Monograph:

Owens, D.H. and Jones, R.P (1977) Quadratic Optimization with Bounded Linear Constraints. Research Report. ACSE Research Report 66 . Department of Automatic Control and Systems Engineering

QUADRATIC OPTIMIZATION WITH BOUNDED LINEAR CONSTRAINTS

by

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

R.P. Jones* B.Sc., M.Sc.(Tech), Grad.I.M.A.

Department of Control Engineering,
University of Sheffield,
Mappin Street,
Sheffield S1 3JD.

December 1977

Research Report No. 66

* Control theory centre, University of Warwick.

Abstract

A recently developed algorithm for the solution of linear constrained differential-algebraic systems is extended to provide a systematic computational procedure for the solution of quadratic optimization problems with bounded linear constraints. The technique has rapid convergence and can be applied to minimum norm solution of algebraic equations, minimum energy control problems and linear quadratic optimal control problems with linear control and state constraints.

1. Introduction

Quadratic optimization problems with bounded linear constraints have played an important role in the understanding of control systems behaviour. The standard problem takes the form of the computation of x_∞ solving

$$||x_\infty||^2 = \inf \{ ||x||^2 \ : Lx = b \} \ , \quad Lx_\infty = b \qquad (1)$$

where x is regarded as a point in a real Hilbert space H (with inner products $<\cdot \ , \cdot>$ and induced norm $||\cdot|| = <\cdot \ , \cdot>^{\frac{1}{2}}$) and L is a bounded linear map from H into a real Hilbert space H_1. The choice of well-defined and rapidly converging algorithms for the solution of this problem has been a major topic of research (Luenberger, 1969). This paper extends the application of a recently developed algorithm (Owens and Jones, 1978) for the calculation of feasible solutions of constrained differential-algebraic systems to the solution of the control problem

$$||x_\infty||^2 = \inf \ \{||x||^2 \ : L_i x = b_i \ , \quad 1 \leqslant i \leqslant m\}$$
$$L_i x_\infty = b_i \ , \qquad 1 \leqslant i \leqslant m \qquad (m \geqslant 2) \qquad (2)$$

where L_i , $1 \leqslant i \leqslant m$, are bounded linear operators mapping H into real Hilbert spaces H_i , $1 \leqslant i \leqslant m$, respectively. Problem (2) can be regarded as a decomposition of (1) by splitting the constraint $Lx = b$ into m simpler constraints $L_i x = b_i$, $1 \leqslant i \leqslant m$, or as an extended version of (1) with $L_1 = L$, $b_1 = b$ and the inclusion of the additional constraints. The techniques proposed are capable of rapid convergence and has application to both finite and infinite-dimensional control problems.

2. Problem Formulation

With the notation of section 1, define the closed and convex linear varieties

$$D_i = \{x \in H \ : \ L_i x = b_i\} \ , \quad 1 \leqslant i \leqslant m \qquad (3)$$

so that (2) takes the form,

$$||x_\infty||^2 = \inf \{||x||^2 : x \in \bigcap_{i=1}^{m} D_i\} \quad , \; x_\infty \in \bigcap_{i=1}^{m} D_i \qquad (4)$$

which is simply a minimum norm problem on the closed, convex set $D_1 \cap D_2 \cap \ldots \cap D_m$. The solution exists and is unique (Luenberger, 1969) and is characterised by the stationary points of a Lagrangian functional. In practice this formulation can present computational problems and an iterative technique is to be preferred.

The following easily proven lemma is fundamental to the discussion:

Lemma

$x_\infty \in D_1 \cap D_2 \cap \ldots \cap D_m$ solves (2) if, and only if, it is orthogonal to the subspace $N(L_1) \cap N(L_2) \cap \ldots \cap N(L_m)$. Moreover x_∞ (if it exists) is unique.

With this in mind the following result provides an iterative solution of (2) as the weak limit of a sequence of optimization problems.

