Case i): If $x_1 = 0$, then $x'(\Pi_0 - \tilde{\Pi}_0)x = x_2'\Pi_{22}x_2 \geq 0$ because $\Pi_0 \geq 0$.

Case ii): If $x_1 \neq 0$, then

$$\begin{aligned} x'(\Pi_0 - \tilde{\Pi}_0)x &= x'\Pi_0 x - \epsilon x_1' x_1 \\ &\geq \lambda_{\min} x_1' x_1 - \epsilon x_1' x_1 \\ &= (\lambda_{\min} - \epsilon) x_1' x_1 > 0. \end{aligned}$$

Theorem 3: Subject to

i) (A, C) is detectable,

ii) $\Pi_0 \geq 0$,

iii) $\overline{E}(A, B) \supseteq N[\Pi_0]$,

then $\lim_{t \to \infty} P(t) = P_S$, where $P(\cdot)$ is the solution of the RE with initial condition $P(0) = \Pi_0$ and P_S is the strong solution of the ARE.

Proof: In view of Lemma 3, it is always possible to find $\tilde{\Pi}_{11} > 0$ such that in the standard basis

$$\tilde{\Pi}_0 = \begin{bmatrix} \tilde{\Pi}_{11} & 0 \\ 0 & 0 \end{bmatrix} \leq \Pi_0.$$

Consider now the reduced-order RE associated with the triple (A_{11}, B_1, C_1). The pair (A_{11}, C_1) is detectable and A_{11} has no (A_{11}, B_1)-unreachable boundary eigenvalue. Then, by Theorem 2, $\tilde{P}_{11}(\cdot)$ converges to P_{11S}, where $\tilde{P}_{11}(t)$ denotes the solution of the reduced-order RE with initial condition $\tilde{P}_{11}(0) = \tilde{\Pi}_{11}$, and P_{11S} is the strong solution of the corresponding reduced-order ARE. Note that

$$P_S = \begin{bmatrix} P_{11S} & 0 \\ 0 & 0 \end{bmatrix}.$$

Therefore, denoting by $\tilde{P}(\cdot)$ the solution of the RE (1) with initial condition $\tilde{P}(0) = \tilde{\Pi}_0$, $\tilde{P}(\cdot)$ converges to P_S.

It is also always possible to find $\overline{\Pi}_0$ such that $\overline{\Pi}_0 \geq \Pi_0$ and $\overline{\Pi}_0 \geq P_S$. Then, by Theorem 1, letting $\overline{P}(\cdot)$ be the solution of the RE (1), with initial condition $\overline{P}(0) = \overline{\Pi}_0$, $\overline{P}(\cdot)$ converges to P_S.

Finally, Lemma 1 entails that $\tilde{P}(t) \leq P(t) \leq \overline{P}(t)$, $t \geq 0$, so that the thesis follows.

Corollary: If (A, C) is detectable and $\Pi_0 > 0$, then $\lim_{t \to \infty} P(t) = P_S$. ■

Theorem 3 improves on existing convergence results in that it handles systems having possibly unreachable boundary eigenvalues. If we restrict our attention to the class of detectable systems with no unreachable boundary eigenvalues, a necessary and sufficient condition for convergence to the strong solution is already available [4]. A comparison of [4] with our Theorem 3 shows that, for detectable systems with a nonnegative Π_0, condition iii) of Theorem 3 is only sufficient. In conclusion, the search for a necessary and sufficient condition for convergence to the strong solution in the case of detectable systems is still an open question.

References

[1] R. E. Kalman, "New methods in Wiener filtering theory," in *Proceedings of the First Symposium on Engineering Applications of Random Function Theory and Probability*, J. Bogdanoff and F. Kozin, Eds. New York: Wiley, 1963, pp. 270-388.

[2] W. M. Wonham, "On a matrix Riccati equation of stochastic control," *SIAM J. Contr.*, vol. 6, pp. 681-698, 1968.

[3] K. Martensson, "On the matrix Riccati equation," *Inform. Sci.*, vol. 3, pp. 17-49, 1971.

[4] F. M. Callier and J. L. Willems, "Criterion for the convergence of the solution of the Riccati differential equation," *IEEE Trans. Automat. Contr.*, vol. AC-26, pp. 1232-1242, 1981.

[5] S. W. Chan, G. C. Goodwin, and K. S. Sin, "Convergence properties of the Riccati difference equation in optimal filtering of nonstabilizable systems," *IEEE Trans. Automat. Contr.*, vol. AC-29, pp. 110-118, 1984.

[6] M. A. Poubelle, I. R. Petersen, M. R. Gevers, and R. R. Bitmead, "A miscellany of results on an equation of Count J. F. Riccati," *IEEE Trans. Automat. Contr.*, vol. AC-31, pp. 651-654, 1986.

[7] C. E. de Souza, M. R. Gevers, and G. C. Goodwin, "Riccati equations in optimal filtering of nonstabilizable systems having singular state transition matrices," *IEEE Trans. Automat. Contr.*, vol. AC-31, pp. 831-838, 1986.

[8] R. R. Bitmead, M. R. Gevers, I. R. Petersen, and R. J. Kaye, "Monotonicity and stabilizability properties of solutions of the Riccati difference equation: Propositions, lemmas, theorems, fallacious conjectures and counterexamples," *Syst. Contr. Lett.*, vol. 5, pp. 309-315, 1985.

[9] P. Lancaster, A. C. M. Ran, and L. Rodman, "Hermitian solutions of the discrete algebraic Ricatti equation," *Int. J. Contr.*, vol. 44, pp. 777-802, 1986.

