IEICE TRANSACTIONS

on Fundamentals of Electronics, Communications and Computer Sciences

DOI:10.1587/transfun.2024MAL0001

Publicized:2024/10/08

This advance publication article will be replaced by the finalized version after proofreading.

A PUBLICATION OF THE ENGINEERING SCIENCES SOCIETY

The Institute of Electronics, Information and Communication Engineers

Kikai-Shinko-Kaikan Bldg., 5-8, Shibakoen 3chome, Minato-ku, TOKYO, 105-0011 JAPAN

LETTER *Special Section on Mathematical Systems Science and its Applications*

Quaternionic Vector Product Hopfield Network

Masaki KOBAYASHI[†], ***Member***

SUMMARY A vector product Hopfield network (VPHN) is a 3-D Hopfield model. A VPHN provides excellent noise tolerance. Since projection rule is not available, it is not applicable for large numbers of data. Vector product appears in the multiplication of quaternions. In this work, the VPHNs are extended to the quaternionic VPHNs (QVPHNs). The projection rule is available for the QVPHNs. We evaluate the QVPHNs by computer simulations and show that they provide excellent noise tolerance.
key words: *Hopfield networks, vector product, high-dimensional neural networks, quaternion*

1. Introduction

In recent years, models of neural networks have been extended to the high-dimensional ones. Complex-valued MLPs (CVMLPs) have been proposed for the 2-D extensions [1]. The hyperbolic-valued MLP is another model of 2-D MLP [2]. CVMLPs were extended to 4-D MLPs using quaternions [3]. In 3-D MLPs, several models have been attempted. Nitta proposed the 3-D MLPs using rotation matrices [4]. He also proposed the 3-D MLPs using vector product [5]. Arena et al. achieved the 3-D rotational MLPs using quaternions [6].

We describe high-dimensional models of Hopfield networks (HNs). Complex-valued HNs (CVHNs) have been used for multistate associative memories [7]-[10]. Hyperbolic-valued HNs (HVHNs) provide better noise tolerance than CVHNs [11]. Quaternion-valued HNs (QVHNs), 4-D HN models, have also been proposed [12]. A few 3-D HN models have been proposed. In the Hopfield version of [6], major learning algorithms, such as the hebbian learning and projection rules, are not applicable [13]. A vector product HN (VPHN) is a Hopfield version of [5], and is expected to provide excellent noise tolerance [14]. However, it has the following difficulties.

1. Projection rule is not available.
2. The resolution factor is limited.

The projection rule is a practical learning algorithm, since it is a one-shot learning algorithm with the large storage capacity. Projection rules for CVHNs, HVHNs, and QVHNs are provided in [10], [11], and [12], respectively. This difficulty is attributed to the operation, that is, algebra of 3-D vectors with vector product is not associative. Since the activation function is defined based on the vertices of regular polyhedrons or platonic solids, the resolution factor is limited to 4, 6, 8, 12, and 20. The present work solves the first difficulty. Vector product appears in multiplication of quaternions. In particular, algebra of quaternions is associative. We extend the VPHNs to quaternionic VPHN (QVPHN) so that projection rule is available. The noise tolerance is evaluated by computer simulations, and it is shown that the QVPHNs provide excellent noise tolerance.

[†]The author is with Graduate School of Informatices, Nagoya University, Nagoya, 464-8601, Japan.

Name	Symbol	Result
scalar product	$\vec{q}_1 \cdot \vec{q}_2$	real number
vector product	$\vec{q}_1 \times \vec{q}_2$	vector
quaternion multiplication	$\vec{q}_1 \circ \vec{q}_2$	quaternion

Table 1 multiplications of vectors

2. Quaternions

A quaternion is represented as

$$q = \hat{q} + \vec{q} \tag{1}$$

using a real number $\hat{q}$ and a 3-D vector $\vec{q}$. For two quaternions $q_1 = \hat{q}_1 + \vec{q}_1$ and $q_2 = \hat{q}_2 + \vec{q}_2$, the addition and multiplications are defined by

$$q_1 + q_2 = (\hat{q}_1 + \hat{q}_2) + (\vec{q}_1 + \vec{q}_2), \tag{2}$$

$$q_1 \circ q_2 = (\hat{q}_1 \hat{q}_2 - \vec{q}_1 \cdot \vec{q}_2) + (\hat{q}_1 \vec{q}_2 + \hat{q}_2 \vec{q}_1 + \vec{q}_1 \times \vec{q}_2). \tag{3}$$

