Regularization of the Trajectories of Dynamical Systems by Adjusting Parameters

Helle Hein and Ülo Lepik

Abstract—A gradient learning method to regulate the trajectories of some nonlinear chaotic systems is proposed. The method is motivated by the gradient descent learning algorithms for neural networks. It is based on two systems: dynamic optimization system and system for finding sensitivities. Numerical results of several examples are presented, which convincingly illustrate the efficiency of the method.

Keywords—Chaos, Dynamical Systems, Learning, Neural Networks

I. INTRODUCTION

THE study of controlling chaos in non-linear deterministic systems has been in focus during the last two decades. The ability to bring chaotic dynamical systems to regular motions is an important subject. The existing chaos control algorithms can be classified mainly into two categories: feedback methods [1 - 4] and non-feedback methods [5, 6]. In [7, 8] the authors showed that adding noise to the excitation frequency of the Duffing system the initially chaotic motion becomes regular.

In the present paper we propose a gradient learning method for adjusting the system parameters so that the trajectory of the system has certain specified properties. In the theory of neural networks the gradient descent learning requires that the performance of a dynamical system is assessable through certain error function which measures the discrepancy between the trajectories of the dynamical system and the desired behavior [9]. During gradient learning the interconnection weights between neurons are iteratively adjusted to reduce the error. In the present work we consider the system coefficients as adjustable parameters to obtain the desired behavior of the dynamical system.

II. PROBLEM STATEMENT

Consider the dynamical system

$$\frac{du}{dt} = F(u, w, I, u^*), \quad t \in \left[0, t_f\right] \tag{1}$$

H. Hein is with the Institute of Computer Science, University of Tartu, Liivi 2, Tartu, 50409, Estonia (phone: 372-7375475; fax:372-375441; e-mail: helle.hein@ut.ee).

Ü. Lepik is with the Institute of Applied Mathematics, University of Tartu, Liivi 2, Tartu, 50409, Estonia (e-mail: ulo.lepik@ut.ee).

with the initial condition $u(t_0) = u_0$. Here $u(t)$ is a N-dimensional row vector, which represents the state variables of the system, w is a J-dimensional array of adjustable parameters, I – the vector of external inputs. The vector $u^*(t)$ is used to represent particular target values for some of the u values.

The problem consists of adjusting parameters w so that the trajectories (1) have certain specified properties. For this purpose the error function

$$E = \int_{t_0}^{t_f} e\left[u(t, w), u^*(t)\right] dt, \tag{2}$$

which measures the discrepancy between the trajectories of (1) and the desired behavior, is introduced. Such parameters w for which $E = min$ are sought. The concrete form of the function e is specified later on.

For calculating the minimum of E main ideas, as gradient learning and backpropagation from the theory of artificial networks [9], are applied.

At first, the gradient

$$\frac{\partial E}{\partial w} = \int_{t_0}^{t_f} \frac{\partial e}{\partial u} \frac{\partial u}{\partial w} dt = \int_{t_0}^{t_f} \frac{\partial e}{\partial u} p \, dt \tag{3}$$

is computed.

Here the notation $p = \partial u / \partial w$ is introduced. So p stands for $N \times J$ matrix (sensitivities), $\partial e / \partial u$ is a N-dimensional vector and $\partial E / \partial w$ a J-dimensional vector.

By differentiating the matrix p we obtain

$$\frac{dp}{dt} = \dot{p} = \frac{d\dot{u}}{dw} = \frac{\partial F}{\partial w} + \frac{\partial F}{\partial u} \frac{\partial u}{\partial w}. \tag{4}$$

Introducing the matrix $L = (\partial F_i / \partial u_j)$ the formulae (4) can be put into the form

$$\dot{p} = Lp + \frac{\partial F}{\partial w}. \tag{5}$$

There are different possibilities for integrating the matrix equation (5) [9]. In this paper the derivative $\dot{p}$ is calculated from the formula (assuming that Δt is small enough)

$$\dot{p} \approx \frac{1}{\Delta t}\left[p(t + \Delta t) - p(t)\right]. \tag{6}$$

After solving the system (5) the gradient $\partial E / \partial w$ can be calculated from (3). New values for the parameters w are found from

$$w_{new} = w_{old} - \alpha \left(\frac{\partial E}{\partial w} \right)_{old}, \qquad (7)$$

where α is the learning rate. This procedure is repeated until the necessary exactness is reached.

