OPTIMAL CONTROL APPLICATIONS & METHODS, VOL. 16, 217–227 (1995)

ITERATIVE DYNAMIC PROGRAMMING FOR MINIMUM ENERGY CONTROL PROBLEMS WITH TIME DELAY*

S. A. DADEBO AND K. B. McAULEY†

Department of Chemical Engineering, Dupuis Hall, Queen's University, Kingston, Ont. K7L 3N6, Canada

SUMMARY

This paper presents the use of iterative dynamic programming employing exact penalty functions for minimum energy control problems. We show that exact continuously non-differentiable penalty functions are superior to continuously differentiable penalty functions in terms of satisfying final state constraints. We also demonstrate that the choice of an appropriate penalty function factor depends on the relative size of the time delay with respect to the final time and on the expected value of the energy consumption. A quadratic approximation (QA) of the delayed variables is much better than a linear approximation (LA) of the same for relatively large time delays. The QA improves the rate of convergence and avoids the formation of 'kinks'.

A more general way of selecting appropriate penalty function factors is given and the results obtained using four illustrative examples of varying complexity corroborate the efficacy of the method.

KEY WORDS iterative dynamic programming; absolute error penalty function; quadratic approximation; time delay systems; minimum energy control

1. INTRODUCTION

The solution of minimum energy control (MEC) problems based on the maximum principle is generally difficult because the transversality condition does not provide the freedom to state an initial condition for the adjoint variables. For systems with time delay the situation is aggravated because the set of differential equations involves both delayed (in state) and advanced (in adjoint) variables. Therefore several methods of solving MEC problems with time delay have been proposed.[1–3] Most of these methods require considerable mathematical effort and expertise and have been limited to linear systems.

We present the use of iterative dynamic programming (IDP) for minimum energy control of any general dynamic system with or without time delay. Luus and Rosen[4] have applied IDP to final state constrained optimal control problems but not to systems with time delay. Recently Dadebo and Luus[5] have shown that IDP can be applied to free endpoint time delay systems to give good results. The purpose of this paper is to employ exact penalty functions in handling final state constraints and to extend the algorithm proposed by Dadebo and Luus[5] to handle minimum energy control problems. However, for time delay systems we investigate the possibility of improving the convergence of the algorithm by using a quadratic approximation rather than a linear approximation of the delayed state profiles which often resulted in 'kinks'.[5]

*Part of this article was presented at the 1993 American Control Conference, San Francisco, CA.
†Author to whom correspondence should be addressed.

CCC 0143–2087/95/030217–11

Received 6 July 1993
Revised 25 July 1994

We investigate the effect of the relative size of the time delay and the energy consumption on the choice of an appropriate penalty function. We also address the problem of how to choose suitable weighting factors for these penalty functions. The size of a suitable penalty function factor in relation to energy consumption, the magnitude of the time delay and convergence are investigated using four examples of varying complexity.

2. MINIMUM ENERGY CONTROL PROBLEM FORMULATION

Consider the system described by the following vector differential equation with a constant time delay τ in one or more of the state variables:

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{f}(\mathbf{x}(t), \mathbf{x}(t-\tau), \mathbf{u}(t)) \tag{1}$$

The initial state profile for each of the delayed states is a known function of time,

$$\mathbf{x}(t) = \phi(t), \quad -\tau \leq t < 0 \tag{2}$$

and the initial condition is specified for each of the states,

$$\mathbf{x}(0) = \mathbf{x}_0 \tag{3}$$

$\mathbf{x}(t)$ is an $(n \times 1)$ state vector and $\mathbf{u}(t)$ is an $(m \times 1)$ control vector which may be bounded,

$$\alpha_j \leq u_j \leq \beta_j, \quad j = 1, 2, \ldots, m \tag{4}$$

For a minimum energy problem the associated performance index to be minimized is

$$I[\mathbf{x}(0), t_f] = \tfrac{1}{2}\int_0^{t_f} \mathbf{u}^{\mathrm{T}}(t)\mathbf{R}\mathbf{u}(t)\, dt \tag{5}$$

