

Solutions to the Inverse LQR Problem With Application to Biological Systems Analysis

M. Cody Priess, Richard Conway, Jongeun Choi, *Member, IEEE*, John M. Popovich, Jr., and Clark Radcliffe

Abstract—In this brief, we present a set of techniques for finding a cost function to the time-invariant linear quadratic regulator (LQR) problem in both continuous- and discrete-time cases. Our methodology is based on the solution to the inverse LQR problem, which can be stated as: does a given controller K describe the solution to a time-invariant LQR problem, and if so, what weights Q and R produce K as the optimal solution? Our motivation for investigating this problem is the analysis of motion goals in biological systems. We first describe an efficient linear matrix inequality (LMI) method for determining a solution to the general case of this inverse LQR problem when both the weighting matrices Q and R are unknown. Our first LMI-based formulation provides a unique solution when it is feasible. In addition, we propose a gradient-based, least-squares minimization method that can be applied to approximate a solution in cases when the LMIs are infeasible. This new method is very useful in practice since the estimated gain matrix K from the noisy experimental data could be perturbed by the estimation error, which may result in the infeasibility of the LMIs. We also provide an LMI minimization problem to find a good initial point for the minimization using the proposed gradient descent algorithm. We then provide a set of examples to illustrate how to apply our approaches to several different types of problems. An important result is the application of the technique to human subject posture control when seated on a moving robot. Results show that we can recover a cost function which may provide a useful insight on the human motor control goal.

Index Terms—Biological system modeling, inverse optimal control problem, system identification.

I. INTRODUCTION

SYSTEMS science has produced a number of useful tools for the analysis of biological systems. The application of systems identification techniques to human motor control problems has allowed researchers to determine physiological parameters, control gains, and control bandwidth for a variety

Manuscript received December 10, 2013; revised April 30, 2014; accepted June 30, 2014. Manuscript received in final form July 24, 2014. This work was supported by the National Center for Complementary and Alternative Medicine, National Institutes of Health, under Grant U19 AT006057. Recommended by Associate Editor C. Lagoa.

M. C. Priess and C. Radcliffe are with the MSU Center for Orthopedic Research, Department of Mechanical Engineering, Michigan State University, East Lansing, MI 48824 USA (e-mail: priessma@msu.edu; radcliff@egr.msu.edu).

R. Conway is with Western Digital Inc., Fremont, CA 94539 USA (e-mail: rconway345@gmail.com).

J. Choi is with the MSU Department of Mechanical Engineering, Department of Electrical and Computer Engineering, MSU Center for Orthopedic Research, Michigan State University, East Lansing, MI 48824 USA (e-mail: jchoi@egr.msu.edu).

J. M. Popovich, Jr., is with the MSU Department of Surgical Specialties, Michigan State University, East Lansing, MI 48824 USA (e-mail: popovi16@msu.edu).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TCST.2014.2343935

of motion tasks [1]–[3]. The goal of biomechanical researchers is often to determine differences which may exist between healthy subjects and those suffering from pain or disease [4]. Even though these differences may be visible in the standard characteristics accessible through system identification (plant parameters, feedback gains, etc.), if one assumes that the motion under analysis results from an optimal control method, it may be possible to additionally determine a cost function that would generate that control in some optimal sense [5]. These cost functions can offer additional relevant information about the system, for example, how much weight does the controller put on the various states as compared to the control effort? Several prior studies have attempted to determine optimality criteria from human motion data in an effort to explain human motion goals [5]–[8]. In contrast to the more general potential cost functions used in these studies, we propose the use of a control theoretic method using the linear quadratic regulator (LQR) framework.

In optimal control theory, the LQR problem [9] is to find the optimal infinite-horizon full state-feedback control law for the continuous-time LTI system

$$\frac{dx(t)}{dt} = Ax(t) + Bu(t) \quad (1)$$

with respect to the cost function

$$J = \int_0^\infty \begin{bmatrix} x(t) \\ u(t) \end{bmatrix}^T \begin{bmatrix} Q & S \\ S^T & R \end{bmatrix} \begin{bmatrix} x(t) \\ u(t) \end{bmatrix} dt \quad (2)$$

where $Q = M^T M \succeq 0$ and $R = R^T \succ 0$. Assume, for the moment, that $S = 0$. Then, assuming that (A, B) is controllable and (A, M) is detectable, the optimal stabilizing control minimizing J is found as

$$u(t) = -Kx(t) \quad (3)$$

where

$$K = R^{-1}B^T P \quad (4)$$

with P being the unique positive-semidefinite solution to the algebraic Riccati equation (ARE)

$$A^T P + PA - PBR^{-1}B^T P + Q = 0. \quad (5)$$

The typical use of the LQR problem in (1)–(5) is the forward result, i.e., to determine the optimal control law K from a given set of weight matrices Q and R . However, the inverse LQR problem has received some attention as well. In general, the inverse problem has been defined by two subproblems [10]. Given a stabilizing control law (3),

P. 1) Determine what necessary and sufficient conditions exist on (K, A, B) such that K is an optimal

control law for a cost of the form in (2) with $Q \succeq 0$ and $R = I$.

P. 2) Determine all Q for some (K, A, B) that satisfy the conditions found in P. 1.

According to Fujii and Narazaki [10], Problem 1 was first addressed with Kalman's investigation of the single-input case [11], which was later extended to the multiinput case in [12]. A necessary and sufficient condition when K is not necessarily stabilizing and R unknown was determined in [13], who also show an analytic solution for recovering $R > 0$ and $Q = Q^T$. However, a feasible Q recovered using this method is not guaranteed to be positive-semidefinite even when the closed-loop system is stable [13].

While further results have been determined for potentially destabilizing controllers [10], only stabilizing controllers are of interest when investigating engineered or biological systems. In addition, it has been found that when the cross-term S is included in the LQR cost function, then trivially any controller K is optimal for some cost function [14]. However, we choose to exclude S from this brief for a number of reasons, it is rarely used in the design of LQR controllers for practical systems, and inverse results that include the cross-term provide less salient information about the control goals than results which separate control and state costs in a straightforward manner (e.g., principle of parsimony). As a result, we will be restricting our focus in the rest of this brief to the stable LQR problem described in (1)–(5) (and later, its discrete-time counterpart) with $S = 0$. Molinari [15] found a necessary and sufficient condition to Problem 1 for some $Q \succeq 0$, $R = I$ when the admissible controls are in the class of control $u(t)$ such that the corresponding state $x(t)$ satisfies $\lim_{t \rightarrow \infty} x(t) = 0$. This result from [15] is stated as follows.

