A Data-driven Riccati Equation

Anders Rantzer   RANTZER@CONTROL.LTH.SE
Lund University, Sweden

Abstract

Certainty equivalence adaptive controllers are analysed using a "data-driven Riccati equation", corresponding to the model-free Bellman equation used in Q-learning. The equation depends quadratically on data correlation matrices. This makes it possible to derive simple sufficient conditions for stability and robustness to unmodeled dynamics in adaptive systems. The paper is concluded by short remarks on how the bounds can be used to quantify the interplay between excitation levels and robustness to unmodeled dynamics.

Keywords: dual control, adaptive control, online learning, linear quadratic control

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) gave conditions for 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), and Sastry and Bodson (2011).

Recently, there has been a renewed interest in analysis of adaptive controllers, driven by progress in statistical machine learning. See Tsiamis et al. (2023) for a review. In parallel, there is also a rapidly developing literature on (off-line) data-driven control. De Persis and Tesi (2019); Markovsky and Dörfler (2021); Berberich et al. (2020).

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); Vinnicombe (2004); Megretski (2004) and more recently in Rantzer (2021); Cederberg et al. (2022); Kjellqvist and Rantzer (2022). However, the disturbances in this paper are assumend to be bounded in terms of past states and inputs. This causality constraint is different from above mentioned references.

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. Given $x \in \mathbb{R}^n$ and $A \in \mathbb{R}^{n \times n}$, the notation $|x|_A^2$ means $x^\top A x$. The expression $\min_K \begin{bmatrix} I \\ K \end{bmatrix}^\top Q \begin{bmatrix} I \\ K \end{bmatrix}$ is equivalent to $Q^{xx} - Q^{xu}(Q^{uu})^{-1}Q^{ux}$ when $Q = \begin{bmatrix} Q^{xx} & Q^{xu} \\ Q^{ux} & Q^{uu} \end{bmatrix} \succ 0$.

3. A Data-driven Riccati Equation

Consider a linear quadratic optimal control problem:

$$\text{Minimize} \sum_{t=0}^{\infty} \left(|x_t|^2 + |u_t|^2 \right)$$

$$\text{subject to } x_{t+1} = Ax_t + Bu_t,$$

with $x_0 \in \mathbb{R}^n$ given and $u_t \in \mathbb{R}^m$. Assuming that the system is stabilizable, the optimal value has the form $|x_0|_P^2$ where $P \in \mathbb{R}^{n \times n}$ can be obtained by solving the Riccati equation

$$P = \min_K \left[I + K^\top K + (A + BK)^\top P (A + BK) \right]. \tag{1}$$

Define $Q := I + \begin{bmatrix} A & B \end{bmatrix}^\top P \begin{bmatrix} A & B \end{bmatrix} \in \mathbb{R}^{(n+m) \times (n+m)}$. Then (1) can be written as

$$Q - I = \begin{bmatrix} A & B \end{bmatrix}^\top \min_K \left(\begin{bmatrix} I \\ K \end{bmatrix}^\top Q \begin{bmatrix} I \\ K \end{bmatrix} \right) \begin{bmatrix} A & B \end{bmatrix} \tag{2}$$

or alternatively

$$\begin{bmatrix} x \\ u \end{bmatrix}^\top (Q - I) \begin{bmatrix} x \\ u \end{bmatrix} = x_+^\top \min_K \left(\begin{bmatrix} I \\ K \end{bmatrix}^\top Q \begin{bmatrix} I \\ K \end{bmatrix} \right) x_+ \tag{3}$$

where $x_+ = Ax + Bu$. The equation (3) is sometimes called *model free*, since it does not include the model parameters (A, B). This makes it possible to collect data points (x, u, x_+) and use (3) to get information about Q. In fact, the total matrix Q can be computed from a trajectory $x_0, u_0, \dots, x_{t-1}, u_{t-1}, x_t$ spanning all directions of (x_t, u_t), using the equation

$$\begin{bmatrix} x_0 & \dots & x_{t-1} \\ u_0 & \dots & u_{t-1} \end{bmatrix}^\top (Q - I) \begin{bmatrix} x_0 & \dots & x_{t-1} \\ u_0 & \dots & u_{t-1} \end{bmatrix} = \begin{bmatrix} x_1 & \dots & x_t \end{bmatrix}^\top \min_K \begin{bmatrix} I \\ K \end{bmatrix}^\top Q \begin{bmatrix} I \\ K \end{bmatrix} \begin{bmatrix} x_1 & \dots & x_t \end{bmatrix}$$