Norm Based Robust Control of State-Constrained Discrete-Time Linear Systems

Mario Sznaier

Abstract—**Most realistic control problems involve some type of constraint. However, to date, all the algorithms that deal with constrained problems assume that the system is perfectly known. On the other hand, during the last decade a considerable amount of time has been spent in the robust control problem. However, in its present form, the robust control theory can address only the idealized situation of completely unconstrained problems. In this note we present a theoretical framework to analyze the stability properties of constrained discrete-time systems under the presence of uncertainty and we show that this formalism provides a unifying approach, including as a particular case the well-known technique of estimating robustness bounds from the solution of a Lyapunov equation. These results are applied to the problem of designing feedback controllers capable of stabilizing a family of systems, while at the same time satisfying state-space constraints.**

I. Introduction

A large class of problems frequently encountered in practice involves the control of linear systems with states restricted to closed convex regions of space. Several methods have been proposed recently to deal with this class of problems (see [1] for a thorough discussion and several examples), but as a rule, all of these schemas assume exact knowledge of the dynamics involved (i.e., exact knowledge of the model). Such an assumption can be too restrictive, ruling out cases where good qualitative models of the plant are available but the numerical values of various parameters are unknown or even change during operation. On the other hand, during the last decade a considerable amount of time has been spent analyzing the question of whether some relevant quantitative properties of a system (most notably asymptotic stability) are preserved under the presence of unknown perturbations. This research effort has led to procedures for designing controllers, termed "robust controllers," capable of achieving desirable properties under various classes of perturbations. However, these design procedures cannot accommodate directly time domain constraints, although some progress has been made recently in this direction [2]-[4].

Manuscript received March 6, 1990; revised November 16, 1990. Paper recommended by Associate Editor, A. L. Tits.

The author is with the Department of Electrical Engineering, University of Central Florida, Orlando, FL 32816-0450.

IEEE Log Number 9107971.

0018-9286/92$03.00

In this note we present a theoretical framework to analyze the stability properties of constrained discrete-time systems under the presence of uncertainties and we apply this framework to the problem of designing feedback controllers capable of stabilizing a family of systems while at the same time satisfying state-space constraints. We believe that the results presented here will provide a useful new approach for addressing more realistic control design problems.

This note is organized as follows. In Section II we introduce the concepts of constrained stability and robust constrained stability and we use these concepts to give a formal definition of the robust constrained stability analysis and robust constrained stability design problems. The analysis problem is studied in Section III where we give necessary and sufficient conditions for constrained robustness and where we show that our approach includes as a special case the well-known technique of estimating robustness bounds from the solution of a Laypunov equation. In Section IV we apply the results of Section III to the design problem and we show that in cases of practical interest our approach yields a well-behaved optimization problem. Finally, in Section V, we summarize our results and we indicate directions for future research.

II. Definitions and Statement of the Problem

In this section we introduce a formal definition of the robust constrained control problem. We begin by introducing the concept of constrained stability.

Definition 1: Consider the linear, time-invariant, discrete-time, unforced system modeled by the difference equation

$$x_{k+1} = Ax_k, \qquad k = 0, 1 \cdots \tag{S}$$

subject to the constraint

$$x \in \mathcal{G} \subseteq R^n \tag{1}$$

where $A \in R^{n*n}$ and where x indicates x is a vector quantity. The system (S) is constraint stable (C-stable) if for *any* point $\tilde{x} \in \mathcal{G}$, the trajectory $x_k(\tilde{x})$ originating in $\tilde{x}$ remains in $\mathcal{G}$ for all k.

We proceed to introduce now a restriction on the class of constraints allowed in our problem. As it will become apparent later, the introduction of this restriction, termed the constraint qualification hypothesis, while not affecting significantly the number of real-world problems that can be handled by our formalism [5], introduces more structure into the problem. This addition structure is used in Lemma 1 to show that the constraints induce a norm in $\mathcal{G}$. In turn, this norm will play a key role in Section III where we derive necessary and sufficient conditions for constrained stability.

A. Constraint Qualification Hypothesis

In this note, we will limit ourselves to constraints of the form

$$x \in \mathcal{G} = \{x \in R^n : (G(x))_i \le \omega_i, \quad i = 1 \cdots p\} \tag{2}$$

where $\omega \in R^p$, $\omega_i > 0$ and where G: $R^n \to R^p$ is a positive-definite sublinear function, i.e., it has the following properties:

$$G(x)_i \ge 0, \qquad i = 1 \cdots p \forall x$$
$$G(x) = 0 \Leftrightarrow x = 0$$
$$G(x+y)_i \le G(x)_i + G(y)_i, \qquad i = 1 \cdots p \forall x, y$$
$$G(\lambda x) = \lambda G(x), 0 \le \lambda \le 1. \tag{3}$$

In the next lemma we show that $G(\cdot)$ induces a norm, and we characterize $\mathcal{G}$ in terms of this norm.

Lemma 1: Let

$$v(x) = \max_{1 \le i \le p} \left\{ \frac{G(x)_i}{\omega_i} \right\} = \| W^{-1} G(x) \|_\infty \stackrel{\Delta}{=} \| x \|_{\mathcal{G}} \tag{4}$$

where $W = \operatorname{diag}(\omega_1, \cdots, \omega_p)$. Then $v(\cdot)$ defines a norm in R^n and the set $\mathcal{G}$ can be characterized as

$$\mathcal{G} = \{x : \|x\|_{\mathcal{G}} \le 1\}. \tag{5}$$

Proof: The proof of the lemma follows by noting that the constraint qualification hypothesis (3) implies that

$$\|x\|_{\mathcal{G}} = \|W^{-1}G(x)\|_\infty \tag{6}$$

satisfies the conditions for a norm in R^n. □

Next, we take into account uncertainty in the dynamics by extending the concept of constrained stability to a family of systems and we define a quantitative way of measuring the "size" of the smallest destabilizing perturbation.

