Monograph:

Banks, S.P. (1985) On the Optimal Control of Nonlinear Systems. Research Report. Acse Report 278 . Dept of Automatic Control and System Engineering. University of Sheffield

ON THE OPTIMAL CONTROL OF NONLINEAR SYSTEMS

by

S.P.Banks

Research Report No. 278

Department of Control Engineering
University of Sheffield
Mappin Street
SHEFFIELD S1 3JD

Abstract

The Lie algebra of tensors on a Hilbert space is used to obtain optimal controls for a class of nonlinear systems.

Keywords : Optimal Control, bilinear systems, nonlinear systems, tensors, Lie algebras.

1. Introduction

In this paper we shall consider the optimal control of a nonlinear analytic system by writing the system in the form of an infinite-dimensional bilinear system. This approach has been used before; see Takata [6], Banks [1] , Banks and Ashtiani [3]. However we shall use the theory of tensors (cf. Greub, [5]) and consider the Lie algebra of tensors generated by the operators in the associated bilinear system. This simplifies considerably the approach taken by Banks and Yew [4], where a 'power series' in the tensor operators is obtained for the optimal control. We show here that particularly simple results hold if the Lie algebra is nilpotent.

In section 2 we shall introduce some simple tensor notation and in section 3 the bilinear realisation of nonlinear systems will be considered. The application of simple Lie algebra theory to the bilinear representation will be discussed in section 4, and finally, in section 5, a very simple example will be considered. The example is chosen mainly for its illustrative clarity rather than its applicability.

2, Notation and Terminology

We shall use the elementary theory of tensors on a Hilbert space, usually written in component form. Thus, a tensor $\Phi \in \bigotimes_{k=1}^{n} \ell^2$, for some n, will be written $\phi_{i_1 \ldots i_n}$ where $0<i_j<\infty$ for $1\leq j\leq n$. If A is a tensor operator in $\mathcal{L}(\bigotimes_{k=1}^{n} \ell^2)$, i.e. a bounded operator from $\bigotimes_{k=1}^{n} \ell^2$ to itself we shall denote its componentwise operation, in an expression of the form $\Psi = A\Phi$, by

$$\psi_{i_1 \ldots i_n} = \sum_{j_1=o}^{\infty} \ldots \sum_{j_1=o}^{\infty} A^{j_1 \ldots j_n}_{i_1 \ldots i_n} \phi_{j_1 \ldots j_n} .$$

(Matrix multiplication then becomes $\psi_i = \sum_{j=1}^{n} A^j_i \phi_i$.) Finally the transpose A' of A is the tensor defined by $(A')^{j_1 \ldots j_n}_{i_1 \ldots i_n} = A^{i_1 \ldots i_n}_{j_1 \ldots j_n}$.

3. Bilinear Realisation of Nonlinear Systems

In this section we shall consider a general (analytic) nonlinear system

$$\dot{x} = f(x,u) \quad , \quad x \in \mathbb{R}^n \tag{3.1}$$

where u is assumed, for simplicity, to be a scalar control. As in Banks [2], we shall associate with this system the 'augmented' system

$$\begin{aligned} \dot{x} &= f(x,u) \\ \dot{u} &= v \end{aligned} \tag{3.2}$$

where we restrict the controls to be differentiable. Now introduce the functions

$$\phi_{i_1 \dots i_n} = x_1^{i_1} \dots x_n^{i_n} u^j . \tag{3.3}$$

We have

$$\frac{d\phi_{i_1 \dots i_n j}}{dt} = \sum_{k=1}^{n} i_k x_1^{i_1} \dots x_k^{i_k - 1} \dots x_n^{i_n} u^j \dot{x}_k + j x_1^{i_1} \dots x_n^{i_n} u^{j-1} \dot{u}$$