Theorem 1: Assume that (A, B) from (1) is controllable. Then, K will be optimal for some $Q \succeq 0$ if and only if:

- 1) $A - BK$ is Hurwitz;
- 2) $T^*(jw)T(jw) - I \succeq 0$, where $T(s) = I + K(I_s - A)^{-1}B$.

For a stabilizing controller K , a necessary and sufficient condition utilizing coprime matrix fraction descriptions was derived in [16]. An approach utilizing convex optimization to find a maximally diagonal $Q \succeq 0$ describing a given stabilizing control law K was proposed in [17]. While $R = I$ in all of these results, if $R > 0$ is known, then any result found for $R = I$ can be determined by a simple coordinate transformation of B and K .

However, a more general case of the inverse LQR problem is still unanswered, if it exists, what set $(Q \succeq 0, R > 0)$ generates a given stabilizing feedback law K ? So far as we are aware, there is no analytical solution to this more general problem, although for K stabilizing, at least one convex optimization formulation exists for determining a feasible solution [18]. This problem is an interesting extension to conventional system identification theory, and has potential uses in both the reverse-engineering of black-box control systems as well as in the analysis of biological control systems. For example, it may be possible to determine underlying motor control goals from analysis of a human subject's feedback gains, which would give clinical researchers an additional way

to quantify and evaluate patients. The objective of this brief is to provide a methodology for determining an inverse LQR solution in both continuous- and discrete-time cases, and, as an example, to apply this method to recover a cost function from a human motor control task.

The contribution of this brief is as follows. We present an linear matrix inequality (LMI)-based formulation similar to that in [18] to determine whether or not for a given stabilizing feedback law K that has been estimated from a set of experimental time-series data, there exists some set (Q, R) for which K is the optimal feedback gain. If such a solution exists, then the LMIs are solved for (Q, R) directly. Our first LMI formulation provides a unique solution when it is feasible, which can be viewed as a regularization of the feasibility formulation given in [18]. If the exact solution does not exist due to the infeasibility of the LMIs, we show how to formulate a gradient descent algorithm based on the derivative of the ARE in order to minimize the difference between the resulting best-fit and experimental feedback gains. This new method is very useful in practice since the estimated gain matrix K from the noisy experimental data could be perturbed by the estimation error, which may result in the infeasibility of the LMIs. Since this minimization using the gradient descent algorithm guarantees only the local optimality of the solution, finding a good initial starting point (or initial guess) for the gradient descent algorithm becomes important. Hence, we also provide an LMI minimization problem to find a good initial point for the minimization using the gradient approach. We then provide examples to illustrate how to apply our approaches to several different types of problems. One important contribution is to apply our proposed technique to the biological data obtained from a seated balance test using a commercial robot with a human subject. This test is designed to investigate the control mechanism of the human subject on an actuated seat. A practical experimental result obtained in this brief shows a proof of concept in human cost function recovery for future clinical research activities. Previous work [19] has appeared without human test data and analysis. The current work also augments previous work with an LMI formulation that provides a good initial point for the gradient descent method.

This brief is organized as follows. In Section II, we formulate the LMI problem describing the inverse LQR solution in both continuous and discrete time. In Section II, we formulate a gradient descent algorithm which can be applied to cases where the LMI problem is infeasible, and show an LMI method for determining a good starting point for this algorithm. In Section III, we demonstrate the use of the LMI method for several feasible example problems. In Section IV, we demonstrate the use of the gradient descent algorithm to solve a problem where the LMI method is infeasible. In Section V, we apply the method to experimentally determine an LQR-type cost function in a human subject. Finally, conclusions are drawn in Section VI.

II. INVERSE LQR PROBLEM

The inverse LQR problem has both the continuous-time formulation (1)–(5) and a formulation for the discrete-time

LTI system

$$x_{k+1} = Ax_k + Bu_k \quad (6)$$

which minimizes the value of the cost function

$$J = \sum_{k=0}^{\infty} \begin{bmatrix} x_k \\ u_k \end{bmatrix}^T \begin{bmatrix} Q & S \\ S^T & R \end{bmatrix} \begin{bmatrix} x_k \\ u_k \end{bmatrix}. \quad (7)$$

Assuming, again, that $S = 0$, the optimal feedback control is

$$u_k = -Kx_k$$

where

$$K = (B^T PB + R)^{-1} B^T PA \quad (8)$$

and P is the unique positive-semidefinite solution to the discrete-time ARE

$$A^T PA - P - (A^T PB)(B^T PB + R)^{-1} B^T PA + Q = 0. \quad (9)$$

We additionally define an auxiliary notation for the solution K to the discrete-time LQR problem as

$$K = \text{DLQR}(A, B, Q, R)$$

and one for the continuous-time LQR problem as

$$K = \text{CLQR}(A, B, Q, R).$$

In formal terms, for the continuous-time (respectively, discrete-time) case, our problem is to determine the weighting matrices

$$\begin{aligned} &(\hat{Q}, \hat{R}) \text{ such that} \\ &\hat{Q} \succeq 0, \hat{R} \succ 0, \hat{K} = K_e \\ &\hat{K} = \text{CLQR}(A, B, \hat{Q}, \hat{R}) \\ &[\text{respectively, } \hat{K} = \text{DLQR}(A, B, \hat{Q}, \hat{R})] \end{aligned} \quad (10)$$

where K_e is the full-state feedback gain matrix determined via a system identification method from the experimental data.