where $\mathbf{R}$ is a diagonal positive definite weighting matrix. The final time t_f is specified. The minimum energy control problem is to find the control policy $\mathbf{u}(t)$ in the given time interval $(0, t_f)$ such that the performance index I is minimized while satisfying the terminal constraint

$$\mathbf{x}(t_f) = \mathbf{x}_f \tag{6}$$

3. APPLICATION OF ITERATIVE DYNAMIC PROGRAMMING

Details of the IDP algorithm for MEC and derivation of the equations for the quadratic approximation of the delayed state profiles required for each grid point integration during the backward optimization step can be found in Reference 6. In order to satisfy the terminal constraints, we construct the augmented performance indices

$$J_1 = I + \sum_{j=1}^{\eta} \theta_j [x_j(t_f) - x_{f_j}]^2, \quad 1 \leq j \leq \eta \leq n \tag{7}$$

$$J_2 = I + \sum_{j=1}^{\eta} \omega_j |x_j(t_f) - x_{f_j}|, \quad 1 \leq j \leq \eta \leq n \tag{8}$$

where η is the number of terminally constrained state variables and θ_j and ω_j are corresponding penalty function factors or weighting factors. The acronyms SEPF (squared error penalty function) and AEPF (absolute error penalty function) corresponding to equations (7) and (8) respectively will be used throughout this paper. Luus and Rosen[4] suggested the use of a relatively small value of θ_j for non-delay systems. Since the relative benefits of the SEPF and AEPF approaches are not well understood, especially for time delay systems, we investigate the effects of the penalty function and weighting factor choices on the convergence properties of the solution. Our goal is to develop a reliable guideline in selecting appropriate weighting factors, since the use of the AEPF has rarely been seen in the optimal control literature. This alternative is easily managed with IDP, whereas other techniques, more often than not, use the quadratic penalty function for mathematical tractability.

The convergence of the IDP algorithm depends on the initial choice of the control region which is contracted after each iteration. If a contraction factor of about 70% is used, the control region will be reduced by at least a factor of 1000 after 20 iterations and thus convergence to the optimum will be attained.[7]

4. NUMERICAL EXAMPLES AND DISCUSSION

All integrations were done on a 386/33 personal computer using a modified Runge–Kutta subroutine of order 4/5 to handle delay terms in the differential equations.[8]

Example 1. SISO linear problems with stable and unstable open loop behaviour

We consider the linear time delay system optimized by Palanisamy and Rao[2] using single-term Walsh series and also by Inoue *et al.*[1] using sensitivity analysis. The system is described by the delay differential equation

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = ax(t) + bx(t-\tau) + u(t), \quad a, b \in \mathbb{R} \tag{9}$$

$$x(t) = 1{\cdot}0, \quad -1 \leqslant t \leqslant 0 \tag{10}$$

with the associated performance index to be minimized,

$$I = \tfrac{1}{2}\int_0^1 u^2(t)\,\mathrm{d}t \quad \text{such that } x(1) = 0{\cdot}0 \tag{11}$$

For this system we consider two cases, system A $(a = 1, b = 1)$ and system B $(a = -1, b = -1)$, which are open loop unstable and stable respectively. The results obtained are given in Table I using the AEPF and a quadratic approximation of the delayed state profile with $\gamma = 0{\cdot}70$, $r = 1{\cdot}0$

Table I. Effect of time delay on energy consumption for Example 1 using $N = 17$, $M = 3$, $P = 10$ and allowing 20 iterations

	Energy consumption I	
Time delay τ	System A $(a = 1, b = 1)$	System B $(a = -1, b = -1)$
0·00	2·04682	0·03737
0·05	2·18606	0·02856
0·10	2·28403	0·02099
0·15	2·35599	0·01461
0·20	2·43097	0·00937

and $\omega = 1\cdot0$. The final state constraint was satisfied in each case. Table II presents a comparison of the results obtained using IDP and other methods in the literature. Clearly IDP gives more reliable results. Table III indicates the effect of the penalty function on the performance index and the degree to which the final state constraints are met. Notice that $x(t_f)$ deviates significantly from zero even for very large values of θ using the SEPF, whereas the AEPF reliably meets the terminal constraint for a wide range of weighting factors. Figure 1 shows the plot of performance index I obtained after 20 iterations versus time delay for system B using linear and quadratic approximations. Clearly a quadratic interpolation of the delayed state becomes necessary for rapid convergence to the optimum when the relative size of the time delay is large.

