International Journal of Modern Nonlinear Theory and Application, 2014, 3, 230-235
Published Online December 2014 in SciRes. http://www.scirp.org/journal/ijmnta
http://dx.doi.org/10.4236/ijmnta.2014.35025

Performance of Suboptimal Controllers for Affine-Quadratic Problems

Ankita Sharma, A. J. Shaiju

Department of Mathematics, Indian Institute of Technology Madras, Chennai, India
Email: ankita.iitm22@gmail.com, ajshaiju@iitm.ac.in

Received 5 November 2014; revised 1 December 2014; accepted 9 December 2014

Copyright © 2014 by authors and Scientific Research Publishing Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY).
http://creativecommons.org/licenses/by/4.0/
Open Access

Abstract

In this article, affine-quadratic control problems are studied. Error bounds are derived for the difference between the performance indices corresponding to the optimal and a class of suboptimal controls. In particular, it is shown that the performance of these suboptimal controls is close to that of the optimal control whenever the error in estimating the costate initial condition is small.

Keywords

Affine-Quadratic Control, Nonlinear Control, Optimal Control, Suboptimal Control

1. Introduction

One of the most active areas in control theory is optimal control and methods to find them [1]-[3]. It has a wide range of practical applications in engineering (Aerospace, Chemical, Mechanical, Electrical), science (Physics, Biology), and economics (see e.g. [4]-[7]). Optimal control theory has been developed for linear systems ([1] [2] [8]) and explicit formulae for computing optimal control inputs are available. However, control of nonlinear systems is much more challenging and obtaining formulae for optimal controls seems in general not possible. This motivated researchers to study various classes of nonlinear control problems separately, and *affine-qudratic problems* is one such class. In a recent paper [9], the optimal control for affine-quadratic problems is obtained in terms of the associated costate. But, in practice, it is difficult to compute the costate (at each time t) as the knowledge of its terminal condition is required.

In this article, we study the affine-quadratic control problem given by ((1), (2)). We note that a method for finding the initial condition for the costate is recently proposed [10]. This allows one to compute the initial costate (at $t = 0$) exactly or approximately. This approximation of the initial costate and the explicit formula for

How to cite this paper: Sharma, A. and Shaiju, A.J. (2014) Performance of Suboptimal Controllers for Affine-Quadratic Problems. *International Journal of Modern Nonlinear Theory and Application*, **3**, 230-235.
http://dx.doi.org/10.4236/ijmnta.2014.35025

optimal control (as in [11]) are shown, in this article, which give rise to suboptimal controls of practical importance. More precisely, our main theorem (Theorem 2) provides an upper bound for the difference in performance between these suboptimal and optimal control.

The article is organized as follows. In Section 2, the affine-quadratic control problem is described. We also explain how to obtain the optimal control in terms of costate. The main (Theorem 2) is proved in Section 3. This theorem provides a method to obtain the costate (without the knowledge of its terminal value) which results in an explicit formula and performance bounds for a class of suboptimal controls.

Notation: For $x=(x_1,x_2,\cdots,x_n)\in\mathbb{R}^n$, $A=(a_{ij})\in\mathbb{R}^{n\times n}$, and $1\le p<\infty$, we use the notation

$$\|x\|_p=\left(\sum_{i=1}^n|x_i|^p\right)^{\frac{1}{p}},\quad \|x\|_\infty=\max\{|x_1|,|x_2|,\cdots,|x_n|\},\quad \|A\|=\max_{1\le i\le n}\left(\sum_{j=1}^n|a_{ij}|\right).$$

2. Problem Description

We consider the affine control system

$$\dot{x}(t)=f(x(t))+g(x(t))u(t);\quad x(0)=x_0,\quad 0\le t\le T, \tag{1}$$

with the quadratic cost functional

$$J(x_0,u(\cdot))=\frac{1}{2}x'(T)Sx(T)+\frac{1}{2}\int_0^T\left(x'(t)Qx(t)+u'(t)Ru(t)\right)\mathrm{d}t. \tag{2}$$

Here $x(t)\in\mathbb{R}^n$ is the state vector, $u(t)\in\mathbb{R}^m$ is the control vector, $f:\mathbb{R}^n\to\mathbb{R}^n$, $g=[g_1,g_2,\cdots,g_m]:\mathbb{R}^n\to\mathbb{R}^{n\times m}$, $S\in\mathbb{R}^{n\times n}$, $Q\in\mathbb{R}^{n\times n}$, $R\in\mathbb{R}^{m\times m}$, and $'$ denotes transposition.