For solving some problems the error function E can be taken in the non-integral form

$$E = e(u_f, u^*), \qquad (8)$$

where $u_f = u(t_f)$ and the vector u^* does not depend upon time. In this case $p = const$ and $\dot{p} = 0$. The matrix $p = \partial u_f / \partial w$ can be evaluated directly from (5) which gives

$$p = -L^{-1} \frac{\partial F}{\partial w}, \qquad (9)$$

where L^{-1} denotes the inverse matrix of L.
According to (3) one can find

$$\frac{\partial E}{\partial w} = \frac{\partial e}{\partial u_f} p \Delta t. \qquad (10)$$

This approach is much simpler as in the case of integral criterion (2), but its disadvantage can be the fact that for the minimization of the error function E only the final values of u_f are used. Therefore one can hope that the integral criterion (2) which takes into account all values of $u(t)$ for $t \in [t_0, t_f]$ guarantees a better rate of convergence.

III. LEARNING FIXED POINTS

Let us consider the case of unforced motion $I = 0$ and choose a point u^* for the fixed point. It is tried to learn the parameters w so that the dynamical system should converge to the target value u^*. For the error function E the functions

$$E = \frac{1}{2} \int_{t_0}^{t_f} [u(t) - u^*]^2 dt \qquad (11)$$

or

$$E = \frac{1}{2} (u_f - u^*)^2 \qquad (12)$$

are taken.
The further course of solution proceeds as indicated in Section II. Since at the fixed point

$$F(u_f, w, u^*) = 0 \qquad (13)$$

it follows from (1) that $(du/dt)_{t=t_f} = 0$ and consequently, the motion is terminated at the point u^*.

IV. LEARNING TO PERIODIC TRAJECTORIES

Let us return to (1). We shall assume now that the external input is periodic with a period T_*: e.g. $I(t_0 + iT_*) = I(t_0)$ $i = 1, 2, 3, \dots$. The trajectory $u(t)$ itself is generally non-periodic. Our aim is to adjust the parameters w so that the trajectory would be periodic with the prescribed period T.

For solving this problem let us choose a point u^* from the trajectory. This trajectory is represented in the phase space.

For simplicity sake the following diagrams (Fig. 1) are plotted in 2-dimensional phase space. For t_f the value $t_f = t_0 + T$ is taken. In general we get an open curve (Fig. 1 a), but in the case of periodical motion the phase trajectory must be a closed curve (Fig. 1 b). If the error criterion (12) is applied then in the case of a periodic solution it must be $E = 0$.
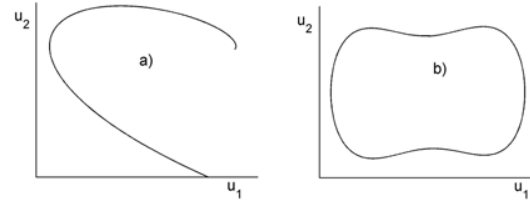

Fig. 1 trajectories in phase space a) non-periodic; b) periodic

Since

$$F[u(t+T), w, I(t+T), u^*] = F[u(t), w, I(t), u^*] \qquad (14)$$

then it follows from (1) that

$$\left(\frac{\partial u}{\partial t} \right)_{t+T} = \left(\frac{\partial u}{\partial t} \right)_t, \qquad (15)$$

Consequently the motion for $t > t_0 + T$ proceeds along the same trajectory as for $t_0 < t < T$ and we get a closed curve. The motion can be also multiple periodic (the case of a double-periodic is plotted in Fig. 2). To clear up this case we have to investigate the behavior of the error function E during one period of motion. If this function has γ minima then the motion is γ-periodic.
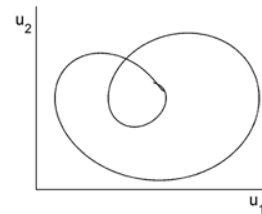

Fig. 2 double-periodic trajectories in phase space

V. RESTRICTIONS ON THE TRAJECTORIES

Here it is demanded that the trajectories found by solution of (1) should be closed between two limit curves

$$u_\gamma^-(t) \le u_\gamma(t) \le u_\gamma^+(t), \quad t \in [t_0, t_1] \qquad (16)$$

where the index γ denotes numbers of the sequence $1, 2, \dots, N$.