Example 2. A linear SISO problem with large time delay and bounded input

Consider the linear time delay system optimized by Liou and Chou[3] and Chyung and Lee.[9] The system is described by

$$\frac{dx(t)}{dt} = -x(t-1) + u(t) \tag{12}$$

$$x(t) = 1, \quad t \in [-1, 0] \tag{13}$$

The problem is to find an optimal control policy $u(t)$ on $[0, 2]$ driving the response $x(t)$ to the origin while minimizing the performance index

$$I = \frac{1}{2}\int_0^2 u^2(t)\,dt \quad \text{such that } x(2) = 0 \tag{14}$$

Here the control is bounded $(u(t) \geqslant 0)$ and measurable on $[0, 2]$. Chyung and Lee[9] obtained the exact solution of the system analytically and reported a performance index of $0\cdot09375$. We obtained a minimum value of $0\cdot10611$ in 20 iterations using $\omega = 10$, $P = 8$, $M = 3$, $N = 17$ and a guess of $u^0 = 0\cdot0$ with $r = 0\cdot12$. The terminal error was $-0\cdot00012$. The convergence, however, was quite slow. For relatively small values of τ the convergence is

Table II. Comparison of present results (boldface) with other methods reported in the literature for Example 1

System	Time delay τ	Method*	Terminal error $x(t_f) - x_f$	Performance index
A	0·0	CS	0·2933	2·0373
$(a = b = 1)$	0·0	SS	0·0078	5·4181
	0·0	STWS	0·0000	2·2381
	0·0	**IDP**	**0·0000**	**2·0468**
B	0·0	SS	−0·0356	0·0373
$(a = b = -1)$	**0·0**	**IDP**	**0·0000**	**0·03737**
	0·1	SS	0·0122	1·4181
	0·1	STWS	0·0000	0·0209
	0·1	**IDP**	**0·00000**	**0·02099**

*CS, conventional synthesis (as reported by Inoue *et al.*[1]); SS, sensitivity synthesis (as reported by Inoue *et al.*[1]); STWS, single-term Walsh series (as reported by Palanisamy and Rao[2]).

Table III. Effect of penalty function factor on I for Example 1, system A ($\tau = 0{\cdot}1$)

Penalty function factor θ, ω	Absolute error penalty function (AEPF)		Squared error penalty function (SEPF)	
	I	$x(t_f)$	I	$x(t_f)$
1·0	2·283936	0·00006498	2·006491	0·340480100
10	2·747395	−0·00000316	2·234977	0·035988570
100	2·934113	0·00000220	2·345831	0·003605908
1000	2·830834	0·00036539	2·868448	0·000457411

very rapid with P as low as 8. However, for $\tau = 1{\cdot}0$, which is one-half of the final time of 2·0, it was necessary to try different values of ω before an appropriate penalty function was obtained. The values of M and N were increased to 5 and 21 respectively. The number of stages chosen was four initially (corresponding to $L = 0{\cdot}5 = 0{\cdot}5\tau$) and doubled to eight after 20 iterations. The variation in the performance index I with the penalty function factor ω is shown in Figure 2. The best value of the performance index obtained was 0·094963 with $x(2) = -0{\cdot}00032774$ using $\omega = 0{\cdot}5$. By increasing the number of grid points to 41, we obtained $I = 0{\cdot}093538$ and $x(2) = -0{\cdot}007464$. The quadratic approximation gives reliable results over a wider range of ω. However, the appropriate choice of ω for relatively large time delays is quite critical. It may be necessary to use a higher-order approximation for such systems with large time delays.