A. Solution via LMIs

For the formulated inverse LQR problem in Sections I and II, there is an associated uniqueness issue; for example, if we multiply \hat{Q} and \hat{R} in (10) by the scalar $\beta > 0$ and find the LQR solution, then the resulting controller gain matrix \hat{K} will be identical no matter what the value of β . Consequently, we expect there to be a manifold of possible solutions (\hat{Q}, \hat{R}) to the inverse problem defined in (10). Therefore, we define the additional criteria that an optimal solution (\hat{Q}, \hat{R}) must minimize the condition number of the weighting matrix

$$\begin{bmatrix} \hat{Q} & 0 \\ 0 & \hat{R} \end{bmatrix},$$

which can be defined explicitly as

$$\begin{aligned} (\hat{Q}, \hat{R}, \hat{\alpha}) &= \arg \min_{(\hat{Q}, \hat{R}) \in (\hat{Q}, \hat{R}), \hat{\alpha}} \hat{\alpha}^2 \\ &\text{such that } I \preceq \begin{bmatrix} \hat{Q} & 0 \\ 0 & \hat{R} \end{bmatrix} \preceq \hat{\alpha}I. \end{aligned} \quad (11)$$

Minimizing the condition number ensures the numerical stability for operations involving (\hat{Q}, \hat{R}) [20]. This will also

lead us to obtain the unique solution to our problem (see Remark 1). Note, however, that (11) will force $\hat{Q} \succ 0$, which is more restrictive than the $\hat{Q} \succeq 0$ requirement in (10) and is, in general, not a necessary condition for the defined LQR problem. In addition, forcing $\hat{Q} \succ 0$ means that (A, M) (with $M = M^T M$) will trivially satisfy the detectability requirement.

The problems defined in (10) and (11) can be written together as a convex optimization problem subject to LMI constraints. For the continuous-time LQR from (1) to (5), the LMI optimization can be written as follows:

$$\begin{aligned} (\hat{Q}, \hat{R}, \hat{P}, \hat{\alpha}) &= \arg \min_{\hat{Q}, \hat{R}, \hat{P}, \hat{\alpha}} \hat{\alpha}^2, \text{ such that} \\ &P \succeq 0 \\ &A^T P + PA - PBK_e + Q = 0 \\ &B^T P - RK_e = 0 \\ &I \preceq \begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix} \preceq \hat{\alpha}I. \end{aligned} \quad (12)$$

For the discrete-time LQR defined in (6)–(9), the problem becomes

$$\begin{aligned} (\hat{Q}, \hat{R}, \hat{P}, \hat{\alpha}) &= \arg \min_{\hat{Q}, \hat{R}, \hat{P}, \hat{\alpha}} \hat{\alpha}^2, \text{ such that} \\ &P \succeq 0 \\ &A^T PA - P - A^T PBK_e + Q = 0 \\ &B^T PA - (B^T PB + R)K_e = 0 \\ &I \preceq \begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix} \preceq \hat{\alpha}I. \end{aligned} \quad (13)$$

Since one of the major applications of the inverse LQR solution presented here is to use the recovered cost matrices to draw some broader conclusions from a control system, it is important that any solution be unique. If multiple solutions to the inverse LQR problem exist for a given system, then multiple cost functions give equivalent descriptions of the controller and no useful conclusions can be found. In that regard, we make the following statement.

Remark 1: Equations (12) and (13) are convex optimization problems with strictly convex objectives [21]. Therefore, if a feasible solution exists that minimizes the objective function, it will be unique [21], [22]. Note that strict convexity of the objective is only a sufficient condition for uniqueness of the solution. Our approach used in (12) and (13) can be viewed as a regularization of the feasibility formulation given in [18], providing a great utility to inverse problem applications. Note also that (12) and (13) can be formulated and solved as semidefinite programs (SDP) by adding the LMI constraint

$$\begin{bmatrix} \gamma & \alpha \\ \alpha & 1 \end{bmatrix} \succeq 0$$

and then minimize γ instead of $\hat{\alpha}^2$.

The feasible solution to the problem in (12) (respectively, in (13)), will satisfy both (10) and (11) simultaneously. Infeasibility implies that there is no solution to the LQR problem such that $\hat{K} = K_e$ while satisfying all constraints.

Previous works [10], [11], [15]–[17] include results that a stabilizing K is optimal relative to some $Q \succeq 0$ for R known, while the inverse problem (10) involves both Q and R unknown. We therefore make Remark 2.

Remark 2: If the LMI problem defined in (12) [respectively, (13)] is feasible, then at least one exact solution to the inverse problem of (10) exists. If the problem is infeasible, then no exact solution exists.

However, if no exact solution exists, then an approximation \hat{K} minimizing the residual error between \hat{K} and K_e can be found via a gradient descent law, which is the main idea in the following section.

B. Solution via Gradient Descent Algorithm

In accordance with Remark 2, if the solution to the LMI problem in (12) [respectively, (13)] is infeasible, then we consider the following minimization problem:

$$\begin{aligned}\hat{\theta} &= \arg \min_{\theta} \|K(\theta) - K_e\|_F^2 \\ K(\theta) &= \text{CLQR}(A, B, Q(\theta), R(\theta)) \\ [\text{respectively, } K(\theta) &= \text{DLQR}(A, B, Q(\theta), R(\theta))] \quad (14)\end{aligned}$$

where $\|\cdot\|_F$ denotes the Frobenius norm, i.e., $\|A\|_F := (\text{trace}(A^T A))^{1/2}$ (for $A \in \mathbb{R}^{a \times b}$), and θ defines the upper triangular entries of the symmetric weighting matrices $Q \in \mathbb{R}^{n \times n}$, $R \in \mathbb{R}^{m \times m}$ as

$$\theta := [Q_{11}, Q_{12}, \dots, Q_{nn}, R_{11}, R_{12}, \dots, R_{mm}]^T.$$

We can find a local minimum to (14) by using the analytical gradient of the cost in (14). For a concise presentation, we first vectorize $K(\theta)$ and K_e such that

$$K^v(\theta) := \text{vec}(K(\theta)), \quad K_e^v := \text{vec}(K_e)$$

where the $\text{vec}(\cdot)$ operator converts an arbitrary matrix $A \in \mathbb{C}^{m \times n}$ into a column vector such that

$$\text{vec}(A) = [a_{11}, \dots, a_{m1}, a_{12}, \dots, a_{m2}, \dots, a_{1n}, \dots, a_{mn}]^T$$

where a_{ij} is the $(i, j)^{\text{th}}$ element of A .