This is essentially equation (3) in Bradtke (1992) and (14) in Rizvi and Lin (2018). The fact that the equation is over-determined for $t > n + m$ makes it natural to multiply from the left and right by

$$\begin{bmatrix} \lambda^t x_0 & \lambda^{t-1} x_1 & \dots & x_{t-1} \\ \lambda^t u_0 & \lambda^{t-1} u_1 & \dots & u_{t-1} \end{bmatrix},$$

where $\lambda \in (0, 1)$ is a forgetting factor. With notation

$$\Sigma_t = \begin{bmatrix} \Sigma_t^{xx} & \Sigma_t^{xu} \\ \Sigma_t^{ux} & \Sigma_t^{uu} \end{bmatrix} = \sum_{k=0}^{t-1} \lambda^{t-1-k} \begin{bmatrix} x_k \\ u_k \end{bmatrix} \begin{bmatrix} x_k \\ u_k \end{bmatrix}^\top + \lambda^t \Sigma_0 \tag{4}$$

$$\hat{\Sigma}_t = \sum_{k=0}^{t-1} \lambda^{t-1-k} x_{k+1} \begin{bmatrix} x_k \\ u_k \end{bmatrix}^\top, \tag{5}$$

where we have added a regularization term $\lambda^t \Sigma_0 \succ 0$ to make Σ_t intervible, this gives a *data driven Riccati equation*

$$\Sigma_t \left(Q_t - I \right) \Sigma_t = \hat{\Sigma}_t^\top \min_K \left(\begin{bmatrix} I \\ K \end{bmatrix}^\top Q_t \begin{bmatrix} I \\ K \end{bmatrix} \right) \hat{\Sigma}_t. \tag{6}$$

The equation (6) makes it possible to estimate the optimal cost matrix Q directly from data. The model-free aspect should not be over-emphasized, since an algebraically equivalent approach is to first estimate A and B according to $\begin{bmatrix} \hat{A}_t & \hat{B}_t \end{bmatrix} := \hat{\Sigma}_t \Sigma_t^{-1}$ and then solve the standard Riccati equation

$$P_t = \min_K \left[I + K^\top K + (\hat{A}_t + \hat{B}_t K)^\top P_t (\hat{A}_t + \hat{B}_t K) \right]$$

in analogy with (1). Nevertheless, the fact that (6) is quadratic in the correlation matrices $\hat{\Sigma}_t$, Σ_t makes it a good starting point for analysis of data-driven controllers.

4. Problem formulation

For the linear system

$$x_{t+1} = A x_t + B u_t + w_t \qquad\qquad t \geq 0 \tag{7}$$

we will analyse controllers of the following form (See Figure 1):

$$\begin{cases} \Sigma_t = \lambda \Sigma_{t-1} + \begin{bmatrix} x_{t-1}^\top & u_{t-1}^\top \end{bmatrix}^\top \begin{bmatrix} x_{t-1}^\top & u_{t-1}^\top \end{bmatrix}, & \Sigma_0 \succ 0 \\ \hat{\Sigma}_t = \lambda \hat{\Sigma}_{t-1} + x_t \begin{bmatrix} x_{t-1}^\top & u_{t-1}^\top \end{bmatrix}, & \hat{\Sigma}_0 = 0 \\ u_t = K_t x_t + \epsilon_t \end{cases} \tag{8}$$

where $\lambda \in (0,1)$ and K_t is obtained from Σ_t and $\hat{\Sigma}_t$ as the minimizing argument in a solution of (6) and ϵ_t is added to provide excitation (exploration). In the adaptive control literature such controllers are known as *certainty equivalence* controllers, since the main feedback term $K_t x_t$ can be obtained by using the model estimate $\begin{bmatrix} \hat{A}_t & \hat{B}_t \end{bmatrix} := \hat{\Sigma}_t \Sigma_t^{-1}$ for optimization, disregarding its uncertainty.