Since f is analytic it has a Taylor series of the form

$$f_k(x,u) = \sum_{\alpha_1=0}^{\infty} \dots \sum_{\alpha_n=0}^{\infty} \sum_{\beta=0}^{\infty} C^k_{\alpha_1 \dots \alpha_n \beta} x_1^{\alpha_1} \dots x_n^{\alpha_n} u^\beta$$

for some constants $C^k_{\alpha_1 \dots \alpha_n \beta}$. Hence

$$\frac{d\phi_{i_1 \dots i_n j}}{dt} = \sum_{k=1}^{n} \sum_{\alpha_1=0}^{\infty} \dots \sum_{\alpha_n=0}^{\infty} \sum_{\beta=0}^{\infty} i_k C^k_{\alpha_1 \dots \alpha_n \beta} x_1^{\alpha_1 + i_1} \dots x_k^{i_k + \alpha_k - 1} \dots x_n^{\alpha_n + i_n} u^{j+\beta}$$

$$+ j x_1^{i_1} \dots x_n^{i_n} u^{j-1}$$

$$= \sum_{k=1}^{n} \sum_{\alpha_1=0}^{\infty} \dots \sum_{\alpha_n=0}^{\infty} \sum_{\beta=0}^{\infty} i_k C^k_{\alpha_1 \dots \alpha_n \beta} \phi_{\alpha_1 + i_1, \dots, \alpha_k + i_k - 1, \dots, \alpha_n + i_n, j+\beta}$$

$$+ j\phi_{i_1 \dots i_n j-1} v . \tag{3.4}$$

We shall denote the infinite dimensional rank (n+1) tensor $\phi_{i_1 \dots i_n j}$ by Φ ; i.e.

$$(\Phi)_{i_1 \dots i_n j} = \phi_{i_1 \dots i_n j} .$$

Equation (3.4) can then be written in the form

$$\frac{d\Phi}{dt} = A\Phi + B\Phi \tag{3.5}$$

where A is the tensor operator defined by

$$(A\Phi)_{i_1 \dots i_n j} = \sum_{k=1}^{n} \sum_{\alpha_1=0}^{\infty} \dots \sum_{\alpha_n=0}^{\infty} \sum_{\beta=0}^{\infty} i_k C^k_{\alpha_1 \dots \alpha_n \beta} \phi_{\alpha_1+i_1, \dots, \alpha_k+i_k-1, \dots, \alpha_n+i_n, j+\beta}$$

and B is the tensor operator defined by

$$(B\Phi)_{i_1 \dots i_n j} = j\phi_{i_1 \dots i_n j-1}$$

Remark 3.1 We must remember that (3.5) is a tensor differential equation and that A and B are operators on a space of tensors.

Remark 3.2 If $u \varepsilon \mathbb{R}^m$, $m>1$, then in an exactly similar way to that given above, we can show that

$$\frac{d\Phi}{dt} = A\Phi + \sum_{i=1}^{m} v_i B_i \Phi \tag{3.6}$$

where

$$\dot{u}_i = v_i$$

and the B_i operators are defined in an obvious way. (In this case

$$\Phi = (\phi_{i_1 \dots i_n j_1 \dots j_m})$$

where

$$\phi_{i_1 \dots i_n j_1 \dots j_m} = x_1^{i_1} \dots x_n^{i_n} \; u_1^{j_1} \dots u_m^{j_m} \quad .)$$

It has been shown by Banks and Ashtiani [3] that with suitable scaling, if the solutions of (3.1) are bounded then we can regard equation (3.5) (or(3.6)) as being defined on the space $\bigotimes_{k=1}^{n+1} \ell^2$, i.e. the tensor product of $n+1$ copies of ℓ^2.

Remark 3.3 The tensor operators A and B can be written in the forms

$$(A\Phi)_{i_1\ldots i_n j} = \sum_{\alpha_1=0}^{\infty} \cdots\cdots \sum_{\alpha_n=0}^{\infty} \sum_{\beta=0}^{\infty} a_{i_1\ldots i_n j}^{\alpha_1\ldots\alpha_n\beta} \phi_{\alpha_1\ldots\alpha_n\beta} \tag{3.7}$$

and

$$(B\Phi)_{i_1\ldots i_n j} = \sum_{\alpha_1=0}^{\infty} \cdots\cdots \sum_{\alpha_n=0}^{\infty} \sum_{\beta=0}^{\infty} b_{i_1\ldots i_n j}^{\alpha_1\ldots\alpha_n\beta} \phi_{\alpha_1\ldots\alpha_n\beta} \tag{3.8}$$

where

$$a_{i_1\ldots i_n j}^{\alpha_1\ldots\alpha_n\beta} = \sum_{k=1}^{n} i_k C^k_{\alpha_1-i_1,\ldots\alpha_k-i_k+1,\ldots,\alpha_n-i_n,\beta-j}$$

and

$$b_{i_1\ldots i_n j}^{\alpha_1\ldots\alpha_n\beta} = j\delta_{i_1}^{\alpha_1}\delta_{i_2}^{\alpha_2}\ldots\delta_{i_n}^{\alpha_n}\delta_{j-1}^{\beta} \quad .$$