Theorem

Suppose that $D_1 \cap D_2 \cap \ldots \cap D_m$ is nonempty and define the sequence of linear varieties

$$K_j = D_{(j-1)\bmod(m-1) + 2} \qquad , \quad j \geq 1 \qquad (5)$$

Then the sequences $T_i \triangleq \{r_1^{(i)}, r_2^{(i)}, \ldots\}$, $i = 1,2$, defined by the relations

$$r_j^{(1)} \in D_1 \quad , \; r_j^{(2)} \in K_j \qquad , \quad j \geq 1$$

$$||r_1^{(1)}||^2 = \min \{||r||^2 \; : \; r \in D_1\}$$

$$||r_\ell^{(2)} - r_\ell^{(1)}||^2 = \min \{ ||r - r_\ell^{(1)}||^2 : r \in K_\ell\} \quad , \; \ell \geq 1$$

$$||s_\ell - r_\ell^{(2)}||^2 = \min \{ ||s - r_\ell^{(2)}||^2 : s \in D_1\} \quad , \; \ell \geq 1$$

$$||s_\ell - r_\ell^{(1)}||^2 \lambda_\ell^* = ||r_\ell^{(2)} - r_\ell^{(1)}||^2 \quad , \; \lambda_\ell^* \geq 1 \quad , \; \ell \geq 1$$

$$1 \leqslant \lambda_\ell \leqslant \lambda_\ell^* \qquad , \ \ell \geqslant 1$$

$$r_{\ell+1}^{(1)} = r_\ell^{(1)} + \lambda_\ell \ \{s_\ell - r_\ell^{(1)}\} \quad , \ \ell \geqslant 1 \tag{6}$$

are well-defined and converge weakly to the unique point $r_\infty \in D_1 \cap \ldots \cap D_m$ solving problem (2). Moreover, if L_i is compact, $\{L_i r_j^{(1)}\}_{j \geqslant 1}$ converges strongly to b_i and, if H is finite-dimensional, both T_1 and T_2 converge strongly to r_∞.

Finally, for all $j \geqslant 1$, $r_j^{(1)} \in D_1$ solves the minimum norm problem

$$||x_\infty||^2 = \min \ \{||x||^2 : x \in D_1 \ , \ L_i x = L_i r_j^{(1)} \ , \ 2 \leqslant i \leqslant m\} \tag{7}$$

Proof

Noting that D_i, $1 \leqslant i \leqslant m$, are closed linear varieties then the seuqences T_1, T_2 are well defined if $\lambda_\infty^* \geqslant 1$ for $\ell \geqslant 1$. If $||s_\ell - r_\ell^{(1)}|| = 0$ for some $\ell \geqslant 1$, then the assumption $D_1 \cap D_2 \cap \ldots \cap D_m \neq \emptyset$ ensures that $||r_\ell^{(2)} - r_\ell^{(1)}|| = 0$ and we may choose $\lambda_\ell^* = 1$. Hence, without loss of generality assume that $||s_\ell - r_\ell^{(1)}|| > 0$ for all $\ell \geqslant 1$. Using standard results (Luenberger, 1969).

$$< r_1^{(1)} , x > = 0 \qquad \forall \ x \in N(L_1) \tag{8}$$

and, for $i \geqslant 1$,

$$< r_i^{(2)} - r_i^{(1)} , x - r_i^{(2)} > = 0 \qquad \forall \ x \in K_i$$

$$< s_i - r_i^{(2)} , x - s_i > = 0 \qquad \forall \ x \in D_1 \tag{9}$$

It follows directly that

$$||r_i^{(2)} - r_i^{(1)}||^2 = ||r_i^{(2)} - s_i||^2 + ||s_i - r_i^{(1)}||^2 \ , \ i \geqslant 1 \tag{10}$$

from which $\lambda_i^* \geqslant 1$.