Definition 2: Consider the system (S). Let the perturbed system (S_Δ) be defined as

$$x_{k+1} = (A + \Delta) x_k \tag{S_Δ}$$

where Δ belong to some perturbation set $\mathcal{D} \subseteq R^{n*n}$. The system (S) is robust constraint stable (RC-stable) with respect to the set $\mathcal{D}$ if (S_Δ) is C-stable for all perturbation matrices $\Delta \in \mathcal{D}$.

Definition 3: Let $\|\cdot\|_{\mathcal{N}}$ be an operator norm defined in the set $\mathcal{D}$, and define the set $B\Delta^{\mathcal{N}}$ as the intersection of $\mathcal{D}$ and the origin centered $\mathcal{N}$-norm unity ball in parameter space, i.e.,

$$B\Delta^{\mathcal{N}} = \{\Delta \in \mathcal{D} : \|\Delta\|_{\mathcal{N}} \le 1\}.$$

The constrained stability measure with respect to the norms $\|\cdot\|_{\mathcal{N}}$ and $\|\cdot\|_{\mathcal{G}}$, $\varrho_{\mathcal{G}}^{\mathcal{N}}$, is defined as

$$\varrho_{\mathcal{G}}^{\mathcal{N}} = \max\{\mu : (S_\Delta) \text{ is } C\text{-stable with respect to } \mu B\Delta^{\mathcal{N}}\}.$$

In the particular case that the induced operator norm $\|\cdot\|_{\mathcal{G}}$ is used in the set $\mathcal{D}$, we will denote the constrained stability measure as $\varrho_{\mathcal{G}}$ and the set $B\Delta^{\mathcal{N}}$ as $B\Delta$.

With the concepts introduced in this section, we are now ready to give a formal definition to our problem:

* *Robust Constrained Stability Analysis Problem:* Given the family of linear time-invariant discrete-time systems represented by (S_Δ), compute the constrained stability measure $\varrho_{\mathcal{G}}^{\mathcal{N}}$.

* *Linear Robust Constrained Stability Design Problem:* Given the family of linear time-invariant discrete-time systems represented by

$$x_{k+1} = (A + \Delta) x_k + Bu_k$$

find a constant feedback matrix F such that for the closed-loop system

$$x_{k+1} = (A + BF + \Delta) x_k \tag{S_{cl_Δ}}$$

the constrained stability measure is maximized.

III. Theoretical Results

In this section we present the basic results that are required to solve the analysis problem. These results will be used in Section IV to solve the design problem. We begin by presenting a necessary and sufficient condition for robust constrained stability of a family of systems. This result is then used to compute the actual value and lower bounds on the constrained stability measure introduced in the last section.

Theorem 1: The system (S) is RC-stable with respect to the set $\mathcal{D}$ iff

$$\|A + \Delta\|_{\mathcal{G}} \le 1 \forall \Delta \in \mathcal{D} \tag{7}$$

where $\|\cdot\|$ denotes the induced operator norm, i.e.,

$$\|A + \Delta\|_{\mathcal{G}} = \max_{\|x\|_{\mathcal{G}} = 1} \{\|(A + \Delta)x\|_{\mathcal{G}}\}. \tag{8}$$

Proof: The proof follows immediately from Definition 1 and

(5) by noting that

$$\|A+\Delta\|_{\mathscr{G}} \leq 1 \Leftrightarrow \|(A+\Delta)x\|_{\mathscr{G}} \leq 1 \forall \|x\|_{\mathscr{G}} \leq 1. \quad (9)$$

□

Remark: Note that if $\|A+\Delta\|_{\mathscr{G}} < 1$ for all $\Delta \in \mathscr{D}$, then $(A+\Delta)$ is a contraction mapping and the system (S_Δ) is asymptotically stable [6].

Corollary 1:

$$\varrho_{\mathscr{G}}^{\mathscr{N}} = \min_{\Delta \in \mathscr{D}} \{\|\Delta\|_{\mathscr{N}} : \|A+\Delta\|_{\mathscr{G}} = 1\}. \quad (10)$$

In the next lemma we introduce a lower bound of the constrained stability measure. In Theorem 2 we show that for unstructured perturbations (i.e., the case where $\mathscr{D} \equiv R^{n*n}$) this lower bound is saturated.

Lemma 2:

$$\varrho_{\mathscr{G}} \geq 1 - \|A\|_{\mathscr{G}}. \quad (11)$$

Proof: Let Δ_1 be such that $\|A+\Delta_1\|_{\mathscr{G}} = 1$. Then

$$1 = \|A+\Delta_1\|_{\mathscr{G}} \leq \|A\|_{\mathscr{G}} + \|\Delta_1\|_{\mathscr{G}} \quad (12)$$

or

$$\|\Delta_1\|_{\mathscr{G}} \geq 1 - \|A\|_{\mathscr{G}}. \quad (13)$$

Hence

$$\varrho_{\mathscr{G}} = \min_{\mathscr{D}} \|\Delta_1\|_{\mathscr{G}} \geq 1 - \|A\|_{\mathscr{G}} \quad (14)$$

□

Theorem 2: For the unstructured perturbation case, i.e., the case where $\mathscr{D} \equiv R^{n*n}$, condition (11) is saturated.