Let us define $e := K^v(\theta) - K_e^v$, then we have $e^T e = \|K^v(\theta) - K_e^v\|^2 = \|K(\theta) - K_e\|_F^2$. We now present a gradient descent law that drives each element θ_i of θ such that the error norm in (14) is decreased as follows:

$$\begin{aligned}\frac{d\theta_i(t)}{dt} &= -\lambda \frac{\partial(e^T e)(t)}{\partial\theta_i(t)} \\ &= -\lambda \left[\frac{\partial e^T(t)}{\partial\theta_i(t)} e(t) + e^T(t) \frac{\partial e(t)}{\partial\theta_i(t)} \right] \quad (15)\end{aligned}$$

where λ is a positive constant that controls the convergence rate. Note that we need to compute

$$\frac{\partial e(t)}{\partial\theta_i(t)} = \frac{\partial K^v(\theta(t))}{\partial\theta_i(t)}. \quad (16)$$

To obtain the value in (16) analytically, we first introduce the notation

$$\dot{\cdot} := \frac{\partial \cdot}{\partial\theta_i(t)}.$$

As it is clear that (K^v, Q, R) are functions of θ and t , we will drop the explicit dependencies as in $(K^v(\theta(t)), Q(\theta(t)), R(\theta(t)))$.

Consider the continuous-time case for a representative presentation. The discrete-time case follows similar steps. Now, we have the following:

$$\frac{\partial e}{\partial\theta_i} = (K^v)' = \text{vec}((R^{-1})'B^T P + R^{-1}B^T P'). \quad (17)$$

The stabilizing solution P to the ARE in (5) is analytic in A, B, M [23], and can therefore be differentiated implicitly with respect to θ . If we take the derivative of the ARE from (5) with respect to θ_i , we arrive at

$$\begin{aligned}0 &= A^T P' + P' A \\ &\quad - [P' B R^{-1} B^T P + P B (R^{-1})' B^T P + P B R^{-1} B^T P'] + Q' \\ &= \bar{A}^T P' + P' \bar{A} + [-P B (R^{-1})' B^T P + Q'] \quad (18)\end{aligned}$$

where $\bar{A} = A - B R^{-1} B^T P$ and $P = P^T \succeq 0$ is the solution to the original ARE in (5). Equation (18) defines a Lyapunov equation that can be solved for P' . By determining (15) for all elements i of $\theta_i(t)$, we find a directional derivative that drives the elements in $\theta(t)$ such that the norm of the error in (14) is minimized. For actual computation of $\theta(t)$, we apply the preceding result in a discrete sense, i.e., we replace the continuous-time $\theta_i(t)$ in (15) with the discrete-time equivalent $\theta_i(k) = \theta_i^k$

$$\theta_i^{k+1} = \theta_i^k - \lambda \frac{\partial(e^T e)}{\partial\theta_i^k} \quad (19)$$

which is iterated for a desired number of iterations. In addition, a projection rule for θ is applied because Q must remain positive-semidefinite and R must remain positive-definite for the LQR solution to exist.

The success of any gradient descent algorithm depends on the quality of the initial starting point (or initial guess). We can exploit the fact that there always exists an exact solution to the inverse LQR problem when $S \neq 0$ [14] to determine a close approximation for the case when $S = 0$. To this end, we consider the following LMI problem:

$$\begin{aligned}(Q_s, S_s, R_s, P_s) &= \arg \min_{Q, S, R, P} \|S\|_F, \text{ such that} \\ P &\succeq 0 \\ B^T P + S^T - R K_e &= 0 \\ A^T P + P A - (P B + \$) K_e + Q &= 0 \\ \begin{bmatrix} Q & S \\ S^T & R \end{bmatrix} &\succ 0. \quad (20)\end{aligned}$$

Additional constraints on Q and R may be included in (20) to improve the quality of the initial point.

Solving this set of LMIs for Q_s, R_s, S_s yields an initial point (Q_s, R_s) which is nearly optimal in the sense that it is based on an exact solution minimizing S . Note that by introducing an additional LMI constraint and decision variable, it would be possible to reformulate (20) as a SDP which can be solved efficiently. Complete details of the algorithm are given in Table I.

III. ILLUSTRATIVE EXAMPLE WITH FEASIBLE SOLUTION

Consider the continuous-time LTI system

$$\frac{dx(t)}{dt} = Ax(t) + Bu(t), \quad y(t) = Cx(t) \quad (21)$$

TABLE I
ALGORITHM FOR THE UPDATE OF θ

Input:	(1) The known system matrices $A \in \mathbb{R}^{n \times n}$ and $B \in \mathbb{R}^{n \times m}$. (2) The estimated feedback gain vector $K_e^v = \text{vec}(K_e) \in \mathbb{R}^{mn}$.
Output:	(1) The optimal output parameter vector $\hat{\theta}$.
1:	solve the set of LMIs in (20) for Q_s and R_s .
2:	let $Q_1 = Q_s$ and $R_1 = R_s$
3:	form θ^1 from Q_1 and R_1
4: for $k = 1, \dots, k_{\text{end}}$ do	
5:	compute $K^v = \text{vec}(LQR(A, B, Q_k, R_k))$
6:	compute P solving $A^T P + PA - PBR_k^{-1}B^T P + Q_k = 0$
7: for $i = 1, \dots, \frac{n(n+1)}{2} + \frac{m(m+1)}{2}$ do	
8:	compute $\bar{A} = A - BR_k^{-1}B^T P$
9:	compute P' solving $\bar{A}^T P' + P' \bar{A} + [-PB(R_k^{-1})'B^T P + Q'_k] = 0$
10:	compute $(K^v)' = \text{vec}((R_k^{-1})'B^T P + R_k^{-1}B^T P')$
11:	compute $e = K^v - K_e^v$
12:	compute $\frac{\partial(e^T e)}{\partial \theta_i^k} = [(K^v)'^T e + e^T (K^v)']$
13:	let the element $\theta_i^{k+1} = \theta_i^k - \lambda \frac{\partial(e^T e)}{\partial \theta_i^k}$
14: end for	
15:	form Q_{k+1} and R_{k+1} from θ^{k+1}
16: if $Q_{k+1} \prec 0$ then	
17:	let $Q_{k+1} = Q_k$
18:	let $\theta_i^{k+1} = \theta_i^k, \forall i = 1, \dots, \frac{n(n+1)}{2}$
19: end if	
20: if $R_{k+1} \preceq 0$ then	
21:	let $R_{k+1} = R_k$
22:	let $\theta_i^{k+1} = \theta_i^k, \forall i = \frac{n(n+1)}{2} + 1, \dots, \frac{n(n+1)}{2} + \frac{m(m+1)}{2}$
23: end if	
24: end for	
25:	let $\hat{K} = \text{LQR}(A, B, Q_{k_{\text{end}}+1}, R_{k_{\text{end}}+1})$