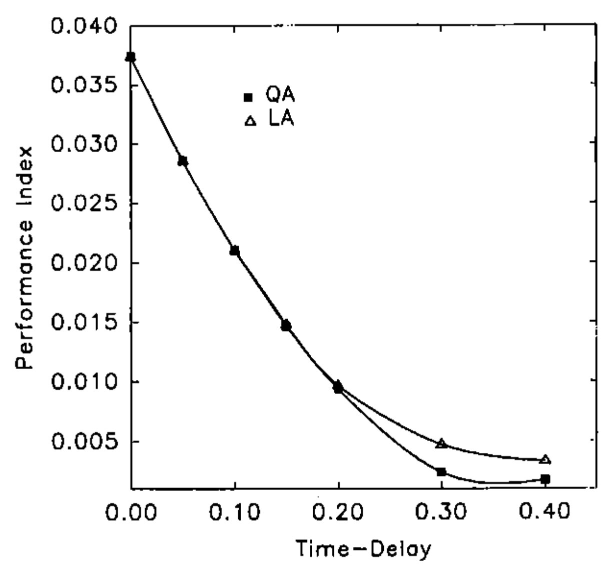

Figure 1. Effect of time delay on I^0 for Example 1

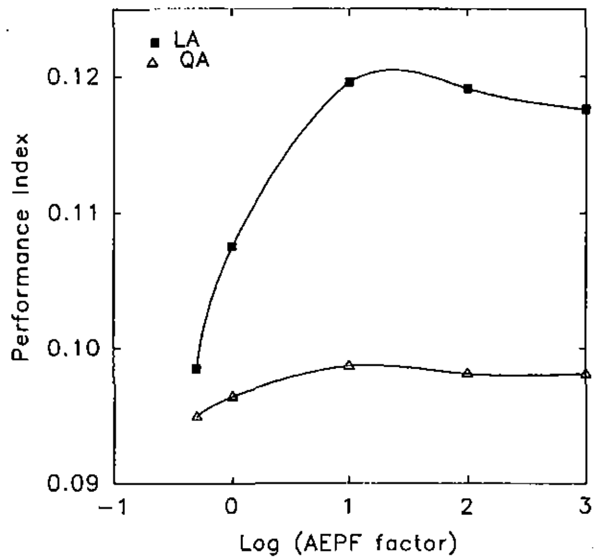

Figure 2. Effect of ω on I^o for Example 2

Example 3. A multivariable non-linear problem

Example 3 is a non-linear problem given by Malek-Zavarei and Jamshidi.[10] The system is described by

$$\frac{dx_1(t)}{dt} = -2x_1(t)x_2(t) + x_2(t-0{\cdot}1) + u_1(t) \tag{15}$$

$$\frac{dx_2(t)}{dt} = -x_2(t) - 2x_1(t)x_2(t-0{\cdot}1) + u_2(t) \tag{16}$$

with the initial state profile

$$x_1(t) = x_2(t) = 1{\cdot}0, \quad -0{\cdot}1 \leq t \leq 0{\cdot}0 \tag{17}$$

and the associated performance index to be minimized is

$$I = \frac{1}{2}\int_0^1 [u_1^2(t) + u_2^2(t)]\, dt \tag{18}$$

We shall impose the constraint that

$$x_1(1) = x_2(1) = 0{\cdot}0 \tag{19}$$

We obtained $I = 0{\cdot}417318$ with $x_1(1) = 0{\cdot}062446$ and $x_2(1) = 0{\cdot}000910$ using $P = 12$, $M = 5$, $N = 21$, $\gamma = 0{\cdot}80$, $\omega_1 = 1$ and $\omega_2 = 1$. Although $x_2(1)$ is acceptable, $x_1(1)$ is too large, so we increased the weighting on $x_1(1)$ to $\omega_1 = 10$. The associated control policy and the state trajectories obtained are shown in Figures 3 and 4 respectively. The performance index more than doubled ($I = 0{\cdot}838197$ with $x_1(1) = 0{\cdot}000027$ and $x_2(1) = 0{\cdot}000569$) to achieve the terminal constraint. Therefore, if the final state constraint on x_1 can be relaxed, a considerable energy saving can be realized and $\omega_1 = \omega_2 = 1$ is an appropriate choice. When $\theta_1 = \theta_2 = 100$ was used, the best choice in the SEPF case, the performance index obtained was $1{\cdot}00359$ after 20