Proof: The proof follows from Lemma 2 by noting that for

$$\Delta^o \triangleq \frac{(1-\|A\|_{\mathscr{G}})A}{\|A\|_{\mathscr{G}}} \quad (15)$$

we have $\|\Delta^o\|_{\mathscr{G}} = 1 - \|A\|_{\mathscr{G}}$ and $\|A+\Delta^o\|_{\mathscr{G}} = 1$ □

A. Quadratic Constraints Case

In this subsection we particularize our theoretical results for the special case where the constraint region is an hyperellipsoid, i.e., the case where

$$G(x) = (x^T P x)^{\frac{1}{2}}, \quad P \in R^{n*n} \text{ positive definite}. \quad (16)$$

We will show that in this case our approach yields a generalization of the well-known technique of estimating the robustness measure by using quadratic based Lyapunov functions (see [7] and references therein) by obtaining robustness bounds previously derived in this context. Moreover, using our approach we will show that in some cases these bounds give the actual value of the constrained stability measure.

Example 1 (Unstructured Perturbation): In this case, Theorem 2 yields $\varrho_{\mathscr{G}} = 1 - \|A\|_{\mathscr{G}}$ where

$$\|A\|_{\mathscr{G}}^2 = \|A\|_{\mathscr{P}}^2 = \max_x \left(\frac{x^T A^T P A x}{x^T P x}\right). \quad (17)$$

Consider now the case where $\varrho_{\mathscr{G}} > 0$. Then, there exists Q positive definite such that

$$A^T P A - P = -Q \quad (18)$$

and

$$\|A\|_{\mathscr{G}}^2 = \max_x \left(1 - \frac{x^T Q x}{x^T P x}\right) \leq 1 - \frac{\sigma_{\min}(Q)}{\sigma_{\max}(P)}. \quad (19)$$

Hence

$$\varrho_{\mathscr{G}} = 1 - \|A\|_{\mathscr{G}} \geq 1 - \left(1 - \frac{\sigma_{\min}(Q)}{\sigma_{\max}(P)}\right)^{\frac{1}{2}}. \quad (20)$$

A common technique in state-space robust analysis is to obtain robustness bounds from equation (18) [8], [9]. This case can be accommodated by our formalism by recognizing the fact that once P is selected, the system becomes effectively constrained to remain within a hyperellipsoidal region. It has been suggested [8], [9] that good robustness bounds can be obtained from (18) when P is selected such that $Q = I$. In this case our approach yields

$$\varrho_{\mathscr{G}} = 1 - \|A\|_{\mathscr{G}} = 1 - \left(1 - \frac{1}{\sigma_{\max}(P)}\right)^{\frac{1}{2}} \quad (21)$$

which coincides with the robustness bound found by Sezer and Siljak [9]. Note, however, that our derivation shows this bound to be exact for the unstructured perturbation case.

Example 2 (Unstructured Perturbation, A Semisimple): Consider the case where A is semisimple, i.e.,

$$A = L^{-1}\Lambda L$$
$$\Lambda = \operatorname{diag}\left\{\begin{pmatrix}\sigma_1 & \omega_1 \\ -\omega_1 & \sigma_1\end{pmatrix}, \cdots, \begin{pmatrix}\sigma_p & \omega_p \\ -\omega_p & \sigma_p\end{pmatrix}, \sigma_{p+1}, \cdots, \sigma_n\right\}. \quad (22)$$

Then, the maximum of the stability measure $\varrho_{\mathscr{G}}$ over all possible positive definite matrices P, is achieved for $P = L^T L$.

Proof: From (17) and (22) we have

$$\begin{aligned}\|A\|_P^2 &= \max_x \left\{\frac{x^T A^T P A x}{x^T P x}\right\} \\ &= \max_x \left\{\frac{x^T L^T L^{-T} A^T L^T L A L^{-1} L x}{x^T L^T L x}\right\} \\ &= \max_{\|y\|_2 = 1} \|LAL^{-1}y\|_2^2 \\ &= \|LAL^{-1}\|_2^2 = \sigma_{\max}^2(\Lambda). \end{aligned} \quad (23)$$

From (22) it follows that

$$\sigma_{\max}(\Lambda) = \max_i |\lambda_i^A| = \rho(A) \quad (24)$$

where λ_i^A denotes the eigenvalues of A and $\rho(\cdot)$ denotes the spectral radius. Since the spectral radius is always *smaller* than any other matrix norm [10] we have that

$$\|A\|_M \geq \rho(A) = \|A\|_{L^T L} \quad (25)$$

and therefore

$$\begin{aligned}\varrho_{L^T L} &= 1 - \|A\|_{L^T L} \\ &\geq \varrho_M = 1 - \|A\|_M \forall M \in R^{n*n}, \quad \text{positive definite}. \end{aligned} \quad (26)$$

□

B. Polyhedral Constraints

In this subsection we consider the case where the region $\mathscr{G}$ is polyhedral, i.e., the case where

$$G(x) = |Gx| \quad (27)$$

where $G \in R^{m*n}$, rank$(G) = n$, and the $|\cdot|$ should be interpreted on a component by component sense. We begin by showing that in this case the induced norm of an operator M, $\|M\|_{\mathscr{G}}$ can be expressed in terms of the infinity norm of an operator H linearly related to M.

Lemma 3: Let $M \in R^{n*n}$ and define $H \stackrel{\Delta}{=} GM(G^TG)^{-1}G^T$. Then

$$\|M\|_{\mathscr{G}} = \|W^{-1}HW\|_\infty. \tag{28}$$

Proof: The proof follows immediately from (6) and (8) by noting that $GM = HG$. □

The results of Lemma 3 can be used to efficiently compute $\varrho_{\mathscr{G}}^{\mathscr{N}}$ as the minimum of the solution of p linear programming problems as follows.