Choosing $x \in D_1 \cap \ldots \cap D_m \subset D_1 \cap K_i$ we obtain

$$||x - r_i^{(1)}||^2 = ||x - r_{i+1}^{(1)}||^2 + ||r_{i+1}^{(1)} - r_i^{(1)}||^2 + 2 < x - r_{i+1}^{(1)}, r_{i+1}^{(1)} - r_i^{(1)} > \tag{11}$$

and, using (9),

$$< x - r_{i+1}^{(1)} \ , \ r_{i+1}^{(1)} - r_i^{(1)} > = \lambda_i < x - r_i^{(1)} - \lambda_i (s_i - r_i^{(1)}), s_i - r_i^{(1)} >$$

$$= \lambda_i \{ <x-r_i^{(1)}, s_i-r_i^{(1)}> - \lambda_i ||s_i-r_i^{(1)}||^2\}$$

$$= \lambda_i\{<x-s_i+s_i-r_i^{(1)}, s_i-r_i^{(2)}+r_i^{(2)}-r_i^{(1)}>-\lambda_i||s_i-r_i^{(1)}||^2\}$$

$$= \lambda_i\{<x-r_i^{(1)}, r_i^{(2)}-r_i^{(1)}> - \lambda_i ||s_i-r_i^{(1)}||^2\}$$

$$= \lambda_i\{<x-r_i^{(2)}+r_i^{(2)}-r_i^{(1)}, r_i^{(2)}-r_i^{(1)}> - \lambda_i||s_i-r_i^{(1)}||^2\}$$

$$= \lambda_i\{ ||r_i^{(2)} - r_i^{(1)}||^2 - \lambda_i ||s_i - r_i^{(1)}||^2 \} \geq 0 \tag{12}$$

equality holding if, and only if, $\lambda_i = \lambda_i^*$. Applying (12) to (11) yields

$$||x - r_i^{(1)}||^2 \geq ||x - r_{i+1}^{(1)}||^2 + ||r_{i+1}^{(1)} - r_i^{(1)}||^2 \tag{13}$$

so that T_1 is a bounded sequence in H and

$$||x - r_1^{(1)}||^2 \geq \sum_{i=1}^{\infty} ||r_{i+1}^{(1)} - r_i^{(1)}||^2 \tag{14}$$

Noting that

$$||r_{i+1}^{(1)} - r_i^{(1)}||^2 + < x - r_{i+1}^{(1)}, r_{i+1}^{(1)} - r_i^{(1)} >$$

$$= < x - r_i^{(1)}, r_{i+1}^{(1)} - r_i^{(1)} >$$

$$= \lambda_i < x - r_i^{(1)}, s_i - r_i^{(1)} >$$

$$= \lambda_i < x - s_i + s_i - r_i^{(1)}, s_i - r_i^{(2)} + r_i^{(2)} - r_i^{(1)} >$$

$$= \lambda_i < x - r_i^{(1)}, r_i^{(2)} - r_i^{(1)} >$$

$$= \lambda_i < x - r_i^{(2)} + r_i^{(2)} - r_i^{(1)}, r_i^{(2)} - r_i^{(1)} >$$

$$= \lambda_i ||r_i^{(2)} - r_i^{(1)}||^2$$

$$\geq ||r_i^{(2)} - r_i^{(1)}||^2 \tag{15}$$

then, combining with (11) and (12) yields

$$||x - r_1^{(1)}||^2 \geq \sum_{i=1}^{\infty} ||r_i^{(2)} - r_i^{(1)}||^2 \tag{16}$$

Let R be the closure of span $\{r_i^{(1)}\}_{i\geq 1}$ in H. As T_1 is bounded in R, it possesses a weakly convergent subsequence $\{r_{i_k}^{(1)}\}_{k\geq 1}$ with a weak limit $r_\infty \in R$. Moreover, (14) indicates that $\{r_{i_k+\ell}^{(1)}\}_{k\geq 0}$ converges weakly to r_∞ for all $\ell \geq 0$. Noting that (equation (16)) $||r_i^{(2)} - r_i^{(1)}|| \to 0$ $(i \to \infty)$, it follows that $\{r_{i_k+\ell}^{(2)}\}_{k\geq 1}$ converges weakly to r_∞. Examination of $< g , r_{i_k+\ell}^{(1)} - r_\infty >$ with $g = L_1^*(b_1 - L_1 r_\infty)$ (whose L_1^* is the adjoint of L_1) yields $L_1 r_\infty = b_1$. In a similar manner, examination of $<g , r_{i_k+\ell}^{(2)} - r_\infty>$ with $g = L_i^*(b_i - L_i r_\infty)$ indicates that $L_i r_\infty = b_i$, $2 \leq i \leq m$, from which $r_\infty \in D_1 \cap D_2 \cap \ldots \cap D_m$.

