

# ESE 680 Final Report: First-Order Methods for Constrained Linear Quadratic Regulator Optimization

Anton Xue

December 19, 2019

## Abstract

This is the final report on our investigation into constrained optimization of linear quadratic regulators (LQR) via policy gradient methods. The primary objectives for this project were: (1) to gain familiarity with literature and mathematical tooling, (2) to identify suitable constraints on the LQR objective function, and (3) to study convergence properties of first-order methods of LQR with such constraints. We present basic observations and findings over the course of this project, as well as outline difficulties in constrained LQR optimization.

## 1 Introduction

First-order policy gradient methods [12] have seen rising prominence in reinforcement learning and control theory for several key reasons: they exploit simple ideas about derivatives that many practitioners are already well-acquainted with; they can be efficiently implemented even when the underlying system may be complex or black-box; and they often work well in practice. The idea is simple: gradient descent (or ascent) directly over some objective (policy's cost or reward) using until some (near) optimal value is reached.

A recently result in mathematical control theory applies such techniques to solving an instance of the LQR optimization problem [5]. For a discrete time linear dynamical system  $x_{t+1} = Ax_t + Bu_t$  with states  $x_t \in \mathbb{R}^n$  and control input  $u_t \in \mathbb{R}^k$ , the objective is to find a linear feedback controller of form  $u_t = -Kx_t$  that minimizes

$$J_{x_0}(K) = \sum_{t=0}^{\infty} x_t^\top Q x_t + u_t^\top R u_t \quad (1)$$

where  $x_0 \in \mathbb{R}^n$  is an initial starting state,  $Q, R \succ 0$  (positive-definite) with appropriate dimensions. Equation 1 is known as the infinite-horizon discrete-time LQR cost function. As one of the earliest inventions of feedback control by Kalman in the 1960's, the LQR framework remains an important tool in the control engineer's toolbox. Despite its relevance, the LQR cost function, especially in the context of finding optimal  $K$ , can be challenging to work with. Provided that one can solve the algebraic Ricatti equation (ARE):

$$P = A^\top P A + Q - A^\top P B (B^\top P B + R)^{-1} B^\top P A \quad (2)$$

for a  $P \succ 0$  the optimal controller  $K^*$  can then be stated as

$$K^* = -(B^\top P B + R)^{-1} B^\top P A \quad (3)$$

More classical approaches involve eigenvalue decomposition with matrix inversion [7] and semidefinite programming [1], among others. However such methods come with drawbacks: they do not parametrize

directly over the policy  $K$ , they do not necessarily target the objective function of interest  $J$ , and they are not easily adaptable to model-free settings where  $A$  and  $B$  may only be interacted with in possibly noisy environments.

Fazel et al [5] demonstrate that it is possible for first-order gradient-based methods to work directly over the LQR cost function. Moreover, such methods are able to achieve convergence to a globally optimal  $K^*$  provided that one begins at a stabilizing  $K_0$ . The ability to run first-order methods to optimize over the LQR cost function has strong implications: it enables LQR optimization problems to fall back on highly-optimized solvers, opens up more angles of attack for model-free LQR problem instances, and introduces new theoretical tooling to analyze the LQR cost function.

Indeed, new ways to find an optimal controller is exciting and important, but what good is a controller if it is not *safe*? In problem settings where a system’s state and control trajectory must avoid “bad” regions, optimality alone does not suffice. In this project we are concerned with constrained optimization of the LQR cost function using first-order methods. Our outline is as follows:

- (1) Review literature and texts related to optimization techniques [3, 13]. As we enter the project with little background in optimization, this is crucial before investigation of more recent results on LQR [5] and optimization [6]. Much of this is relevant foundations summarized in Section 2.
- (2) Investigate different types of constraints on the objective function, both hard and soft, that are amenable to first-order methods. We detail our approaches, findings, and challenges in Section 3.
- (3) Discuss fundamental challenges to first-order LQR optimization, in particular NP-Hardness issues that arise when constraining the controller  $K$ . These are deferred to Section 4.