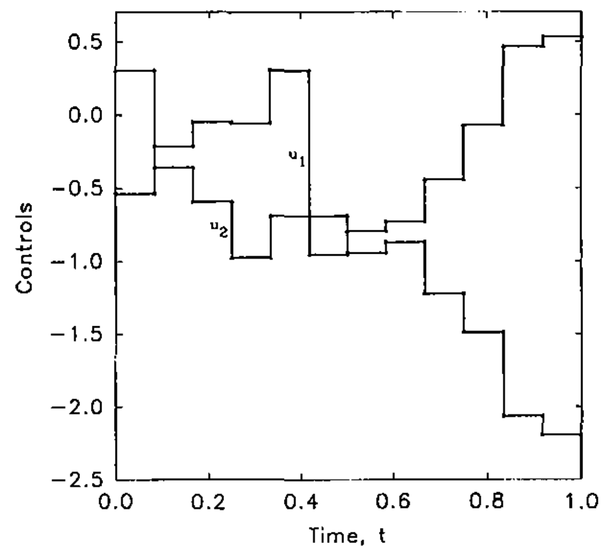

Figure 3. Control policy for Example 3 with 12 stages

iterations (with $x_1(1)=0{\cdot}000154$ and $x_2(1)=0{\cdot}005339$), illustrating that the choice of both the penalty function and weighting parameters affects the convergence of the IDP scheme.

Example 4. A nuclear reactor problem

Consider the nuclear reactor system with a single group of delayed neutrons described by Sage and White[11] and subsequently by Van Dooren[12] and Luus and Rosen:[4]

$$\frac{\mathrm{d}x_1(t)}{\mathrm{d}t}=1000[u(t)-0{\cdot}0064]x_1(t)+0{\cdot}1x_2(t) \tag{20}$$

$$\frac{\mathrm{d}x_2(t)}{\mathrm{d}t}=6{\cdot}4x_1(t)+0{\cdot}1x_2(t) \tag{21}$$

with $x_1(0)=0{\cdot}5$, $x_2(0)=32$ and a desired final state constraint $x_1(1)=5{\cdot}0$. Here $x_1(t)$ and $x_2(t)$ refer to the neutron flux density and the precursor concentration respectively. The problem is to find the reactivity (control) $u(t)|u(t)\geqslant 0$ which will drive the neutron flux density from its initial state to the desired final state $x_1(1)=5{\cdot}0$ with minimum control effort:

$$I=\tfrac{1}{2}\int_0^1 u^2(t)\,\mathrm{d}t \tag{22}$$

This problem has been studied quite extensively by Luus and Rosen using IDP (with the SEPF), who seemed to have had some difficulty in selecting a suitable value of θ and resorted to trying several values. They concluded that if θ is reasonably small, then the solution converges to the optimum readily. Using the AEPF, we note that if ω is chosen to be of the same magnitude as the expected performance index, then convergence to the global optimum occurs. Using $P=10$, $N=17$, $M=3$, $\omega=1\times10^{-5}$ and allowing 25 iterations, we obtained

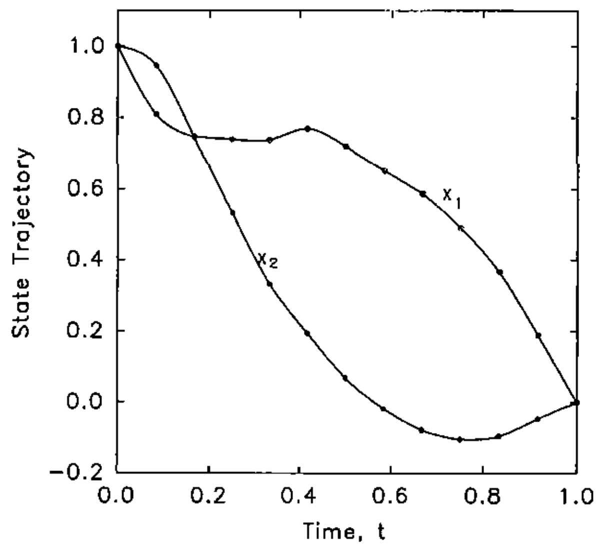

Figure 4. Evolution of states for Example 3 using $\omega_1 = 10$ and $\omega_2 = 1$