where

$$A = \begin{bmatrix} 0 & 1 & 7 & 9 \\ 4 & -8 & -5 & -3 \\ 8 & -7 & 7 & -6 \\ 10 & -5 & -5 & -5 \end{bmatrix}$$

$$B = \begin{bmatrix} 2 & 2 & 5 & -9 \\ -1 & 1 & 5 & -9 \\ -3 & 9 & -3 & 1 \\ 7 & -4 & 1 & 6 \end{bmatrix}$$

$$C = I_4. \quad (22)$$

In contrast to (1), we include an output equation so that measurement noise can be considered later.

The system in (22) is unstable, with two eigenvalues in the open right-half plane. However, the pair (A, B) is controllable, allowing us to design an LQR controller with weights

$$Q_0 = \begin{bmatrix} 13.9 & -1.32 & 3.9 & 2.65 \\ -1.32 & 7.35 & 3.72 & -2.27 \\ 3.9 & 3.72 & 5.28 & 0.373 \\ 2.65 & -2.27 & 0.373 & 4.33 \end{bmatrix}$$

$$R_0 = \begin{bmatrix} 10.6 & 0.302 & 6.56 & -1.83 \\ 0.302 & 4.69 & 2.06 & 3.87 \\ 6.56 & 2.06 & 7.17 & -0.266 \\ -1.83 & 3.87 & -0.266 & 6.88 \end{bmatrix}. \quad (23)$$

The feedback gain matrix K_0 resulting from the solution of the ARE in (4) and (5) with $(Q, R) = (Q_0, R_0)$ from (23) is

$$K_0 = \begin{bmatrix} 2.362 & -1.32 & 1.186 & 1.836 \\ 4.293 & -0.8472 & 3.789 & 1.538 \\ -2.252 & 1.339 & -2.081 & -1.312 \\ -2.707 & -0.06996 & -2.062 & -0.7263 \end{bmatrix}.$$

Q_0 and R_0 in this case were chosen to minimize the condition number of the weighting matrix used to produce K_0 . The closed-loop system is then

$$\frac{dx(t)}{dt} = (A - BK_0)x(t) + BR(t), \quad y(t) = Cx(t) \quad (24)$$

that is $u(t) = -K_0x(t) + r(t)$. The system in (24) is simulated by computing the response at a finite number of sampling points t_k . The input $r(t)$ is then computed as a zero-order hold of a sampled input $r(t_k)$, which we realize as an independent identically distributed (i.i.d.) Gaussian white noise process with $r(t_k) \sim \mathcal{N}(0, 10I_4)$.

Suppose that a noisy version of the sampled state response of the true system $A - BK_0$ is available and denoted by $\tilde{y}(t_k) = y(t_k) + w(t_k)$, where $w(t_k)$ is the measurement noise that is realized by the i.i.d. Gaussian white noise process $w(t_k) \sim \mathcal{N}(0, \Sigma)$ with $\Sigma = I_4$. Let $y^*(t_k)$ be the sampled state response of the estimated system $A - BK^*$, and T be the total number of samples k in the measured responses. Then, if the system matrices (A, B, C) are known, the feedback gain matrix can be estimated by a maximum-likelihood estimator (MLE) via minimizing the least-squares error

$$K_e = \arg \min_{K^*} J(K^*)$$

$$J(K^*) = \sum_{k=0}^{T-1} (y^*(t_k) - \tilde{y}(t_k))^T \Sigma^{-1} (y^*(t_k) - \tilde{y}(t_k)). \quad (25)$$

We apply MATLAB's numerical optimization algorithm *fminsearch* to the problem of (25), with A, B known and recover

$$K_e = \begin{bmatrix} 2.376 & -1.328 & 1.188 & 1.847 \\ 4.294 & -0.8621 & 3.77 & 1.509 \\ -2.279 & 1.366 & -2.067 & -1.323 \\ -2.7 & -0.06092 & -2.036 & -0.7217 \end{bmatrix} \approx K_0.$$

Measurement noise and other disturbances may perturb the estimation K_e away from K_0 . However, under the conditions we have specified, we make Remark 3.

Remark 3: In general, for a given set of system matrices (A, B, C) , K_e minimizing the cost function in (25) is the MLE of K_0 [24], and under mild conditions (e.g., identifiability [24]), has a limit of K_0 w. p. 1, as the length of the experiment T goes to ∞ [25].

Once we have recovered K_e we can solve the convex LMI optimization problem defined in (12) efficiently using the SeDuMi [26] package with the YALMIP modeling toolbox [27] in MATLAB. The optimal estimated weights (\hat{Q}, \hat{R})

are found to be

$$\hat{Q} = \begin{bmatrix} 13.63 & -1.184 & 4.065 & 2.723 \\ -1.184 & 7.334 & 3.497 & -2.449 \\ 4.065 & 3.497 & 4.924 & 0.1462 \\ 2.723 & -2.449 & 0.1462 & 4.395 \end{bmatrix}$$

$$\hat{R} = \begin{bmatrix} 10.55 & 0.3451 & 6.513 & -1.717 \\ 0.3451 & 4.326 & 1.811 & 3.514 \\ 6.513 & 1.811 & 6.93 & -0.4812 \\ -1.717 & 3.514 & -0.4812 & 6.555 \end{bmatrix}.$$

Because the original weights Q_0 and R_0 were designed to minimize the condition number of the overall weighting matrix, the recovered \hat{Q} and \hat{R} match the original weights closely. In general, however, the condition number minimization in (12) and (13) means that the recovered \hat{Q} and \hat{R} may not be numerically similar to Q_0 and R_0 , but will produce an equivalent controller in the forward LQR problem (assuming that K_e estimates K_0 accurately).