## 2 Background

Here we present background information relevant to the investigation of first-order methods on the LQR cost function. This section is a partial summary of information obtained from literature [5, 6] and pedagogical texts [3, 13]. Additionally, we present relevant notation and definitions.

### 2.1 Gradient Descent

For a differentiable real-valued function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ , gradient descent is a simple iterative method that seeks to find local minima of the function. Starting at some initial point  $x_0 \in \mathbb{R}^n$ , gradient descent generates a sequence of improvements with the update rule

$$x_{k+1} = x_k - \alpha_k \nabla f(x_k) \quad (4)$$

for some possibly adaptive step size  $\alpha_k > 0$ . With sufficiently strong assumptions on  $f$ , gradient descent will converge to a local minima. A common assumption is  $f$  convex with Lipschitz continuity on  $\nabla f$ . This allows gradient descent variants to achieve a linear convergence rate in the sense of

$$f(x_k) - f^* \leq c^k [f(x_0) - f^*] \quad (5)$$

for some contraction factor  $c < 1$  that is dependent on  $f$  and the specific technique applied. In other words, gradient descent techniques are able to achieve an exponential approximation to the optimal value in a linear number of iterations.

## 2.2 The Polyak-Łojasiewicz Condition

Unsurprisingly, convexity of  $f$  is not the weakest condition necessary for gradient descent to achieve a linear convergence rate. As early as the 1960's Polyak [10] and Łojasiewicz [8] independently established sufficient condition on  $f$  that enables gradient descent to achieve linear convergence. This is further explored in [6] and is dubbed the Polyak-Łojasiewicz condition, alternatively known as the PL condition, PL inequality, or PL for short.

**Definition 1** (PL Inequality [6, 8, 10]). *A differentiable  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is said to satisfy the PL inequality if for some  $\mu > 0$*

$$\frac{1}{2\mu} \|\nabla f(x)\|^2 \geq f(x) - f(x_p) \quad (6)$$

for all  $x \in \text{dom } f$ , where  $x_p$  is the projection of  $x$  onto the optimal level set  $\mathcal{X}^*$ .

**Theorem 1** (Linear Convergence Under PL Inequality [6, 8, 10]). *Suppose  $f$  satisfies the PL inequality with constant  $\mu$  and has  $L$ -Lipschitz gradient, then gradient descent with a step-size of  $1/L$  in the sense of*

$$x_{k+1} = x_k - \frac{1}{L} \nabla f(x_k) \quad (7)$$

has a global linear convergence rate

$$f(x_k) - f^* \leq \left(1 - \frac{\mu}{L}\right)^k [f(x_0) - f^*] \quad (8)$$

Karimi et al also establish the equivalence of the PL inequality and a condition known as error-bound, or EB for short. This equivalent condition may be easier to intuit about the statement of the PL inequality.

**Definition 2** (Error Bound [6]). *A differentiable  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is said to satisfy the error bound condition if there exists some  $\beta > 0$  such that*

$$\langle \nabla f(x), x_p - x \rangle \geq \beta \|x_p - x\|^2 \quad (9)$$

for all  $x \in \text{dom } f$ , where  $x_p$  is the projection of  $x$  onto the optimal level set  $\mathcal{X}^*$ .

EB states that the growth of  $f$  away from an optimal point  $x_p$  is at least linear. An implication here is that a function that satisfies EB (and thus the PL inequality) is lower-bounded below by a quadratic. That is, there exists  $\alpha > 0$  such that

$$f(x) - f(x_p) \geq \alpha \|x_p - x\|^2 \quad (10)$$

Furthermore, functions that satisfy the PL inequality also have an important property known as *invexity*.

**Definition 3** (Invexity [6]). *A differentiable  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is invex if there exists a vector-valued  $\eta$  such that*

$$f(y) \geq f(x) + \nabla f(x)^\top \eta(x, y) \quad (11)$$

for all  $x, y \in \text{dom } f$ .