$I = 1{\cdot}787 \times 10^{-5}$ with $x_1(1) = 5{\cdot}00000$. This result is important because Luus and Rosen could not get that close to the optimum using an SEPF with $M = 3$. Using $P = 20$, $M = 21$, $N = 21$ and $\omega = 1 \times 10^{-5}$, we obtained $I = 1{\cdot}769 \times 10^{-5}$ with $x_1(1) = 5{\cdot}00000$, which compares very favourably with Van Dooren's value of $1{\cdot}767 \times 10^{-5}$. There was no difficulty in converging to the optimum using the AEPF approach and the final state constraint $x_1(1)$ was exactly 5·0 as required. Figures 5 and 6 show the evolution of the states and the corresponding control policy respectively.

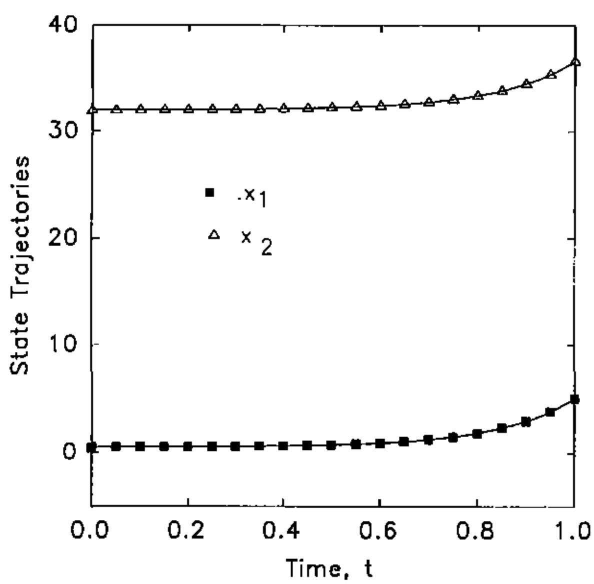

Figure 5. Evolution of states for Example 4 using $P = 20$

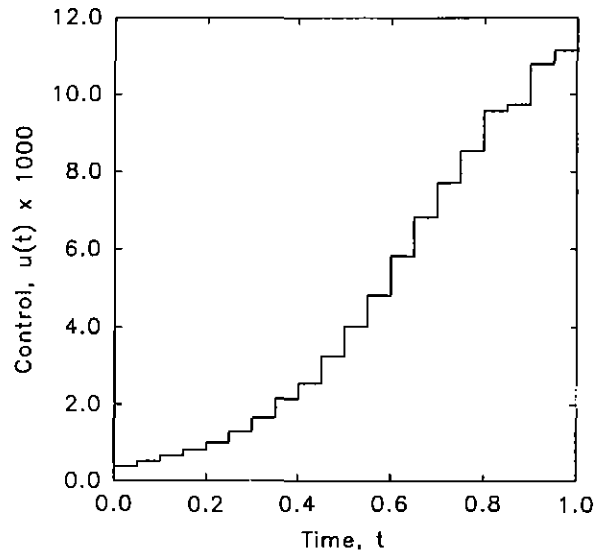

Figure 6. Piecewise constant control policy for Example 4 using $P = 20$

5. CONCLUSIONS

Iterative dynamic programming provides a reliable method for solving the minimum energy control problem without having to solve a two-point boundary value problem. For relatively small time delays (less than 20% of the final time) a linear approximation is adequate. However, for larger time delays a quadratic approximation leads to faster convergence towards the optimum.

Final state constraints can be enforced implicitly by augmenting the minimum energy performance index with either a squared error penalty function (SEPF) or an absolute error penalty function (AEPF). The superiority of exact penalty functions has been confirmed and the AEPF approach gives reliable results with rapid convergence over a wider range of penalty function factors than does the SEPF. The most appropriate choice of the weighting factor ω when using the AEPF is similar in magnitude to the expected minimum energy. This choice of ω is generally smaller than the corresponding θ when an SEPF is desired. In situations where the final constraint requirement far outweighs the energy consumption, the AEPF approach is highly recommended. On the other hand, depending on the problem, some trade-off between quality or final state specification and energy consumption may be necessary especially when one or more of the terminal constraints can be relaxed. An advantage of the IDP approach is that the procedure can easily handle both linear and non-linear time-varying systems.