IV. ILLUSTRATIVE EXAMPLE WITH INFEASIBLE SOLUTION

In the case when the LMI problem in (12) is infeasible, we apply the gradient descent algorithm outlined in Table I. Consider the system of (21) with

$$A = \begin{bmatrix} 100 & 0 & -1 \\ 0 & 0.1 & 50 \\ 0.333 & 10 & 0 \end{bmatrix}$$

$$B = \begin{bmatrix} -1 & 0 & 10 \\ 1 & 1 & 0 \\ 0.1 & -20 & 4 \end{bmatrix}$$

$$C = I_3.$$

Design a controller via pole-placement such that the closed-loop poles are $\lambda_1 = -90$, $\lambda_2 = -20$, $\lambda_3 = -10$. This results in the feedback gain matrix

$$K_0 = \begin{bmatrix} -3.69 & 20.1 & 49.3 \\ 3.69 & 0.00244 & 0.712 \\ 18.6 & 2.01 & 4.83 \end{bmatrix}.$$

The closed-loop system is again simulated at a number of sampling points t_k with an additive i.i.d. measurement noise realization of $w(t_k) \sim \mathcal{N}(0, I)$ such that $\tilde{y}(t_k) = y(t_k) + w(t_k)$. If we again use a sampled input sequence which is a realization of the process $r(t_k) \sim \mathcal{N}(0, 10I_3)$, the gain matrix K_e recovered via system identification with (A, B, C) known is

$$K_e = \begin{bmatrix} -3.47 & 20.2 & 49.3 \\ 3.7 & 0.0519 & 0.714 \\ 18.7 & 2.21 & 4.83 \end{bmatrix} \approx K_0.$$

However, for this system, numerical computation shows that the solution to the LMI problem in (12) is infeasible. Applying the gradient descent algorithm from (19) using the initial point derived from a solution to (20)

$$Q_s = \begin{bmatrix} 71.4 & 96.4 & 62.8 \\ 96.4 & 302 & 8.62 \\ 62.8 & 8.62 & 1550 \end{bmatrix}, R_s = \begin{bmatrix} 1.33 & -1.49 & -3.19 \\ -1.49 & 387 & -77 \\ -3.19 & -77 & 53.3 \end{bmatrix}$$

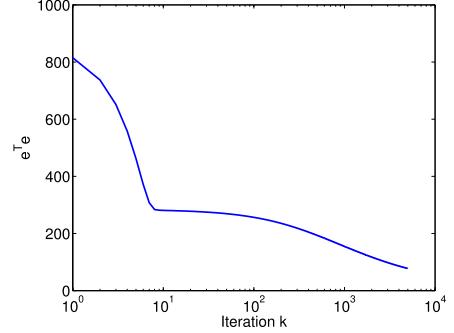


Fig. 1. Convergence of the error norm during the gradient descent process.



Fig. 2. Example setup for seated balance testing using the Mikrolar R-3000 Rotopod.

yields

$$\hat{K} = \begin{bmatrix} -3.75 & 27.6 & 45.3 \\ 3.93 & -0.276 & -1.7 \\ 19.7 & 2.98 & 4.57 \end{bmatrix}$$

which has a residual cost of $e^T e = 77.24$ after 5000 iterations (Fig. 1). The final optimal weighting matrices formed from \hat{K} were

$$\hat{Q} = \begin{bmatrix} 71.4 & 96.4 & 62.7 \\ 96.4 & 302 & 8.76 \\ 62.7 & 8.76 & 1550 \end{bmatrix}, \hat{R} = \begin{bmatrix} 1 & 7.09 & -5.05 \\ 7.09 & 388 & -76.2 \\ -5.05 & -76.2 & 53.1 \end{bmatrix}.$$

Even though we were not able to recover $\hat{K} = K_0$, the gradient descent algorithm has produced a locally optimal estimate minimizing the residual error $e^T e$ despite the imperfect initial guess θ_0 .

V. EXPERIMENTAL HUMAN COST FUNCTION RECOVERY

We have developed an experimental setup for identification of the human response during an upright seated balance task (Fig. 2), and to which our inverse LQR solution method can be applied. One subject volunteered for this portion of the experiment and the testing was designated as nonregulated research by the MSU Institutional Review Board. The subject was seated on a hexapod robot (R-3000 Rotopod, Mikrolar Inc., Hampton, NH), which was used to apply rigid position disturbances to the subject's lower body about a lateral

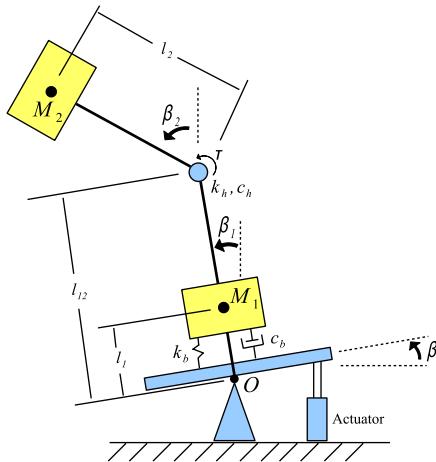


Fig. 3. Seated-balance model of the human subject. The values of the given parameters are adapted from [3] and [28].

bending axis centered on the pelvis. In order to calculate kinematics, LED markers (Visualeyez Motion Capture System, Phoenix Technologies Inc., Burnaby Canada) were attached to the subject (on the trunk and sacrum) and the robotic platform. Experimental kinematic data were sampled at 100 samples/s. During the trial, the subject was given the goal of keeping the upper body as close to vertical as possible. We constructed a rigid-body model of this task (Fig. 3) similar to the seated balance model developed in [3], and having parameter values adapted from [3] and [28]. In this model, control authority from the spine muscles is represented by the torque, τ , acting about the L4 vertebrae. The lower body has mass M_1 at a distance l_1 from the actuator pivot point, and moment of inertia J_1 about the center of mass. The pivot point at the L4 vertebrae is a distance l_{12} from the actuator pivot. The upper body has mass M_2 at a distance l_2 from the L4 pivot, with moment of inertia J_2 about the center of mass. The spine has some intrinsic stiffness k_h and intrinsic damping c_h . Because the actuator provides a rigid disturbance β_i , we have modeled the interaction of the seat and lower body through the soft gluteal tissues which have stiffness k_b and damping c_b , which would be fitted to the experimental data later.