A function is invex if and only if every stationary point of  $f$  is a global minima:  $\nabla f(x) = 0$  implies  $x$  is optimal. Furthermore, when  $\eta(x, y) = y - x$ , convexity is invexity as a special case.

## 2.3 Proximal Methods

Sometimes it is desirable to run constrained optimization. These can be encoded in problems of form:

$$\arg \min_{x \in \mathbb{R}^n} F(x) = f(x) + g(x) \quad (12)$$

where  $g$  is an extended real-valued function assumed to be convex but not necessarily smooth, while  $f$  is differentiable along with different assumptions (such as convex) depending on problem domain. The perspective here should be that:  $f$  is the target function to be minimized while  $g$  acts as the penalty. Most naively, we can put more positive weight on  $g$  for values that *more violate* where we would like the optimal solution to be. A small trick allows hard constraints to also be encoded: if  $g = I_C$  is an indicator function over the convex set  $C$  taking on values in  $\{0, +\infty\}$ , this is equivalent to

$$\arg \min_{x \in \mathbb{R}^n} f(x) + I_C(x) = \arg \min_{x \in C} f(x) \quad (13)$$

supposing that  $f < \infty$ . Additionally suppose that such convex  $C$  is a set in which we would like the optimal value to lie in, projected gradient descent is first-order technique that entails applying a “projection” step after each gradient step:

$$x_k \in C \xrightarrow{\text{gradient step}} x'_k \notin C \xrightarrow{\text{projection step}} x_{k+1} \in C \quad (14)$$

The proximal operator is a generalization of the projection operator in Hilbert spaces, and is defined with respect to an extended real-valued function  $g$  as follows:

$$\text{prox}_{g,\lambda}(x) = \arg \min_{z \in \text{dom } h} g(z) + \frac{1}{2\lambda} \|z - x\|^2 \quad (15)$$

To reiterate, projection is a special case with  $g = I_C$ . For proximal (generalized projected) gradient descent the update rule is

$$x_{k+1} = \text{prox}_{g,\alpha_k}(x_k - \alpha_k \nabla f(x_k)) \quad (16)$$

There are several useful interpretations of proximal methods [9, 11], a particularly nice one being that this is gradient descent with  $\nabla f$  on a Moreau-Yosida regularization of  $g$  with parameter of  $\alpha_k$  updated at each step  $k$ . Classical analysis of proximal gradient descent enjoys a linear convergence rate when  $f$  is assumed to be convex (although the projection step itself may be difficult). Karimi et al relax the convexity assumption on  $f$  and yield the following result.

**Definition 4** (Proximal-PL Inequality [6]). *Consider the problem presented in Equation 12 where  $f$  is differentiable with  $L$ -Lipschitz continuous gradient and  $g$  is simple and convex but potentially non-smooth. The proximal-PL inequality is satisfied if there exists a  $\mu > 0$  such that*

$$\frac{1}{2\mu} \mathcal{D}_g(x, L) \geq F(x) - F^* \quad (17)$$

where

$$\mathcal{D}_g(x, \alpha) = -2\alpha \min_{y \in \text{dom } g} \left\{ \langle \nabla f(x), y - x \rangle + \frac{\alpha}{2} \|y - x\|^2 + g(y) - g(x) \right\} \quad (18)$$

in which case we say that  $F$  satisfies the proximal-PL inequality.