(The coefficients $C^k_{i_1 i_2\ldots i_n j}$ are interpreted as 0 if any of the indices $i_{1,\ldots,}\, i_n, j$ are zero).

4. Optimal Control of Nonlinear Systems

Consider now the nonlinear control system

$$\dot{x} = f(x,u) \tag{4.1}$$

together with the simple quadratic cost function

$$J_x(u) = \int_0^T u^2\, dt + x'(T)Fx(T) \tag{4.2}$$

Now replace (4.1) by the augmented system (3.2), i.e.

$$\dot{x} = f(x,u) \tag{4.3}$$

$$\dot{u} = v$$

We have seen that this system is equivalent to a system of the form

$$\dot{\Phi} = A\Phi + vB\Phi \tag{4.4}$$

where $\Phi \in \bigotimes_{k=1}^{n+1} \ell^2$ and $A,B \in \mathcal{L}(\bigotimes_{k=1}^{n+1} \ell^2)$. Since we are now using v as a control rather than u we must replace the cost function (.42) by

$$J_{\Phi}(v) = \int_0^T v^2\, dt + \Phi'(T)\Gamma\Phi(T) \tag{4.5}$$

where $\Gamma \in \mathcal{L}(\bigotimes_{k=1}^{n+1} \ell^2)$ is defined by

$$\Gamma^{\alpha_1 \ldots \alpha_n \beta}_{i_1 \ldots i_n j} = \sum_{\ell=1}^{n} \sum_{k=1}^{n} F^k_\ell \, \delta^{\alpha_1}_o \cdots \delta^{\alpha_k}_1 \cdots \delta^{\alpha_n}_o \, \delta^\beta_o \delta^o_{i_1} \cdots \delta^1_{i_\ell} \cdots \delta^o_{i_n} \delta^o_j + \delta^{\alpha_1}_o \cdots \delta^{\alpha_n}_o \, \delta^o_{i_1} \cdots \delta^o_{i_n} \delta^\beta_1 \delta^1_j \tag{4.6}$$

where the last term is chosen to weight u^2 at time T.

We shall now consider the problem (4.4) with the cost function (4.5) to replace (4.1) and (4.2). Although the two problems are not equivalent, it seems reasonable to minimise the control u at the final time and the square of $\dot{u}$ over $[0, T]$.

The Hamiltonian for this problem is

$$H = v^2 + \Lambda' ((A\Phi + vB\Phi)$$

and so we obtain the equations

$$\dot{\Lambda}' = -(\Lambda' A + v\Lambda' B) \quad , \qquad \Lambda(T) = \Gamma \, \Phi(T) \tag{4.7}$$

$$\dot{\Phi} = A\Phi + vB\Phi \tag{4.8}$$

and

$$2v = -\Lambda' B \Phi \tag{4.9}$$

Since $\dot{v} = \frac{-d}{dt}(\Lambda' B \Phi) = -\Lambda' [B, A] \Phi$, we have that the optional control is a constant if B commutes with A (as tensor operators). More generally we have

<u>Lemma 4.1</u> For any constant tensor operator X we have

$$\frac{d}{dt}(\Lambda' X \Phi) = \Lambda' [X, A+vB] \, \Phi \tag{4.10}$$

<u>Proof</u> From (4.7) and (4.8) we have

$$\frac{d}{dt}(\Lambda' X \Phi) = \dot{\Lambda}' X\Phi + \Lambda' X \dot{\Phi} = -(\Lambda' A + v\Lambda' B) X\Phi + \Lambda' X (A+vB)\Phi$$
$$= \Lambda' [X, A+vB] \Phi. \quad \square$$