We will give conditions that guarantee stability and performance (in terms of $\sum_t (|x_t|^2 + |u_t|^2)$) for the closed loop system. Obviously, this cannot be done without some prior constraints on the model (such as stabilizability). Many alternative formulations have been considered in the literature, Konstantinov et al. (1993),(Mania et al., 2019, Assumption 2), but the following simple version will be enough for our purposes. *Given any $\beta > 1$, we restrict the model parameters to be in the set $\mathcal{M}_\beta$ consisting of all pairs (A, B) such that the Riccati equation (2) has a solution Q with $I \preceq Q \preceq \beta^2 I$.*

We also need to put assumptions on the disturbance w. For this purpose, introduce the notation

$$\begin{bmatrix} \Sigma_t^{wx} & \Sigma_t^{wu} \end{bmatrix} := \sum_{k=0}^{t-1} \lambda^{t-1-k} w_k \begin{bmatrix} x_k^\top & u_k^\top \end{bmatrix} - \lambda^t \begin{bmatrix} A & B \end{bmatrix} \Sigma_0 \tag{9}$$

and notice that $\begin{bmatrix} \Sigma_t^{wx} & \Sigma_t^{wu} \end{bmatrix} = \hat{\Sigma}_t - \begin{bmatrix} A & B \end{bmatrix} \Sigma_t$. In statistical system analysis, w_t is often modeled as white noise with fixed amplitude. In this paper, we will instead think of w_t as unmodeled dynamics, which makes it natural to bound Σ_t^{wx} and Σ_t^{wu} in terms of Σ_t^{xx} and Σ_t^{uu}.
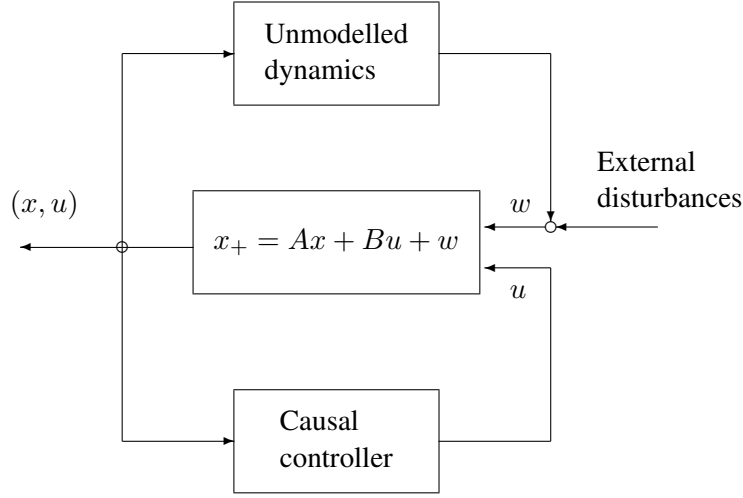
Figure 1: The system (7) is connected to a causal controller defined by (8). The disturbance w could represent a combination of external disturbances and unmodeled dynamics.

5. Main Results

Our first and main theorem is quantifying the effect in a single time-step of using the feedback gain K_t obtained from the data-driven Riccati equation instead of the optimal gain from the standard Riccati equation:

Theorem 1 *Consider $\beta, \rho \in \mathbb{R}_+$ with $2\beta^2\rho(\rho + 2) < 1$, a matrix $\hat{\Sigma}_t$ and a positive definite matrix Σ_t. Let $(A, B) \in \mathcal{M}_\beta$ with P satisfying Riccati equation (1), while*

$$\left\| \begin{bmatrix} A & B \end{bmatrix} - \hat{\Sigma}_t \Sigma_t^{-1} \right\| \le \rho. \tag{10}$$

