Preprints of the 21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

Minimax Adaptive Control for State Matrix with Unknown Sign

Anders Rantzer

Lund University, Sweden (e-mail: rantzer@control.lth.se).

Abstract: For linear time-invariant systems having a state matrix with uncertain sign, we formulate and solve a minimax adaptive control problem as a zero sum dynamic game. Explicit expressions for the optimal value function and the optimal control law are given in terms of a Riccati equation. The optimal control law is adaptive in the sense that past data is used to estimate the uncertain sign for prediction of future dynamics. Once the sign has been estimated, the controller behaves like standard H_∞ optimal state feedback.

Keywords: Adaptive control, linear systems, robust control, game theory

1. INTRODUCTION

The history of adaptive control dates back at least to aircraft autopilot development in the 1950s. Following the landmark paper Åström and Wittenmark [1973], a surge of research activity during the 1970s derived conditions for convergence, stability, robustness and performance under various assumptions. For example, Ljung [1977] analysed adaptive algorithms using averaging, Goodwin et al. [1981] derived an algorithm that gives mean square stability with probability one, while Guo [1995] analyzed the optimal asymptotic rate of convergence. On the other hand, conditions that may cause instability were studied in Egardt [1979], Ioannou and Kokotovic [1984] and Rohrs et al. [1985]. Altogether, the subject has a rich history documented in numerous textbooks, such as Åström and Wittenmark [2013], Goodwin and Sin [2014], Sastry and Bodson [2011] and Astolfi et al. [2007]. In this paper, the focus is on worst-case models for disturbances and uncertain parameters, as discussed in Cusumano and Poolla [1988], Sun and Ioannou [1987], Megretski and Rantzer [2003]. The "minimax adaptive" paradigm was introduced for linear systems in Didinsky and Basar [1994] and nonlinear systems in Pan and Basar [1998].

The outline of the paper is as follows: Sections 2 introduces notation. Section 3 states the problem and reformulates it as a zero sum dynamic game on standard form. The main results are presented in section 4 together with an example. Proofs are given in section 5, followed by concluding remarks in section 6.

⋆ The author is a member of the excellence center ELLIIT. Financial support was obtained from the Swedish Research Council and the European Research Council (ERC) Advanced Grant No.834142 (ScalableControl). The work was also partially supported by the Wallenberg AI, Autonomous Systems and Software Program (WASP) funded by the Knut and Alice Wallenberg Foundation.

2. NOTATION

The set of $n \times m$ matrices with real coefficients is denoted $\mathbb{R}^{n\times m}$. The transpose of a matrix A is denoted $A^\top$. For a symmetric matrix $A \in \mathbb{R}^{n\times n}$, we write $A \succ 0$ to say that A is positive definite, while $A \succeq 0$ means positive semi-definite. For $A, B \in \mathbb{R}^{n\times m}$, the expression $\langle A, B\rangle$ denotes the trace of $A^\top B$. Given $x \in \mathbb{R}^n$ and $A \in \mathbb{R}^{n\times n}$, the notation $|x|^2_A$ means $x^\top Ax$. Similarly, given $B \in \mathbb{R}^{m\times n}$ and $A \in \mathbb{R}^{n\times n}$, the trace of $B^\top AB$ is denoted $\|B\|^2_A$. For $y \in \mathbb{R}$, define $\text{sat}(y)$ to be 1 if $y > 1$, -1 if $y < -1$ and otherwise equal to y.

3. MINIMAX ADAPTIVE CONTROL

This paper is devoted to the following problem:

Let $Q \in \mathbb{R}^{n\times n}$ and $R \in \mathbb{R}^{m\times m}$ be positive definite matrices and let $B \in \mathbb{R}^{n\times m}$. Given $A \in \mathbb{R}^{n\times n}$, and a number $\gamma > 0$, find, if possible, a control law μ that for every initial state x_0 attains the infimum

$$\inf_{\mu} \sup_{w,i,N} \sum_{t=0}^{N} \left(|x_t|^2_Q + |u_t|^2_R - \gamma^2 |w_t|^2\right), \tag{1}$$

where $i \in \{-1, 1\}$, $w_t \in \mathbb{R}^n$, $N \geq 0$ and the sequences x and u are generated according to

$$x_{t+1} = iAx_t + Bu_t + w_t \qquad t \geq 0 \tag{2}$$
$$u_t = \mu_t(x_0, \ldots, x_t, u_0, \ldots, u_{t-1}). \tag{3}$$

The problem can be viewed as a dynamic game, where the μ-player tries to minimize the cost, while the (w, i)-player tries to maximize it. If it wasn't for the parameter i, this would be the standard game formulation of H_∞ optimal control Basar and Bernhard [1995]. In our formulation, the maximizing player can choose not only w, but also the parameter i. This parameter is unknown, but constant, so an optimal feedback law tends to "learn" the value of i in the beginning, in order to exploit this knowledge later. Such nonlinear adaptive controllers can stabilize and optimize the behavior also when no linear controller can simultaneously stabilize (2) for both $i = 1$ and $i = -1$.

