DIRECT STATE FEEDBACK OPTIMAL CONTROL OF A DOUBLE INTEGRATOR PLANT IMPLEMENTED BY AN ARTIFICIAL NEURAL NETWORK

Xavier Matieni, Stephen J. Dodds and Sin W. Lee
School of Computing, Information Technology and Engineering, University of East London
xamat@hotmail.com, stephen.dodds@spacecon.co.uk, s.w.lee@uel.ac.uk

Abstract: The purpose of this paper is to assess the capability of an artificial neural network (ANN) to implement a nonlinear state feedback optimal control law for a double integrator plant. In this case, the cost function to be minimised is the settling time subject to control saturation constraints. The reason for selection of this cost function is that the control law is known in the analytical form and this will be used to form a benchmark. The ultimate aim is to apply the method to form a new direct state feedback optimal position control law for mechanisms in which the frictional energy loss is minimised. An analytical solution is not available in this case so first the time optimal control law is studied to enable straightforward comparison on the ANN and directly implemented closed loop control laws. Since Pontryagin's method will be used to compute the optimal state trajectories for the ANN training in the future investigation of the minimum energy loss control, this method is applied to derive the time optimal double integrator state trajectories to illustrate the method. Furthermore, a modification of the time optimal control law is made that avoids the control chatter following a position change that would occur if a practical implementation of the basic control law, which is bang-bang, were to be attempted. Training the ANN with state and control data could be inaccurate due to the discontinuity of the control law on the switching boundary in the state space. This problem is overcome by the authors by instead training the ANN with state and switching function data, as the switching function is nonlinear but continuous, the control function, i.e., the function relating the switching function output to the control variable, being externally implemented. The simulations confirm that the ANN can be trained to accurately reproduce the time optimal control.

1. Introduction:

The celebrated minimum (or maximum) principle of Pontryagin has been used to tackle various optimal control problems in the past for finding the "best" or the "optimal" solution according to a selected cost function while respecting control saturation constraints. The overall objective of the research programme to which this paper contributes is to minimise the frictional energy loss in motion control systems employing electric drives. In this case an analytical solution in the form of a nonlinear state feedback control law has not yet been derived and does not appear to be mathematically tractable. This would therefore be one of many problems in which Pontryagin's method would be used to compute optimal control trajectories off line and implement them open-loop in real time. This has worked sometimes in process control applications but is impracticable in high dynamic motion control applications. Hence the approach taken by the authors is to use the optimal trajectories computed by Pontryagin's method to train an artificial neural network to produce the optimal controls when presented with the corresponding plant state variables and reference inputs. Then the ANN provides the means of implementing closed loop optimal control whose feedback structure affords a degree of robustness against external disturbances and plant modelling errors. As one of the first steps in achieving this goal, this paper is focused on establishing how well an ANN can implement a known nonlinear optimal state feedback control law, i.e., the time optimal control of a double integrator plant, an application example being the large angle slewing control about one axis of a rigid body spacecraft. The reason for choosing this example is that the closed loop behaviour is very well known and can be computed independently of the ANN to form a benchmark for performance assessment of the corresponding ANN based control. Although the closed loop time optimal control of a double integrator plant is known, it is derived via Pontryagin's method in this paper to show how the computed optimal trajectories are used to train the ANN.

2. Problem statement:

The double integrator plant is governed by the state differential equations

$$\begin{aligned}\dot{x}_1(t) &= x_2(t)\\ \dot{x}_2(t) &= bu(t)\end{aligned} \qquad (1)$$

where x_1 and x_2 are the state variables, b is a constant input coefficient and u is the control variable subject to the control saturation constraint

$$-u_{\max} \le u \le +u_{\max} \qquad (2)$$

This plant has its title due to the fact that $x_1(t) = \int x_2(t)\,dt$ and $x_2(t) = b\int u(t)\,dt$. The classical problem is to 'move' the plant to the origin, $(x_1, x_2) = (0,0)$, of the state space from a given initial state, $\mathbf{x}_0 = [x_1(t_0), x_2(t_0)]$ in the shortest possible time. In other words, the problem under consideration is the time optimal control one for the case where the terminal position is $\mathbf{x}_1 = [x_1(t_1), x_2(t_1)] = (0,0)$.