Now consider the Lie algebra generated by the tensor operators A,B; i.e. the set of linear combinations of all brackets generated by A and B. Denote this Lie albegra by T(A,B) and denote the Lie algebra of all tensor operators (of rank 2(n+1)) by T. In order to consider finite-dimensional Lie algebras we shall truncate the bilinear representation (3.5) of the original nonlinear system (3.2) so that the tensors involved have indices ranging from 0 to, say, ℓ. Since the system is analytic and we are controlling the system in some bounded region, as ℓ increases we shall obtain an accurate representation of the nonlinear system. We shall denote the corresponding Lie algebras by $T_\ell(A,B)$ and T_ℓ. However, to avoid further confusing subscripts we shall not write A_ℓ, B_ℓ etc. for the truncated tensors, but use the same notation A and B, leaving it to the context to indicate which is intended.

Since T_ℓ is a finite-dimentional Lie algebra (of dimension $\ell^{2(n+1)}$) and $T_\ell(A,B)$ is a subalgebra of T_ℓ it follows that $T_\ell(A,B)$ is a finite-dimensional Lie algebra of dimension $m_\ell \leq \ell^{2(n+1)}$. Let $X_1, \ldots, X_{m_\ell}$ be a basis $T_\ell(A,B)$ and define

$$\begin{aligned}
v_1 &= -2v = \Lambda' B \Phi \\
v_2 &= \Lambda' X_1 \Phi \\
&\vdots \\
v_{m_\ell+1} &= \Lambda' X_m \Phi.
\end{aligned}$$

Thus,

$$\begin{aligned}
\dot{v}_1 &= \Lambda' [B,A] \Phi \\
\dot{v}_2 &= \Lambda' [X_1, A+vB] \Phi \\
&\vdots \\
\dot{v}_{m_\ell+1} &= \Lambda' [X_m, A+vB] \Phi
\end{aligned}$$

Now, each term of the form $[X_i, A]$, $[X_i, B]$ belongs to $T_\ell(A,B)$ and so we may write

$$[X_i, A] = \sum_{j=1}^{m_\ell} \alpha_{ij} X_j, \quad [X_i, B] = \sum_{j=1}^{m_\ell} \beta_{ij} X_j$$

for some constants α_{ij}, β_{ij}. Similarly, $[B,A] \varepsilon T_\ell(A,B)$ so

$$[B, A] = \sum_{j=1}^{m} b_j X_j ,$$

for some constants b_j. Substituting into (4.11) we have

$$\begin{aligned} \dot{v}_1 &= \sum_{j=1}^{m_\ell} b_j v_{j+1} \\ \dot{v}_i &= \sum_{j=1}^{m_\ell} \alpha_{ij} v_{j+1} + \frac{v_1}{2} \sum_{j=1}^{m_\ell} \beta_{ij} v_{j+1} \qquad 2 \leq i \leq m+1 \end{aligned} \tag{4.12}$$

Write

$$\mu = (v_1, v_2, \ldots, v_{m_\ell + 1})^T$$

Then (4.12) is of the form

$$\dot{\mu} = f(\mu)$$

for some function f. Suppose μ_o is a guess at $\mu(0)$. Then solving (4.12) (numerically) gives $\mu(t)$, $t \varepsilon [0,T]$, which can then be used to find x(t) from

$$\dot{\Phi} = A\Phi + \mu_1 B\Phi \quad , \; \Phi(0) = \Phi_o$$

(again numerically). Then the cost functional becomes

$$J(\mu_o) = \int_o^T \mu_1^2(t;\mu_o)dt + \Phi'(T;\mu_o)\Gamma\Phi\,(T;\mu_o)$$

which is a function of $m_\ell + 1$ variables and can be optimised numerically.

If the algebra $T_\ell(A,B)$ is nilpotent certain simplifications can be made. Recall that a Lie algebra L is nilpotent if

$$(\mathrm{Ad}\ L)^k = 0$$

for some $k>o$ where $(\mathrm{AdL})X = [L,X]$, $X \varepsilon L$.

