Hopfield neural network: The hyperbolic tangent and the piecewise-linear activation functions

Amanda C. Mathias, Paulo C. Rech

Departamento de Física, Universidade do Estado de Santa Catarina, 89219-710 Joinville, Brazil

A R T I C L E   I N F O

Article history:
Received 12 December 2011
Received in revised form 20 June 2012
Accepted 22 June 2012

Keywords:
Hopfield neural network
Parameter-space
Lyapunov exponents

A B S T R A C T

This paper reports two-dimensional parameter-space plots for both, the hyperbolic tangent and the piecewise-linear neuron activation functions of a three-dimensional Hopfield neural network. The plots obtained using both neuron activation functions are compared, and we show that similar features are present on them. The occurrence of self-organized periodic structures embedded in chaotic regions is verified for the two cases.

© 2012 Elsevier Ltd. All rights reserved.

1. Introduction

A Hopfield-type neural network (Hopfield, 1984) is a significant model in artificial neurocomputing (Korner, Kupper, Rahman, & Shkuro, 2007). It is a continuous-time nonlinear dynamical system described by a set of \( n \) autonomous first-order ordinary differential equations given by

\[
\dot{x}_i = -\frac{x_i}{R_i} + \sum_{j=1}^{n} w_{ij} v_j + I_i, \quad i = 1, 2, \ldots, n, \tag{1}
\]

where \( v_j = f_j(x_j) \), \( x_i \) represent dynamical variables, \( C_i, R_i, \) and \( I_i \) are parameters, and \( w_{ij} \) are the elements of an \( n \times n \) matrix, called weight matrix or connectivity matrix, which describes the strength of the connections between the \( n \) neurons. The neuron activation function \( f_j(x_j) \) is a bounded monotonic differentiable function usually represented by any smooth function.

In this paper we investigate the dynamics of a low-dimensional form of (1). Namely, a system composed of three neurons and whose behavior depends on two parameters, \( \alpha \) and \( \beta \), given by

\[
\begin{align*}
\dot{x}_1 &= -x_1 + 1.5f_1(x_1) + 2.9f_2(x_2) + \alpha f_3(x_3), \\
\dot{x}_2 &= -x_2 + \beta f_2(x_2) + 1.18f_3(x_3), \\
\dot{x}_3 &= -x_3 + 2.977f_1(x_1) - 22f_2(x_2) + 0.47f_3(x_3).
\end{align*} \tag{2}
\]

The parameter \( \alpha \) denotes the connection weight between third and first neurons, while the parameter \( \beta \) is related with the connection weight between first and second neurons.

To obtain system (2) we have considered in system (1) the weight matrix equal to

\[
\begin{pmatrix}
1.5 & 2.9 & \alpha \\
\beta & 1.18 & 0 \\
2.977 & -22 & 0.47
\end{pmatrix},
\]

\( n = 3, C_i = I_i = 0 \) and \( R_i = 1 \) for \( i = 1, 2, 3 \). Our choice of parameters to compose the weight matrix above was motivated by the original work of Huang and Huang (2011). Two neuron activation functions \( f_j(x_j) \) are considered in our study:

- The hyperbolic tangent function

\[
f_j(x) = \tanh(x), \tag{3}
\]

- The piecewise-linear function

\[
f_j(x) = \begin{cases} 
-1 & \text{if } x_i \leq -1, \\
x_i & \text{if } -1 < x_i \leq 1, \\
1 & \text{if } x_i > 1, 
\end{cases} \tag{4}
\]

plotted in Fig. 1(b).

Recently Huang and Huang (2011) present several results concerning system (2), with activation function (3) and \( \alpha \) kept fixed at 0.8, involving mainly Lyapunov exponents, bifurcation diagrams, and attractors in the phase-space. Periodic and chaotic attractors were reported by these authors, as a function of the parameter \( \beta \). Therefore, the investigation realized concerning system (2) and reported by Huang and Huang (2011) considered only the points located along a straight line \( \alpha = 0.8 \) in a two-dimensional \((\alpha, \beta)\) parameter-space. Also recently a three-dimensional Hopfield network with the piecewise-linear activation function (4) was presented by Zheng, Tang, and Zhang (2010).

In this case, again, the dynamics of the network was investigated...
by varying only one parameter, this time using bifurcation diagrams, Lyapunov exponents, and phase-space portraits. The authors showed that the system displays chaotic behavior for some well chosen parameter values.

Our main goal in this work is to investigate the two-dimensional \((\alpha, \beta)\) parameter-space of system (2), considering its dynamical behavior for a very large number of points, namely \(250 \times 10^3\), and comparing the dynamics of both Hopfield networks, with the hyperbolic tangent activation function (3) and with the piecewise-linear activation function (4). For this purpose, both \(\alpha\) and \(\beta\) parameters are simultaneously varied in each case, and Lyapunov exponents are sequentially calculated.

