrix} a_1 & a_2 & a_3 & a_4 \\ a_2 & a_1 & a_3 & a_4 \\ a_3 & a_3 & a_1 & a_4 \\ a_4 & a_4 & a_4 & a_1 \end{pmatrix} \quad \text{and} \quad \Sigma_3 = \begin{pmatrix} a_1 & a_2 & a_2 & a_3 & a_4 \\ a_2 & a_1 & a_2 & a_3 & a_4 \\ a_2 & a_2 & a_1 & a_3 & a_4 \\ a_3 & a_3 & a_3 & a_1 & a_2 \\ a_4 & a_4 & a_4 & a_2 & a_1 \end{pmatrix}.$$

When the underlying distribution is not normal, one needs to find a positive semidefinite matrix having the required pattern and closest to a given sample covariance matrix [11]. For instance, if the covariance matrix has the pattern Σ_2, then one needs to solve the following semidefinite programming problem

$$\text{(SDP)} \quad \text{minimize} \quad \left\| \sum_{i=1}^{4} B_i (x)_i - \bar{B} \right\|_F^2$$
$$\text{subject to} \quad \sum_{i=1}^{4} B_i (x)_i \geqslant 0,$$

where $\bar{B}$ is the given sample covariance matrix, and

$$B_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad B_2 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$B_3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \qquad B_4 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}.$$

In practice, some numerical methods actually solve the following perturbed problem [11]

$$\text{(P)} \quad \text{minimize} \quad \left\| \sum_{i=1}^{4} B_i (x)_i - \bar{B} \right\|_F^2$$
$$\text{subject to} \quad \sum_{i=1}^{4} B_i (x)_i - P \geqslant 0.$$

Letting $(\hat{x})_1 = 1$, and $(\hat{x})_i = 0$ for $i = 2, 3, 4$, we have $B(\hat{x}) = I$ and thus a constant τ satisfying Assumption 1 is $\tau = 1$. From Theorem 2.1, we know that $d_\infty(S(P), S(Q)) \leqslant \|P - Q\|$, i.e., the Hausdorff distance between the feasible sets of these problems is bounded by $\|P - Q\|$. Now let

$$D = \begin{pmatrix} 4 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 6 \end{pmatrix} \quad \text{and} \quad \bar{c} = \begin{pmatrix} \sum_{i=1}^{4} \bar{B}(i,i) \\ 2\bar{B}(1,2) \\ 2(\bar{B}(1,3) + \bar{B}(2,3)) \\ 2\sum_{i=1}^{3} \bar{B}(i,4) \end{pmatrix},$$

we have $f(x) = \| \sum_{i=1}^{4} B_i(x)_i - \bar{B}\|_F^2 = x^T Dx - 2\bar{c}^T x + \|\bar{B}\|_F^2$, $\nabla f(x) = 2Dx - 2\bar{c}$, and the α and β satisfying Assumption 2 are $\alpha = 2$ and $\beta = 2$. Using $z = \hat{x}$ as a feasible solution to the unperturbed problem, one can easily verify that the constant μ defined in Lemma 2.2 is $\mu = 8 + 2\|\bar{c}\|$, and the Lipschitz constant $L_{\bar{\mu}}$ defined in Theorem 2.3 is $L_{\bar{\mu}} = 120 + 26\|\bar{c}\|$. From Theorem 2.3, we know that if $\max\{\|P\|, \|Q\|\} \leqslant (240 + 52\|\bar{c}\|)^{-1}$, then

$$|V(P) - V(Q)| \leqslant (120 + 26\|\bar{c}\|)\|P - Q\|$$

and

$$\|X(P) - X(Q)\| \leqslant \left((120 + 26\|\bar{c}\|)^{1/2} + (120 + 26\|\bar{c}\|)^{-1/2}\right) \|P - Q\|^{1/2}.$$

In particular, for a suboptimal solution obtained by setting $P = \epsilon I$, if $\epsilon \leqslant (240 + 52\|\bar{c}\|)^{-1}$, then

$$|V(\epsilon I) - V(0)| \leqslant (120 + 26\|\bar{c}\|)\epsilon$$

and

$$\|X(\epsilon I) - X(0)\| \leqslant \left((120 + 26\|\bar{c}\|)^{1/2} + (120 + 26\|\bar{c}\|)^{-1/2}\right) \epsilon^{1/2}.$$

Similarly, one can show that $\tau = 1$ for the patterns Σ_1 and Σ_3, and the bounds for the feasible sets, optimal values, and optimal solutions between (SDP) and (P) can be computed easily. For a general (SDP), one can find the bounds given by Theorems 2.1, 2.2 and 2.3 from a vector satisfying Assumption 1 and a feasible solution to the unperturbed problem.

There are numerous sensitivity analysis results for convex programming focusing on the existence of quantitative bounds. Because of the special structure of the problems, we are able to compute these bounds in terms of the problem data. It is not difficult to extend most of the results in this paper to convex programming problems where recession functions associated with the constraint systems satisfy the Slater condition.

Acknowledgments

We thank the referees for their helpful suggestions, which helped us to improve the quality and the presentation of the paper.

References

1. Overton, M. (1992), Large-scale optimization of eigenvalues, *SIAM J. Optimization* 2: 88–120.
2. Vandenberghe, L. and Boyd, S. (1996), Semidefinite programming, *SIAM Review* 38: 49–95
3. Burke, J. and Tseng, P. (1996), A unified analysis of Hoffman's bound via Fenchel duality, *SIAM J. Optimization* 6: 265–282.
4. Luo, Z.Q. and Luo, X.D. (1994), Extension of Hoffman's error bound to polynomial systems, *SIAM J. Optimization* 4: 383–392.
5. Luo, Z.Q. and Pang, J.S. (1995), Error bounds for analytic systems and their applications, *Mathematical Programming* 67: 1–28.
6. Lewis, A.S. and Pang, J.S. (1997), Error bounds for convex inequality systems, in J.P. Crouzeix, J.-E. Martinez-Legaz and M. Volle (eds.), *Generalized Convexity, Generalized Monotonicity: Recent Results, Proceedings of the Fifth Symposium on Generalized Convexity*, Luminy, June 1996, Kluwer Academic Publishers, pp. 75–100.
7. Mangasarian, O.L. (1985), A condition number for differentiable convex inequalities, *Mathematics of Operations Research* 10: 175–179.
8. Luo, Z.Q. and Tseng, P. (1994), Perturbation analysis of a condition number for linear systems, *SIAM J. Matrix Anal. Appl.* 15: 636–660.
9. Deng, S. (1997), Computable error bounds for convex inequality systems in reflexive Banach Spaces, *SIAM J. Optimization* 7: 274–279.
10. Fletcher, R. (1985), Semidefinite matrix constraints in optimization, *SIAM J. Control and Opt.* 23: 493–513.
11. Hu, H. and Olkin, I. (1991), A numerical procedure for finding the positive definite matrix closest to a patterned matrix, *Statistics and Probability Letters* 12: 511–515.
12. Hu, H. and Wang, Q. (1989), On approximate solutions of infinite systems of linear inequalities, *Linear Algebra and its Applications* 114/115: 429–438.
13. Rockafellar, R.T. (1970), *Convex Analysis*, Princeton University Press, Princeton, New Jersey.
14. Warga, J. (1992), A necessary and sufficient condition for a constrained minimum *SIAM J. Optimization* 2: 665–667.
15. Attouch, H. and Wets, R.J.-B. (1993), Quantitative stability of variational systems: III ϵ-approximate solutions, *Mathematical Programming* 61: 197–214.