White Rose Research Online
eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk
Universities of Leeds, Sheffield and York

Deposited via The University of Sheffield.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/id/eprint/86206/

Monograph:
Owens, D.H. (1975) Descent Algorithm for the Optimization of Bilinear Systems. Research Report. ACSE Research Report 30 . Department of Automatic Control and Systems Engineering

Reuse
Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

UNIVERSITY OF LEEDS
University of Sheffield
UNIVERSITY of York

DESCENT ALGORITHM FOR THE OPTIMIZATION

OF BILINEAR SYSTEMS

by

D. H. Owens, B.Sc.,A.R.C.S.,Ph.D.

Lecturer in the Department of Control Engineering
University of Sheffield
Mappin Street
Sheffield S1 3JD

May 1975

Research Report No. 30

Abstract

A simple identity is given which forms the basis of a descent algorithm for the optimization of bilinear systems with quadratic performance index. The technique does not require linearization of the system equations.

The dynamics of xenon-135 transient power effects during power manoeuvres in a thermal nuclear power reactor are described by a set of bilinear ordinary differential equations[1]. Recent work[1] on the systematic control analysis of such systems has demonstrated that the problem of maximization of control rod margins can be formulated as an optimization problem with terminal quadratic performance index. This letter presents a generalization of the previous results[1,2] to a general bilinear system with a general form of quadratic performance index and outlines a numerical approach to the solution of the optimization problem.

Consider a system described by the state equations

$$\dot{x}(t) = A(t)x(t) + B(t)u(t) + C(x(t),u(t),t) + D(t) \qquad \ldots(1)$$

$$x(o) = x_o$$

where $x(t)\varepsilon R^n$, $u(t)\varepsilon R^\ell$, $A(t),D(t)$ and $B(t)$ are matrix functions of time which, for practical purposes are assumed to be piecewise continuous, and $C(x,u,t)$ is continuous in x,u,t and bilinear in x,u. The performance criterion for the system is

$$J(u) = \tfrac{1}{2}\langle x(T) - z,\ F(x(T) - z)\rangle$$

$$+ \tfrac{1}{2}\int_o^T\{\langle x(t) - r(t),\ Q(t)(x(t) - r(t))\rangle$$

$$+\tfrac{1}{2}\langle u(t) - v(t),\ R(t)(u(t) - v(t))\rangle\}dt \qquad \ldots(2)$$

where T is fixed, $z\varepsilon R^n$, $r(t)$ and $v(t)$ are desired state and control trajectories, F is a constant symmetric, positive semidefinite matrix and $Q(t)$, $R(t)$ are symmetric positive semi-definite and positive definite matrices respectively whose elements are piecewise continuous functions of time.

Let $u_o(t)$, $u_1(t)$ be admissible controllers generating state trajectories $x_o(t)$ and $x_1(t)$ respectively, then I claim that

$$J(u_1)-J(u_o) = \int_o^T\{H[x_1(t),u_1(t),p(t),t]-H[x_1(t),u_o(t),p(t),t]\}dt$$

$$+ \tfrac{1}{2}\langle x_1(T)-x_o(T),\ F(x_1(T)-x_o(T))\rangle$$

$$+ \tfrac{1}{2}\int_o^T\langle x_1(t)-x_o(t),\ Q(t)(x_1(t)-x_o(t))\rangle dt \qquad \ldots(3)$$

where the Hamiltonian function

$$H[x,p,u,t] = \tfrac{1}{2}\langle x-r(t),Q(t)(x-r(t))\rangle + \tfrac{1}{2}\langle u-v(t),R(t)(u-v(t))\rangle$$

$$+ \langle p,\ B(t)u + C(x,u,t)\rangle \qquad \ldots(4)$$

and the costate $p(t)$ is the solution of

$$\dot{p}(t) = -A^T(t)p(t) - \left.\frac{\partial C}{\partial x}\right|_{u_o(t)}^{T} p(t) - Q(t)\{x_o(t) - r(t)\} \qquad \ldots(5)$$

$$p(T) = F(x_o(T) - z) \qquad \ldots(6)$$

To prove the above statement, consider the identity

$$\int_o^T\{\langle\dot{p}(t),x_1(t)-x_o(t)\rangle+\langle p(t),\dot{x}_1(t)-\dot{x}_o(t)\rangle\}dt = \Big[\langle p(t),x_1(t)-x_o(t)\rangle\Big]_0^T \qquad \ldots(7)$$