To accommodate the uncertainty in i when deciding u_t, it is sufficient for the controller to consider historical data collected in the matrix

$$Z_t = \sum_{\tau=0}^{t-1} \begin{bmatrix} Bu_\tau - x_{\tau+1} \\ x_\tau \end{bmatrix} \begin{bmatrix} Bu_\tau - x_{\tau+1} \\ x_\tau \end{bmatrix}^\top, \tag{4}$$

since this gives $\| [I \quad iA]^\top \|_{Z_t}^2 = \sum_{\tau=0}^{t-1} |w_\tau|^2$.

In fact, our problem can be reformulated as follows:

Given $Q \succ 0, R \succ 0, \gamma > 0$ and a system

$$\begin{cases} x_{t+1} = v_t \\ Z_{t+1} = Z_t + \begin{bmatrix} Bu_t - v_t \\ x_t \end{bmatrix} \begin{bmatrix} Bu_t - v_t \\ x_t \end{bmatrix}^\top, \quad Z_0 = 0, \end{cases} \tag{5}$$

find, if possible, a control law

$$u_t = \eta(x_t, Z_t) \tag{6}$$

that attains the infimum

$$\inf_\eta \sup_{v,N} \left[\sum_{t=0}^{N} \left(|x_t|_Q^2 + |u_t|_R^2 \right) - \gamma^2 \min_i \| [I \quad iA]^\top \|_{Z_{N+1}}^2 \right] \tag{7}$$

when x, u, Z are generated from v and x_0 using (5)-(6).

In this formulation, the unknown sign i does not appear in the dynamics, only in the penalty of the final state. As a consequence, no past states are needed in the control law (6), only the state (x_t, Z_t). In fact, the problem is a standard zero-sum dynamic game Basar and Olsder [1999], which can be addressed by dynamic programming. Hence we define the Bellman operator $V \mapsto \mathcal{F}V$ by

$$\mathcal{F}V(x, Z) :=$$

$$\min_u \max_v \left\{ |x|_Q^2 + |u|_R^2 + V\left(v, Z + \begin{bmatrix} Bu - v \\ x \end{bmatrix} \begin{bmatrix} Bu - v \\ x \end{bmatrix}^\top \right) \right\}$$

and conclude this section by stating the following:

Theorem 1. Given A, B, Q, R, define the operator $\mathcal{F}$ as above and $V_0, V_1, V_2 \ldots$ according to the iteration

$$V_0(x, Z) = -\gamma^2 \min_{i=\pm 1} \| [I \quad iA]^\top \|_Z^2 \tag{8}$$

$$V_{k+1}(x, Z) = \mathcal{F}V_k(x, Z) \tag{9}$$

The expressions (1) and (7) have finite values if and only if the sequence $\{V_k(x, 0)\}_{k=0}^\infty$ is upper bounded, in which case the limit $V_* := \lim_{k\to\infty} V_k$ exists and $V_*(x_0, 0)$ is equal to the values of (1) and (7). Defining $\eta(x, Z)$ as the minimizing value of u in the expression for $\mathcal{F}V_*(x, Z)$ gives an optimal η for (7), while the control law μ defined by

$$\mu_t(x_0, \ldots, x_t, u_0, \ldots, u_{t-1}) = \eta\left(x_t, \sum_{\tau=0}^{t-1} \begin{bmatrix} Bu_\tau - x_{\tau+1} \\ x_\tau \end{bmatrix} \begin{bmatrix} Bu_\tau - x_{\tau+1} \\ x_\tau \end{bmatrix}^\top \right) \tag{10}$$

is optimal for (1).