From (12) with $\lambda_i = \lambda_i^*$ we obtain

$$<x - r_{i+1}^{(1)} , r_{i+1}^{(1)} - r_i^{(1)} > = 0 , \forall x \in \bigcap_{j=1}^{m} D_j , i \geq 1 \tag{17}$$

so that

$$<x , s_i - r_i^{(1)}> = 0 , \forall x \in \bigcap_{j=1}^{m} N(L_j) , i \geq 1 \tag{18}$$

In particular, independent of the choice of λ_i in the specified range

$$<x , r_{i+1}^{(1)} - r_i^{(1)}> = 0 \quad \forall x \in \bigcap_{j=1}^{m} N(L_j) , \quad i \geq 1 \tag{19}$$

Using (8), it follows by induction that R is contained in the orthogonal complement of $N(L_1) \cap \ldots \cap N(L_m)$ in H. Applying the lemma, r_∞ solves problem (2) and is unique. In particular, every weakly convergent subsequence of T_1 converges weakly to r_∞ i.e. T_1 converges weakly to r_∞.

By definition $r_j^{(1)} \in R$, for all $j \geq 1$, and hence solves problem (7) by the lemma. The theorem is now proven by noting that compact operators map weakly convergent sequences into strongly convergent sequences and that weak convergence and strong convergence coincide in finite dimensional spaces.

Q.E.D.

In practical terms the result generates a sequence $T_1 = \{r_j^{(1)}\}_{j\geq 1}$ in D_1 converging weakly to the unqiue solution of (2). It has the interesting property that, for each $j \geq 1$, the point $r_j^{(1)}$ is the unique solution of the approximate optimization problem (7). In particular, if $L_2, L_3, \ldots L_m$ are compact, $r_j(1)$ solves problem (7) and, for sufficiently large $j \geq 1$, $L_i r_j^{(1)}$ lies in an arbitrarily small neighbourhood of b_i (in the norm topology of H_i), $2 \leq i \leq m$. In this sense, $r_j^{(1)}$, $j \geq 1$, can be regarded as the unique solution of the approximation (7) to problem (2).

Finally, the flexibility in the choice of λ_i, $i \geq 1$, inherent in (6) will be of great practical significance in the sense that λ_i can be regarded as an accelerating extrapolation factor for the algorithm (Owens and Jones, 1978) if $\lambda_i \simeq \lambda_i^* \gg 1$. Alternatively, the choice of $1 \leq \lambda_i << \lambda_i^*$ will prevent the growth of numerical errors if they are a problem.

3. Illustrative Applications

The abstract formulation of the algorithm described in section 2 leaves open the possibility of its application to a wide variety of finite and infinite-dimensional optimization problems. For simplicity, attention is restricted to the case m = 2 when (equation 5) $K_j = D_2$ for all $j \geq 1$.

3.1 Finite Dimensional Optimization

Consider the solution of the optimization problem in $H = R^n$ (with suitable inner product and norm)

$$||x_\infty||^2 = \inf\{\, ||x||^2 : Lx = b \,\} \tag{20}$$

where $L : R^n \to R^{n_1+n_2}$, $n_1 + n_2 \leq n$. By decomposing L into block matrices Lx = b can be written in the form

$$\begin{pmatrix} L_1 \\ L_2 \end{pmatrix} x = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \tag{21}$$

where $L_i : R^n \to R^{n_i}$, i = 1,2. Application of the algorithm (6) yields the sequence,

$$r_1^{(1)} = L_1^+ b_1$$

$$r_\ell^{(2)} = r_\ell^{(1)} + L_2^+ \{ b_2 - L_2 r_\ell^{(1)} \}$$

$$s_\ell = r_\ell^{(2)} + L_1^+ \{ b_1 - L_1 r_\ell^{(2)} \}$$

$$r_{\ell+1}^{(1)} = r_\ell^{(1)} + \lambda_\ell \{ s_\ell - r_\ell^{(1)} \}$$

$$1 \leqslant \lambda_\ell \leqslant \frac{||r_\ell^{(2)} - r_\ell^{(1)}||^2}{||s_\ell - r_\ell^{(1)}||^2} \quad , \ell \geqslant 1 \qquad (22)$$

where L_i^+ is the pseudo-inverse of L_i , $i = 1,2$. Moreover , $\{r_i^{(1)}\}_{i \geqslant 1}$ tends strongly to the unique solution $r_\infty \in D_1 \cap D_2$ solving (20).