3. Time optimal control Derivation:

3.1. Application of Pontryagin's method:

The Hamiltonian function in this case has the form

$$H = p_1 x_2^2 + p_2 bu \qquad (3)$$

where p_1 and p_2 are the co-state variables satisfying the equations:

$$\begin{cases}\dfrac{dp_1}{dt} = -\dfrac{\partial H}{dx_1} = 0,\\[2mm] \dfrac{dp_2}{dt} = -\dfrac{\partial H}{dx_2} = -p_1\end{cases} \qquad (4)$$

Hence, $p_1 = c_1$ and $p_2 = c_2 - c_1 t$ where c_1 and c_2 are arbitrary constants of integration.

The control which minimises the Hamiltonian is given by

$$\begin{aligned}u_{opt}(t) &= -u_{max} sgn(p_2)\\ &= -u_{max} sgn[c_2 - c_1 t]\end{aligned} \qquad (5)$$

It follows that every optimal control $u_{opt}(t)$, $t_0 \le t \le t_1$, is a piecewise constant function that takes on the values $\pm u_{max}$ and has at most two intervals on which it is constant (since the linear function $c_2 - c_1 t$ changes sign at most once on the interval $t_0 \le t \le t_1$.

3.2. Closed loop time optimal control by ANN training:

Suppose at this stage an analytical solution in the form of a closed loop time optimal control law is not available. Then the authors would adopt the following approach to obtain a practicable closed loop optimal control. For each initial state, $[x_1(t_0), x_2(t_0)]$, the corresponding initial co-state, $[p_1(t_0), p_2(t_0)] = (c_1, c_2)$ (if

$t_0 = 0$) would have to be found for which the computed state trajectory, $[x_1(t), x_2(t)]$, obtained by computing the numerical solution of (1), (4) and (5), reaches the origin, $(0,0)$, of the state space. The result would be the optimal control trajectory, $[x_1(t), x_2(t)]$, and the corresponding optimal control, $u_{opt}(t)$. The ANN would then be trained to produce $u_{opt}(t_i)$ when presented with $[x_1(t_i), x_2(t_i)]$, for a selected set of points, $i = 1, 2, \ldots, N$, this being repeated for many different initial states, $[x_1(t_0), x_2(t_0)]$, so that all the selected points spanned the operational region of the state space for the particular application.

3.3. Closed loop time optimal control law:

To form a benchmark for testing the closed loop ANN based time optimal controller described in subsection 3.2, an analytical solution to the problem of closed loop time optimal control of the double integrator plant will now be derived following the method of Dodds (2002). Using the information obtained by the application of Pontryagin's method, the closed loop time optimal control may be found from the solution of the state trajectory differential equations obtained from (1) as follows:

$$\frac{\dot{x}_1}{\dot{x}_2} = \frac{x_2}{bu} \Rightarrow \boxed{\frac{dx_1}{dx_2} = \frac{x_2}{bu}} \qquad (6)$$

For any constant value of u, this equation may be solved analytically by the method of separation of variables. Cross-multiplying (6) yields

$$dx_1 = \frac{1}{bu} x_2 dx_2$$

Integrating then yields:

$$\int dx_1 = \frac{1}{bu}\int x_2 dx_2 \Rightarrow x_1 = \frac{1}{2bu} x_2^2 + A \qquad (7)$$

where A is a constant of integration.

The closed loop control law is derived by determining a switching boundary in the phase plane (a term used to describe the state space of a second order plant in which one state variable is the derivative of the other) that divides it into two distinctive regions, one in which the positive control value is applied and the other in which the negative one is applied. The implementation of this boundary closes the loop on the plant. The particular boundary is chosen for which at most one control switch occurs when implementing this closed loop system.