ACKNOWLEDGEMENTS

Financial support from NSERC and the School of Graduate Studies and Research, Queen's University is gratefully acknowledged. The authors wish to express their gratitude to Professor Rein Luus, Department of Chemical Engineering and Applied Chemistry, University of Toronto, for his helpful suggestions.

APPENDIX: NOMENCLATURE

$\mathbf{f}$	vector of n functions of $\mathbf{x}(t)$, $\mathbf{x}(t-\tau)$ and $\mathbf{u}(t)$
I	performance index to be minimized
J	augmented performance index
k	index used to denote time stage
L	length of time stage
m	dimension of the control vector
M	odd number of allowable values of each control variable at each time stage
n	dimension of state vector
N	odd number of grid points for $\mathbf{x}$
P	number of time stages chosen
$\mathbf{r}$	region allowed for control
$\mathbf{R}$	diagonal weighting matrix
t	time
$\mathbf{u}$	$(m \times 1)$ control vector
$\mathbf{x}$	$(n \times 1)$ state vector
$\mathbf{x}_0$	initial condition of state vector

Greek letters

α	lower bound on control
β	upper bound on control
γ	amount by which control region is contracted after each iteration
η	number of final state constrained variables
θ	squared error penalty function factor
τ	time delay
ϕ	initial state profile
ω	absolute error penalty function factor

Subscripts

f	final
i	index for entry in vector
k	index for time at stage k
0	initial

Superscripts

(j)	iteration index
o	optimal
T	transpose

Acronyms

AEPF	absolute error penalty function
CS	conventional synthesis
IDP	iterative dynamic programming
LA	linear approximation

QA	quadratic approximation
SEPF	squared error penalty function
SISO	single-input/single-output
SS	sensitivity synthesis
STWS	single-term Walsh series

REFERENCES

1. Inoue, K., H. Akashi and Y. Sawaragi, 'Sensitivity approaches to optimization of linear systems with time delay', *Automatica*, **7**, 671–679 (1971).
2. Palanisamy, K. R. and G. P. Rao, 'Minimum energy control of time-delay systems via Walsh functions', *Optim. control appl. methods*, **4**, 213–226 (1983).
3. Liou, C.-T. and Y.-S. Chou, 'Minimum energy control of time delay systems via piecewise linear polynomial functions', *Int. J. Control*, **46**, 1595–1603 (1987).
4. Luus, R. and O. Rosen, 'Application of dynamic programming to final state constrained problems', *Ind. Eng. Chem. Res.*, **30**, 1525–1530 (1991).
5. Dadebo, S. and R. Luus, 'Optimal control of time delay systems by dynamic programming', *Optim. control appl. methods*, **13**, 29–41 (1992).
6. Dadebo, S. A. and K. B. McAuley, 'Minimum energy control of time-delay systems via iterative dynamic programming', *Proc. 1993 Am. Control Conf.*, San Francisco, CA, Vol. 2, pp. 1261–1265 (1993).
7. Dadebo, S. A. and K. B. McAuley, 'Dynamic optimization of constrained chemical engineering problems using dynamic programming', *Comput. Chem. Eng.*, **19**, (5) 513–525 (1995).
8. Hairer, E., S. P. Norsett and G. Wanner, *Solving Ordinary Differential Equations I*, Springer, Berlin, 1987.
9. Chyung, D. H. and E. B. Lee, 'Linear optimal systems with time-delay', *SIAM J. Control*, **4**, 548–575 (1966).
10. Malek-Zavarei, M. and M. Jamshidi, *Time-Delay Systems; Analysis, Optimization and Applications*, North-Holland, Amsterdam, 1987.
11. Sage, A. P. and C. C. White, *Optimum Systems Control*, Prentice-Hall, Englewood Cliffs, NJ, 1977, pp. 91–317.
12. Van Dooren, R., 'A Chebyshev technique applied to a controlled nuclear reactor system', *Optim. control appl. methods*, **10**, 285–291 (1989).