Lemma 4: Let $\varrho_i^{\mathscr{N}}$ be the solution of the following optimization problem

$$\varrho_i^{\mathscr{N}} = \min_{\Delta \in \mathscr{D}} \{ \|\Delta\|_{\mathscr{N}}: \|W^{-1}(H + \Delta H)W\|_1^{(i)} \geq 1 \} \tag{29}$$

where $\|M\|_1^{(i)}$ indicates the l_1 norm of the ith row of the matrix M and where H and ΔH are defined as in Lemma 3. Then

$$\varrho_{\mathscr{G}}^{\mathscr{N}} = \min_{1 \leq i \leq p} \{\varrho_i^{\mathscr{N}}\}. \tag{30}$$

Proof: Assume that the lemma is false and that there exist $\tilde{\varrho}$ and $\tilde{\Delta}$ such that

$$\|A + \tilde{\Delta}\|_{\mathscr{G}} = 1; \|\tilde{\Delta}\|_{\mathscr{N}} = \tilde{\varrho} < \varrho_{\mathscr{G}}^{\mathscr{N}}. \tag{31}$$

Since $\|A + \tilde{\Delta}\|_{\mathscr{G}} = 1$ there exists i^o such that $\|W^{-1}(H + \tilde{\Delta}H)W\|_1^{(i^o)} = 1$, $\|W^{-1}(H + \tilde{\Delta}H)W\|_1^{(j)} \leq 1$, $j \neq i^o$, but this implies (29) that $\varrho_{i^o}^{\mathscr{N}} \leq \tilde{\varrho}$ which contradicts (31). □

Example 3 (Unstructured Perturbation): Consider the following case:

$$A = \begin{pmatrix} 0.8 & 0.5 \\ -0.0208 & 0.5083 \end{pmatrix} G = \begin{pmatrix} 1.0 & 2.0 \\ -1.5 & 2.0 \end{pmatrix} \omega = \begin{pmatrix} 5.0 \\ 10.0 \end{pmatrix}. \tag{32}$$

Then, from the definition of H, we have that

$$H = \begin{pmatrix} 0.7583 & 0.0 \\ 0.4167 & 0.55 \end{pmatrix}, \|A\|_{\mathscr{G}} = 0.7583 \tag{33}$$

and, from Lemma 4,

$$\varrho_i = \min_{\|\Delta\|_{\mathscr{G}}} \left\{ \|\Delta\|_{\mathscr{G}}: \sum_{j=1}^{2} \frac{|H + \Delta|_{ij}\omega_j}{\omega_i} = 1 \right\} \quad i = 1, 2. \tag{34}$$

Casting the problems (34) into a linear programming form and solving we have that

$$\varrho_1 = 0.2417,\ \varrho_2 = 0.2417 \text{ and } \varrho_{\mathscr{G}} = \min_{1 \leq i \leq 2} \varrho_i = 0.2417.$$

Note that in this case $\varrho_{\mathscr{G}} = 1 - \|A\|_{\mathscr{G}} = 0.2417$ as shown in Theorem 2.

IV. Application to Robust Controllers Design

In this section we apply our formalism to solve the linear robust constrained stability design problem introduced in Section II. From Theorem 1 it follows that a full state feedback matrix F such that the constrained stability measure $\varrho_{\mathscr{G}}^{\mathscr{N}}$ of the closed-loop system is maximized can be selected by solving the following max-min problem:

$$\max_F \left\{ \min_{\Delta \in \mathscr{D}} \|\Delta\|_{\mathscr{N}} \right\} \tag{35}$$

subject to

$$\|A + BF + \Delta\|_{\mathscr{G}} = 1.$$

Define

$$\varrho_{\mathscr{G}}^{\mathscr{N}}(F) \stackrel{\Delta}{=} \min_{\Delta \in \mathscr{D}} \{ \|\Delta\|_{\mathscr{N}}: \|A + BF + \Delta\|_{\mathscr{G}} = 1 \} \tag{36}$$

then (35) is equivalent to the following optimization problem:

$$\max_F \{ \varrho_{\mathscr{G}}^{\mathscr{N}}(F) \}. \tag{37}$$

Note that since the function defined is (36) is in general nondifferentiable, nonsmooth optimization techniques must be used to solve (37). Moreover, in general nothing can be stated about the existence of the local maxima of (36). Hence, a general nonsmooth optimization algorithm could conceivably get trapped at local extrema. However, in the next theorem we show that for a case of practical interest, (37) reduces to the well-behaved problem of finding the maximum of a concave function.

Theorem 3: Consider the particular case where $\mathscr{D}$ is a cone with vertex at the origin, (i.e., $\Delta \in \mathscr{D} \Leftrightarrow \lambda\Delta \in \mathscr{D}$; $\lambda \geq 0$). Then $\varrho_{\mathscr{G}}^{\mathscr{N}}(F)$ is a concave function.

The proof of the theorem is given in the Appendix. Note that the class of sets considered in this theorem includes as a particular case set of the form

$$\mathscr{D} = \left\{ \Delta: \Delta = \sum_1^m \mu_i E_i;\ \mu_i \geq 0,\ E_i \text{ given} \right\} \tag{38}$$

which has been the object of much interest lately ([11]–[13]) and references therein).