The two phase portraits generated by (7) for $u = +u_{max}$ and $u = -u_{max}$ are sketched in Fig. 1.

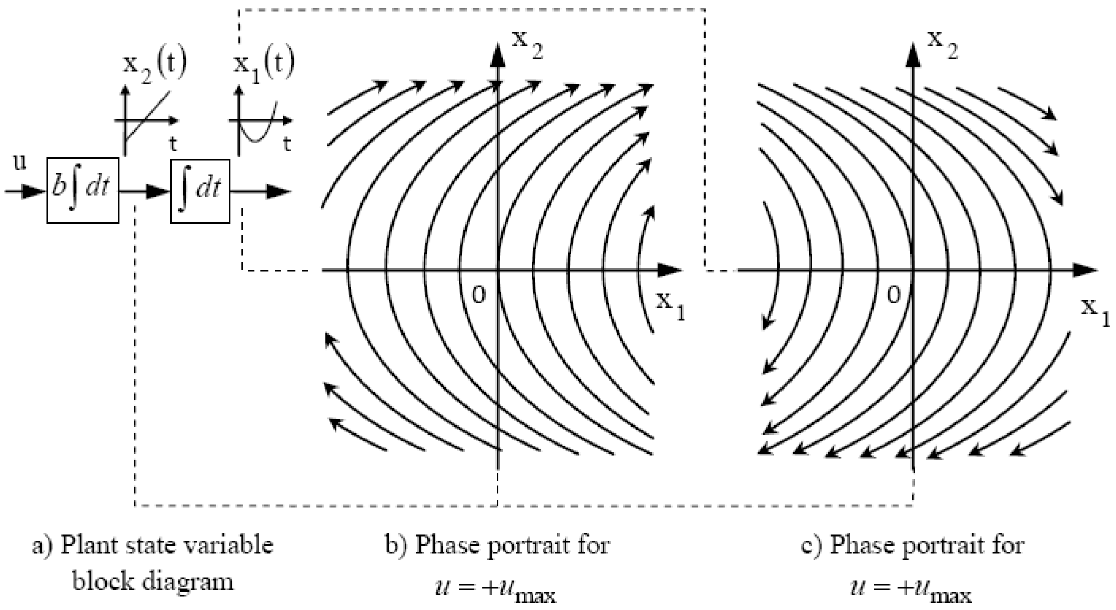

Fig.1: Phase portraits for bang-bang control of a double integrator plant.

It is evident that the required switching boundary is the one comprising the two parabolic state trajectory segments that terminate at the origin of the phase plain, as shown in Fig. 2 and this yields the closed loop phase portrait sketched in Fig. 3 that is seen to have at most one control switch for an arbitrary initial state, as required.

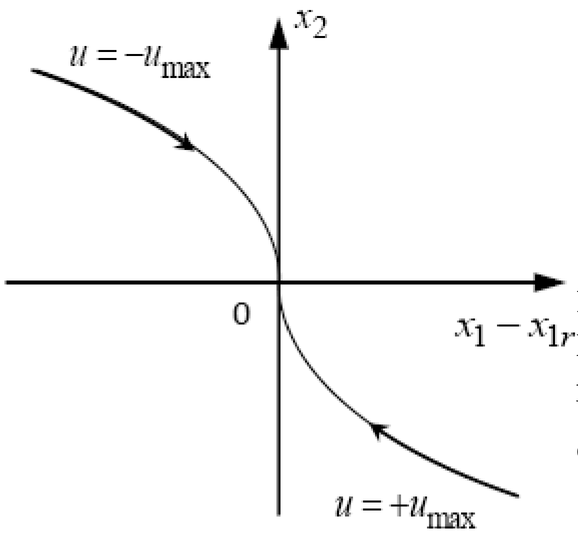

Fig. 2: Time optimal switching boundary formed from two state trajectories that terminate at the origin.