**Theorem 2** (Convergence of Proximal-PL Gradient Descent). *Consider the problem presented in Equation 12 where  $f$  has  $L$ -Lipschitz continuous gradient,  $F$  has a non-empty solution set,  $g$  is convex, and  $F$  satisfies the proximal-PL inequality with constant  $\mu$ . Then a proximal gradient method with a step size of  $1/L$*

$$x_{k+1} = \arg \min_{y \in \text{dom } g} \left\{ \langle \nabla f(x), y - x_k \rangle + \frac{L}{2} \|y - x_k\|^2 + g(y) - g(x_k) \right\} \quad (19)$$

*converges linearly to the optimal  $F^*$*

$$F(x_k) - F^* \leq \left(1 - \frac{\mu}{L}\right)^k [F(x_0) - F^*] \quad (20)$$

## 2.4 Recent Results on LQR

A recent result by Fazel et al [5] demonstrates that the LQR cost function (1) satisfies the PL inequality and (2) an “almost smoothness” condition. Together, these are sufficient conditions for gradient descent to achieve global convergence to the minimum at a linear rate provided that the initial controller  $K_0$  is stable: when  $\|A - BK_0\| < 1$  under the operator norm. Our notation so far has largely followed that of Fazel et al’s, although we use  $J$  instead of  $C$  for the LQR cost function due to compensate for handwriting legibility, and we introduce additional definitions and results when relevant. To begin, let  $P_K \succ 0$  be the solution to:

$$P_K = Q + K^\top R K + (A - BK)^\top P_K (A - BK) \quad (21)$$

which allows us to then express the LQR cost function as  $J_{x_0}(K) = x_0^\top P_K x_0$ . As Fazel et al assume deterministic dynamics except for the initial state  $x_0$ , the (un-normalized) state correlation matrix is

$$\Sigma_K = \mathbb{E}_{x_0 \sim \mathcal{D}} \sum_{t=0}^{\infty} x_t x_t^\top \quad (22)$$

where the trajectory is generated with respect to the controller  $K$ . A key contribution of Fazel et al is an explicit characterization of the gradient of  $J$  (the Hessian is additionally investigated and characterized in follow-up work [4]).

**Theorem 3** (LQR Gradient Expression [5]).

$$\nabla J(K) = 2E_K \Sigma_K, \quad E_K = (R + B^\top P_K B)K - B^\top P_K A \quad (23)$$

There are minute technical details that should we remark here: Fazel et al work over an initial distribution of states  $\mathcal{D}$ , but as their analysis does not rely on assumptions about  $\mathcal{D}$  we can make a trajectory completely deterministic to start on some  $x_0$  by taking  $\mathcal{D}$  to be the Dirac distribution. Our work builds on top of Fazel et al with a few simplifications: among which we assume that some initial state  $x_0$  is fixed and simply write  $J(K)$  instead of  $J_{x_0}(K)$ .

For the result of  $J$ ’s PL inequality, Fazel et al use the term “gradient domination” (Lemma 11 [5]), which is stated as follows:

**Theorem 4** (Gradient Domination (PL Inequality) [5]). *Let  $K^*$  be an optimal policy. Suppose that  $K$  has finite cost and  $\mu > 0$ , it holds that*

$$J(K) - J(K^*) \leq \frac{\|\Sigma_{K^*}\|}{\mu^2 \sigma_{\min}(R)} \text{Tr}(\nabla J(K)^\top \nabla J(K)), \quad \mu < \sigma_{\min}(\Sigma_K) \quad (24)$$

for a lower-bound:

$$\frac{\mu}{\|R + B^\top P_K B\|} \text{Tr}(E_K^\top E_k) \leq J(K) - J(K^*) \quad (25)$$

Note that the  $\mu$  used here is not necessarily the same  $\mu$  for the PL inequality statement of Equation 6. In any case, gradient domination (PL inequality) alone is not sufficient to guarantee linear convergence of gradient-based methods. Recall that an additional assumption of Lipschitz gradient is often assumed. The result on “almost-smoothness” is the substitute here.