At the present time, we are investigating several methods of solving (37), and a future paper is planned to report the results. In this note, we will limit ourselves to the restricted case of unstructured perturbations. In this case, from Theorem 2 we have that $\varrho_{\mathscr{G}} = 1 - \|A + BF\|_{\mathscr{G}}$. Hence, (37) reduces to solving the following convex minimization problem:

$$F = \operatorname*{argmax}_F \varrho_{\mathscr{G}} = \operatorname*{argmin}_F \|A + BF\|_{\mathscr{G}}. \tag{39}$$

In the remainder of this section, we will indicate how problem (39) can be solved for the particular cases of quadratic and polyhedral constraints. We begin by considering quadratic constraints.

A. Quadratic Constraints Case

In this case (39) can be solved using standard results on matrix dilations [14]. Let $P = L^TL$ and assume that rank$(B) = m$. Then, since the 2-norm is invariant under orthonormal transformations we have that

$$\|A + BF\|_{L^TL} = \|L(A + BF)L^{-1}\|_2 = \|\tilde{A} + \tilde{B}\tilde{F}\|_2 \tag{40}$$

where $\tilde{A} = QLA(QL)^{-1}$, $\tilde{B} = QLB$, $\tilde{F} = F(QL)^{-1}$, and where Q is an orthonormal matrix such that

$$\tilde{B} = \begin{pmatrix} B_1 \\ O \end{pmatrix},\ B_1 \text{ invertible}. \tag{41}$$

Then

$$\tilde{A} + \tilde{B}\tilde{F} = \begin{pmatrix} A_1 + B_1\tilde{F} \\ A_2 \end{pmatrix} \tag{42}$$

and it follows that the optimal F is such that $A_1 + B_1\tilde{F} = 0$, i.e., $\tilde{F}^o = -B_1^{-1}\tilde{A}_1$ and that $\min_F \|A + BF\|_{L^TL} = \|A_2\|_2 = \sigma_{\max}(A_2)$.

Example 4: Consider the system

$$A = \begin{pmatrix} 0.8 & 0.5 \\ -0.4 & 1.2 \end{pmatrix} B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$P = \begin{pmatrix} 1.0 & 0.0 \\ 0.0 & 1.0 \end{pmatrix} \omega = 1.$$

A direct application of (42) yields

$$F^o = (0.4 \quad -1.2),\ \|A + BF^o\|_I = 0.9434,\ \varrho_I = 0.0566.$$

B. Polyhedral Constraints

When the constraints are polyhedral, (39) can be cast in the following format:

$$\min_F \epsilon \tag{43}$$

subject to

$$\| A + BF \|_{\mathscr{G}} \leq \epsilon. \tag{44}$$

By using (28), the inequalities (43) can be transformed into

$$|G(A + BF)(G^T G)^{-1} G^T| \omega \leq \epsilon\omega. \tag{45}$$

The optimization problem defined by (43) and (45) can be cast into a linear programming problem and solved using the simplex method. Note that a similar design algorithm was proposed by Vassilaki *et al.* [15], although in their case the goal was to find admissible linear controllers for systems under polyhedral constraints, without taking into account robustness considerations.

Example 5: Consider the following system:

$$A = \begin{pmatrix} 0.8 & 0.5 \\ -0.4 & 1.2 \end{pmatrix} \quad B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$G = \begin{pmatrix} 1.0 & 2.0 \\ -1.5 & 2.0 \end{pmatrix} \quad \omega = \begin{pmatrix} 5.0 \\ 10.0 \end{pmatrix}.$$

Using linear programming we get

$$F = (0.3792 \quad -0.6917)$$

$$A_{cl} = \begin{pmatrix} 0.8 & 0.5 \\ -0.0208 & 0.5083 \end{pmatrix}, \ \mathrm{eig}\,(A_{cl}) = \begin{pmatrix} 0.55 \\ 0.7583 \end{pmatrix}$$

$$\| A_{cl} \|_{\mathscr{G}} = 0.7583 = \rho(A)$$

where A_{cl} denotes the closed-loop matrix, $\mathrm{eig}\,(A_{cl})$ its eigenvalues, and $\rho(A)$ its spectral radius. Hence, we have

$$\varrho_{\mathscr{G}} = 1 - \| A_{cl} \|_{\mathscr{G}} = 0.2417.$$

V. Conclusions

The ultimate objective in control design can perhaps be summarized as [2]: "achieve acceptable performance under perhaps substantial system uncertainty and under design constraints." This statement looks deceptively simple, but up-to-date design techniques focus either only on the uncertainty issue or only on the constraint satisfaction issue. In this note we presented a theoretical framework capable of simultaneously addressing both issues. Since most physically generated constraints have a natural expression in time domain, our analysis focuses on state-space robustness analysis.

In Section II, we introduced the concept of robust constrained stability and we introduced a quantity, the constrained stability measure, that measures the "size" of the smallest destabilizing perturbation. In Section III we presented necessary and sufficient conditions guaranteeing constrained robust stability and we showed that our formalism provides a unifying approach, including as a particular case the well-known technique of estimating robustness bounds from the solution of a Lyapunov equation. Finally, in Section IV, we considered the design problem. There, we showed that a full state feedback matrix that maximizes the stability measure of the closed-loop system can be found as the solution of a game-like problem. Although the properties of this problem are still unknown for the general case, we proved that in a specific case that has been the object of much attention lately, it leads to the well behaved problem of finding the maximum of a concave function. Finally, we considered the particular case of unstructured perturbations and we showed that in this case the problem reduces to the simpler case of finding the minimum of a convex (albeit perhaps nondifferentiable) function.

We believe that the results presented here will provide a valuable new approach to the problems of robust controllers analysis and design for linear systems. Further, since our approach is based purely upon time-domain analysis, we have reasons to believe the theory could be extended to encompass nonlinear systems in a much more direct fashion than other currently used techniques.