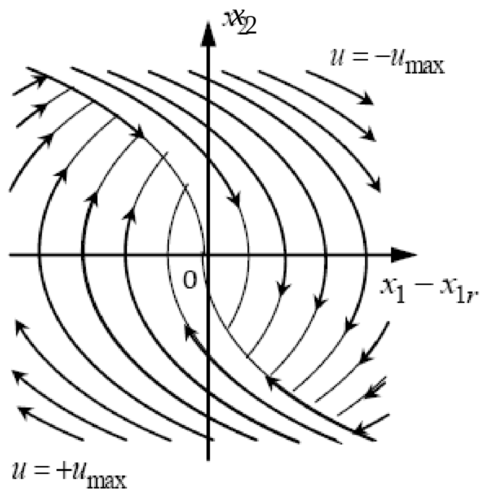

Fig. 3: Closed loop time optimal phase portrait.

According to (7) and Fig. 2, the switching boundary equation can be determined by the combination of the following equations:

$$x_1 = \begin{cases} -\dfrac{1}{2bu_{max}} x_2^2 & \rightarrow x_2 > 0 \\ +\dfrac{1}{2bu_{max}} x_2^2 & \rightarrow x_2 < 0 \end{cases}$$

which yields:

$$x_1 = -\frac{1}{2bu} x_2^2 sgn[x_2] \quad (8)$$

where $sgn[x_2] \triangleq \begin{cases} +1 \; for \; x_2 > 0 \\ 0 \; for \; x_2 = 0 \\ -1 \; for \; x_2 < 0 \end{cases}$.

Since $x_2^2 sgn(x_2) = x_2|x_2|$, then the switching boundary equation may be written as follows:

$$x_1 = -\frac{1}{2bu_{max}} x_2|x_2| \quad (9)$$

By inspection of Fig. 2(b), the corresponding equation for u, i.e., the required control law, may be found as follows.

If $x_1 + \frac{1}{2bu_{max}} x_2|x_2| > 0$ then u=-u_{max} (10a)

If $x_1 + \frac{1}{2bu_{max}} x_2|x_2| < 0$ then u=+u_{max} (10b)

Finally equations (10a) and (10b) can be combined into one equation for the time optimal control law:

$$\boxed{u = -u_{max} sgn[x_1 + \frac{1}{2bu_{max}} x_2|x_2|]} \quad (11)$$

3.4. Elimination of control chatter:

The time optimal control law (11) has the problem of limit cycling about the origin with digital implementation: Instead of reaching the origin and stopping there, the control holds its last computed value until the next state sample and therefore repeatedly overshoots the origin of the phase plane. This problem is similar in nature to the control chatter experienced with sliding mode control (Utkin, 1992), which is traditionally overcome by the boundary layer method. So this method will be applied here. First, control law (11) may be separated into a switching function,

$$S(x_1, x_2, x_{1r}) = x_1 - x_{1r} + \left[\frac{1}{2bu_{\max}}\right] x_2|x_2| \quad (12a)$$

and a control function

$$u = -u_{\max} \operatorname{sgn}\left[S(x_1, x_2, x_{1r})\right] \quad (12b)$$

Then the switching boundary is replaced by a boundary layer, i.e., a region straddling the original switching boundary within which the control undergoes a smooth transition between $-u_{\max}$ and $+u_{\max}$ between its edges. This is implemented by replacing the signum function of (12b) with the saturation function as follows.

$$u = -u_{\max}\, sat\left[S(x_1, x_2, x_{1r}), K/u_{\max}\right] \quad (13a)$$

where $sat(x, q) \overset{\Delta}{=} \begin{cases} qx \text{ for} |qx| < 1 \\ \operatorname{sgn}(x) \text{ for} |qx| \geq 1 \end{cases}$.