The nonlinear equations of motion were derived using Lagrange's equation using a state-space representation $x = [\beta_1 \dot{\beta}_1 \beta_2 \dot{\beta}_2]^T$, and linearized about the operating point $x = [0 \ 0 \ 0 \ 0]^T$ to form the plant model G . We assume that the sampled system outputs are a noisy direct measure of the states, i.e., $\tilde{y}(t_k) = Cy(t_k) + w(t_k)$, with $C = I_4$ and $w(t_k)$ a realization of the i.i.d. process $w(t_k) \sim \mathcal{N}(0, \Sigma)$ with $\Sigma = \sigma^2 I_4$. We presumed the existence of a full-state feedback controller $K = [K_1 \ K_2 \ K_3 \ K_4]$ which would produce the voluntary input torque τ via $\tau = -Ky$.

We fit the six unknown model parameters $\zeta := [K_1, \dots, K_4, k_b, c_b]^T$ by using MATLAB's *fminsearchbnd* function to perform the minimization

$$\zeta_e = \arg \min_{\zeta^*} \sum_{t=0}^{T-1} (y^*(t_k) - \tilde{y}(t_k))^T W (y^*(k) - \tilde{y}(t_k))$$

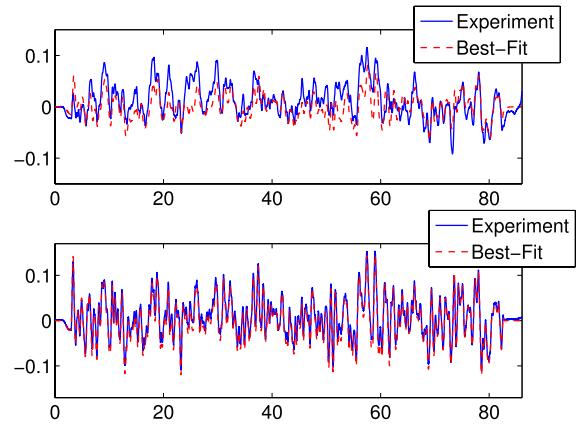


Fig. 4. Plot of the experimental and best-fit angular responses. Top: upper body angle versus time. Bottom: lower body angle versus time.

TABLE II
OPTIMAL ESTIMATED PARAMETERS ζ_e OF THE HUMAN SUBJECT

Parameter	Value
$K_{1,e}$	-1.8×10^5
$K_{2,e}$	-5.573×10^3
$K_{3,e}$	2.295×10^5
$K_{4,e}$	7.37×10^4
$k_{b,e}$	$2.958 \times 10^4 \text{ Nm/rad}$
$c_{b,e}$	$8.09 \times 10^3 \text{ Nms/rad}$

where $W = \text{diag}(1, 0, 1, 0)$, k is the sample index, T is the number of samples in the experiment, $y^*(t_k)$ is the sampled state response of the estimated system using ζ^* , and $\tilde{y}(t_k)$ is the sampled response of the experimentally measurable system states. The resulting best-fit and experimental responses for β_1 and β_2 are shown in Fig. 4. The estimated parameters ζ_e are shown in Table II. The estimated feedback controller K_e was formed as $K_e = [K_{1,e}, \dots, K_{4,e}]$.

The values $k_{b,e}$ and $c_{b,e}$ were incorporated into the A matrix of the plant G to form A_e , and the columns of B associated with β_i and $\dot{\beta}_i$ were removed to form B_e . The inverse LQR procedure of (12) was then applied to determine \hat{Q} and \hat{R} such that $K_e = \text{CLQR}(A_e, B_e, \hat{Q}, \hat{R})$. In this case, a feasible solution to the LMI problem could not be found. The gradient descent method of Algorithm I was applied, recovering

$$\hat{Q} = 10^{10} \times \begin{bmatrix} 1.693 & 0.2328 & -0.1532 & 0.09889 \\ 0.2328 & 1.794 & 0.1507 & 0.571 \\ -0.1532 & 0.1507 & 5.247 & -0.0426 \\ 0.09889 & 0.571 & -0.0426 & 0.5411 \end{bmatrix}$$

$$\hat{R} = 1.00$$

with residual cost $\|\hat{K} - K_e\|_F^2 = 24.514$.

This (\hat{Q}, \hat{R}) meets all the conditions defined in (12). Notice that the diagonal element associated with the upper body angle (\hat{Q}_{33}) provides the largest single contribution to the cost. Further, the relative weights of \hat{Q} and \hat{R} suggest that the system states are penalized much more heavily than the control effort in the cost function. However, it is not immediately clear from \hat{Q} whether linear combinations of the states may offer a more salient picture of how the cost is distributed. Therefore, we apply a similarity transform to x and \hat{Q} such

that \tilde{Q} is diagonal and operates on \tilde{x} , which is a vector of linear combinations of the elements in x . If we let V be an orthogonal matrix whose columns are the eigenvectors of \hat{Q} , and Λ the square diagonal matrix whose diagonal elements are the corresponding eigenvalues of \hat{Q} , then $\tilde{Q} = \Lambda$, and $\tilde{x} = V^T x$. This similarity transformation will satisfy the equality $x^T \hat{Q} x = \tilde{x}^T \tilde{Q} \tilde{x}$. For the experimental \hat{Q} found above

$$V = \begin{bmatrix} -0.003621 & -0.874 & -0.4842 & -0.04044 \\ -0.362 & 0.4511 & -0.8148 & 0.03988 \\ 0.01901 & -0.05254 & 0.01132 & 0.9984 \\ 0.932 & 0.1729 & -0.3186 & -0.005035 \end{bmatrix}$$

$$\tilde{Q} = 10^{10} \times \text{diag}(0.3181, 1.544, 2.153, 5.259).$$

Note that the last column in V , which is the eigenvector of \hat{Q} that corresponds to the largest eigenvalue in \tilde{Q} , will produce a coordinate in \tilde{x} that is a linear combination of the body angles and rates (i.e., $\tilde{x}_4 = -0.04044\beta_1 + 0.03988\dot{\beta}_1 + 0.9984\beta_2 - 0.005035\dot{\beta}_2$), with a highest weight on the upper body angle β_2 . If we consider only the largest eigenvalue, \tilde{x}_4 is reasonably consistent with the motion goal given to the subject (minimize $|\beta_2|$). However, a quantitative clinical study would have to be performed to draw any scientific conclusions.