Then (6) has a solution with minimizer K_t and

$$\frac{1}{1 - 2\beta^2\rho(\rho + 2)} P \succeq I + K_t^\top K_t + (A + BK_t)^\top P(A + BK_t). \tag{11}$$

Remark 1. In analysis of (7), the assumption (10) can be written as $\left\| \begin{bmatrix} \Sigma_t^{wx} & \Sigma_t^{wu} \end{bmatrix} \Sigma_t^{-1} \right\| \le \rho$. This inequality tends to fail for small values of t, but proper choice of excitation and regularization together with sufficiently small disturbances makes it valid after collection of enough data. $\qquad\square$

Remark 2. Notice that the optimal control law $\bar{K}$ from the standard Riccati equation satisfies

$$P = I + \bar{K}^\top \bar{K} + (A + B\bar{K})^\top P(A + B\bar{K}) \tag{12}$$

while K_t from the data driven Riccati equation gives

$$P_t = I + K_t^\top K_t + (\hat{A}_t + \hat{B}_t K_t)^\top P_t(\hat{A}_t + \hat{B}_t K_t)$$

4

The purpose of Theorem 1 is to show what happens to the storage function $x^\top P x$ when we use K_t instead of the optimal $\bar{K}$ to control the system. As expected, the difference vanishes as $\rho \to 0$. Larger values of β open up for systems that are harder to stabilizable, which means that ρ needs to be smaller and more accurate data is needed. This is reflected in the left hand side of (11). $\qquad \square$

The optimal closed loop system obtained with $u_t = \bar{K} x_t$ is stable, with the Lyapunov function $|x_t|_P^2$ decays according to $|x_{t+1}|_P^2 - |x_t|_P^2 \leq -|x_t|^2 - |\bar{K} x_t|^2$. The strict decay rate gives robustness to deviations from optimality. Hence, as long as the difference between the left hand sides of (11) and (12) is not larger than the identity matrix, stability is retained. This is exploited in to get the following corollary.

Corollary 2 *Consider* $(A, B) \in \mathcal{M}_\beta$ *with* P *being a positive definite solution to the Riccati equation* (1)*, while* β, ρ, γ *are positive numbers. Let*

$$x_{t+1} = A x_t + B u_t + w_t, \qquad\qquad u_t = K_t x_t + \epsilon_t$$

and suppose that $\Sigma_t, \hat{\Sigma}_t$ *are defined by* (4)-(5)*, while* (10) *holds for* $t \geq t_0$*. For each* t*, let* Q_t *be a solution of* (6) *with* K_t *being a minimizing argument. Define*

$$\alpha := \beta^2 + \frac{1}{1 - \beta^2/\gamma^2} \left(1 - \frac{\beta^2}{1 - 2\beta^2 \rho(\rho + 2)} \right).$$

If $\alpha > 0$*, then*

$$\sum_{t=t_0}^{T-1} \left(|x_t|^2 + |K_t x_t|^2 \right) \leq \alpha^{-1} |x_{t_0}|_P^2 + \frac{\gamma^2}{\alpha} \sum_{t=t_0}^{T-1} |B \epsilon_t + w_t|^2. \tag{13}$$

6. Concluding remarks

The results above are the first steps towards a proof that a properly designed certainty equivalence controller can robustly stabilize the set $\mathcal{M}_\beta$ for arbitrarily large values of β. Here robustness means tolerance to a certain amount of unmodeled dynamics. The argument, with further details given in Rantzer (2023), has the following steps:

1. Pick any $\beta > 1$.

2. Select $\rho > 0$ small enough to make $[1 - 2\beta^2 \rho(\rho + 2)]^{-1} < 1 + \beta^{-2}$. Then Corollary 2 proves finite gain from $B\epsilon + w$ to x, provided that $\left\| \begin{bmatrix} \Sigma_t^{wx} & \Sigma_t^{wu} \end{bmatrix} \Sigma_t^{-1} \right\| \leq \rho$.