<u>Lemma 4.2.</u> If $X \in (\mathrm{AdT}_\ell(A,B))^\ell B$ then $\frac{d}{dt}(\Lambda' X\Phi) = \Lambda' Y\Phi + v\Lambda' Z\Phi$ where $Y, Z \varepsilon (\mathrm{AdT}_\ell(A,B))^{\ell+1}B$. Hence if $T_\ell(A,B)$ is nilpotent and $(\mathrm{Ad}\ T_\ell(A,B))^k = 0$ then $\Lambda' X\Phi$ = constant for any

$$X \varepsilon (\mathrm{AdT}_\ell(A,B))^{k-1} .$$

<u>Proof.</u> This follows from lemma 4.1, since

$$\begin{aligned} \frac{d}{dt}(\Lambda' X\Phi) &= \Lambda' [X, A+vB]\Phi \\ &= -\Lambda'(\mathrm{AdA})X\Phi - v\Lambda'(\mathrm{AdB})X\,\Phi. \;\square \end{aligned}$$

It is then easy to see that equations (4.11) can be written in the form

$$\dot{v}_1 = b'\mu$$

$$\dot{\mu} = \Gamma\mu + \frac{v_1}{2}\Delta\mu,$$

where $b = (b_1, \ldots, b_m)'$ and Γ, Δ are nilpotent matrices.

For example, if $(\mathrm{Ad}\, T_\ell(A,B))^2 = 0$, then

$$2\dot{v} = \Lambda'[B,A]\Phi$$

and

$$2\ddot{v} = \Lambda'\left[[B,A], A+uB\right]\Phi$$

so $\ddot{v} = 0$ and the optimal control is of the form $v^* = c_1 + c_2 t$.

5. Example

To illustrate the theory developed above we consider the simple system

$$\dot{x} = x^2 u$$

$$J(u) = \int_0^T u^2 dt + x^2(T)$$

Then we have

$$\dot{x} = x^2 u$$

$$\dot{u} = v$$

and putting

$$\phi_{ij} = x^i u^j$$

it follows that

$$\begin{aligned} \dot{\phi}_{ij} &= ix^{i-1}u^j\dot{x} + jx^i u^{j-1}\dot{u} \\ &= ix^{i+1}u^{j+1} + jx^i u^{j-1} v \\ &= i\phi_{i+1j} + j\phi_{ij-1} v \end{aligned}$$

Hence,

$$A_{ij}^{k\ell} = i\delta_{i+1}^{k}\delta_{j+1}^{\ell}, \quad B_{ij}^{k\ell} = j\delta_i^k \delta_{j-1}^{\ell},$$

and so

$$A_{ij}^{k\ell} B_{kl}^{mn} = i(j+1)\delta_{i+1}^{m}\delta_j^n, \quad B_{ij}^{k\ell}A_{k\ell}^{mn} = ij\delta_{i+1}^{m}\delta_j^n$$

Thus,

$$C_{ij}^{k\ell} \triangleq [A,B]_{ij}^{k\ell} = i\delta_{i+1}^{k}\delta_j^{\ell}.$$

However, $[C,A] = [C,B] = 0$ and so the optimal control u^* is of the form

$$\dot{u}^* = c_1 t + c_2$$

i.e. $$u^* = c_1 t^2/_2 + c_2 t + c_3$$ for some constants c_1, c_2, c_3.

6. Conclusions

In this paper we have derived a general method for reducing the optimal control of certain nonlinear systems to simple numerical optimisation. The application of the method to the study of controllability, stabilizability etc. of nonlinear systems would also be desirable and this will be considered in future papers.

7. References

[1] S. P. Banks, On the generation of infinite-dimensional bilinear systems and Volterra series, Internat. J. Systems Sci., 16 (1985),145-160

[2] S. P. Banks, On the equivalent bilinearisation of nonlinear controlled delay systems, submitted.

[3] S. P. Banks and A. Ashtiani, Linear and bilinear infinite-dimensional representations of finite-dimensional nonlinear systems, to appear in Internat. J. Systems Sci.

[4] S. P. Banks and M. K. Yew, On a class of suboptimal controls for infinite-dimensional bilinear systems, Systems & Control Lett., 5 (1985) 327-333.

[5] W. Greub, Multilinear Algebra, 2nd Ed. (Springer, Berlin-New York, 1978).

[6] H. Takata. Transformation of a nonlinear system into an augmented linear system, IEEE Trans. autom. Control, 24, 736-741.