Perhaps the most severe limitation to the theory in its present form, arises from the fact that the design procedure is limited to constant linear feedback. However, it is clear that only a fraction of the feasible constrained problems admits a constant linear feedback solution. It is our goal to extend the theory to include the nonlinear, optimization-based controllers that were the subject of [5].

Appendix
Proof of Theorem 3

The following preliminary lemma is introduced without proof.

Lemma 5: Let $\rho_1 > 0$, $\rho_2 > 0$, and $0 \leq \lambda \leq 1$ be given numbers and assume that $\mathscr{D}$ is a cone with vertex at the origin. Consider the following sets:

$$\rho_1 B\Delta = \{\Delta \in \mathscr{D}: \|\Delta\|_{\mathscr{N}} \leq \rho_1\}$$

$$\rho_2 B\Delta = \{\Delta \in \mathscr{D}: \|\Delta\|_{\mathscr{N}} \leq \rho_2\}$$

$$\rho B\Delta = \left\{\Delta \in \mathscr{D}: \|\Delta\|_{\mathscr{N}} \leq \rho \overset{\Delta}{=} \lambda\rho_1 + (1-\lambda)\rho_2\right\}. \tag{A1}$$

Then

$$\rho B\Delta \subseteq \lambda\rho_1 B\Delta + (1-\lambda)\rho_2 B\Delta.$$

Proof of Theorem 3: Given two matrices F_1 and F_2, consider a convex linear combination $F = \lambda F_1 + (1-\lambda)F_2$. Then

$$\begin{aligned} &\max_{\Delta \in \rho B\Delta} \| A + BF + \Delta \|_{\mathscr{G}} \\ &\quad \leq \max_{\substack{\Delta_1 \in \rho_1 B\Delta \\ \Delta_2 \in \rho_2 B\Delta}} \|\lambda(A + BF_1 + \Delta_1) + (1-\lambda)(A + BF_2 + \Delta_2)_{\mathscr{G}} \\ &\quad \leq \lambda \max_{\Delta_1 \in \rho_1 B\Delta} \| A + BF_1 + \Delta_1 \|_{\mathscr{G}} \\ &\qquad + (1-\lambda) \max_{\Delta_2 \in \rho_2 B\Delta} \| A + BF_2 + \Delta_2 \|_{\mathscr{G}}. \end{aligned} \tag{A2}$$

Consider now the case where $\rho_1 = \varrho_{\mathscr{G}}^{\mathscr{N}}(F_1)$ and $\rho_2 = \varrho_{\mathscr{G}}^{\mathscr{N}}(F_2)$. Then it follows from the definition of $\varrho_{\mathscr{G}}^{\mathscr{N}}$ that both maximizations on the right-hand side of (A2) yield 1 and therefore

$$\max_{\Delta \in \rho B\Delta} \| A + BF + \Delta \|_{\mathscr{G}} \leq 1. \tag{A3}$$

Hence, from the definition of $\varrho_{\mathscr{G}}^{\mathscr{N}}$:

$$\varrho_{\mathscr{G}}^{\mathscr{N}}[\lambda F_1 + (1-\lambda)F_2] \geq \varrho = \lambda\varrho_{\mathscr{G}}^{\mathscr{N}}(F_1) + (1-\lambda)\varrho_{\mathscr{G}}^{\mathscr{N}}(F_2). \quad \square$$

Acknowledgment

The author is grateful to Prof. M. Damborg at the University of Washington and to H. Rotstein and Prof. A. Sideris at Caltech for many suggestions.

References

[1] M. Sznaier, "Suboptimal feedback control of constrained linear systems," Ph.D. dissertation, Univ. Washington, 1989.

[2] J. W. Helton and A. Sideris, "Frequency response algorithms for H_∞ optimization with time domain constraints," *IEEE Trans. Automat. Contr.*, vol. 34, no. 4, pp. 427–434, Apr. 1989.

[3] S. Boyd *et al.*, "A new CAD method and associated architectures for linear controllers," *IEEE Trans. Automat. Contr.*, vol. 33, no. 3, pp. 268–283, Mar. 1988.

[4] E. Polak and S. Salcudean, "On the design of linear multivariable feedback systems via constrained nondifferentiable optimization in H_∞ spaces," *IEEE Trans. Automat. Contr.*, vol. 34, no. 3, pp. 268-276, Mar. 1989.
[5] M. Sznaier and M. J. Damborg, "Heuristically enhanced feedback control of constrained discrete time linear systems," *Automatica*, vol. 26, no. 3, pp. 521-532, May 1990.
[6] R. E. Kalman and J. E. Bertram, "Control systems analysis and design via the second method of Lyapunov, Part II, Discrete time systems," *ASME J. Basic Eng.*, June 1960, pp. 394-400.
[7] D. D. Siljak, "Parameter space methods for robust control design: A guided tour," *IEEE Trans. Automat. Contr.*, vol. 34, no. 7, pp. 674-687, July 1989.
[8] R. V. Patel and M. Toda, "Quantitative measures of robustness for multivariable systems," in *Proc. Joint Automat. Contr. Conf.*, San Francisco, CA, 1980, paper TP-8A.
[9] M. E. Sezer and D. D. Siljak, "Robust stability of discrete systems," *Int. J. Contr.*, vol. 48, no. 5, pp. 2055-2063, Nov. 1988.
[10] G. R. Belitskii and Yu. I. Lyubich, "Matrix norms and their applications," in *Operator Theory: Advances and Applications, Vol. 36*, Berlin: Birkhäuser-Verlag, 1988.
[11] R. K. Yedavalli, "Improved measures of stability robustness for linear state space models," *IEEE Trans. Automat. Contr.*, vol. AC-30, pp. 557-579, June 1985.
[12] K. Zhou and P. P. Khargonekar, "Stability robustness bounds for linear state-space models with structured uncertainty," *IEEE Trans. Automat. Contr.*, vol. AC-32, no. 7, pp. 621-623, July 1987.
[13] L. H. Keel *et al.*, "Robust control with structured perturbations," *IEEE Trans. Automat. Contr.*, vol. 33, no. 1, pp. 68-77, Jan. 1988.
[14] J. C. Doyle, *Lecture Notes in Advances in Multivariable Control*, ONR/Honeywell Workshop, Minneapolis, MN, 1984.
[15] M. Vassilaki *et al.*, "Feedback control of discrete-time systems under state and control constraints," *Int. J. Contr.*, vol. 47, no. 6, pp. 1727-1735, 1988.