**Theorem 5** (Almost Smoothness [5]).  *$J(K)$  satisfies*

$$J(K') - J(K) = -2 \text{Tr}(\Sigma_{K'} \Delta^\top E_K) + \text{Tr}(\Sigma_{K'} \Delta^\top (R + B^\top P_K B) \Delta), \quad \Delta = K - K' \quad (26)$$

We remark that the sign of  $\Delta$  is important. A good amount of perturbation analysis is done to show that  $\Sigma_{K'} \approx \Sigma_K + O(\|\Delta\|)$ , which allows the term  $2 \text{Tr}(\Sigma_{K'} \Delta^\top E_K)$  to behave like  $\text{Tr}(\Delta^\top \nabla J(K))$ . Together, these are able to yield a convergence result for gradient descent:

**Theorem 6** (Global Convergence of Gradient Descent Methods [5]). *Suppose that  $J(K_0) < \infty$  and  $\mu > 0$ , then for an appropriate step size  $\eta$  where*

$$\eta = \text{poly} \left( \frac{\mu \sigma_{\min}(Q)}{J(K_0)}, \frac{1}{\|A\|}, \frac{1}{\|B\|}, \frac{1}{\|R\|}, \sigma_{\min}(R) \right) \quad (27)$$

and for

$$N \geq \frac{\|\Sigma_{K^*}\|}{\mu} \log \left( \frac{J(K_0) - J(K^*)}{\varepsilon} \right) \text{poly} \left( \frac{J(K_0)}{\mu \sigma_{\min}(Q)}, \|A\|, \|B\|, \|R\|, \frac{1}{\sigma_{\min}(R)} \right) \quad (28)$$

gradient descent enjoys the performance bound of  $J(K_N) - J(K^*) \leq \varepsilon$ .

It should be noted that Fazel et al provide additional analysis for Gauss-Newton, natural gradient, and stochastic gradient descent methods, but they were not our primary targets of investigation for this project.

### 3 Approaches to Constrained LQR Optimization

An important step in constrained LQR optimization is picking *which* constraints to use. Such constraints should ideally be (1) reasonable and (2) amenable to gradient descent. Here we present a few of the constraints attempted and document basic results and challenges encountered during the investigation.

#### 3.1 Convex and Quadratic (Hard) Constraints on States

A basic form of hard-constrained optimization would look as follows

$$\min_K J(K), \quad \text{s.t. } x_t \in C \text{ for all } t \quad (29)$$

where  $C$  is a convex set, possibly a polytope. In other words, we must find some controller  $K$  such that each state of the trajectory  $x_t \in C$  for infinite time. A similar constrained optimization problem was of form

$$\min_K J(K), \quad \text{s.t. } \sum_{t=0}^{\infty} x_t^\top S x_t \leq M < \infty \quad (30)$$

for some  $S \succ 0$ , which requires that  $x_t$  converge to the origin sufficiently fast. The primary challenge with these two constraint forms is that we are performing gradient descent on a function parametrized by  $K$ , and it is unclear how techniques such as projected / proximal gradient descent would apply to relate  $K$  to elements  $x_t$  of the trajectory. Slight modifications to Equation 30 yields the optimization problem

$$\min_K J(K), \quad \text{s.t. } x_t^\top S x_t \leq M < \infty \text{ for all } t \quad (31)$$

with  $S \succ 0$ . That is, instead of requiring a sum over all  $x_t$  terms to converge, we simply require that each  $x_t$  be sufficiently close to the origin. Although projected gradient methods still appear to be out of our immediate reach, we are nevertheless able to derive two basic results presented below.

**Theorem 7** (Small Operator Norm Implies Bounded). *If  $S \succ 0$ ,  $\|x_0\| \leq 1$ , and*

$$\|(A - BK)^t\|^2 \leq \frac{M}{\lambda_{\max}(S)} \quad \text{for all } t, \quad (32)$$

*then  $x_t^\top S x_t \leq M$  for all  $t$ .*

*Proof.* Suppose the assumptions hold, then for any  $t \geq 0$ :

$$x_t^\top S x_t = x_0^\top ((A - BK)^t)^\top S (A - BK)^t x_0 \quad (33)$$

$$\leq \|(A - BK)^t\|^2 x_0^\top S x_0 \quad (34)$$

$$\leq \frac{M}{\lambda_{\max}(S)} x_0^\top S x_0 \quad (35)$$

$$\leq M \quad (36)$$

□

**Theorem 8** (Bounded Implies Small Operator Norm). *If  $S \succ 0$ ,  $\|x_0\| \leq 1$ , and  $x_t^\top S x_t \leq M$  for all  $t$ , then*

$$\|(A - BK)^t\|^2 \leq \frac{M}{\lambda_{\min}(S)} \quad \text{for all } t. \quad (37)$$