3.2 Quadratic Optimal Control with State and Control Constraints

Consider the iterative solution of the linear quadratic optimal control problem (T fixed),

$$\dot{x}(t) = Ax(t) + Bu(t) \ , \ y(t) = Cx(t) \ , \ x(o) = x_o$$

$$J = \tfrac{1}{2} \int_0^T \{ < y(t) \ , \ y(t) > + < u(t), u(t) > \} dt \qquad (23)$$

subject to equality constraints either of the form, (d(t) known)

$$Eu(t) + Fy(t) = d(t) \qquad , \quad o \leqslant t \leqslant T \qquad (24)$$

or of the form (d fixed in R^p)

$$\int_0^T \{ Eu(t) + Fy(t) \} \ dt = d \qquad (25)$$

where A,B, are nxn, nxℓ (possibly time-varying) matrices respectively, C is an mxn matrix and E,F are pxℓ, pxm matrices respectively. Consider, for example, the problem generated by (23) and (25) in the Hilbert space $H = L_2^m [0,T] \times L_2^\ell [0,T]$ with norm generated by the performance criterion. In this case, assuming that A,B are constant matrices for simplicity, L_1 is the map

$$L_1 : (y,u) \to y(t) - \int_o^t C \, e^{A(t-s)} \, B \, u(s)ds \qquad (26)$$

from H into $H_1 \triangleq L_2^m [0,T]$ and $b_1 = C \, e^{At} x(o)$. The map L_2 is defined by

$$L_2 : (y,u) \to \int_o^T \{Eu(t) + Fy(t)\} \ dt \qquad (27)$$

with range in $H_2 = R^p$ and $b_2 = d$

In computational terms, defining $r_j^{(i)} = \{y_j^{(i)}(t) \ , \ u_j^{(i)}(t)\}$, $i=1,2$, $j \geqslant 1$, then $r_j^{(1)}$ is the solution pair to the optimization problem defined by (23) with no constraints and hence take the form $u_1^{(1)}(t) = - B^T K(t) x_1^{(1)}(t)$ where K(t) is the solution of a matrix Riccati equation. The iterates $r_j^{(2)}$ solve

$$\min \{\tfrac{1}{2} \int_0^T \{<y(t)-y_j^{(1)}(t), y(t)-y_j^{(1)}(t)>+<u(t)-u_j^{(1)}(t), u(t)-u_j^{(1)}(t)>\} dt$$
$$\text{subject to} \quad \int_0^T \{Eu(t) + Fy(t)\} \, dt = d \tag{28}$$

which is a simple algebraic exercise. The iterate s_ℓ is the solution of the problem

$$\min \{\tfrac{1}{2} \int_0^T \{<y(t)-y_j^{(2)}(t), y(t)-y_j^{(2)}(t)>+<u(t)-u_j^{(2)}(t), u(t)-u_j^{(2)}(t)>\} dt$$
$$\text{subject to the state equations (23)} \tag{29}$$

The solution to this problem takes the form $u(t) = u_j^{(2)} - B^T\{K(t)x(t)+g(t)\}$ where $g(t)$ is the solution of a well-defined vector differential equation. In practical terms this analysis illustrates that the algorithm is generated by the Riccati matrix of the unconstrained optimal control problem plus sequential integration of the state equations and the auxiliary equations defining $g(t)$. That is, the algorithm can be easily implemented using standard optimization routines.