Throughout this paper, it is assumed that S, Q are positive semidefinite, R is positive definite, the functions f, g are continuously differentiable with bounded derivatives, the control input $u(\cdot)$ is chosen from the admissible control space $\mathcal{U}=L^1([0,T];\mathbb{R}^m)$.

Under these assumptions, for each admissible control $u(\cdot)\in\mathcal{U}$ there exist a unique solution (trajectory) of the control system (1) denoted by $\phi_u(t;x_0)$.

The value function of the control problem given by (1), (2), is defined as

$$V(x_0)=\inf_{u(\cdot)\in\mathcal{U}}J(x_0,u(\cdot)).$$

A control input $u^*(\cdot)\in\mathcal{U}$ is optimal (for (x_0)) if

$$J(x_0,u^*(\cdot))=V(x_0).$$

Similarly a control input $u^\epsilon(\cdot)$ is ϵ-optimal (for (x_0)) if

$$J(x_0,u^\epsilon(\cdot))\le V(x_0)+\varepsilon.$$

Given x_0, the optimal control problem is to find a control $u^*(\cdot)$ which minimizes the cost functional $J(x_0,u(\cdot))$. The Hamiltonian associated with the optimal control problem (1), (2), is given as

$$H(x,u,\lambda)=\frac{1}{2}(x'Qx+u'Ru)+\lambda'(f(x)+g(x)u), \tag{3}$$

where $\lambda\in\mathbb{R}^n$ is the adjoint vector.

To derive an expression for the optimal control $u^*(\cdot)$ (for (x_0)), it is convenient to introduce the adjoint system:

$$\dot{\lambda}^*(t)=-\frac{\partial H}{\partial x}(x^*(t),u^*(t),\lambda^*(t));\lambda^*(T)=Sx^*(T),\quad 0\le t\le T. \tag{4}$$

Here $x^*(t)=\phi_{u^*}(t;x_0)$. We now state the Pontryagin's Minimum Principle (PMP) for the affine-quadratic control system (1), (2), which provides a set of necessary conditions for $u^*(\cdot)$ to be optimal [12].

Theorem 1 [PMP] *Let $x_0 \in \mathbb{R}^n$, $u^*(\cdot)\in\mathcal{U}$, and $x^*(t)=\phi_{u^*}(t;x_0)$. Also let $\lambda^*(t)$ be the adjoint vector corresponding to $u^*(t)$ and x_0, as given by the Equation* (4). *Then for a control input $u^*(\cdot)$ to be optimal for (x_0), it is necessary that the map*

$$u \to H\left(x^*(t),u,\lambda^*(t)\right),$$

attains minimum at $u=u^*(t)$, for a.e. $0\le t\le T$.

Corollary 1 *Let $x_0\in\mathbb{R}^n$, $u^*(\cdot)\in\mathcal{U}$, and $x^*(t)=\phi_{u^*}(t;x_0)$. Also let $\lambda^*(t)$ be the adjoint vector corresponding to $u^*(t)$ and x_0, as given by the Equation* (4). *Then the optimal control* (*for (x_0)*) *is*

$$u^*(t)=-R^{-1}g'\left(x^*(t)\right)\lambda^*(t). \tag{5}$$

Proof. The proof follows immediately from the above theorem. □

Now to obtain $\lambda^*(t)$ (in (5)) in terms of $x^*(t)$, we solve the coupled systems given in (1) and (4) together with the initial conditions $x^*(0)=x_0$ and $\lambda^*(0)=\lambda_0$ respectively.

In general, solving this coupled system and finding a closed form solution $\lambda^*(t)$ is very difficult. However it may be easier to find λ_0 approximately. Such an approximation $\hat{\lambda}_0$ will lead to the associated adjoint state $\hat{\lambda}(t)$ and admissible control $\hat{u}(t)=-R^{-1}g'\left(\hat{x}(t)\right)\hat{\lambda}(t)$. In the next section, we provide bounds for the difference between the performance indices corresponding to $u^*(\cdot)$ and $\hat{u}(\cdot)$.

3. Performance of Suboptimal Controllers

In this section, we prove the main result.