If $u_\gamma^-(t) = u_\gamma^+(t)$ for $\forall t \in [t_0, t_1]$ we get the case where the system follows a prescribed trajectory over the time interval $t \in [t_0, t_1]$; this case was discussed in [10].

The parameter w will be perturbed until inequalities (16) are satisfied. For that purpose we introduce the function

$$\varphi_\gamma(t) = \begin{cases} u_\gamma(t) - u_\gamma^+(t), & \text{if } u_\gamma(t) > u_\gamma^+(t), \\ u_\gamma(t) - u_\gamma^-(t), & \text{if } u_\gamma(t) < u_\gamma^-(t), \\ 0, & \text{elsewhere} \end{cases} \quad (17)$$

and take the error function in the form

$$e(t) = \frac{1}{2}\sum_\gamma \varphi_\gamma^2(t). \quad (18)$$

In (18) the summation is carried out only for the values of γ, which occurred in the inequalities (16). Since

$$\frac{\partial e}{\partial u_\gamma} = \varphi_\gamma \frac{\partial \varphi_\gamma}{\partial u_\gamma} = \varphi_\gamma, \quad (19)$$

the equation (3) obtains the form

$$\frac{\partial E}{\partial w} = \int_{t_0}^{t_1} \varphi_\gamma p \, dt \quad (p = \frac{\partial u_\gamma}{\partial w}). \quad (20)$$

Differential equation for calculating p remains valid if we understand under φ_γ a γ^*-dimensional vector, p - a $\gamma^* \times J$ - dimensional matrix and $L = (\partial F_\gamma / \partial u_y)$ a $\gamma^* \times \gamma^*$ - dimensional matrix. The subsequent course of solution proceeds as shown in Section II.

VI. APPLICATION OF THE OPTIMAL CONTROL THEORY

The problems considered in the preceding sections can be solved also with the aid of optimal control theory, but generally the solutions are more complicated. By this reason we confine ourselves to the solution of the following problem. Solve the dynamical system

$$\frac{du_1}{dt} = u_2,$$

$$\frac{du_2}{dt} = F(u_1, u_2, u_1^*, u_2^*) + I \quad (21)$$

with the initial conditions $u_1(0) = u_{10}$, $u_2(0) = u_{20}$. The function $I = I(t)$ describes the external force; u_1^*, u_2^* are the prescribed target coordinates. The problem is to find such a control $I = I(t)$ which transfers the system to the point (u_1^*, u_2^*) with the minimal time T. It is assumed that the motion is terminated at T, so that $(du_1/dt)_{t=T} = (du_2/dt)_{t=T} = 0$.

This problem is solved by means of the optimal control theory. We have to minimize the functional

$$J = T + \frac{1}{2}\left\{\left[u_1(T) - u_1^*\right]^2 + \left[u_2(T) - u_2^*\right]^2\right\} \quad (22)$$

satisfying by it equations (21). This is a Bolza problem [11] which can be transferred to a Lagrange problem by taking

$$K = \int_0^T \frac{dJ}{dt} dt. \quad (23)$$

Since

$$\frac{dJ}{dt} = 1 + \left[u_1(t) - u_1^*\right]\dot{u}_1(t) + \left[u_2(t) - u_2^*\right]\dot{u}_2(t) = $$
$$1 + \left[u_1(t) - u_1^*\right]u_2(t) + \left[u_2(t) - u_2^*\right][F(t) + I(t)], \quad (24)$$

we get the Lagrange problem

$$K = \int_0^T \left[1 + (u_1 - u_1^*)u_2 + (u_2 - u_2^*)(F + I)\right]dt = \min. \quad (25)$$

Next we introduce the Hamiltonian

$$H(u_1, u_2, \psi_1, \psi_2, I) = -1 - (u_1 - u_1^*)u_2 - $$
$$(u_2 - u_2^*)(F + I) + \psi_1 u_2 + \psi_2(F + I). \quad (26)$$

Here $\psi_1(t)$, $\psi_2(t)$ are the adjoint variables which are computed from the equations

$$\frac{d\psi_1}{dt} = -\frac{\partial H}{\partial u_1}, \quad \frac{d\psi_2}{dt} = -\frac{\partial H}{\partial u_2}. \quad (27)$$