Theorem 2. With notation as in Theorem 1, suppose that (1) has a finite value and let $V_* := \lim_{k\to\infty} V_k$. Then the Riccati equation

$$|x|_P^2 = \min_u \max_v \left\{ |x|_Q^2 + |u|_R^2 - \gamma^2 |Ax + Bu - v|^2 + |v|_P^2 \right\} \tag{11}$$

has a solution $0 \prec P \prec \gamma^2 I$ and the sequence defined by

$$\bar{V}_0(x, Z) = |x|_P^2 - \gamma^2 \min_{i=\pm 1} \| [I \quad iA]^\top \|_Z^2 \tag{12}$$

$$\bar{V}_{k+1}(x, Z) = \mathcal{F}\bar{V}_k(x, Z). \tag{13}$$

satisfies $\bar{V}_0 \le \bar{V}_1 \le \cdots \le \lim_{k\to\infty} \bar{V}_k = V_*$.

4. AN EXPLICIT OPTIMAL CONTROL LAW

The following result, Theorem 3, specifies a minimax optimal adaptive controller on explicit form for a range of γ-values. It is followed by Theorem 4, which gives a lower bound on the values of γ for which a solution exists.

Theorem 3. Given $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{m\times m}$ and some positive definite $Q \in \mathbb{R}^{n\times n}$, $R \in \mathbb{R}^{m\times m}$, assume that (11) has a solution $0 \prec P \prec \gamma^2 I$, with minimizing argument $u = -Kx$. Define $T := Q + A^\top (P^{-1} - \gamma^{-2} I)^{-1} A$. and suppose that $T \prec \gamma^2 I$, while

$$Q + K^\top \left[R + B^\top (T^{-1} - \gamma^{-2} I)^{-1} B \right] K \preceq 2T - P + \gamma^2 A^\top A. \tag{14}$$

Then (1) has a finite value and the optimal control law

$$u_t = \text{sat}\left(\frac{\gamma^2 \sum_{\tau=0}^{t-1} (Bu_\tau - x_{\tau+1})^\top A x_\tau}{|x_t|_{T-P}^2} \right) Kx. \tag{15}$$

Moreover, define the sequence $\{\bar{V}_k\}_{k=0}^\infty$ by (12)-(13) and let $Y := \gamma^2 [I \ 0] Z [0 \ I]^\top$. Then

$$\bar{V}_1(x, Z) = \bar{V}_2(x, Z) = \cdots =$$

$$= \begin{cases} |x|_P^2 - \gamma^2 \min_{i=\pm 1} \| [I \ iA]^\top \|_Z^2 & \text{if } |\langle A, Y\rangle| \ge |x|_{T-P}^2 \\ |x|_T^2 - \gamma^2 \| \operatorname{diag}\{I, A\}^\top \|_Z^2 + \langle A, Y\rangle^2 |x|_{T-P}^{-2} & \text{otherwise.} \end{cases} \tag{16}$$

Theorem 4. With A, B, P, Q, R, T as in Theorem 3, (1) has no finite value unless $0 \prec P \prec \gamma^2 I$ and $T \preceq \gamma^2 I$.

Remark 1. The intuition behind the optimal control law in Theorem 3 is simple: The cases $u_t = Kx_t$ and $u_t = -Kx_t$ describe the situation when historical data collected in the expression $\sum_{\tau=0}^{t-1} (Bu_\tau - x_{\tau+1})^\top A x_\tau$ is rich enough to make a reliable estimate about the uncertain parameter i. This estimate is then used as truth and the corresponding H_∞ state feedback control law is applied. In the intermediate case, the historical data does not give a conclusive answer, so the controller gain is down-scaled accordingly.

Example 1. Consider now the case $n = m = Q = R = A = B = 1$. First of all, Theorem 4 shows that the game has no finite value unless $\gamma \ge 2.01$. On the other hand, Theorem 3 gives an optimal strategy for the dynamic game (1) whenever $\gamma \ge 2.1851$. Specifically, consider the case $\gamma = 2.1851$. Solving the Riccati equation gives $P = 1.7308$, which is clearly in the interval $[0, \gamma^2]$. It follows that $T = 3.7150$ and condition (14) marginally holds. For larger γ, the margin would be bigger.

An exact expression for the value function V_* is now given by the formula Theorem 3, which shows that

$$V_*\left(x, \begin{bmatrix} z_{11} & z_{12} \\ z_{12} & z_{22} \end{bmatrix} \right)$$

$$= \begin{cases} 1.73x^2 - 4.77(z_{11} + z_{22} - 2|z_{12}|) & \text{if } |z_{12}| \ge 0.42x^2 \\ 3.72x^2 - 4.77(z_{11} + z_{22}) + \dfrac{11.49 z_{12}^2}{x^2} & \text{otherwise} \end{cases}$$

and the optimal control law is

$$u_t = \text{sat}\left(\frac{2.41 \sum_{\tau=0}^{t-1} (u_\tau - x_{\tau+1}) x_\tau}{x_t^2} \right) 0.73 x_t.$$