*Proof.* Suppose the assumptions hold, then for any  $t \geq 0$ :

$$\begin{aligned} M &\geq x_t^\top S x_t \\ &= \text{Tr}(x_0^\top ((A - BK)^t)^\top S (A - BK)^t x_0) \\ &= \lambda_{\min}(S) \text{Tr}(((A - BK)^t)^\top (A - BK)^t) \\ &= \lambda_{\min}(S) \|(A - BK)^t\|^2 \end{aligned}$$

□

### 3.2 Infinite-Horizon Quadratic (Soft) Constraints

We further considered a soft constraint of form

$$\min_K J(K) + \alpha \sum_{t=0}^{\infty} x_t^\top S x_t \quad (38)$$

where  $\alpha > 0$  and  $S \succ 0$ . Unfortunately this makes the problem too easy: one can rewrite this into a new LQR cost function

$$J(K) + \alpha \sum_{t=0}^{\infty} x_t^\top S x_t = \sum_{t=0}^{\infty} x_t^\top (Q + \alpha S) x_t + u_t^\top R u_t = \tilde{J}(K) \quad (39)$$

where we may define a new  $\tilde{Q} = Q + \alpha S$ . Plugging this new  $\tilde{Q}$  for  $Q$ , and  $\tilde{J}$  for  $J$  in Fazel's results would make all the analyses go through. We therefore seek more interesting cases of soft constraints.

### 3.3 Proximal-PL Inequality-Based (Soft) Constraints

Yet another approach was attempting to apply proximal gradient methods to the LQR cost function. That is, if a soft penalty function  $G$  is found such that  $J + G$  satisfies the proximal-PL inequality, then proximal gradient descent approaches would converge at a linear rate.

At some point due to an **algebraic mistake** it was thought that if  $J$  satisfies the restricted secant inequality (RSI) and  $G$  was an indicator function on a convex set, then these would be sufficient conditions to ensure that convergence happens at a linear rate.

**Definition 5** (Restricted Secant Inequality [6]). *A differentiable function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is said to satisfy RSI if there exists  $\mu > 0$  such that*

$$\langle \nabla f(x), x - x_p \rangle \geq \mu \|x - x_p\|^2 \quad (40)$$

for all  $x \in \text{dom } f$ , where  $x_p$  is the projection of  $x$  onto the optimal level set  $\mathcal{X}^*$ .

Supposing that we only assume  $g$  is convex, the weakest condition for  $f$  that Karimi et al derived that would suffice for  $F = f + g$  to satisfy the proximal-PL inequality was that  $f$  needed to be strongly convex away from the optimal level set. We were also not able to derive any weaker conditions. It is unknown if  $f$  satisfying RSI is strong enough to imply that  $f + g$  satisfies the proximal-PL inequality. Nevertheless, this mistake motivated attempts to show whether  $J$  satisfies RSI. To do this, it would suffice to show that for some  $\mu > 0$ :

$$\text{Tr}(\nabla J(K)^\top \Delta) \geq \mu \text{Tr}(\Delta^\top \Delta), \quad \Delta = K - K' \quad (41)$$

We present two main approaches towards this goal and discuss their difficulties and limitations.