From the state and costate boundary conditions the right-hand-side becomes, after some manipulation

$$\langle F(x_o(T)-z),x_1(T)-x_o(T)\rangle \equiv \tfrac{1}{2}\langle x_1(T)-z,\ F(x_1(T)-z)\rangle$$

$$- \tfrac{1}{2}\langle x_o(T)-z,F(x_o(T)-z)\rangle$$

$$- \tfrac{1}{2}\langle x_1(T)-x_o(T),F(x_1(T)-x_o(T))\rangle \qquad \ldots(8)$$

The left-hand-side of eqn (7), after substitution for the state and costate derivatives, becomes

$$\int_0^T \{<-A^T(t)p(t) - \frac{\partial C}{\partial x}\Big|_{u_o(t)}^T p(t) - Q(t)(x_o(t)-r(t)),\ x_1(t)-x_o(t)>$$

$$+ <p(t),\ A(t)(x_1(t)-x_o(t)) + B(t)\ (u_1(t)-u_o(t))$$

$$+ C(x_1(t),u_1(t),t) - C(x_o(t),u_o(t),t)>\}dt$$

$$= \int_0^T\{<p(t),B(t)(u_1(t)-u_o(t))+C(x_1(t),u_1(t),t)-C(x_o(t),u_o(t),t)-\frac{\partial C}{\partial x}\Big|_{u_o(t)}(x_1(t)-x_o(t)>$$

$$- <Q(t)(x_o(t)-r(t)),x_1(t)-x_o(t)>\}dt \qquad ...(9)$$

The result now follows by combining equations (7)-(9), noting that

$$C(x_1(t),u_1(t),t) - C(x_o(t),u_o(t),t) - \frac{\partial C}{\partial x}\Big|_{u_o(t)}(x_1(t)-x_o(t))$$

$$\equiv C(x_1(t),u_1(t),t) - C(x_1(t),u_o(t),t) \qquad ...(10)$$

applying a similar identity to equation (8) to $<Q(t)(x_o(t)-r(t)),x_1(t)-x_o(t)>$ and rearranging the resulting expression.

Relation (3) has direct application in numerical solution of the optimization problem. If $u(t)$ is constrained to lie in a restraint set $\Omega(t)$ and $u_1(t)$, $0\leqslant t\leqslant T$, is a solution of the algebraic minimization problem

$$\min_{u\varepsilon\Omega(t)} \{H[x_1(t),p(t),u,t]-H[x_1(t),p(t),u_o(t),t]+\tfrac{1}{2}\lambda<u-u_o(t),u-u_o(t)>\} \qquad ...(11)$$

then, for all large positive λ, we have $J(u_1)<J(u_o)$. An unusual point in equation (11) is that the descent direction is a function of the new state $x_1(t)$. This is in direct contrast to standard gradient type methods where, due to linearization, the descent direction depends only upon the previous iterate. This does not introduce any practical numerical

problems however, as can be seen from the following simple computational procedure, and, a priori, one expects good results as the approach uses information on the new state trajectory in the choice of the new controller.

STEP 1: Choose an initial admissible controller $u_o(t)$ and calculate the state trajectory $x_o(t)$.

STEP 2: Calculate the costate trajectory $p(t)$ from equations (5),(6).

STEP 3: Integrate the state equations using, at each step, the controller calculated from (11). As $u_1(t)$ depend upon the trajectory $x_1(t)$, $u_1(t)$ is calculated at time t in the simulation.

STEP 4: If $J(u_1)<J(u_o)$, set $u_o = u_1$ and go to step two. If $J(u_1)\geq J(u_o)$ increase λ and return to step three.

The suggested procedure has been applied[(1),(2)] with some success in the nuclear field where in some cases the performance criterion was reduced to 1/100th the value corresponding to the initial control guess in 6-8 iterations, for sets of 10-12 ordinary bilinear equations.

Identity (3) has a parallel, proved in a similar manner,

$$J(u_1)-J(u_o) = \int_0^T\{H[x_o(t),p(t),u_1(t),t]-H[x_o(t),p(t),u_o(t),t]\}dt$$

$$+ \tfrac{1}{2}<x_1(T)-x_o(T),\ F(x_1(T)-x_o(T))>$$

$$+ \tfrac{1}{2}\int_0^T<x_1(t)-x_o(t),Q(t)(x_1(t)-x_o(t))>dt \qquad ...(12)$$

where the costate $p(t)$ is the solution of the equations

$$\dot{p}(t) = -A^T(t)p(t) - \left.\frac{\partial C}{\partial x}\right|_{u_1(t)}^{T} p(t) - Q(t)\{x_o(t)-r(t)\} \qquad ...(13)$$

$$p(T) = F(x_o(T) - z) \qquad ...(14)$$

This identity leads to a similar algorithm but the new controller $u_1(t)$ must now be computed at each step of the integration of the costate equations as the solution of the problem,

$$\min_{u \in \Omega(t)} \{H[x_o(t),p(t),u,t]-H[x_o(t),p(t),u_o(t),t]+\tfrac{1}{2}\lambda\langle u-u_o(t),u-u_o(t)\rangle\} \quad \ldots(15)$$

The major problem in this case is that, if an iteration is unsuccessful, it is necessary to integrate the costate equations again to generate $u_1(t)$ for the increased value of λ. This feature tends to increase computational times.

References

(1) D. H. OWENS: 'Optimization of Xenon Control Margins during Power Manoeuvres', Atomic Energy Establishment Winfrith, Report AEEW-R 868, April 1973, (HMSO).

(2) D. H. OWENS: 'Multivariable Control Analysis of Distributed Parameter Nuclear Reactors', Ph.D. Thesis, Imperial College, London University, April 1973.

SHEFFIELD UNIV
APPLIED SCIENCE
LIBRARY