□

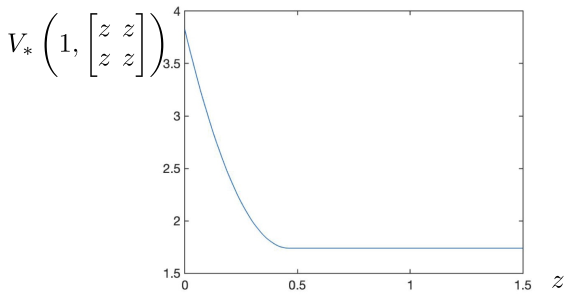

Fig. 1. Optimal value function plot for Example 1. The variable z represents information about A. The cost is maximal for $z = 0$ (no information). For $|z| \geq 0.42$, the cost is the same as if the value of A was known.

5. THE MAIN PROOFS.

Proof of Theorem 1. First note that $V_1 \geq V_0$, so the sequence $V_0, V_1, V_2, \dots$ is monotonically non-decreasing.

For any fixed $N \geq 0$, the value of (1) is bounded below by the expression

$$\inf_{\mu} \sup_{w,i} \sum_{t=0}^{N} \left(|x_t|_Q^2 + |u_t|_R^2 - \gamma^2 |w_t|^2\right), \tag{17}$$

where $i \in \{-1, 1\}$, $w_t \in \mathbb{R}^n$ and the sequences x and u are generated according to (2)-(3). The value of (17) grows monotonically with N and (1) is obtained in the limit. A change of variables with $v_t := x_{t+1}$ and Z_t given by (4) shows that (17) is equal to

$$\inf_{\mu} \sup_{v} \left[-\gamma^2 \min_i \left\| \begin{bmatrix} I & iA \end{bmatrix}^\top \right\|_{Z_{N+1}}^2 + \sum_{t=0}^{N} \left(|x_t|_Q^2 + |u_t|_R^2\right) \right] \tag{18}$$

where x, Z, u are generated by (5) combined with (3). Standard dynamic programming shows that the value of (18) is $V_{N+1}(x_0, 0)$, where V_k is defined by (8)-(9). This proves that (1) has a finite value if and only if the sequence $\{V_k(x, 0)\}_{k=0}^{\infty}$ is upper bounded, in which case the limit $V_* := \lim_{k\to\infty} V_k$ exists and $V_*(x_0, 0)$ is equal to the value of (1).

If (7) is finite, then (18) is bounded above by (7), so also $V_* := \lim_{k\to\infty} V_k$ is finite. Conversely, if V_* is finite, we may define $\eta(x, Z)$ as a minimizing value of u in the expression for $\mathcal{F}V_*(x, Z)$. Then define the sequence $W_0, W_1, W_2, \dots$ recursively by $W_0 = V_0$ and

$$\begin{aligned} W_{k+1}(x, Z) = \max_v \Bigg\{ & |x|_Q^2 + |\eta(x, Z)|_R^2 \\ & + W_k\left(v, Z + \begin{bmatrix} B\eta(x, Z) - v \\ x \end{bmatrix} \begin{bmatrix} B\eta(x, Z) - v \\ x \end{bmatrix}^\top \right) \Bigg\}. \end{aligned}$$

By dynamic programming,

$$\begin{aligned} & W_N(x, 0) \\ & = \sup_v \left[-\gamma^2 \min_i \left\| \begin{bmatrix} I & iA \end{bmatrix}^\top \right\|_{Z_{N+1}}^2 + \sum_{t=0}^{N} \left(|x_t|_Q^2 + |u_t|_R^2\right) \right], \end{aligned}$$

where x, Z, u are generated by (5) combined with (6). Hence (7) is bounded above by $\lim_{k\to\infty} W_k(x_0, 0)$. The definitions of V_k and W_k give by induction $V_* \geq W_k \geq V_k$ for all k, so $\lim_{k\to\infty} W_k = V_*$. This proves that the value of (7) equals $V_*(x_0, 0)$ and η is a minimizing argument.