Theorem 2 Consider the affine-quadratic control problem (1), (2). Let $x_0\in\mathbb{R}^n$, $u^*(\cdot)$ be the optimal control as given in (5), $x^*(t)=\phi_{u^*}(t;x_0)$, and $\lambda^*(\cdot)$ be the adjoint vector corresponding to $u^*(\cdot)$ and x_0. Also let $\hat{u}(\cdot)$ be a suboptimal control and $\left(\hat{x}(\cdot),\hat{\lambda}(\cdot)\right)$ be the solution of the coupled system ((1), (4)) with initial condition $\left(x_0,\hat{\lambda}_0\right)$. Then

$$\left|J\left(x_0,u^*(\cdot)\right)-J\left(x_0,\hat{u}(\cdot)\right)\right|\le n\|S\|\epsilon_1 k_1+n\|Q\|\epsilon_1 k_1 T+nC\epsilon_2 k_2 T+\frac{1}{2}nK\epsilon_1 k_2^2 T,$$

where

$$\epsilon_1 := \sup_{0\le t\le T}\left\|x^*(t)-\hat{x}(t)\right\|_\infty,$$
$$\epsilon_2 := \sup_{0\le t\le T}\left\|\lambda^*(t)-\hat{\lambda}(t)\right\|_\infty,$$
$$k_1 := \sup\max_{0\le t\le T}\left\{\left\|x^*(t)\right\|_\infty,\left\|\hat{x}(t)\right\|_\infty\right\},$$
$$k_2 := \sup\max_{0\le t\le T}\left\{\left\|\lambda^*(t)\right\|_\infty,\left\|\hat{\lambda}(t)\right\|_\infty\right\}.$$

The constant C depends only on the matrix function g and the constant K depends only on its gradient.

Proof. Note that

$$\left|J\left(x_0,u^*(\cdot)\right)-J\left(x_0,\hat{u}(\cdot)\right)\right|=$$
$$\left|\left\{\frac{1}{2}x^{*\prime}(T)Sx^*(T)+\frac{1}{2}\int_0^T\left(x^{*\prime}(t)Qx^*(t)+u^{*\prime}(t)Ru^*(t)\right)\mathrm{d}t\right\}-\left\{\frac{1}{2}\hat{x}'(T)S\hat{x}(T)+\frac{1}{2}\int_0^T\left(\hat{x}'(t)Q\hat{x}(t)+\hat{u}'(t)R\hat{u}(t)\right)\mathrm{d}t\right\}\right|. \tag{6}$$

From R.H.S. of (6), we first consider the term

$$\frac{1}{2}\left|\int_0^T\left(x^{*\prime}(t)Qx^*(t)-\hat{x}'(t)Q\hat{x}(t)\right)\mathrm{d}t\right|\le\frac{1}{2}\int_0^T\left|x^{*\prime}(t)Qx^*(t)-\hat{x}'(t)Q\hat{x}(t)\right|\mathrm{d}t.$$

By adding and subtracting $\hat{x}'(t)Qx^*(t)$ inside the integral, we get

$$\begin{aligned}
&\frac{1}{2}\int_0^T\left|x^{*\prime}(t)Qx^*(t)-\hat{x}'(t)Q\hat{x}(t)\right|\mathrm{d}t\\
&\le\frac{1}{2}\int_0^T\left|\left\{x^{*\prime}(t)-\hat{x}'(t)\right\}Qx^*(t)\right|\mathrm{d}t+\frac{1}{2}\int_0^T\left|\hat{x}'(t)\left\{Q\left(x^*(t)-\hat{x}(t)\right)\right\}\right|\mathrm{d}t\\
&\le\frac{1}{2}\int_0^T\left\|x^{*\prime}(t)-\hat{x}'(t)\right\|_2\left\|Qx^*(t)\right\|_2\mathrm{d}t+\frac{1}{2}\int_0^T\left\|\hat{x}'(t)\right\|_2\left\|Q\left(x^*(t)-\hat{x}(t)\right)\right\|_2\mathrm{d}t\\
&\qquad\qquad\left(\text{using Cauchy-Schwarz inequality}\right)\\
&\le\frac{n}{2}\int_0^T\left\|x^{*\prime}(t)-\hat{x}'(t)\right\|_\infty\left\|Qx^*(t)\right\|_\infty\mathrm{d}t+\frac{n}{2}\int_0^T\left\|\hat{x}'(t)\right\|_\infty\left\|Q\left(x^*(t)-\hat{x}(t)\right)\right\|_\infty\mathrm{d}t\\
&\qquad\qquad\left(\text{using }\|x\|_2\le\sqrt{n}\|x\|_\infty\text{ for all }x\in\mathbb{R}^n\right)\\
&\le\frac{n}{2}\|Q\|\epsilon_1\int_0^T\left\|x^*(t)\right\|_\infty\mathrm{d}t+\frac{n}{2}\|Q\|\epsilon_1\int_0^T\left\|\hat{x}'(t)\right\|_\infty\mathrm{d}t\\
&\le\frac{n}{2}\|Q\|\epsilon_1k_1T+\frac{n}{2}\|Q\|\epsilon_1k_1T\\
&=n\|Q\|\epsilon_1k_1T.
\end{aligned}$$