3. Design the excitation term ϵ_t to bound the impact of Σ_t^{-1}.

4. Introduce a bound on w_t to limit the size of $\begin{bmatrix} \Sigma_t^{wx} & \Sigma_t^{wu} \end{bmatrix}$.

Acknowledgments

References

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.

Julian Berberich, Johannes Köhler, Matthias A Müller, and Frank Allgöwer. Data-driven model predictive control with stability and robustness guarantees. *IEEE Transactions on Automatic Control*, 66(4):1702–1717, 2020.

Steven Bradtke. Reinforcement learning applied to linear quadratic regulation. *Advances in neural information processing systems*, 5, 1992.

Roger W Brockett. *Finite dimensional linear systems*. SIAM, 2015.

Daniel Cederberg, Anders Hansson, and Anders Rantzer. Synthesis of minimax adaptive controller for a finite set of linear systems. In *2022 IEEE 61st Conference on Decision and Control (CDC)*, pages 1380–1384. IEEE, 2022.

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.

Claudio De Persis and Pietro Tesi. Formulas for data-driven control: Stabilization, optimality, and robustness. *IEEE Transactions on Automatic Control*, 65(3):909–924, 2019.

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.

Olle Kjellqvist and Anders Rantzer. Minimax adaptive estimation for finite sets of linear systems. In *2022 American Control Conference (ACC)*, pages 260–265. IEEE, 2022.

Michail M Konstantinov, P Hr Petkov, and Nikolai D Christov. Perturbation analysis of the discrete riccati equation. *Kybernetika*, 29(1):18–29, 1993.

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

Horia Mania, Stephen Tu, and Benjamin Recht. Certainty equivalence is efficient for linear quadratic control. *Advances in Neural Information Processing Systems*, 32, 2019.

Ivan Markovsky and Florian Dörfler. Behavioral systems theory in data-driven analysis, signal processing, and control. *Annual Reviews in Control*, 52:42–64, 2021.

A Megretski. A nonlinear dynamical game interpretation of adaptive l2 control: Performance limitations and suboptimal controllers. In *Proceedings of 16th International Symposium on Mathematical Theory of Networks and Systems (MTNS2004)*, Leuven, 2004.

Anders Rantzer. Minimax adaptive control for a finite set of linear systems. In *Learning for Dynamics and Control*, pages 893–904. PMLR, 2021.

Anders Rantzer. Linear quadratic dual control. *arXiv preprint arXiv:2312.06014*, 2023.

Syed Ali Asad Rizvi and Zongli Lin. Output feedback q-learning control for the discrete-time linear quadratic regulator problem. *IEEE transactions on neural networks and learning systems*, 30(5): 1523–1536, 2018.

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.

Anastasios Tsiamis, Ingvar Ziemann, Nikolai Matni, and George J Pappas. Statistical learning theory for control: A finite-sample perspective. *IEEE Control Systems Magazine*, 43(6):67–97, 2023.

Glenn Vinnicombe. Examples and counterexamples in finite l2-gain adaptive control. In *Proceedings of 16th International Symposium on Mathematical Theory of Networks and Systems (MTNS2004)*, Leuven, 2004.

7. Appendix: Proofs

Lemma 3 *Consider positive numbers β and ρ, matrices $\hat{\Sigma}$, $\tilde{\Sigma}$ and positive semi-definite matrices Σ, P, Q. Suppose that $(\hat{\Sigma} - \tilde{\Sigma})^\top P(\hat{\Sigma} - \tilde{\Sigma}) = \Sigma(Q - I)\Sigma$ and $\tilde{\Sigma}^\top \tilde{\Sigma} \preceq \rho^2 \Sigma^2$. If $I \preceq Q \preceq \beta^2 I$, then $\hat{\Sigma}^\top P \hat{\Sigma} \preceq \Sigma Q \Sigma + \left(\beta^2 \rho(\rho + 2) - 1\right) \Sigma^2$.*