3.3 Simultaneous Control of Independent Processes

Consider a system described by the relations, $i = 1,2$,

$$\dot{x}_i(t) = A_i x_i(t) + B_i u(t) \quad , \quad x_i(o) = x_{io} \in R^{n_i}$$
$$y_i(t) = C_i x_i(t) \in R^{m_i} \quad , \quad u(t) \in R^\ell \tag{30}$$

with performance criterion (T fixed)

$$J = \tfrac{1}{2} \int_0^T \{<y_1(t), y_1(t)> + <y_2(t) y_2(t)> + <u(t), u(t)>\} dt \tag{31}$$

Such a model could arise in the simultaneous optimal control of independent processes with a common input $u(t)$. Alternatively the composite model generated by (30) can be regarded as a partial modal decomposition of an ℓ-input, (m_1+m_2) output model of a process of state dimension $n = n_1 + n_2$.

The problem can be formulated in the Hilbert space $H = L_2^{m_1}[0,T] \times L_2^{m_2}[0,T] \times L_2^\ell [0,T]$ of triples (y_1, y_2, u) with norm specified by the performance criterion (31). The spaces H_i, $i = 1,2$ required are $H_i = L_2^{m_i}[0,T]$, $i = 1,2$ and $L_i : H \to H_i$ are defined by the relations

$$L_i : (y_1,y_2,u) \to y_i(t) - \int_o^t C_i \, e^{A_i(t-s)} B_i u(s)\, ds \qquad (32)$$

with $b_i = C_i \, e^{A_i t} \, x_i(o), i = 1,2.$

Comparing the direct matrix Ricatti solution of (30)-(31) with Ricatti solution of the individual problem inherent in the technique defined by Theorem 1, it is easily seen that computer storage requirements can be halved if n_i, i = 1,2, are large i.e. the algorithm provides a decomposition technique for the solution of large scale problems of this type.

3.4 Linear Optimal Control with Quadratic Criterion and Linear Terminal Constraints

Consider the optimal control problem

$$\dot{x}(t) = Ax(t) + Bu(t) \quad , \quad x(o) = x_o \in R^n \quad , \; u(t) \in R^\ell$$
$$y(t) = Cx(t) \quad , \quad y(t) \in R^m \qquad (33)$$

with performance criterion, (T fixed)

$$J = \tfrac{1}{2} < y(T), \, y(T) > + \tfrac{1}{2} \int_0^T < u(t), u(t) > dt \qquad (34)$$

subject to the terminal constraints

$$E \; y(T) = f \in R^p \qquad (35)$$

and the bracketed terms in (34) are suitable inner products in R^m, R^ℓ respectively. In this case, take $H = R^m \times L_2^\ell \, [0,T]$ with norm specified by (34), $H_1 = R^m$ and $H_2 = R^p$ where

$$L_1 : (y,u) \to y - C\int_0^T e^{A(T-S)} Bu(s)\, ds$$
$$L_2 \; ; \; (y,u) \to Ey \qquad (36)$$

and $b_1 = C \, e^{AT} \, x(o)$, $b_2 = f$. Noting that L_2 is compact, it follows directly that E y(T) converge strongly to f. If, after k iterations, the algorithms is terminated and $E \; y(T) = f_k$ with $||f_k - f|| < \varepsilon$, then (y,u) is the unique solution of the problem defined by (33), (34) with the constraint $E \; y(T) = f_k$. If ε is small, then the solution of this approximate optimization problem will, in practice, be a useful approximate solution to the original problem.

3.5 Minimum Energy Control

Consider the minimum energy problem defined by equation (33), the performance criterion

$$J = \tfrac{1}{2} \int_0^T < u(t) , u(t) > dt \tag{37}$$

and the constraint

$$x(T) = x_f \tag{38}$$

This problem can be transformed into a linear optimal control problem with quadratic criterion and linear terminal constraints (section 3.4) by augmenting J by the additional term $\frac{1}{2} < x(T), x(T)>$ and choosing $C = E = I_n$ and $f = x_f$. Applying the proposed algorithm to the augmented problem generates a sequence $(x_j(T), u_j)$ converging weakly to a weak limit (x_f, u^*) where u^* solves the minimum energy problem (33),(37),(38). Moreover $x_j(T)$ converges strongly to x_f.

