Neural Networks 100 (2018) 84–94

Contents lists available at ScienceDirect

Neural Networks

journal homepage: www.elsevier.com/locate/neunet

Max-plus and min-plus projection autoassociative morphological memories and their compositions for pattern classification

Alex Santana dos Santos [a], Marcos Eduardo Valle [b,*]

[a] Exact and Technological Science Center, Federal University of the Recôncavo of Bahia, Rua Rui Barbosa, 710, Centro - Cruz das Almas-BA CEP 44.380-000, Brazil

[b] Department of Applied Mathematics, University of Campinas, Rua Sérgio Buarque de Holanda, 651, Campinas-SP CEP 13083-859, Brazil

HIGHLIGHTS

- Robust and computationally efficient high-capacity associative memory models.
- New models project the input onto the set of max-plus or min-plus combinations.
- New autoassociative morphological memories obtained by their compositions.
- Excellent performance on problems with a large number of classes and features.

ARTICLE INFO

Article history:
Received 20 September 2017
Received in revised form 13 December 2017
Accepted 26 January 2018
Available online 3 February 2018

Keywords:
Associative memory
Lattice computing
Minimax algebra
Morphological neural network
Pattern classification

ABSTRACT

Autoassociative morphological memories (AMMs) are robust and computationally efficient memory models with unlimited storage capacity. In this paper, we present the max-plus and min-plus projection autoassociative morphological memories (PAMMs) as well as their compositions. Briefly, the max-plus PAMM yields the largest max-plus combination of the stored vectors which is less than or equal to the input. Dually, the vector recalled by the min-plus PAMM corresponds to the smallest min-plus combination which is larger than or equal to the input. Apart from unlimited absolute storage capacity and one step retrieval, PAMMs and their compositions exhibit an excellent noise tolerance. Furthermore, the new memories yielded quite promising results in classification problems with a large number of features and classes.

1. Introduction

Associative memories (AMs), which are models inspired by the human brain ability to store and recall information by associations, have been extensively investigated since the advent of Hopfield network in the early 1980s (Hassoun & Watta, 1997; Hopfield, 1982). Besides the biological motivation, associative memory models have been applied, for instance, for pattern classification (Chyzhyk & Graña, 2015; Esmi, Sussner, Bustince, & Fernandez, 2015; Valle & de Souza, 2016; Zhang, Huang, Huang, & Zhang, 2005), time-series prediction (Sussner & Schuster, 2013; Valle & Sussner, 2011), image processing and analysis (Grana & Chyzhyk, 2016; Lechuga-S, Valdiviezo-N, & Urcid, 2014; Valle, 2014a; Valle & de Souza, 2015), and optimization (Hopfield & Tank, 1985; Serpen, 2008).

An AM model is usually classified as either autoassociative or heteroassociative (Hassoun & Watta, 1997). An autoassociative memory, such as the Hopfield network (Hopfield, 1982), is designed for the storage of a finite set of vectors $\{\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^k\}$. In contrast, a heteroassociative memory is designed for the storage of a set of association pairs $\{(\mathbf{x}^1, \mathbf{y}^1), \ldots, (\mathbf{x}^k, \mathbf{y}^k)\}$. The bidirectional associative memory of Kosko is an example of a heteroassociative memory (Kosko, 1988). In this paper, we focus only on autoassociative memories.

Desired properties of an autoassociative memory include high storage capacity, tolerance to noisy or partial inputs, fast retrieval of a stored item, and few spurious memories (Hassoun & Watta, 1997). Differently from many traditional autoassociative memories, such as the Hopfield network, the autoassociative morphological memories (AMMs) introduced by Ritter and Sussner in the middle 1990s, exhibit unlimited absolute storage capacity and fast retrieval of a stored vector (Ritter & Sussner, 1996; Ritter, Sussner, & de Leon, 1998). Moreover, since they are based on lattice operations, they are computationally cheaper than many traditional autoassociative models. As to the error correction capability, the

* Corresponding author.
E-mail addresses: assantos@ufrb.edu.br (A.S. dos Santos), valle@ime.unicamp.br (M.E. Valle).

https://doi.org/10.1016/j.neunet.2018.01.013

original AMMs exhibit excellent tolerance to either dilative or erosive noise. Real-world applications of AMMs include classification of schizophrenia patients (Chyzhyk & Graña, 2015), restoration of historical documents (Lechuga-S et al., 2014; Valdiviezo, Urcid, & Lechuga, 2017), and hyperspectral image analysis (Grana & Chyzhyk, 2016).