VI. CONCLUSION

In this brief, we described a comprehensive methodology for determining a cost function to the time-invariant LQR problem in both continuous- and discrete-time cases. Our results have potential application not only to the determination of human control cost, but also to the reverse-engineering of black-box controllers, and offer a new dimension of information (control design cost function) beyond that available using traditional system identification techniques. A set of several numerical problems and an experimental result with a human subject on a seated balance testing apparatus successfully demonstrate that our proposed method is able to determine a salient measure of control performance weights from experimental data. We plan to use this methodology in the future to more comprehensively evaluate human postural control and determine if consistent features or control goals can be extracted from the resulting cost functions.

REFERENCES

- [1] G. C. Peng, T. C. Hain, and B. W. Peterson, "A dynamical model for reflex activated head movements in the horizontal plane," *Biol. Cybern.*, vol. 75, no. 4, pp. 309–319, 1996.
- [2] A. D. Goodworth and R. J. Peterka, "Contribution of sensorimotor integration to spinal stabilization in humans," *J. Neurophysiol.*, vol. 102, no. 1, pp. 496–512, 2009.

- [3] N. P. Reeves, J. Cholewicki, and K. S. Narendra, "Effects of reflex delays on postural control during unstable seated balance," *J. Biomech.*, vol. 42, no. 2, pp. 164–170, 2009.
- [4] N. Abaid, P. Cappa, E. Palermo, M. Petrarca, and M. Porfiri, "Gait detection in children with and without hemiplegia using single-axis wearable gyroscopes," *PLoS One*, vol. 8, no. 9, p. e73152, 2013.
- [5] K. Mombaur, A. Truong, and J.-P. Laumond, "From human to humanoid locomotion—An inverse optimal control approach," *Auto. Robots*, vol. 28, no. 3, pp. 369–383, 2010.
- [6] G. Arechavaleta, J.-P. Laumond, H. Hicheur, and A. Berthoz, "An optimality principle governing human walking," *IEEE Trans. Robot.*, vol. 24, no. 1, pp. 5–14, Feb. 2008.
- [7] B. Berret, E. Chiovetto, F. Nori, and T. Pozzo, "Evidence for composite cost functions in arm movement planning: An inverse optimal control approach," *PLoS Comput. Biol.*, vol. 7, no. 10, p. e1002183, 2011.
- [8] X. Niu, M. L. Latash, and V. M. Zatsiorsky, "Reproducibility and variability of the cost functions reconstructed from experimental recordings in multifinger prehension," *J. Motor Behavior*, vol. 44, no. 2, pp. 69–85, 2012.
- [9] P. J. Antsaklis and A. N. Michel, *Linear Systems*. Boston, MA, USA: Birkhäuser, 2005.
- [10] T. Fujii and M. Narazaki, "A complete optimality condition in the inverse problem of optimal control," *SIAM J. Control Optim.*, vol. 22, no. 2, pp. 327–341, 1984.
- [11] R. E. Kalman, "When is a linear control system optimal?" *J. Basic Eng.*, vol. 86, no. 1, pp. 51–60, 1964.
- [12] B. D. Anderson, "The inverse problem of optimal control," DTIC Document, Fort Belvoir, VA, USA, Tech. Rep. 6560-3, 1966.
- [13] A. Jameson and E. Kreindler, "Inverse problem of linear optimal control," *SIAM J. Control*, vol. 11, no. 1, pp. 1–19, 1973.
- [14] E. Kreindler and A. Jameson, "Optimality of linear control systems," *IEEE Trans. Autom. Control*, vol. 17, no. 3, pp. 349–351, Jun. 1972.
- [15] B. Molinari, "The stable regulator problem and its inverse," *IEEE Trans. Autom. Control*, vol. 18, no. 5, pp. 454–459, Oct. 1973.
- [16] H. Kong, G. Goodwin, and M. Seron, "A revisit to inverse optimality of linear systems," *Int. J. Control.*, vol. 85, no. 10, pp. 1506–1514, 2012.
- [17] F. Aioun, A. Heniche, and H. Bourles, "Maximally-diagonal solution to the inverse LQR problem," in *Proc. 33rd IEEE Conf. Decision Control*, vol. 2, Dec. 1994, pp. 1445–1446.
- [18] S. Boyd, L. E. Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*, vol. 15. Philadelphia, PA, USA: SIAM, 1987.
- [19] M. C. Priess, J. Choi, and C. Radcliffe, "Determining human control intent using inverse LQR solutions," in *Proc. ASME Dyn. Syst. Control Conf.*, Oct. 2013.
- [20] E. Cheney and D. Kincaid, *Numerical Mathematics and Computing*. Stamford, CT, USA: Cengage Learning, 2012.
- [21] S. P. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
- [22] J. Dattorro, *Convex Optimization and Euclidean Distance Geometry*. New York, NY, USA: Meboo Publishing, 2005.
- [23] D. F. Delchamps, "Analytic stabilization and the algebraic Riccati equation," in *Proc. 22nd IEEE Conf. Decision Control*, vol. 22, Dec. 1983, pp. 1396–1401.
- [24] L. Ljung, *System Identification: Theory for the User*. Upper Saddle River, NJ, USA: Prentice-Hall, 1999.
- [25] H. Hjalmarsson, "From experiment design to closed-loop control," *Automatica*, vol. 41, no. 3, pp. 393–438, 2005.
- [26] I. Polik. (Jun. 2010). *Sedumi Home Page*. [Online]. Available: <http://sedumi.ie.lehigh.edu/>
- [27] J. Löfberg. (Nov. 2012). *Yalmip Wiki Main/Home Page*. [Online]. Available: <http://users.isy.liu.se/johanl/yalmip/>
- [28] V. M. Zatsiorsky, *Kinetics of Human Motion*. Windsor, ON, Canada: Human Kinetics, 2002.