The symbol $\circ$ is introduced to distinguish quaternion multiplication from scalar and vector products. From (3), the quaternion multiplication of vectors is

$$\vec{q}_1 \circ \vec{q}_2 = -\vec{q}_1 \cdot \vec{q}_2 + \vec{q}_1 \times \vec{q}_2. \tag{4}$$

Three multiplications of vectors are listed in Table 2. $\mathrm{Re}(q) = \hat{q}$ is referred to as the real part of q, and the equality

$$\mathrm{Re}\,(q_1 \circ q_2) = \mathrm{Re}\,(q_2 \circ q_1) = -\vec{q}_1 \cdot \vec{q}_2 \tag{5}$$

holds. Algebra of quaternions is a noncommutative field, that is, distributive and associative laws are satisfied, and all elements except 0 are invertible. The conjugate of q is defined by

$$\overline{q} = \hat{q} - \vec{q}, \tag{6}$$

then, the important equalities

$$\overline{q_1 \circ q_2} = \overline{q_2} \circ \overline{q_1} \tag{7}$$

$$\mathrm{Re}(\overline{q}) = \mathrm{Re}(q) \quad (8)$$

hold.

3. Quaternionic Vector Product Hopfield Neural Networks

In the QVPHNs, the neuron states and weights are 3-D vectors and quaternions, respectively. The state of neuron i and weight from neuron j to neuron i are denoted by $\vec{v}_i$ and w_{ij}, respectively. The weights have to satisfy

$$\overline{w_{ij}} = w_{ji}, \quad (9)$$
$$w_{ii} = 0. \quad (10)$$

A VPHN is a restricted QVPHN by $\hat{w}_{ij} = 0$, and a QVPHN is an extension of a VPHN. From (7)-(9), the following equality is obtained.

$$\mathrm{Re}\left(\vec{v}_i \circ w_{ij} \circ \vec{v}_j\right) = \mathrm{Re}\left(\overline{\vec{v}_i \circ w_{ij} \circ \vec{v}_j}\right) \quad (11)$$
$$= \mathrm{Re}\left((-\vec{v}_j) \circ \overline{w_{ij}} \circ (-\vec{v}_i)\right) \quad (12)$$
$$= \mathrm{Re}\left(\vec{v}_j \circ w_{ji} \circ \vec{v}_i\right) \quad (13)$$

Let V be the set of neuron states. We employ V given by [14], that is, V is vertices of a regular polyhedron inscribed inside the unit sphere. Thus, the cardinality of V, that is the resolution factor, is restricted to $K = 4, 6, 8, 12$ and 20. The weighted sum input to neuron i is defined as

$$S_i = \sum_{j=1}^{N} w_{ij} \circ \vec{v}_j, \quad (14)$$

where N is the number of neurons. The activation function is defined as

$$f(S_i) = \arg\min_{\vec{v}\in V} \mathrm{Re}\left(\vec{v} \circ S_i\right) = \arg\max_{\vec{v}\in V} \vec{v} \cdot \vec{S}_i. \quad (15)$$

We should note that the activation function ignores the real part of S_i. We suppose that the neuron state maintains if multiple $\vec{v}$'s maximize $\vec{v} \cdot \vec{S}_i$. For $\vec{u}_i = f(S_i) \neq \vec{v}_i$, the definition (15) implies

$$\mathrm{Re}\left(\vec{u}_i \circ S_i\right) < \mathrm{Re}\left(\vec{v}_i \circ S_i\right). \quad (16)$$

The energy of QVPHN is defined as

$$E = \frac{1}{2}\sum_{i,j} \mathrm{Re}\left(\vec{v}_i \circ w_{ij} \circ \vec{v}_j\right). \quad (17)$$

We prove that E decreases using (13) and (16). Suppose that neuron k is updated from $\vec{v}_k$ to $\vec{u}_k \neq \vec{v}_k$, then the energy change is

$$\Delta E = \frac{1}{2}\sum_{i\neq k} \mathrm{Re}\left(\vec{u}_k \circ w_{ki} \circ \vec{v}_i + \vec{v}_i \circ w_{ik} \circ \vec{u}_k\right)$$
$$-\frac{1}{2}\sum_{i\neq k} \mathrm{Re}\left(\vec{v}_k \circ w_{ki} \circ \vec{v}_i + \vec{v}_i \circ w_{ik} \circ \vec{v}_k\right) \quad (18)$$
$$= \sum_{i\neq k} \mathrm{Re}\left(\vec{u}_k \circ w_{ki} \circ \vec{v}_i\right)$$
$$- \sum_{i\neq k} \mathrm{Re}\left(\vec{v}_k \circ w_{ki} \circ \vec{v}_i\right) \quad (19)$$
$$= \mathrm{Re}\left(\vec{u}_k \circ S_k\right) - \mathrm{Re}\left(\vec{v}_k \circ S_k\right) < 0. \quad (20)$$

Thus, a QVPHN converges to a fixed point.