In more detail, defining $r_j^{(i)} = (x_j^{(i)}(T), u_j^{(i)}) \in H = R^n \times L_2^{\ell} [0,T]$, i=1,2, $j \geq 1$, and assuming that the norms in (34) are Euclidean norms for simplicity, then $r_1^{(1)}$ is defined by the relation,

$$u_1^{(1)}(t) = -B^T K(t) x_1^{(1)}(t) \quad , \quad o \leq t \leq T \tag{39}$$

where K(t) is the unique solution of the matrix Ricatti equation

$$\dot{K}(t) + K(t) A + A^T K(t) - K(t) BB^T K(t) = 0$$
$$K(T) = I_n \tag{40}$$

Moreover, for $j \geq 1$,

$$r_j^{(2)} = (x_f , u_j^{(1)}) \tag{41}$$

Given $r_j^{(2)}$, the iterate s_j is the solution pair minimizing

$$||s_j - r_j^{(2)}||^2 = \tfrac{1}{2}<x(T)-x_f, x(T)-x_f> + \tfrac{1}{2} \int_0^T < u(t)-u_j^{(1)}(t), u(t)-u_j^{(1)}(t)>dt \tag{42}$$

subject to the state equation constraints (33). Applying the Minimum Principle,

$$u(t) = u_j^{(1)}(t) - B^T p(t)$$

$$\dot{p}(t) = - A^T p(t)$$

$$p(T) = x(T) - x_f \tag{43}$$

or, equivalently,

$$u(t) = u_j^{(1)}(t) - B^T \{K(t)\ x(t) + g(t)\} \tag{44}$$

where g(t) is the uniquely defined solution of the vector differential equations

$$\dot{g}(t) = - A^T g(t) + K(t)\ B\ B^T g(t) - K(t)\ B\ u_j^{(1)}(t)$$

$$g(T) = - x_f \tag{45}$$

Implementation of the algorithm simply requires computation of K(t) and sequential integration of the state equations and equation (45).

Finally it is noted that the augmentation of the performance criterion is similar in concept to the penalty function method (Luenberger, 1969). In contrast, however, the proposed algorithm uses a constant penalty function, a modified criterion (equation (42)) at each iteration and is guaranteed convergence from the initial condition (39).

4. Illustrative Examples

4.1 Minimum Energy Control of an Integrator

Consider the simple problem of choosing a control policy to transfer the scalar system $\dot{x}(t) = u(t)$ from the initial state $x(o) = o$ to the final state $x(1) = 1$ and minimizing the performance criterion

$$J(u) = \int_o^1 \tfrac{1}{2}\ u^2(t)dt \tag{46}$$

Using the formulation of section 3.5, consider the equivalent problem of minimizing

$$||(x(1),u)||^2 = \tfrac{1}{2}\ (x(1))^2 + \tfrac{1}{2}\int_o^1 u^2(t)dt \tag{47}$$

subject to the state equations and the terminal constraint $x(1) = 1$. Using the notation, $r_j^{(i)} = (x_j^{(i)}(1), u_j^{(i)})$, $i = 1,2$, $j \geq 1$, then $u_j^{(1)}(t)$ is given by (39) where K(t) solve the Ricatti equation

$$\dot{K}(t) = (K(t))^2 \quad , \quad K(1) = 1 \tag{48}$$

That is $K(t) = 1/(2-t)$ and it is easily shown that $u_1^{(1)}(t) \equiv 0$, $x_1^{(1)}(1) = 0$ and hence that $x_1^{(2)}(1) = 1$, $u_1^{(2)}(t) \equiv 0$. The point s_1 is obtained from (44) where g(t) is the solution of

$$\dot{g}(t) = \frac{1}{2-t} g(t) \quad , \quad g(1) = -1 \tag{49}$$

so that $g(t) = -1/(2-t)$. Substitution into the state equations yields $s_1 = (\frac{1}{2},\frac{1}{2})$ and hence $\lambda_1^* = 2$ giving $r_2^{(1)} = r_1^{(1)} + \lambda_1^* (s_1 - r_1^{(1)}) = (1,1)$ $\in D_1 \cap D_2$ and the problem is solved in two iterations.