Despite the notable features listed in the previous paragraph, the AMM models of Ritter and Sussner have a large number of spurious memories (Ritter & Gader, 2006; Sussner & Valle, 2006a). In 2014, Valle proposed an AMM, called max-plus projection autoassociative morphological memory (max-plus PAMM), which has less spurious memories than the original AMMs (Valle, 2014b). Briefly, the max-plus PAMM projects the input vector onto the set of all max-plus combinations of the stored vectors. Like the original AMMs, the max-plus PAMM exhibits unlimited absolute storage capacity and fast retrieval of a stored vector. As to the computational complexity, the max-plus PAMM does not require synthesizing a synaptic weight matrix. Furthermore, since the max-plus PAMM has less spurious memories, it exhibits a better tolerance with respect to dilative noise than the corresponding original AMM (Valle, 2014b).

In this paper, we present the dual of the max-plus PAMM, called min-plus PAMM. Also, we introduce four new autoassociative memories which are obtained by combining the max-plus and min-plus PAMMs. The new memories models, like the original AMMs, are all defined in terms of lattice-based operations from minimax algebra (Cuninghame-Green, 1979, 1995). Therefore, they belong to the lattice computing (LC) paradigm which, according to Kaburlasos and Kehagias (2014), Kaburlasos and Papakostas (2015) and Kaburlasos, Papadakis, and Papakostas (2013), is defined as an evolving collection of tools and mathematical modeling methodologies with the capacity to process lattice-ordered data *per se*, including logic values, numbers, sets, symbols, graphs, etc. We would like to point that the broad class of the fuzzy associative morphological memories (Sussner & Valle, 2008; Valle & Sussner, 2011), the theta-fuzzy associative memories (Esmi et al., 2015), and dendritic lattice associative memories (Ritter, Chyzhyk, Urcid, & Graña, 2012; Urcid, Ritter, & Valdiviezo-N, 2011, 2012) are also examples of associative memory models that belong to the LC paradigm.

Similar to the original AMMs and the new PAMM models, the four novel memory models exhibit unlimited absolute storage capacity and fast retrieval of a stored vector. Furthermore, they may exhibit better noise tolerance than the original AMMs and the new PAMM models. In this paper, we also investigate the noise tolerance of the new models theoretically. In addition, we address the application of the new models in classification problems from the literature (Alcala-Fdez, Fernández, Luengo, Derrac, & García, 2011; Lichman, 2013).

The paper is organized as follows: Next section provides the mathematical background necessary for dealing with morphological neural networks. Section 3 briefly reviews the original AMMs of Ritter and Sussner. The max-plus and min-plus PAMMs are discussed in Section 4 while their compositions are investigated in Section 5. In Section 6 we evaluate the performance of the proposed models in classification problems. The paper finishes with some concluding remarks in Section 7 and an appendix with the proofs of theorems.

2. Some mathematical background

The memory models considered in this paper are described by lattice-based operations borrowed from minimax algebra, a mathematical structure motivated by problems from scheduling theory, graph theory, and dynamic programming (Cuninghame-Green, 1979). Roughly speaking, the minimax algebra is developed in a mathematical structure obtained by enriching a complete lattice with two group operations (Sussner & Valle, 2006a). For the purposes of this paper, however, we consider the totally ordered field of real numbers (which is not a complete lattice) as the mathematical background. The supremum and the infimum of a bounded set $X \subseteq \mathbb{R}$ are denoted respectively by the symbols $\bigvee X$ and $\bigwedge X$. In case $X = \{x_1, \dots, x_n\} \subseteq \mathbb{R}$ is a finite set, the operations of computing the maximum and the minimum are written as $\bigvee_{j=1}^n x_j$ and $\bigwedge_{j=1}^n x_j$, respectively.