Projection rule for QVHNs is available for QVPHNs [12]. Let $\left\{\mathbf{v}^p = (\vec{v}_1^p, \vec{v}_2^p, \cdots, \vec{v}_N^p)^{\mathrm{T}}\right\}_{p=1}^{P}$ be the training set, where we should note that the vectors $\vec{v}_i^p$'s are regarded as quaternions. P is the number of training patterns. We describe the projection rule;

$$Z = \left(\mathbf{v}^1, \mathbf{v}^2, \cdots, \mathbf{v}^P\right), \quad (21)$$
$$X = Z(Z^*Z)^{-1}Z^*, \quad (22)$$
$$W = X - \mathrm{diag}X \quad (23)$$

Z is the (N, P) quaternion matrix, referred to as the training matrix, whose column vectors are training vectors. Z^* is the Hermitian transpose of Z. The (i, j) element of W is employed by w_{ij}. Projection rule requires the existence of $(Z^*Z)^{-1}$ and the storage capacity is N. $XZ = Z(Z^*Z)^{-1}Z^*Z = Z$ implies $X\mathbf{v}^p = \mathbf{v}^p$. Thus, the projection rule fixes the training vectors. The diagonal elements of X correspond to the self-feedbacks of neurons. The removal of self-feedbacks is often effective for noise tolerance.

4. Computer Simulation

We conduct computer simulations to evaluate the noise tolerance of QVPHNs. It is compared with the noise tolerance of CVHNs and HVHNs. The CVHN is a prototype of multistate Hopfield network. The HVHN is a major model which provides much better noise tolerance than the CVHN. The HVHNs employ noise robust projection rule [11]. The number of neurons is fixed to $N = 200$. Since the resolution factor of QVPHNs is limited to 4, 6, 8, 12, and 20, it is fixed to the largest one 20. For 100 training pattern sets generated randomly, 100 trials are attempted; total of trials is 10000. The procedure of trial is described below.

1. A pattern is randomly selected from the training patterns.
2. Impulsive noise is added on the selected training pattern. The noise rate is denoted as r.
3. If the original training pattern is retrieved, the trial is regarded as successful.

The simulation results are shown in Fig. 1. For $P > 60$, the results of CVHNs and HVHNs are identical. The results of QVPHNs are much better than those of the others. The reasons why the VPHNs provide better performance than the CVHNs was discussed in [14], that is, the VPHNs have no rotational invariance and the neuron states are assigned to 3-D space. In CVHNs, a training vector $\mathbf{v}^p$ is a complex vector and the rotated patterns $\alpha\mathbf{v}^p$ with $|\alpha| = 1$ are also fixed vectors. This fact is referred to as rotational invariance. Rotational invariance increases pseudomemories which deteriorate the noise tolerance. The reasons for VPHNs would