Proof. For every z with appropriate dimension

$$
\begin{aligned}
|\hat{\Sigma}z|_P^2 &\leq \left(|(\hat{\Sigma} - \tilde{\Sigma})z|_P + |\tilde{\Sigma}z|_P\right)^2 \\
&= \left(|\Sigma z|_{Q-I} + |\tilde{\Sigma}z|_P\right)^2 \\
&\leq \left(|\Sigma z|_{Q-I} + \beta\rho|\Sigma z|\right)^2 \\
&= |\Sigma z|_{Q-I}^2 + \beta^2 \rho^2 |\Sigma z|^2 + 2|\Sigma z|_{Q-I}\beta\rho|\Sigma z| \\
&\leq |\Sigma z|_{Q-I}^2 + \beta^2 \rho^2 |\Sigma z|^2 + 2\beta^2 \rho|\Sigma z|^2 \\
&= |\Sigma z|_Q^2 + \left(\beta^2 \rho(\rho + 2) - 1\right) |\Sigma z|^2.
\end{aligned}
$$

Here the first inequality is the triangle inequality, the second follows as $P \preceq \beta^2 I$ and $\tilde{\Sigma}^\top \tilde{\Sigma} \preceq \rho^2 \Sigma^2$, while the third inequality follows from the assumption $I \preceq Q \preceq \beta^2 I$. This completes the proof. □

Lemma 4 *Consider matrices $A, B, \bar{K}$ and $\bar{Q} \succeq I$. Suppose*

$$
[A \quad B]^\top \begin{bmatrix} I \\ \bar{K} \end{bmatrix}^\top \bar{Q} \begin{bmatrix} I \\ \bar{K} \end{bmatrix} [A \quad B] \preceq \bar{Q} - I. \tag{14}
$$

Then there exists a unique Q with $I \preceq Q \preceq \bar{Q}$ such that

$$
[A \quad B]^\top \min_K \left(\begin{bmatrix} I \\ K \end{bmatrix}^\top Q \begin{bmatrix} I \\ K \end{bmatrix}\right) [A \quad B] = Q - I. \tag{15}
$$

Proof. The matrix

$$
\bar{P} := [I \quad \bar{K}^\top] Q [I \quad \bar{K}^\top]^\top
$$

satisfies

$$
\bar{P} = (A + B\bar{K})^\top \bar{P}(A + B\bar{K}) + I + \bar{K}^\top \bar{K},
$$

so $P_0 = \bar{P}$ and the Riccati iteration

$$
P_{k+1} = \min_K \left[(A + B\bar{K})^\top P_k(A + B\bar{K}) + I + \bar{K}^\top \bar{K}\right]
$$

gives a decreasing sequence with limit $P_\infty \succeq I$. Setting

$$
Q := [A \quad B]^\top P_\infty [A \quad B]
$$

proves existence. Stabilizing solutions to the Riccati equation are unique (Brockett, 2015, Ch. 3), so also Q must be unique. □

Proof of Theorem 1. By definition, the assumption $(A, B) \in \mathcal{M}$ implies existence of Q satisfying (2) and $I \preceq Q \preceq \beta^2 I$. Let $\bar{K}$ be the corresponding minimizing argument in (2). Define

$$P := \begin{bmatrix} I & \bar{K}^\top \end{bmatrix} Q \begin{bmatrix} I & \bar{K}^\top \end{bmatrix}^\top$$
$$\tilde{\Sigma} = \hat{\Sigma}_t - \begin{bmatrix} A & B \end{bmatrix} \Sigma_t.$$

Then (2) gives $(\hat{\Sigma}_t - \tilde{\Sigma}_t)^\top P(\hat{\Sigma}_t - \tilde{\Sigma}_t) = \Sigma_t(Q - I)\Sigma_t$, so Lemma 3 implies

$$\hat{\Sigma}_t^\top P \hat{\Sigma}_t \preceq \Sigma_t(Q - \check{\alpha}I)\Sigma_t,$$

where $\check{\alpha} := 1 - \beta^2 \rho(\rho + 2)$. It follows that (14) holds with $\bar{Q} = \check{\alpha}^{-1}Q$, so Lemma 4 proves existence of Q_t (and K_t) satisfying (6) and $I \preceq Q_t \preceq \check{\alpha}^{-1}Q$.