According to the maximum principle of Pontryagin [11] such a control $I(t)$ should be taken for which $H = max$. By physical reasons the control (external force) must be bounded: $|I(t) \le I_{max}|$. It follows from (26) that there are only two possibilities to maximize the Hamiltonian:

$$I = I_0 \quad if \quad u_2 - u_2^* + \psi_2 > 0,$$
$$I = -I_0 \quad if \quad u_2 - u_2^* + \psi_2 < 0. \quad (28)$$

So we see that the control function I is discontinuous ("bang-bang" type). In the case of a continuous external force solution of the problem (25) does not exist. From here follows that e.g., in the case of the Duffing equation

$$\ddot{u} + a\dot{u} + bu + cu^3 = f\cos wt \quad (f \ne 0) \quad (29)$$

regardless of the values of the parameters a, b, c, f there do not exist any fixed points to which the trajectories converge.

VII. COMPUTER SIMULATIONS

Let us consider some examples. For fixed point learning we consider the Lorenz system and the Duffing equation.

1. Lorenz equations
A Lorenz chaotic system can be presented as

$$\dot{x} = -\sigma(x - y),$$
$$\dot{y} = rx - y - xz, \quad (30)$$
$$\dot{z} = xy - bz.$$

Here $u_1 = x$, $u_2 = y$, $u_3 = z$ are the state variables and dot denotes the derivative with respect to time. The parameters σ, r and b are real positive parameters and the system produces chaotic behavior when $\sigma = 10$, $r = 28$ and $b = 8/3$. The fixed points of the Lorenz system are $\left(\sqrt{b(r-1)}, \sqrt{b(r-1)}, r-1\right)$, $\left(-\sqrt{b(r-1)}, -\sqrt{b(r-1)}, r-1\right)$ and $(0, 0, 0)$. Analysis of the stability of these points and their chaotic characteristics can be found in [12].

The control objective is to drive the system to the desired fixed point $F_*(x_*, y_*, z_*)$. The forth-order Runge-Kutta method with step 0.02 was used to integrate the systems numerically.

The simulation results for $\sigma = 10$, $r = 28$, $b = 8/3$ and $F_*(4, 4, 6)$ are shown in Fig. 3. The initial condition was taken as $(x_0, y_0, z_0) = (2, 2, 2)$ and the learning rate as $\alpha = 0.2$.
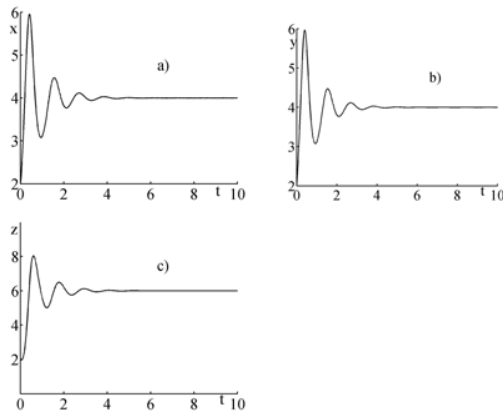

Fig. 3 time response of the Lorenz system after gradient descent learning ($\sigma = 66.38$, $r = 7$, $b = 2.6667$)

For the error function the formula (12) was taken; the dynamic tunneling technique [13, 14] and learning rate adjustment were also applied.

2. Duffing equation

Let us consider the Duffing equation without external excitation:

$$\dot{x} = y,$$
$$\dot{y} = -ay - bx - cx^3. \tag{31}$$

The system has three fixed points with coordinates $(0, 0)$, $\left(\sqrt{-b/c}, 0\right)$ and $\left(-\sqrt{-b/c}, 0\right)$, if $b < 0$ and one fixed point $(0, 0)$, if $b > 0$. The results of computer simulations are presented in Fig. 4 where the desired fixed point was taken $(2, 0)$ and the corresponding calculated parameter values were $a = 3.28$, $b = -1.56$, $c = 0.39$.
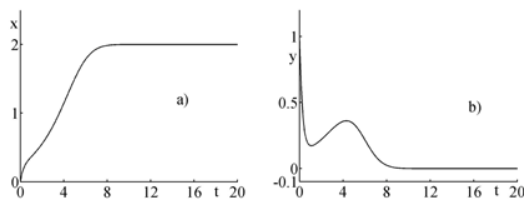