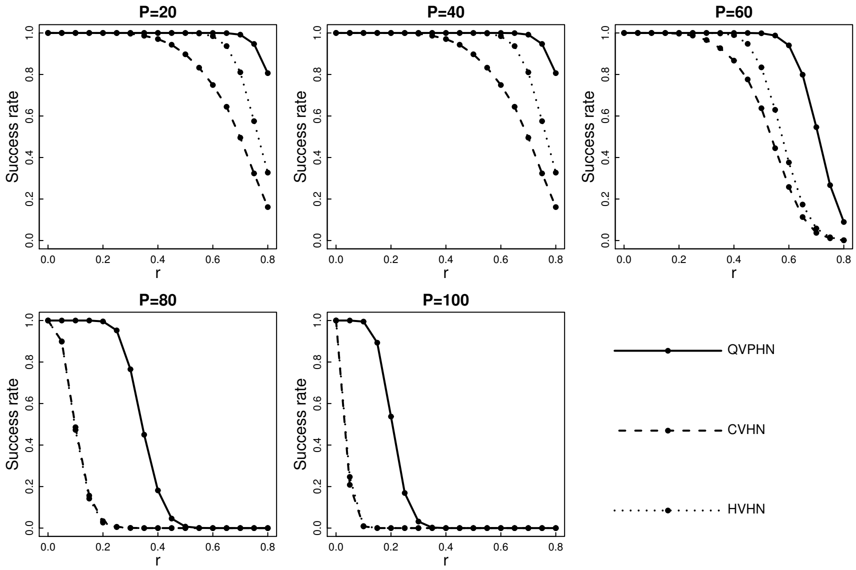

Fig. 1 The noise tolerance of QVPHN, CVHN, and HVHN was compared by computer simulations. The number of neurons and resolution factor were $N = 200$ and $K = 20$, respectively.

be applicable to the QVPHNs.

5. Conclusion

A VPHN is a 3-D Hopfield model using vector product. It provides excellent noise tolerance, though it has two difficulties for practical applications, projection rule and resolution factors. One is that projection rule is not available for the VPHNs. To process large number of training data, projection rule is necessary. The other is the limited resolution factor. In the present work, a VPHN is extended to a QVPHN, and the first problem is solved by extending non-associative algebra of 3-D vectors to associative algebra of quaternions. Computer simulations show that a QVPHN provides excellent noise tolerance. The restriction to resolution factor is the remaining difficulty to apply the QVPHNs to real data, and the activation functions should be improved.

In the present paper, the proposed model is studied for high-dimensional HN models. However, application models are not limited to HNs, we should also note that the present idea would be applicable to MLP models. Then, an extension of [5] would be obtained.

References

[1] T. Nitta, "An extension of the back-propagation algorithm to complex numbers," Neural Networks, vol. 10, no. 8, pp. 1392-1415, 1997.
[2] S.Buchholz, and G.Sommer, "A hyperbolic multilayer perceptron," Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks, pp. 129-133, 2000.
[3] T.Nitta, "An extension of the back-propagation algorithm to quaternions," Proceedings of International Conference on Neural Information Processing, vol. 1, pp. 247-250, 1996.
[4] T.Nitta, "Three-dimensional vector valued neural network and its generalization ability," Neural Information Processing – Letters and Reviews, vol. 10, pp. 237-242, 2006.
[5] T.Nitta, "A back-propagation algorithm for neural networks based on 3d vector product," Proceedings of IEEE/INNS International Joint Conference on Neural Networks, vol. 1, pp. 589-592, 1993.
[6] P.Arena, L.Fortuna, G.Muscato, and M.G.Xibilia, "Multilayer perceptrons to approximate quaternion valued functions," Neural Networks, vol. 10, no. 2, pp. 335-342, 1997.
[7] S.Jankowski, A.Lozowski, and J.M.Zurada, "Complex-valued multistate neural associative memory," IEEE Transactions on Neural Networks, vol. 7, no. 6, pp. 1491-149, 1996.
[8] M.K.Muezzinoglu, C.Guzelis, and J.M.Zurada, "A new design method for the complex-valued multistate hopfield associative memory," IEEE Transactions on Neural Networks, vol. 14, no. 4, pp. 891-899., 2003.
[9] G.Tanaka, and K.Aihara, "Complex-valued multistate associative memory with nonlinear multilevel functions for gray-level image reconstruction," IEEE Transactions on Neural Networks, vlol. 20, no. 9, pp. 1463-1473, 2009.
[10] M. Kobayashi, "Two-Level Complex-Valued Hopfield Neural Networks," IEEE Transactions on Neural Networks and Learning Systems, vol. 32, no. 5, pp. 2274-2278, 2021.
[11] M.Kobayashi, "Noise robust projection rule for hyperbolic Hopfield neural networks," IEEE Transactions on Neural Networks and Learning Systems, vol. 31, no. 1, pp. 352-356, 2019.
[12] M.Kobayashi, "Quaternionic Hopfield neural networks with twin-multistate activation function," Neurocomputing, vol. 267, pp. 304-310, 2017.
[13] M. Kobayashi, "Three-Dimensional Quaternionic Hopfield Neural Networks," IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, vol. 100, no. 7, pp. 1575-1577, 2017.
[14] M.Kobayashi, "Multistate vector product Hopfield neural networks," Neurocomputing, 272 (2018) 425-431.