Also, the velocity feedback term, $x_2|x_2|/(2bu_{\max})$ in (12a) diminishes much faster than the position error, $x_1 - x_{1r}$ along the switching boundary as the origin of the phase plane is approached and therefore affords insufficient damping. The position error would overshoot the constant reference input and thereafter oscillate about the required value and the state trajectory would travel around the origin of the phase plane spiralling in relatively slowly. This further problem is overcome here by adding a linear velocity feedback term to the switching function (12a) as follows.

$$S(x_1, x_2, x_{1r}) = x_1 - x_{1r} + \left[\frac{1}{2bu_{\max}}\right] x_2|x_2| + T_c x_2 \quad (13b)$$

The system then becomes a sliding mode control system with a nonlinear switching boundary that gives near-time-optimal behaviour for large position demands and closely approximates stable first order behaviour with time constant, T_c, as the origin of the phase plane is approached.

4. ANN training using the near-time-optimal benchmark control system:

For this investigation, rather than use Pontryagin's method directly, the ANN will be trained using the states and controls yielded by a simulation of the closed loop system that will be used as the benchmark, based on the near-time-optimal control law (13). The control loop structure is shown in Fig. 4.

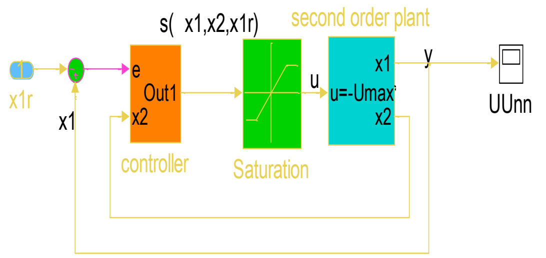

Fig. 4: The structure of the near time optimal control system used as the benchmark for assessing the performance of the ANN based controller.

Table 1 shows the data used to train the ANN obtained from a simulation of the system of Fig. 4 with $x_{1r} = 1$ and $[x_1(0), x_2(0)] = (0,0)$. The constant plant parameter is taken as $b = 1$ in this preliminary investigation.

Since control law (13) yields a very sharp transition between $u = -u_{\max}$ and $u = +u_{\max}$, across the boundary layer, the ANN will not be used to reproduce control values directly to avoid potential inaccuracies of the inherent curve fitting process. Instead, the ANN will be trained to reproduce the S values yielded by (13b) and shown in Table 1, as the switching function is smooth without sharp transitions. Then the control function (13a) will be implemented external to the ANN.

Table 1: Data used for the ANN training

x_1	x_{1r}	x_2	u	u_{nn}	error
1	0	0	10	10.00144	-0.00144
-1	0	0	-10	-10.0004	0.000364
1	0	1.646	-10	-10.0004	0.000364
1	0.107	1.338	-10	-10.0004	0.000364
-1	-0.107	-1.338	10	10.00144	-0.00144
1	0.150	0.816	-10	-10.0004	0.000364
1	0.194	0.109	-10	-9.9961	-0.00384
-1	-0.194	-0.109	10	9.997237	0.002763
1	0.2	0	10	10.00144	-0.00144
1	0	0	10	10.00144	-0.00144
-1	0	0	-10	-10.0004	0.000364

Table 1: Data used for the ANN training

Fig. 5 shows the Simulink block diagram of the training process.

Fig. 5: ANN controller being trained in parallel with the benchmark controller.

5. Results:

The system variables for the step responses of the benchmark and ANN based controllers are compared in Figs. 6, 7 and 8. The corresponding control error of the ANN based controller relative to the benchmark controller is shown in Fig. 9.

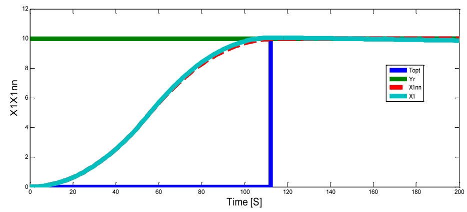

Fig. 6 position $x_1$vs. x_{1nn}

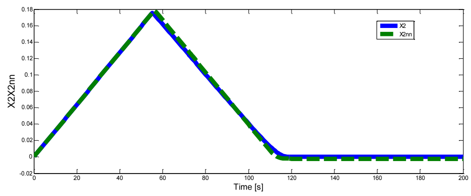

Fig. 7: Superimposed position and velocity responses obtained with the benchmark and ANN based controllers.