Given two matrices $A \in \mathbb{R}^{n\times k}$ and $B \in \mathbb{R}^{k\times m}$, the *max-product* and the *min-product* of A by B, denoted respectively by $C = A \boxdot B \in \mathbb{R}^{n\times m}$ and $D = A \boxtimes B \in \mathbb{R}^{n\times m}$, are given by the following equations for all indexes i and j:

$$c_{ij} = \bigvee_{\xi=1}^{k}(a_{i\xi} + b_{\xi j}) \quad \text{and} \quad d_{ij} = \bigwedge_{\xi=1}^{k}(a_{i\xi} + b_{\xi j}). \tag{1}$$

Note that the max-product satisfies

$$A \boxdot (B + \alpha) = (A \boxdot B) + \alpha, \quad \forall \alpha \in \mathbb{R}. \tag{2}$$

Here, $B + \alpha$ is the matrix obtained by adding α to each entry of B. Similarly, we have

$$A \boxtimes (B + \alpha) = (A \boxtimes B) + \alpha, \quad \forall \alpha \in \mathbb{R}. \tag{3}$$

In words, both lattice-based products are invariant under vertical translations.

The conjugate of $A \in \mathbb{R}^{n\times k}$ is the matrix $A^* \in \mathbb{R}^{k\times n}$ whose entries satisfy

$$a^*_{ij} = -a_{ji}, \quad \forall i, j. \tag{4}$$

Note that $(A^*)^* = A$ for any matrix A. The conjugate can be used to establish the following identities concerning the min-product and the max-product:

$$(A \boxtimes B)^* = B^* \boxdot A^* \quad \text{and} \quad (A \boxdot B)^* = B^* \boxtimes A^*. \tag{5}$$

Apart from (5), the lattice-based matrix operations are related by means of the following adjunction relationship for matrices $A \in \mathbb{R}^{n\times k}$, $B \in \mathbb{R}^{k\times m}$, and $C \in \mathbb{R}^{n\times m}$:

$$A \boxdot B \leq C \Leftrightarrow B \leq A^* \boxtimes C \Leftrightarrow A \leq C \boxtimes B^*. \tag{6}$$

In analogy to the notion of linear combination, a *max-plus combination* of vectors from a set $\mathcal{X} = \{\mathbf{x}^1, \dots, \mathbf{x}^k\} \subseteq \mathbb{R}^n$ is any vector $\mathbf{a} \in \mathbb{R}^n$ of the form

$$\mathbf{a} = \bigvee_{\xi=1}^{k}(\alpha_\xi + \mathbf{x}^\xi), \quad \alpha_\xi \in \mathbb{R}. \tag{7}$$

In words, $\mathbf{a}$ is the maximum of vertical translations of $\mathbf{x}^1, \dots, \mathbf{x}^k$. The set of all max-plus combinations of vectors from $\mathcal{X}$ is denoted by $\mathfrak{A}(\mathcal{X})$, i.e.,

$$\mathfrak{A}(\mathcal{X}) = \left\{ \mathbf{a} \in \mathbb{R}^n : \mathbf{a} = \bigvee_{\xi=1}^{k}(\alpha_\xi + \mathbf{x}^\xi), \alpha_\xi \in \mathbb{R} \right\}. \tag{8}$$

Note that $\mathbf{a} \in \mathfrak{A}(\mathcal{X})$ if and only if $\mathbf{a} = X \boxdot \boldsymbol{\alpha}$ for some $\boldsymbol{\alpha} = [\alpha_1, \dots, \alpha_k]^T \in \mathbb{R}^n$, where $X = [\mathbf{x}^1, \dots, \mathbf{x}^k] \in \mathbb{R}^{n\times k}$ is the matrix whose columns corresponds to the vectors of $\mathcal{X}$.

Dually, a min-plus combination of $\mathbf{x}^1, \dots, \mathbf{x}^k$ is any vector $\mathbf{b} \in \mathbb{R}^n$ defined by

$$\mathbf{b} = \bigwedge_{\xi=1}^{k}(\beta_\xi + \mathbf{x}^\xi), \quad \beta_\xi \in \mathbb{R}. \tag{9}$$