Next define

$$P_t := \begin{bmatrix} I & K_t^\top \end{bmatrix} Q \begin{bmatrix} I & K_t^\top \end{bmatrix}^\top$$
$$\bar{\Sigma}_t := \begin{bmatrix} A & B \end{bmatrix} \Sigma_t.$$

Then the inequalities $(\bar{\Sigma}_t + \tilde{\Sigma}_t)^\top P_t(\bar{\Sigma}_t + \tilde{\Sigma}_t) = \Sigma_t(Q_t - I)\Sigma_t$ and $Q_t \preceq \alpha^{-1}\beta^2 I$ allow application of Lemma 3 to get

$$\bar{\Sigma}_t^\top P_t \bar{\Sigma}_t = \Sigma_t(Q_t - \hat{\alpha}I)\Sigma_t, \tag{16}$$

with $\hat{\alpha} := 1 - \check{\alpha}^{-1}\beta^2 \rho(\rho + 2)$. It follows from Lemma 4 that $Q \preceq \hat{\alpha}^{-1}Q_t$, so

$$
\begin{aligned}
&(A + BK_t)^\top P(A + BK_t) \\
&= \begin{bmatrix} I \\ K_t \end{bmatrix}^\top \begin{bmatrix} A & B \end{bmatrix}^\top \begin{bmatrix} I \\ \bar{K} \end{bmatrix}^\top Q \begin{bmatrix} I \\ \bar{K} \end{bmatrix} \begin{bmatrix} A & B \end{bmatrix} \begin{bmatrix} I \\ K_t \end{bmatrix} \\
&= \begin{bmatrix} I \\ K_t \end{bmatrix}^\top (Q - I) \begin{bmatrix} I \\ K_t \end{bmatrix} \\
&\preceq \begin{bmatrix} I \\ K_t \end{bmatrix}^\top (\hat{\alpha}^{-1}Q_t - I) \begin{bmatrix} I \\ K_t \end{bmatrix} \\
&= \hat{\alpha}^{-1}P_t - (I + K_t^\top K_t) \\
&\preceq \hat{\alpha}^{-1}\check{\alpha}^{-1}P - (I + K_t^\top K_t) \\
&= \begin{bmatrix} 1 - 2\beta^2 \rho(\rho + 2) \end{bmatrix}^{-1} P - (I + K_t^\top K_t)
\end{aligned}
$$

and the proof is complete. $\qquad\square$

Proof of Corollary 2.

$$\max_{w_t} \left(|x_{t+1}|_P^2 - \gamma^2 |B\epsilon_t + w_t|^2 \right)$$

$$= \max_{w_t} \left(|(A + BK_t)x_t + w_t|_P^2 - \gamma^2 |w_t|^2 \right)$$

$$= |(A + BK_t)x_t|_{P(I - \gamma^{-2}P)^{-1}}^2$$

$$\leq \left(1 - \beta^2/\gamma^2 \right)^{-1} |(A + BK_t)x_t|_P^2$$

$$\leq \frac{1}{1 - \beta^2/\gamma^2} \left(\frac{|x_t|_P^2}{1 - 2\beta^2\rho(\rho + 2)} - |x_t|^2 - |K_t x_t|^2 \right)$$

$$\leq |x_t|_P^2 + \left(\frac{1}{(1 - \beta^2/\gamma^2)(1 - 2\beta^2\rho(\rho + 2))} - 1 \right) \beta^2 |x_t|^2 - \frac{|x_t|^2 + |K_t x_t|^2}{1 - \beta^2/\gamma^2}$$

$$\leq |x_t|_P^2 - \alpha(|x_t|^2 + |K_t x_t|^2)$$

Summing over $t_0 \leq t \leq T - 1$ gives the desired result. $\qquad \square$