Therefore

$$\frac{1}{2}\left|\int_0^T\left(x^{*\prime}(t)Qx^*(t)-\hat{x}'(t)Q\hat{x}(t)\right)\mathrm{d}t\right|\le n\|Q\|\epsilon_1k_1T. \tag{7}$$

From R.H.S. of (6), we next consider the term

$$\frac{1}{2}\left|x^{*\prime}(T)Sx^*(T)-\hat{x}'(T)S\hat{x}(T)\right|.$$

In a similar manner (as for (7)), we have

$$\frac{1}{2}\left|x^{*\prime}(T)Sx^*(T)-\hat{x}'(T)S\hat{x}(T)\right|\le n\|S\|\epsilon_1k_1. \tag{8}$$

From R.H.S. of (6), we next consider the term

$$\frac{1}{2}\left|\int_0^T\left(u^{*\prime}(t)Ru^*(t)-\hat{u}'(t)R\hat{u}(t)\right)\mathrm{d}t\right|\le\frac{1}{2}\int_0^T\left|u^{*\prime}(t)Ru^*(t)-\hat{u}'(t)R\hat{u}(t)\right|\mathrm{d}t.$$

Let us have

$$\begin{aligned}
&\left|u^{*\prime}(t)Ru^*(t)-\hat{u}'(t)R\hat{u}(t)\right|\\
&=\left|\left(-R^{-1}g'\left(x^*(t)\right)\lambda^*(t)\right)'R\left(-R^{-1}g'\left(x^*(t)\right)\lambda^*(t)\right)-\left(-R^{-1}g'\left(\hat{x}(t)\right)\hat{\lambda}(t)\right)'R\left(-R^{-1}g'\left(\hat{x}(t)\right)\hat{\lambda}(t)\right)\right|.
\end{aligned}$$

In the above term, put the $n\times n$ matrix $g\left(x^*(t)\right)\left(R^{-1}\right)'g'\left(x^*(t)\right)$ as $h\left(x^*(t)\right)$ and the $n\times n$ matrix

$g(\hat{x}(t))(R^{-1})' g'(\hat{x}(t))$ as $h(\hat{x}(t))$ for each $0 \le t \le T$. Then we have,

$$
\begin{aligned}
&\left|u^{*\prime}(t) R u^*(t) - \hat{u}'(t) R \hat{u}(t)\right| \\
&= \left|(\lambda^*(t))' h(x^*(t)) \lambda^*(t) - (\hat{\lambda}(t))' h(\hat{x}(t)) \hat{\lambda}(t)\right| \\
&= \left|(\lambda^*(t))' h(x^*(t))(\lambda^*(t) - \hat{\lambda}(t)) + (\lambda^*(t))' h(x^*(t)) \hat{\lambda}(t) - (\hat{\lambda}(t))' h(\hat{x}(t)) \hat{\lambda}(t)\right| \\
&\qquad\qquad \left(\text{by adding and subtracting the term } (\lambda^*(t))' h(x^*(t)) \hat{\lambda}(t)\right) \\
&= \left|(\lambda^*(t))' h(x^*(t))(\lambda^*(t) - \hat{\lambda}(t)) + (\lambda^*(t))' (h(x^*(t)) - h(\hat{x}(t))) \hat{\lambda}(t) + \left((\lambda^*(t))' - (\hat{\lambda}(t))'\right) h(\hat{x}(t)) \hat{\lambda}(t)\right| \\
&\qquad\qquad \left(\text{again by adding and subtracting the term } (\lambda^*(t))' h(\hat{x}(t)) \hat{\lambda}(t)\right) \\
&\le \left|(\lambda^*(t))' h(x^*(t))(\lambda^*(t) - \hat{\lambda}(t))\right| + \left|(\lambda^*(t))' (h(x^*(t)) - h(\hat{x}(t))) \hat{\lambda}(t)\right| + \left|\left((\lambda^*(t))' - (\hat{\lambda}(t))'\right) h(\hat{x}(t)) \hat{\lambda}(t)\right| \\
&\le \left\|(\lambda^*(t))'\right\|_2 \left\|h(x^*(t))(\lambda^*(t) - \hat{\lambda}(t))\right\|_2 + \left\|(\lambda^*(t))'\right\|_2 \left\|(h(x^*(t)) - h(\hat{x}(t))) \hat{\lambda}(t)\right\|_2 \\
&\qquad + \left\|(\lambda^*(t))' - (\hat{\lambda}(t))'\right\|_2 \left\|h(\hat{x}(t)) \hat{\lambda}(t)\right\|_2, \quad (\text{using Cauchy-Schwarz inequality}).
\end{aligned}
$$