Fig. 4 time response of the Duffing system after gradient descent learning ($a = 3.28$, $b = -1.56$, $c = 0.39$)

For trajectory restriction let us consider the Duffing equation with external excitation $I = f\cos(\omega t)$:

$$\dot{x} = y,$$
$$\dot{y} = -ay - bx - cx^3 + f\cos(\omega t). \tag{32}$$

The parameter values are taken as follows: $a = 0.1$, $b = 1$; $c = 0$, $f = 15$; $\omega = 1$. The trajectory of state variable x with and initial conditions $x(0) = 0$, $y(0) = 1$ is presented in Fig. 5a. The gradient learning has been executed with parameters $\gamma = 1$ and $x_1^+(t) = 50$, $x_1^-(t) = -50$. The trajectory after learning is presented in Fig. 5b with calculated parameters $a = 0.2977$, $b = 1.0301$; $c = 0$, $f = 15.0036$.
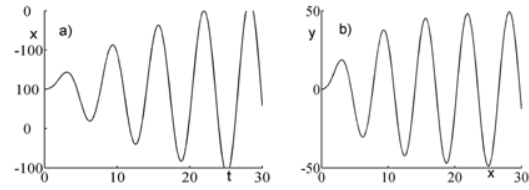

Fig. 5 time response of the Duffing system in the case of restrictions to the trajectory: a) before learning b) after learning

VIII. CONCLUSION

The gradient learning method for adjusting system parameters so that the system trajectory exhibits some prescribed properties, is developed. Simulation results show that the proposed method is able to control dynamical systems. The system can be learned to drive to the prescribed fixed point and restrictions can be posed to the trajectory.

REFERENCES

[1] E. Ott, C. Grebogy, J. A. Yorke, "Controlling chaos," *Phys. Rev. Lett.*, vol. 64, pp. 1196- 1199, 1990.
[2] J. Singer, Y. Z. Wang, H. H. Bau, „Controlling chaotic system," *Phys. Rev. Lett.*, vol. 66, pp. 1123- 1125, 1991.
[3] K. Pyragas, "Continuous control of chaos by self-controlling feedback," *Phys. Lett.*, A, vol. 170, pp. 421-428, 1992.
[4] C. C. Hwang, R. F. Fung, J. Y. Hsieh, W. J. Li, „A nonlinear feedback control of the Lorenz equation," *Int. J. Eng. Sci.*, vol. 37, pp 1893-1900, 1999.
[5] Y. Braiman, I. Goldhirsh, "Taming chaotic dynamic with weak periodic perturbations," *Phys. Rev. Lett.*, vol. 66, pp. 2545- 2548, 1991.
[6] T. Kapitaniak, "The loss of chaos in a quasiperiodically-forced nonlinear oscillator," *Int. J. Bifurcation and Chaos*, vol. 1, pp357-362, 1991.
[7] Ü. Lepik, H. Hein, "On response of nonlinear oscillators with random frequency of excitation," *J. Sound and Vibration*, vol. 288, pp. 275-292, 2005.
[8] H. Hein, Ü. Lepik, "Response of nonlinear oscillators with random frequency of excitation," revisited., *Journal of Sound and Vibration*, vol. 301, pp. 1040-1049, 2007.
[9] P. Baldi, "Gradient descent learning algorithm overview: A general dynamical perspective," *IEEE Transactions on neural networks*, vol. 6, no. 1, January 1995.
[10] I. P. Marino, J. Miguez, "An approximate gradient-descent method for joint parameter estimation and synchronization of coupled chaotic systems," *Phys. Lett.*, A, vol. 351, pp. 262-267, 2006.
[11] A. Bryson, H. Yu-Chi, *Applied optimal control*, New York: Wiley, 1975.
[12] C. Sparrow, *The Lorenz equations: Bifurcations, chaos and strange attractors*, Berlin: Springer, 1982.
[13] Y. Yao, "Dynamic tunneling algorithm for global optimization," *IEEE Transactions on systems, Man and Cybernetics*, vol. 19, no. 5, pp. 1222-1230, October, 1989.
[14] P. RoyChowdhury, Y. P. Singh and R. A. Chansarkar, "Dynamic tunneling technique for efficient training of multilayer perceptrons," *IEEE Transactions on neural networks*, vol. 10, no. 1, pp.48-55, January, 1999.