Proof of Theorem 2. Suppose that (1) has a finite value. By Theorem 1, this implies that the sequence $\{V_k(x, 0)\}_{k=0}^{\infty}$ defined by (8)-(9) is upper bounded. Define

$$\begin{aligned} V_k^+(x, Z) &:= |x|_{P_k}^2 - \gamma^2 \left\| \begin{bmatrix} I & A \end{bmatrix}^\top \right\|_Z^2 \\ V_k^-(x, Z) &:= |x|_{P_k}^2 - \gamma^2 \left\| \begin{bmatrix} I & -A \end{bmatrix}^\top \right\|_Z^2, \end{aligned}$$

where $P_0 = 0$ and P_k is given by the Riccati recursion

$$|x|_{P_{k+1}}^2 = \min_u \max_v \left\{ |x|_Q^2 + |u|_R^2 - \gamma^2 |Ax + Bu - v|^2 + |v|_{P_k}^2 \right\}.$$

Then $V_k(x, Z) \geq \max\{V_k^+(x, Z), V_k^-(x, Z)\}$ for all k. This is trivial for $k = 0$ and follows by induction for $k > 0$, since $\mathcal{F}V_{k+1}^+ = V_k^+$ and $\mathcal{F}V_{k+1}^- = V_k^-$. In the limit, it follows that the limit $P = \lim_{k\to\infty} P_k$ exists and

$$V_*(x, Z) \geq \bar{V}_0(x, Z). \tag{19}$$

Repeated application of $\mathcal{F}$ gives $V_* = \lim_{k\to\infty} \bar{V}_k$.

Before proving Theorem 3 and Theorem 4, consider first a more limited problem:

$$\min_u \max_{i\in\{-1,1\}} \left\{ |x|_Q^2 + |u|_R^2 + |iAx + Bu|_S^2 - 2\langle iA, Y\rangle \right\}$$

where Y is an arbitrary matrix parameter. In other words: The problem is to find a control signal u to minimize a worst case quadratic cost for $\pm A$, with Y representing prior knowledge. The solution is given by the following lemma:

Lemma 5. Given A, B, P, Q, R, S, suppose that

$$|x|_P^2 = \min_u \left\{ |x|_Q^2 + |u|_R^2 + |Ax + Bu|_S^2 \right\}, \tag{20}$$

where the minimizing u is given by $u = -Kx$ with

$$K := (R + B^\top SB)^{-1} B^\top SA. \tag{21}$$

Put $T := Q + A^\top SA$. Then

$$\begin{aligned} & \min_u \max_{i\in\{-1,1\}} \left\{ |x|_Q^2 + |u|_R^2 + |iAx + Bu|_S^2 - 2\langle iA, Y\rangle \right\} \\ & = \begin{cases} |x|_P^2 + 2|\langle A, Y\rangle| & \text{if } |\langle A, Y\rangle| \geq |x|_{T-P}^2 \\ |x|_T^2 + \langle A, Y\rangle^2 |x|_{T-P}^{-2} & \text{otherwise} \end{cases} \\ & = \max_{|\theta|\leq 1} \left\{ |x|_T^2 - \theta^2 |x|_{T-P}^2 + 2\theta \langle A, Y\rangle \right\} \end{aligned}$$

and the minimizing value of u is

$$\hat{u} = \mathrm{sat}\left(\frac{\langle A, Y\rangle}{|x|_{T-P}^2} \right) Kx \tag{22}$$

Proof. The definition of K gives

$$B^\top SA = (R + B^\top SB)K.$$

Multiplication by $K^\top$ from the left, and application of the identity

$$P = Q + K^\top RK + (A - BK)^\top S(A - BK)$$

gives

$$K^\top B^\top SA = A^\top SBK = K^\top (R + B^\top SB)K = T - P.$$

The minimax theorem for convex-concave functions gives

$$\min_u \max_{i\in\{-1,1\}} \left\{ |x|_Q^2 + |u|_R^2 + |iAx + Bu|_S^2 - 2\langle iA, Y\rangle \right\}$$

$$\max_\theta \min_u \sum_{i\in\{-1,1\}} \theta_i \left\{ |x|_Q^2 + |u|_R^2 + |iAx + Bu|_S^2 - 2\langle iA, Y\rangle \right\}$$

where $\theta \in [-1, 1]$, $\theta_{-1} = (1 + \theta)/2$ and $\theta_1 = (1 - \theta)/2$. If $\langle A, Y\rangle \geq |x|_{T-P}^2$, the maximum over θ is attained by $\theta = 1$ and the value is $|x|_P^2 + 2\langle A, Y\rangle$. On the other hand, if $\langle A, Y\rangle \leq -|x|_{T-P}^2$, the maximum is given by $\theta = -1$ and the value is $|x|_P^2 - 2\langle A, Y\rangle$. Finally, if $|\langle A, Y\rangle| < |x|_{T-P}^2$, the optimal value of θ is in the interior of the interval $(-1, 1)$ and determined by

$$\begin{aligned}|Ax - Bu|_S^2 + 2\langle A, Y\rangle &= |Ax + Bu|_S^2 - 2\langle A, Y\rangle \\ \langle A, Y\rangle &= x^\top A^\top SBu \\ &= \theta x^\top A^\top SBKx \\ &= \theta |x|_{T-P}^2.\end{aligned}$$