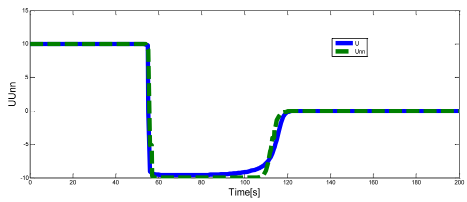

Fig. 8: Control inputs

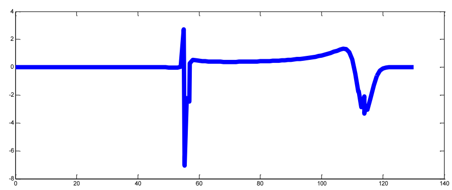

Fig. 9 error between u and u_{nn}

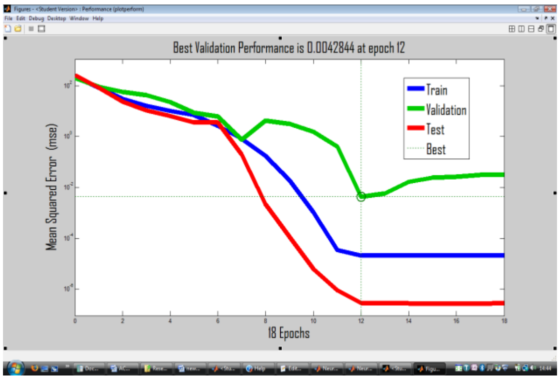

Fig. 10: Mean –square errors during the ANN training.

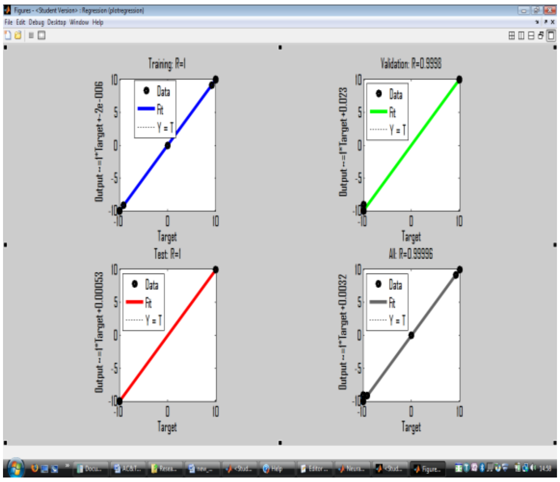

Fig. 11: Regression plot.

6. Conclusion and recommendations:

The ANN based controller was able to accurately reproduce the step response of the benchmark controller with a unit reference input. This is a strong indication that the approach should be successful for other nonlinear state feedback optimal controllers. First, however, further tests should be carried out with the near-time-optimal controller. More than one step response should be used to train the ANN, with

different valued step reference inputs, both positive and negative. Then the ability of the ANN based controller to reproduce those step responses, and intermediate ones, without further training should be assessed by simulation. Then, the degree of robustness against external disturbances and plant parameter mismatching, i.e., the constant, b, should be tested by simulation.
Following this, the authors' next step is to use Pontryagin's method to generate state trajectories for the mechanical load of a motion control system subject to friction for the position control that minimises the frictional energy loss. These trajectories will be used to train an ANN, thereby forming a closed loop state feedback optimal controller whose robustness properties will be assessed.

7. References:

Pontryagin L. S., (1986), the *mathematical theory of optimal processes*, Gordon Breach Science publisher.

Dodds, S. J., (2002), *computer based control*, Lecture notes MSc in Computer Systems Engineering.

Athans M. & Falb, P. L., (1966), *Optimal Control: an Introduction to the Theory and its Applications*, Dover.

Shinners S. M., (1992) *Modern Control System Theory and Design*, john Wiley and Sons.

Utkin, V. I., *Sliding Modes in Control Optimization*, Springer-Verlag, 1992, ISBN 9780387535166.