#### 3.3.1 RSI Proof Approach 1

Observe that RSI requires both  $\nabla J(K)^\top \Delta$  on one side of the inequality and  $\Delta^\top \Delta$  on the other side. The challenge here was manipulating algebraic inequalities and identities in order to get to this form. We primarily leveraged existing ones established in Fazel et al due to the current lack of comfort with complex derivations. A particularly useful one was the almost-smoothness condition, which, for  $K'$  sufficiently close



to  $K$  such that  $\Sigma_{K'} \approx \Sigma_K + O(\|\Delta\|)$ , would allow us use Lemma 12 of Fazel et al [5] to (wishfully) write:

$$\text{Tr}(\nabla J^\top \Delta) = 2 \text{Tr}(\Sigma_K \Delta^\top E_K) \quad (42)$$

$$\approx 2 \text{Tr}(\Sigma_{K'} \Delta^\top E_K) - 2 \text{Tr}(O(\|\Delta\|) \Delta^\top E_K) \quad (43)$$

$$\approx 2 \text{Tr}(\Sigma_{K'} \Delta^\top E_K) \quad (44)$$

$$= \text{Tr}(\Sigma_{K'} \Delta^\top (R + B^\top P_K B) \Delta) + J(K) - J(K') \quad (45)$$

$$\geq \text{Tr}(\Sigma_{K'} \Delta^\top (R + B^\top P_K B) \Delta) \quad (46)$$

$$\geq \sigma_{\min}(\Sigma_{K'}) \sigma_{\min}(R) \text{Tr}(\Delta^\top \Delta) \quad (47)$$

The difficulty in this approach is that we were not able to show that

$$\text{Tr}(\Sigma_{K'} E_K) \geq \text{Tr}(\Sigma_K E_K) \quad (48)$$

for all  $K'$  obtained via a gradient descent step from  $K$ . If this is indeed the case, then the proof would go through. Unfortunately, we were not able to work out a proof of this given their perturbation analysis, which are given in terms of norm-bounds that ignore signs which would be important here.

### 3.3.2 RSI Proof Approach 2

Since the gradient update rule presented in Fazel et al takes form  $K' = K - \eta \nabla J(K)$ , this implies that

$$\text{Tr}(\nabla J(K)^\top \Delta) = \text{Tr}(\nabla J(K)^\top \eta \nabla J(K)) = \eta \|\nabla J(K)\|^2 \quad (49)$$

Thus under this rule it would suffice to show that

$$\text{Tr}(\nabla J(K)^\top \nabla J(K)) \geq \beta \text{Tr}(\Delta^\top \Delta) \quad (50)$$

for some  $\beta > 0$ . In other words, we would need  $\nabla J(K)^\top \nabla J(K)$  on one side of an inequality and  $\Delta^\top \Delta$  on the other wise, which changes the proof goal slightly. One such derivation process in Fazel et al that did contain some of the mechanisms that *should* make this go through existed in their proof of gradient domination (recall: PL inequality) of  $J$  (Lemma 11 [5]). Unfortunately, we were not able to manipulate the terms in a way to make this work.

## 4 Discussion

Here we discuss the LQR optimization problem as a whole. We discuss some fundamental theoretical limitations in addition to other properties of the LQR cost function.

### 4.1 Inherent Challenges and Fundamental Limits of First-Order Methods

A fundamental result from computational complexity [2] states that the decision problem of whether there exists a feedback gain matrix  $K$  such that  $A - BK$  is stable, and that  $K$  lies in an interval family of matrices, is NP-Hard. To see how this relates to LQR, observe that using the observation that  $x_t = (A - BK)^t x_0$ , this is akin to requiring that  $J(K) < \infty$ . This means that unless  $P = NP$ , there is no efficient (weakly polynomial) time algorithm that performs constrained optimization on LQR while ensuring  $K$  is bounded in an interval of matrices.

Because first-order methods aim to have linear convergence rate. This means that we cannot have first-order methods over LQR optimization that achieve an exponential convergence rate that is linear in iteration

number close towards the optimal solution while  $K$  is constrained. As a corollary, this also rules out the efficiency of techniques such as barrier functions and other interior point methods that aim for polynomial iteration convergences.

## 4.2 Properties of the LQR Cost Function

Despite its difficulties,  $J$  has enjoyed several theoretical analyses, many which are found in [4].