This gives $u = |x|_{T-P}^{-2}\langle A, Y\rangle Kx$ and the value

$$\begin{aligned}&\sum_i \theta_i \left\{|x|_Q^2 + |u|_R^2 + |A_i x + Bu|_S^2 - 2\langle A_i, Y\rangle\right\} \\ &= |x|_Q^2 + |u|_R^2 + |Ax|_S^2 + |Bu|_S^2 \\ &= |x|_Q^2 + |Ax|_S^2 + |\theta Kx|_{R+B^\top SB}^2 \\ &= |x|_T^2 + \theta^2 |x|_{T-P}^2 \\ &= |x|_T^2 + \langle A, Y\rangle^2 |x|_{T-P}^{-2}.\end{aligned}$$

Equipped with Lemma 5, we are now ready to prove the main results:

Proof of Theorem 3. Putting $S := (P^{-1} - \gamma^{-2} I)^{-1}$ and eliminating v from the definition of P gives (20) and (21). We will first prove the expression (16) for $\bar{V}_1$. With

$$Z := \begin{bmatrix} Z_{vv} & \gamma^{-2} Y \\ \gamma^{-2} Y^\top & Z_{xx}\end{bmatrix},$$

we have

$$\begin{aligned}&\bar{V}_1(x, Z) + \gamma^2 \operatorname{trace}(Z_{vv}) + \gamma^2 \operatorname{trace}(A Z_{xx} A^\top) - |x|_Q^2 \\ &= \inf_u \sup_v \left\{|u|_R^2 + \bar{V}_0\left(v, \gamma^{-2}\begin{bmatrix} 0 & Y \\ Y^\top & 0\end{bmatrix} + \begin{bmatrix} Bu - v \\ x\end{bmatrix}\begin{bmatrix} Bu - v \\ x\end{bmatrix}^\top\right)\right\} \\ &= \min_u \max_{v,i} \left\{|u|_R^2 + |v|_P^2 - \gamma^2 |iAx + Bu - v|^2 - 2\langle iA, Y\rangle\right\} \\ &= \min_u \max_i \left\{|u|_R^2 + |iAx + Bu|_S^2 - 2\langle iA, Y\rangle\right\}.\end{aligned}$$

Hence Lemma 5 gives

$$\begin{aligned}&\bar{V}_1(x, Z) \\ &= \begin{cases} |x|_P^2 - \gamma^2 \min_{i=\pm 1} \left\| \begin{bmatrix} I & iA\end{bmatrix}^\top \right\|_Z^2 & \text{if } |\langle A, Y\rangle| \geq |x|_{T-P}^2 \\ |x|_T^2 - \gamma^2 \left\| \operatorname{diag}\{I, A\}^\top \right\|_Z^2 + \langle A, Y\rangle^2 |x|_{T-P}^{-2} & \text{otherwise,}\end{cases}\end{aligned}$$

which is the desired expression for $\bar{V}_1$.

The next step will be to prove that $\mathcal{F}\bar{V}_1 \leq \bar{V}_1$ (which implies $\mathcal{F}\bar{V}_1 = \bar{V}_1$). Define

$$\begin{aligned}&W_{\Theta,\theta}(x, Z) \\ &:= |x|_T^2 - \Theta |x|_{T-P}^2 + 2\theta\langle A, Y\rangle - \gamma^2 \left\| \operatorname{diag}\{I, A\}^\top \right\|_Z^2.\end{aligned}$$

and notice that Lemma 5 also gives

$$\bar{V}_1(x, Z) = \max_{\theta^2 \leq \Theta \leq 1} W_{\Theta,\theta}(x, Z).$$

Let $\hat{u}$ be defined by (22) and note that $\mathcal{F}\bar{V}_1(x, Z)$ is bounded above by

$$\max_{\Theta,\theta,v}\left\{|x|_Q^2 + |\hat{u}|_R^2 + W_{\Theta,\theta}\left(v, Z + \begin{bmatrix} B\hat{u} - v \\ x\end{bmatrix}\begin{bmatrix} B\hat{u} - v \\ x\end{bmatrix}^\top\right)\right\}.$$