Now using assumption on the matrix function g, we have that the matrix function h is continuously differentiable and has bounded derivatives. Therefore

$$
\begin{aligned}
\left\|h(x^*(t)) - h(\hat{x}(t))\right\| &= \max_i \left(\sum_j \left|h_{ij}(x^*(t)) - h_{ij}(\hat{x}(t))\right|\right) \\
&\le \max_i \left(\sum_j k_{ij}\right) \left\|x^*(t) - \hat{x}(t)\right\|_\infty, \quad \text{where } k_{ij} = \sup_{0 \le t \le T} \left\{\left\|\left(\frac{\partial h_{ij}(x(t))}{\partial x}\right)\right\|_1\right\} \\
&\le K\epsilon_1, \quad K = \max_i \left(\sum_j k_{ij}\right).
\end{aligned}
$$

Using this and following the procedure as for the inequality (7), we get

$$
\left|u^{*\prime}(t) R u^*(t) - \hat{u}'(t) R \hat{u}(t)\right| \le 2nC\epsilon_2 k_2 + nK\epsilon_1 k_2^2,
$$

$$
\left(\text{where non-negative real no. } C \text{ is the bound for the matrix } h(x(t))\right).
$$

Therefore

$$
\frac{1}{2}\left|\int_0^T \left(u^{*\prime}(t) R u^*(t) - \hat{u}'(t) R \hat{u}(t)\right) \mathrm{d}t\right| \le nC\epsilon_2 k_2 T + \frac{1}{2} nK\epsilon_1 k_2^2 T. \tag{9}
$$

Hence the result follows by the inequalities (7), (8), and (9). □

Remark 3 It follows from the previous theorem that $J(x_0, \hat{u}(\cdot)) \to J(x_0, u^*(\cdot))$, when $(\epsilon_1, \epsilon_2) \to (0,0)$. This implies that $\hat{u}(\cdot) = -R^{-1} g'(\hat{x}(\cdot)) \hat{\lambda}(\cdot)$ is a good suboptimal control when $\hat{\lambda}_0$ is a good approximation of λ_0. We emphasize the fact that $\hat{\lambda}(t)$ (and hence $\hat{u}(t)$) can be computed at each time t as $\hat{\lambda}_0$ is known.

References

[1] Anderson, B.D.O. and Moore, J.B. (1989) Optimal Control: Linear Quadratic Methods. Prentice-Hall, Inc., Upper Saddle River.

[2] Zhou, K., Doyle, J.C. and Glover, K. (1996) Robust and Optimal Control. Prentice-Hall, Inc., Upper Saddle River.

[3] Bardi, M. and Dolcetta, I.C. (2008) Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations. Birkhauser, Boston.

[4] Tang, L., Zhao, L.D. and Guo, J. (2009) Research on Pricing Policies for Seasonal Goods Based on Optimal Control Theory. *ICIC Express Letters*, **3**, 1333-1338.

[5] Garrard, W.L. and Jordan, J.M. (1977) Design of Nonlinear Automatic Flight Control Systems. *Automatica*, **13**, 497-505. http://dx.doi.org/10.1016/0005-1098(77)90070-X

[6] Manousiouthakis, V. and Chmielewski, D.J. (2002) On Constrained Infinite-Time Nonlinear Optimal Control. *Chemical Engineering Science*, **57**, 105-114. http://dx.doi.org/10.1016/S0009-2509(01)00359-1

[7] Notsu, T., Konishi, M. and Imai, J. (2008) Optimal Water Cooling Control for Plate Rolling. *International Journal of Innovative Computing*, *Information and Control*, **4**, 3169-3181.

[8] Kalman, R.E. (1960) Contributions to the Theory of Optimal Control. *Matematica Mexicana*, **5**, 102-119.

[9] Effati, S. and Nik, H.S. (2011) Solving a Class of Linear and Non-Linear Optimal Control Problems by Homotopy Perturbation Method. *IMA Journal of Mathematical Control and Information*, **28**, 539-553. http://dx.doi.org/10.1093/imamci/dnr018

[10] Sharma, A. and Shaiju, A.J. (2014) Solution of Affine-Quadratic Control Problems. *Proceedings of the* 19*th WC-IFAC*, Cape-Town.

[11] Jajarmi, A., Pariz, N., Kamyad, A.V. and Effati, S. (2011) A Novel Series Representation Approach to Solve a Class of Nonlinear Optimal Control Problems. *International Journal of Innovative Computing*, *Information and Control*, **7**, 1413-1425.

[12] Pontryagin, L.S., *et al*. (1962) The Mathematical Theory of Optimal Processes. John Wiley and Sons, Inc., New York.