**Lemma 9** (From [4]). *The sublevel sets of  $\{J < \alpha\}$  are compact for all  $\alpha > 0$ .*

**Lemma 10** (From [4]). *Let  $Y$  be the solution to the Lyapunov matrix equation*

$$(A - BK)^\top Y (A - BK) + \Sigma_K = Y \quad (51)$$

*then the Hessian of the LQR cost function is characterized by*

$$\nabla^2 J(K) [E, E] = 2\langle (RE + B^\top P_K B E)Y, E \rangle - 4\langle B^\top (P'_K(K) [E])(A - BK)Y, E \rangle \quad (52)$$

*where  $E \in \mathbb{R}^{m \times n}$  and  $P'_K(K) [E]$  is the action of differential of the map  $K \mapsto P_K(K)$ .*

**Lemma 11** (From [4]). *Gradient flow along  $J$  by taking the initial value problem*

$$\dot{K}(t) = -\nabla J(K), \quad K(t_0) = K_0 \quad (53)$$

*is able to achieve a convergence rate of*

$$J(K_t) - J(K^*) \leq e^{-\alpha t} [J(K_0) - J(K^*)] \quad (54)$$

*provide that  $K_0$  is stable, and  $\alpha > 0$  is a constant determined by  $A$ ,  $B$ ,  $Q$ ,  $R$ , and  $K_0$ .*

**Lemma 12** (From [4]). *Discretization of gradient flow is possible.*

## 5 Conclusion

In this project we set out to study first-order methods of constrained LQR optimization. This report outlines some of the literature and texts relevant to further study, as well as presents our attempts at deriving basic results. Furthermore, we discuss fundamental limitations of constrained LQR optimization using first-order methods due to the existence of NP-Hardness. The primary value of this project was the author's increased familiarity with literature and mathematical techniques employed in this domain.

## References

- [1] Venkataramanan Balakrishnan and Lieven Vandenbergh. Semidefinite programming duality and linear time-invariant systems. *IEEE Transactions on Automatic Control*, 48(1):30–41, 2003.
- [2] Vincent Blondel and John N Tsitsiklis. Np-hardness of some linear control design problems. *SIAM Journal on Control and Optimization*, 35(6):2118–2127, 1997.
- [3] Stephen Boyd and Lieven Vandenbergh. *Convex optimization*. Cambridge university press, 2004.
- [4] Jingjing Bu, Afshin Mesbahi, Maryam Fazel, and Mehran Mesbahi. Lqr through the lens of first order methods: Discrete-time case. *arXiv preprint arXiv:1907.08921*, 2019.

- [5] Maryam Fazel, Rong Ge, Sham M. Kakade, and Mehran Mesbahi. Global convergence of policy gradient methods for the linear quadratic regulator, 2018.
- [6] Hamed Karimi, Julie Nutini, and Mark Schmidt. Linear convergence of gradient and proximal-gradient methods under the polyak-łojasiewicz condition. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 795–811. Springer, 2016.
- [7] Peter Lancaster and Leiba Rodman. *Algebraic riccati equations*. Clarendon press, 1995.
- [8] Stanisław Łojasiewicz. Sur le probleme de la division. 1961.
- [9] Neal Parikh and Stephen Boyd. Proximal algorithms. *Foundations and Trends® in Optimization*, 1(3):127–239, 2014.
- [10] B.T. Polyak. Gradient methods for the minimisation of functionals. *USSR Computational Mathematics and Mathematical Physics*, 3(4):864 – 878, 1963.
- [11] R Tyrrell Rockafellar. *Convex analysis*, volume 28. Princeton university press, 1970.
- [12] Richard S Sutton, David A McAllester, Satinder P Singh, and Yishay Mansour. Policy gradient methods for reinforcement learning with function approximation. In *Advances in neural information processing systems*, pages 1057–1063, 2000.
- [13] Ryan Tibshirani. Machine learning 10-725, convex optimization: Fall 2018, carnegie mellon university, 2018.