We will now show that the maximum over Θ, θ, v is bounded above by $\bar{V}_1(x, Z)$. Let $X := (T^{-1} - \gamma^{-2} I)^{-1}$ and consider first the case $\Theta = 0$:

$$\begin{aligned}&\max_v \left\{|\hat{u}|_R^2 + W_{0,0}\left(v, \gamma^{-2}\begin{bmatrix} 0 & Y \\ Y^\top & 0\end{bmatrix} + \begin{bmatrix} B\hat{u} - v \\ x\end{bmatrix}\begin{bmatrix} B\hat{u} - v \\ x\end{bmatrix}^\top\right)\right\} \\ &= \max_v \left\{|\hat{u}|_R^2 + |v|_T^2 - \gamma^2 |B\hat{u} - v|^2 - \gamma^2 |Ax|^2\right\} \\ &= |\hat{u}|_R^2 + |B\hat{u}|_X^2 - \gamma^2 |Ax|^2 \\ &= x^\top \left[-\gamma^2 A^\top A + K^\top (R + B^\top XB) K \operatorname{sat}\left(\frac{\langle A, Y\rangle^2}{|x|_{T-P}^4}\right)\right] x \\ &\leq x^\top \left[-\gamma^2 A^\top A + (2T - P - Q + \gamma^2 A^\top A) \operatorname{sat}\left(\frac{\langle A, Y\rangle^2}{|x|_{T-P}^4}\right)\right] x \\ &\leq |x|_{P-Q}^2 + 2|\langle A, Y\rangle| \\ &\leq \bar{V}_1\left(x, \gamma^{-2}\begin{bmatrix} 0 & Y \\ Y^\top & 0\end{bmatrix}\right) - |x|_Q^2\end{aligned}$$

Next consider $\Theta = 1$:

$$\begin{aligned}&\max_{v, |\theta| \leq 1}\left\{|\hat{u}|_R^2 + W_{1,\theta}\left(v, \gamma^{-2}\begin{bmatrix} 0 & Y \\ Y^\top & 0\end{bmatrix} + \begin{bmatrix} B\hat{u} - v \\ x\end{bmatrix}\begin{bmatrix} B\hat{u} - v \\ x\end{bmatrix}^\top\right)\right\} \\ &= \max_v \max_{\theta = \pm 1}\left\{|\hat{u}|_R^2 + |v|_P^2 - \gamma^2 |\theta A + B\hat{u} - v|^2 - 2\langle \theta A, Y\rangle\right\} \\ &= \max_{\theta = \pm 1}\left\{|\hat{u}|_R^2 + |\theta A + B\hat{u}|_S^2 - 2\langle \theta A, Y\rangle\right\} \\ &= \bar{V}_1\left(x, \gamma^{-2}\begin{bmatrix} 0 & Y \\ Y^\top & 0\end{bmatrix}\right) - |x|_Q^2\end{aligned}$$

Notice that $W_{\Theta,\theta}$ is linear in Θ. Maximization over v and θ gives an expression that is convex in Θ. Hence, any bound that holds for $\Theta = 0$ and $\Theta = 1$ must be valid for all $\Theta \in [0, 1]$. Subtracting $\gamma^2 \left\| \operatorname{diag}\{I, A\}^\top \right\|_Z^2$ from all terms gives the desired inequality

$$\begin{aligned}&\max_{\Theta,\theta,v}\left\{|x|_Q^2 + |\hat{u}|_R^2 + W_{\Theta,\theta}\left(v, Z + \begin{bmatrix} B\hat{u} - v \\ x\end{bmatrix}\begin{bmatrix} B\hat{u} - v \\ x\end{bmatrix}^\top\right)\right\} \\ &\leq \bar{V}_1(x, Z) \qquad (23)\end{aligned}$$

and the proof that $\mathcal{F}\bar{V}_1 = \bar{V}_1$ is complete. It follows trivially that $\bar{V}_k = \bar{V}_1$ for $k > 1$ and $\hat{u}$ defines the optimal control law.

Proof of Theorem 4. Inserting the bound (19) into the right hand side of the Bellman equation gives

$$\begin{aligned}&V_*(x, Z) - |x|_Q^2 \\ &\geq \mathcal{F}\bar{V}_0(x, Z) - |x|_Q^2 \\ &= \min_u \max_{v,i}\left\{|u|_R^2 + |v|_P^2 - \gamma^2 |iAx + Bu - v|^2 - \gamma^2 \left\| \begin{bmatrix} I & iA\end{bmatrix}^\top \right\|_Z^2\right\} \\ &= \min_u \max_i \left\{|u|_R^2 + |iAx + Bu|_S^2 - \gamma^2 \left\| \begin{bmatrix} I & iA\end{bmatrix}^\top \right\|_Z^2\right\} \\ &\geq |x|_{T-Q}^2 - \gamma^2 \left\| \operatorname{diag}\{I, A\}^\top \right\|_Z^2\end{aligned}$$