2. Parameter-space plots

Fig. 2 shows plots displaying interesting pieces of the \((\alpha, \beta)\) parameter-space of the Hopfield network (2) working with both neuron activation, the hyperbolic tangent function (3) in Fig. 2(a), and the piecewise-linear function (4) in Fig. 2(b). Both plots show dynamical behaviors of the Hopfield neural network (2), where we identify an intricate mixture of chaotic and periodic regions, represented by yellow–red and black, respectively. They indicate how variations in the connection weight \(\alpha\) between third and first neurons, and \(\beta\), between first and second neurons, affect the dynamical behavior of the network, in each case. Each of the plots in Fig. 2 can be interpreted as presenting a chaotic region, in yellow–red, with several periodic regions in black, embedded in it. In other words, regardless of case, as the parameters \(\alpha\) and \(\beta\) are varied, we can observe regions on the parameter-space where periodic windows appear embedded in a chaotic region.

Plots like those in Fig. 2, as well as all the similar plots that appear forward, are obtained by computing the largest Lyapunov exponent on a \(500 \times 500\) grid of parameters \((\alpha, \beta)\). This means that for each plot \(250 \times 10^3\) Lyapunov exponents were calculated. In each of the \(250 \times 10^3\) procedures, system (2) was integrated using a fourth-order Runge–Kutta algorithm with a step size equal to 0.01, and considering \(500 \times 10^3\) steps to compute each Lyapunov exponent. Integrations were performed along lines of constant parameter \(\beta\), starting always at the lower value of the parameter \(\alpha\), from the initial condition \(P_0 = (x_{10}, x_{20}, x_{30}) = (0.15, -1.12, 7.25)\), and computing Lyapunov exponents subsequently. While computing Lyapunov exponents along lines of constant \(\beta\), the initial condition \(P = P_0\) was utilized only to begin integrations at the small value of \(\alpha\). To begin integrations for each increased value of \(\alpha\), we use the last value of \(P\) obtained with the anterior value of \(\alpha\), as the initial condition for the newly increased \(\alpha\), i.e., the attractor was followed along lines of fixed \(\beta\).

It is clear that by proceeding in this way we lose some information about the coexistence of multiple attractors. However, by comparing figures constructed varying parameters along different directions, we learn to identify the regions more likely to have multistability. Of course, we can only be certain of it after computing explicitly basins of attraction for fixed parameters in such regions. As is known, the construction of basins of attraction to map attractors coexistence requires much more computer time than for computing parameter-space plots. An analysis based on basins of attraction for system (2) will be the subject of a future work.
Colors in plots of Fig. 2 are associated with the magnitude of the largest Lyapunov exponent, as shown for each in the scale of the corresponding column at right. White for more negative, black for zero, and red for more positive. Indeed, a negative exponent is indicated by a continuously changing white–black scale, while a positive exponent is indicated by a continuously changing yellow–red scale. As is known (Wiggins, 2003), a negative largest Lyapunov exponent indicates a stable equilibrium point, a zero largest Lyapunov exponent indicates a stable periodic attractor (or a quasi-periodic attractor, when the second largest Lyapunov exponent is also equal to zero), and a chaotic attractor has a positive largest Lyapunov exponent.

At first sight, the distribution of chaos and periodicity in both panels in Fig. 2 are not similar. It is apparent that the global structure of a parameter-space has nothing to do with the global structure of the other. Roughly speaking, Fig. 2(a) is “very different” from Fig. 2(b). The ranges of the largest Lyapunov exponent are different, despite the positive values that characterize chaos falling within the same order of magnitude.

Our interest now is to investigate smaller regions of both parameter-spaces in order to find possible similarities between the two Hopfield networks. For this purpose, we produce an enlargement in four carefully chosen regions on Fig. 2, more specifically the areas enclosed by the boxes A and B in Fig. 2(a), C and D in Fig. 2(b). The result is shown in plots of Fig. 3.

Numbers written in black regions of the plots in Fig. 3 refer to the lower period (henceforth referred as period) of the respective region, once bifurcations may occur when we move from the center to the periphery of each periodic structure. Period here is assumed as being the number of local maxima of the variable $x_i$ in one complete orbit on the attractor.

Some features and similarities concerning the periodic structures embedded in the chaotic region of the $(\alpha, \beta)$ parameter-space in plots of Fig. 3 deserve prominence, and we discuss it in continuation. In Fig. 3(a) and (c), which are magnifications of the boxed regions A in Fig. 2(a) and C in Fig. 2(b), respectively, can be seen an arrangement of periodic structures in spirals (Albuquerque & Rech, 2012; Gallas, 2010). These spirals are associated to different periodic behaviors, and have a focal point $Q$ located roughly at $(\alpha, \beta) = (0.923724, -0.743591)$ in Fig. 3(a), and at $(\alpha, \beta) = (1.33013, -0.768204)$ in Fig. 3(c). Clearly visible is one of these spirals (of a probably infinite number) in both plots, that coil up clockwise around the focal point $Q$, where the individual spirals initiate or terminate, in each case. By walking clockwise along this clearly visible spiral we can observe an organization of the periods as 5, 7, 7, 9. This result can be interpreted as showing period-adding sequences with the increment factor of 2, insofar as each spiral is covered in the direction of the point $Q$. With the scale used in Fig. 3(a), we can still distinguish a less visible spiral, with periods organized as 5, 6, 7, therefore in a sequence whose period is increased this time by a factor 1. As we know, a similar arrangement of spirals was described for the first time in the chaotic region of a simple electronic circuit (Bonatto & Gallas, 2008), and more recently in the chaotic region of a semiconductor laser (Freire & Gallas, 2010). Also recently, the spirals were detected experimentally (Stoop, Bennet, & Uwate, 2010).