To illustrate the accelerating effect of the use of λ_1^* consider the choice of $\lambda_i = 1$, $i \geq 1$. It is easily shown by induction that $r_j^{(1)} = (1 - \frac{1}{2}^j , 1 - \frac{1}{2}^j)$ so that $||r_{j+1}^{(1)} - (1,1)|| = \frac{1}{2} ||r_j^{(1)} - (1,1)||$, $j \geq 1$, and the convergence is geometric.

4.2 Minimum Energy Control of a 2nd Order Integrator Plant

Consider now, the second order system

$$\begin{aligned} \dot{x}_1(t) &= x_2(t) \\ \dot{x}_2(t) &= u(t) \end{aligned} \tag{50}$$

where it is desired to transfer the state $x(t) = [x_1(t), x_2(t)]^T$ from the initial condition

$$x(o) = |0,0|^T, \tag{51}$$

to the final condition

$$x(1) = |1,0|^T, \tag{52}$$

whilst minimising the energy functional given by (46). Again, with the formulation of section 3.5, this is equivalent to minimizing the norm

$$||(x(1),u)||^2 = \frac{1}{2}\mu x^T(1)x(1) + \frac{1}{2}\int_o^1 u^2(t)dt \ , \ \mu > o \qquad (53)$$

subject to the state equation (50), with associated initial condition (51) and the terminal equality constraint (52). Note that an arbitrary positive weighting factor μ has been introduced into the norm which, intuitively, can be chosen to improve the conditioning of the algorithm.

A discrete formulation of the problem (100 time steps) was solved numerically via the method outlined in section 3.5. With $\mu = 1.0 \times 10^4$ the algorithm gave a control u(t) which had converged to within 1% of the exact value of 6-12t in 3 iterations. Convergence data for this case is given in Table 1. As the value of μ was reduced the algorithm converged more slowly, exhibitng similar convergence in 7 iterations as shown in Table 2 for the case of $\mu = 1.0 \times 10^3$ and eventually displaying an oscillatory behaviour with μ=1.0 x 10^2 as described in Table 3. It is interesting to note from Tables 2 and 3 that in each case the extrapolation factor took on one of only two alternate values. Finally, the algorithm suffered from numerical problems for values of $\mu > 1.0 \times 10^5$ as was revealed by the extrapolation factor having become less than the theoretical minimum of 1.0.

References

LuenbergerD.G.: 1969, 'Optimization using vector space methods', John Wiley & Sons, Inc.

Owens D.H., Jones R.P.: 1978, Int. Jrnl. Control, to appear.

ITERATION (i)	$\|\|r_i^{(2)} - r_i^{(1)}\|\|$	$\|\|s_i - r_i^{(2)}\|\|$	λ_i^*
1	0.707×10^2	0.228×10^2	1.12
2	0.408×10	0.769	1.04
3	0.219	0.706×10^{-1}	1.12

Table 1 : $\mu = 1.0 \times 10^4$

ITERATION (i)	$\|\|r_i^{(2)} - r_i^{(1)}\|\|$	$\|\|s_i - r_i^{(2)}\|\|$	λ_i^*
1	0.224×10^2	0.156×10^2	1.94
2	0.958×10	0.419×10	1.24
3	0.261×10	0.182×10	1.94
4	0.112×10	0.490	1.24
5	0.306	0.213	1.94
6	0.131	0.570×10^{-1}	1.24
7	0.357×10^{-1}	0.249×10^{-1}	1.94

Table 2 : $\mu = 1.0 \times 10^3$

ITERATION (i)	$\|\|r_i^{(2)} - r_i^{(1)}\|\|$	$\|\|s_i - r_i^{(2)}\|\|$	λ_i^*
1	0.707×10	0.643×10	5.8
2	0.849×10	0.630×10	2.23
3	0.392×10	0.356×10	5.8
4	0.470×10	0.349×10	2.23
5	0.217×10	0.197×10	5.8

Table 3 : $\mu = 1.0 \times 10^2$