where the second inequality follows from Lemma 5. Inserting the new bound $V_*(x, Z) \geq |x|_T^2 - \gamma^2 \left\| \operatorname{diag}\{I, A\}^\top \right\|_Z^2$ into the Bellman equation in the same way gives

$$\begin{aligned}&V_*(x, 0) \\ &\geq \min_u \max_{v,i}\left\{|x|_Q^2 + |u|_R^2 + |v|_T^2 - \gamma^2 |Bu - v|^2 - \gamma^2 |Ax|^2\right\}\end{aligned}$$

The last inequality shows that $T \preceq \gamma^2 I$, so the proof is complete.

6. CONCLUDING REMARKS

In this paper, we have formulated a control problem for uncertain linear systems as a zero-sum dynamic game. The solution is remarkable for two reasons:

(1) The dynamic programming formulation has an explicit solution in terms of a Riccati equation.

(2) The resulting optimal controller is adaptive: It reduces the aggressiveness of the controller until until enough data has been collected to get a parameter estimate that can be confidently trusted.

The results are likely to be extendable to many other uncertainty structures. The case of uncertain input matrix B will be particularly important, since the controller then needs to make active exploration in order to collect enough data for the exploitation phase.

REFERENCES

Alessandro Astolfi, Dimitrios Karagiannis, and Romeo Ortega. *Nonlinear and adaptive control with applications.* Springer Science & Business Media, 2007.

Karl Johan Åström and Björn Wittenmark. On self-tuning regulators. *Automatica*, 9:185–199, January 1973.

Karl Johan Åström and Björn Wittenmark. *Adaptive control.* Courier Corporation, 2013.

T. Basar and P. Bernhard. *H_∞-Optimal Control and Related Minimax Design Problems —A Dynamic Game Approach.* Birkhauser, 1995.

Tamer Basar and Geert Jan Olsder. *Dynamic noncooperative game theory*, volume 23. Siam, 1999.

Salvatore J Cusumano and Kameshwar Poolla. Nonlinear feedback vs. linear feedback for robust stabilization. In *Decision and Control, 1988., Proceedings of the 27th IEEE Conference on*, pages 1776–1780. IEEE, 1988.

Garry Didinsky and Tamer Basar. Minimax adaptive control of uncertain plants. In *Proceedings of 1994 33rd IEEE Conference on Decision and Control*, volume 3, pages 2839–2844. IEEE, 1994.

Bo Egardt. *Stability of Adaptive Controllers.* Springer-Verlag, Berlin, FRG, January 1979.

Graham C Goodwin and Kwai Sang Sin. *Adaptive filtering prediction and control.* Courier Corporation, 2014.

Graham C Goodwin, Peter J Ramadge, and Peter E Caines. Discrete time stochastic adaptive control. *SIAM Journal on Control and Optim.*, 19(6):829–853, 1981.

Lei Guo. Convergence and logarithm laws of self-tuning regulators. *Automatica*, 31(3):435–450, 1995.

Petros A Ioannou and Petar V Kokotovic. Instability analysis and improvement of robustness of adaptive control. *Automatica*, 20(5):583–594, 1984.

Lennart Ljung. Analysis of recursive stochastic algorithms. *IEEE Transactions on Automatic Control*, AC-22:551–575, January 1977.

A. Megretski and A. Rantzer. Bounds on the optimal ℓ_2-gain in adaptive control of a first order linear system. Technical Report ISRN IML-R–41-02/03–SE+spring, Institut Mittag-Leffler, The Royal Swedish Academy of Sciences, Stockholm, Sweden, January 2003.

Zigang Pan and Tamer Basar. Adaptive controller design for tracking and disturbance attenuation in parametric strict-feedback nonlinear systems. *IEEE Transactions on Automatic Control*, 43(8):1066–1083, 1998.

C Rohrs, Lena Valavani, Michael Athans, and Gunter Stein. Robustness of continuous-time adaptive control algorithms in the presence of unmodeled dynamics. *IEEE Transactions on Automatic Control*, 30(9):881–889, 1985.

Shankar Sastry and Marc Bodson. *Adaptive control: stability, convergence and robustness.* Courier Corporation, 2011.

J Sun and PA Ioannou. The theory and design of robust adaptive controllers. *Automatica*, pages 19–24, 1987.