Another type of organization of the periodic structures, embedded in the chaotic region of the $(\alpha, \beta)$ parameter-space of the Hopfield neural network (2) with both activation functions, can be seen in Fig. 3(b) and (d), which are magnifications of the boxed regions $B$ and $D$ in Fig. 2(a) and (b), respectively. Now as we move from right to left in Fig. 3(b) or (d), we cross a probably infinite set of periodic structures that get smaller and smaller, and accumulate in the border of the black period-3 region on the left. These periodic structures are organized in two mixed sets of period-adding bifurcations: (i) the first set, represented by the sequence $4 \rightarrow 7 \rightarrow 10 \rightarrow 13 \rightarrow 16 \rightarrow \ldots$, therefore started at period-4 with increase by a factor 3 every bifurcation, and the second set, (ii) with a sequence $11 \rightarrow 17 \rightarrow 23 \rightarrow 29 \rightarrow 35 \rightarrow \ldots$, this last started at period-11 with increase by a factor 6 every bifurcation. In terms of the complete sequence, the periodic structures embedded in the chaotic region of Fig. 3(b) and (d), and that accumulate in the border of the period-3 region at left side, are organized as

\[4 \rightarrow 11 \rightarrow 7 \rightarrow 17 \rightarrow 10 \rightarrow 23 \rightarrow 13 \rightarrow 29 \rightarrow 16 \rightarrow 35 \rightarrow \ldots\]
At first sight, any regularity in the sequence (5) does not exist, i.e., given a member it is impossible to anticipate the previous and the next. However, if we look at the sequence (5) with more attention, some regularity can be discerned. Note that the member at the position $2n$, $n = 1, 2, \ldots$ can be obtained by adding the members occupying both the positions, $2n - 1$ and $2n + 1$. In other words, each number in an even position of the sequence is the sum of the previous and the next number. For example, the member in the position 8, equal to 29 and corresponding to $n = 4$, is obtained by adding the members in the positions 7 and 9, namely the numbers 13 and 16. The sequence (5) is a particular case of a more general sequence given by

\[(i) \rightarrow (2i + k) \rightarrow (i + k) \rightarrow (2i + 3k) \rightarrow (i + 2k) \rightarrow (2i + 5k) \rightarrow \cdots, \tag{6}\]

where $i, k$ are integers, for $i = 4, k = 3$, whose general term is given by

\[a(m) = \begin{cases} 3m + 2 & \text{if } m \text{ is odd}, \\ \frac{1}{2}(3m + 2) & \text{if } m \text{ is even}, \end{cases} \tag{7}\]

with $m \geq 2$.

Thus, we see that although there is agreement of the dynamics observed in the parameter-space for both networks when “small” specific regions like those shown in plots of Fig. 3 are considered, when a “large” parameter-space is taken into account this agreement ceases to exist, as shown in plots of Fig. 2.

3. Summary

In this paper we have investigated a Hopfield-type neural network with three neurons, modeled by a set of three autonomous nonlinear first-order ordinary differential equations depending on two parameters, $\alpha$ and $\beta$. We have reported two-dimensional parameter-space plots concerned with both, the hyperbolic tangent and the piecewise-linear activation functions considered in the network. More specifically, by contrasting the dynamics displayed in each case, we observe that: (i) Although the piecewise-linear activation function can be seen as a good approximation for the hyperbolic tangent function (see Fig. 1), the parameter-space plots are rather different over wide regions (see Fig. 2). (ii) In spite of the difference mentioned above, sets of periodic structures embedded in the chaotic region, previously reported in other continuous-time models (Albuquerque, Rubinger, & Rech, 2008; Bonatto & Gallas, 2008; Freire & Gallas, 2010; Gallas, 2010; Rech, 2010; Stegemann, Albuquerque, & Rech, 2010), are present in both parameter-spaces. (iii) There are regions in both parameter-spaces, where these periodic structures appear self-organized in period-adding sequences (see Fig. 3(b) and (d)), or in an arrangement of spirals associated to different periodic behaviors, and that coil up clockwise around a focal point (see Fig. 3(a) and (c)). This configuration of the periodic structures in spirals was also recently reported for other continuous-time models (Albuquerque & Rech, 2012; Gallas, 2010).

Acknowledgments

The authors thank Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Brazil, for financial support.

References