Local Convergence Analysis of Conjugate Gradient Methods for Solving Algebraic Riccati Equations

Ali R. Ghavimi, Charles Kenney, and Alan J. Laub

Abstract—**Necessary and sufficient conditions are given for local convergence of the conjugate gradient (CG) method for solving symmetric and nonsymmetric algebraic Riccati equations. For these problems, the Frobenius norm of the residual matrix is minimized via the CG method, and convergence in a neighborhood of the solution is predicated on the positive definiteness of the associated Hessian matrix. For the nonsymmetric case, the Hessian eigenvalues are determined by the squares of the singular values of the closed-loop Sylvester operator. In the symmetric case, the Hessian eigenvalues are closely related to the squares of the closed-loop Lyapunov singular values. In particular, the Hessian is positive definite if and only if the associated operator is nonsingular. The invertibility of these operators can be expressed as a noncancellation condition on the eigenvalues of the closed-loop matrices. For example, the stability of the closed-loop matrix, for the positive semidefinite Riccati solution, ensures the invertibility of the Lyapunov operator and hence the convergence of the CG method in a neighborhood of that solution.**

Manuscript received April 23, 1990; revised December 3, 1990. This work was supported in part by the National Science Foundation and the Air Force Office of Scientific Research under Grant ECS87-18897, by the National Science Foundation under Grant DMS88-00817, and by the Air Force Office of Scientific Research under Contract AFOSR-91-0240.

The authors are with the Department of Electrical and Computer Engineering, University of California, Santa Barbara, CA 93106-9560.

IEEE Log Number 9107972.

I. Introduction

When minimizing a scalar function f via the conjugate gradient (CG) method, local convergence is equivalent to the Hessian of f being positive definite at the point of minimization [1]. Moreover, Kantorovich-type error bounds show that the speed of convergence is intimately connected to the Hessian eigenvalue structure.

Our main result is that for the problem of minimizing the Frobenius norm of the Riccati residual matrix, the Hessian matrix is positive definite if and only if the associated closed-loop operator (Sylvester or Lyapunov) is nonsingular. Further, there is a close correspondence between the squares of the singular values of the closed-loop operator and the Hessian eigenvalues. These results are presented in the next two sections and are followed by a discussion of general matrix problems. The remainder of this section is devoted to a synopsis of the CG method.

General CG methods for minimizing smooth functions $f\colon \mathbb{R}^k \to \mathbb{R}$, extend the classical CG theory for quadratic functions and can take many forms. In the following, we use the CG method described by Luenberger in [1]. Letting ∇f denote the gradient of f, the general CG algorithm is then defined as follows.

Step 1: Given $x_0 \in \mathbb{R}^k$, compute $g_0 = \nabla f(x_0)$ and set $d_0 = -g_0$.

Step 2: For $i = 0, 1, \cdots, k-1$:

a) Find α_i minimizing $f(x_i + \alpha d_i)$ over all $\alpha \in \mathbb{R}$.

b) Set $x_{i+1} = x_i + \alpha_i d_i$ and $g_{i+1} = \nabla f(x_{i+1})$.

c) Unless $i = k-1$, set $d_{i+1} = -g_{i+1} + \beta_i d_i$ where β_i is a constant whose choice is described below.

Step 3: If $\nabla f(x_k)$ is sufficiently small in norm then stop, otherwise replace x_0 by x_k and go back to Step 1.

The constants β_i in Step 2 can be chosen in a variety of ways, including

$$\beta_i = \frac{(g_{i+1} - g_i)^T g_{i+1}}{(g_{i+1} - g_i)^T d_i} \tag{1}$$

due to Hestenes and Stiefel [2],

$$\beta_i = \frac{g_{i+1}^T g_{i+1}}{g_i^T g_i} \tag{2}$$

due to Fletcher and Reeves [3], and

$$\beta_i = \frac{(g_{i+1} - g_i)^T g_{i+1}}{g_i^T g_i} \tag{3}$$

due to Polak and Ribière [4]. All three expressions reduce to the standard formula for β_i in the quadratic case. However, for nonquadratic minimization, numerical experiments indicate that the Polak–Ribière method gives better results [5].

In order to discuss the speed of convergence of the CG method, suppose that the objective function f has a local minimum at x_* with associated Hessian matrix, $Q = (\partial^2 f / \partial x_i \, \partial x_j)$. Assume that Q is positive definite at x_* with eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k > 0$. For the error function

$$e(x) = \frac{1}{2}(x - x_*)^T Q (x - x_*) \tag{4}$$

the standard Kantorovich-type error bound for the quadratic case is [1]

$$e(x_{i+1}) \leq \left[\frac{\lambda_{i+1} - \lambda_k}{\lambda_{i+1} + \lambda_k}\right]^2 e(x_i). \tag{5}$$