

Properties of an Invariant Set of Weights of Perceptrons

Charlotte Yuk-Fan Ho, Bingo Wing-Kuen Ling, Muhammad H U Nasir, Hak-Keung Lam and Herbert H. C. Iu

Abstract—In this paper, the dynamics of weights of perceptrons are investigated based on the perceptron training algorithm. In particular, the condition that the system map is not injective is derived. Based on the derived condition, an invariant set that results to a bijective invariant map is characterized. Also, it is shown that some weights outside the invariant set will be moved to the invariant set. Hence, the invariant set is attracting. Computer numerical simulation results on various perceptrons with exhibiting various behaviors, such as fixed point behaviors, limit cycle behaviors and chaotic behaviors, are illustrated.

I. INTRODUCTION

PERCEPTRONS are single neurons that produce either “1” or “-1” at the output via applying the single bit quantization operation on the inner product of the weights and the inputs of the perceptrons. Because the number of the possible values of the output of the perceptrons is finite and the implementation cost of the perceptrons is low, perceptrons are widely employed in many pattern recognition systems [1].

The weights of the perceptrons are usually obtained via the perceptron training algorithm [2]. As the update rule of the weights of the perceptrons is based on the functional values of the single bit quantization function of the inner product of the weights and the inputs of the perceptrons, in which the inputs of the perceptrons are periodically time varying with the period equal to the number of the training feature vectors, the dynamics of the weights of the perceptrons are actually governed by a nonlinear time varying map that maps from the set of a high dimensional real space to itself. By downsampling the weights of the perceptrons with the sampling rate equal to the number of the training feature vectors, the mapping from the set of the high dimensional real space to itself becomes a nonlinear time invariant map. An invariant set of the weights of the perceptrons is defined as a set of the downsampled weights that maps to itself. However, as the perceptrons are nonlinear time varying systems, the shape of an invariant set of the weights of the perceptrons could be very irregular. Up to now, very few results have been reported on the characterization of an invariant set of the weights of the perceptrons. In this paper, properties of an invariant set of the weights of the perceptrons are studied.

It is worth noting that conventional perceptrons are usually operated with linear separable training feature vectors because the perceptron convergence theorem [2] guarantees that the weights of the perceptrons will converge

to a fixed point in the phase space when the set of the training feature vectors is linearly separable. However, an invariant set of the weights of the perceptrons may also exist even though the perceptrons exhibit limit cycle and chaotic behaviors. Although some works have been done on the investigation of the limit cycle behaviors of the weights of the perceptrons [5], these results are not applied if the weights of the perceptrons exhibit chaotic behaviors.

The outline of this paper is as follows. Notations used throughout this paper are introduced in Section II. In Section III, properties of an invariant set of the weights of the perceptrons are discussed. In Section IV, computer numerical simulation results on various perceptrons with exhibiting various behaviors, such as fixed point behaviors, limit cycle behaviors and chaotic behaviors, are illustrated. Finally, a conclusion is drawn in Section V.

II. NOTATIONS

Denote N as the number of the training feature vectors and d as the dimension of these training feature vectors. Denote the elements in the training feature vectors as $x_i(k)$ for $i=1,2,\dots,d$ and for $k=0,1,\dots,N-1$. Define $\mathbf{x}(k) \equiv [1, x_1(k), \dots, x_d(k)]^T$ for $k=0,1,\dots,N-1$. In this paper, we assume that $\mathbf{x}(k) \neq \mathbf{0}$ for $k=0,1,\dots,N-1$. Define $\mathbf{x}(Nn+k) \equiv \mathbf{x}(k) \forall n \in \mathbb{Z} \setminus \{0\}$ and for $k=0,1,\dots,N-1$. Denote the weights of the perceptron as $w_i(n)$ for $i=1,2,\dots,d$ and $\forall n \in \mathbb{Z}$. Denote the threshold of the perceptron as $w_0(n) \forall n \in \mathbb{Z}$ and the activation function of the perceptron as $Q(z) \equiv \begin{cases} 1 & z \geq 0 \\ -1 & z < 0 \end{cases}$. Define $\mathbf{w}(n) \equiv [w_0(n), w_1(n), \dots, w_d(n)]^T \forall n \in \mathbb{Z}$ and denote the output of the perceptron as $y(n) \forall n \in \mathbb{Z}$, then $y(n) = Q(\mathbf{w}^T(n)\mathbf{x}(n)) \forall n \in \mathbb{Z}$. Denote the desired output of the perceptron corresponding to $\mathbf{x}(n)$ as $t(n) \forall n \in \mathbb{Z}$. Assume that the perceptron training algorithm [2] is employed for the training, so the forward dynamics of the weights of the perceptron is governed by the following equation:

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \frac{t(n) - Q(\mathbf{w}^T(n)\mathbf{x}(n))}{2} \mathbf{x}(n) \quad \forall n \in \mathbb{Z}. \quad (1)$$

Define $\tilde{\mathfrak{F}}_k^F : \mathfrak{R}^{d+1} \rightarrow \mathfrak{R}^{d+1}$ such that $\tilde{\mathfrak{F}}_k^F(\mathbf{w}(k)) \equiv \mathbf{w}(k+1) \forall k \in \mathbb{Z}$ and $\tilde{\mathfrak{F}}_k^B : \mathfrak{R}^{d+1} \rightarrow \mathfrak{R}^{d+1}$ such that

$$\tilde{\mathfrak{Z}}_k^B(\mathbf{w}(k)) \equiv \mathbf{w}(k) - \frac{t(k-1) - Q(\mathbf{w}^T(k)\mathbf{x}(k-1))}{2} \mathbf{x}(k-1) \quad \forall k \in Z. \quad (2)$$

Denote $\mathbf{w}'(0)$ as another initial weight. Define $y'(j) \equiv Q(\mathbf{w}'^T(j)\mathbf{x}(j))$ and $\mathbf{w}'(j+1) \equiv \tilde{\mathfrak{Z}}_j^F(\mathbf{w}'(j)) \quad \forall j \in Z$. Suppose that $\exists k \in Z$ such that $\mathbf{w}'(k) = \mathbf{w}(k)$.

A set S is called an invariant set under a system map T if $T(S) = S$. Denote the absolute value of a real number as $|\cdot|$ and the 2-norm of a vector as $\|\mathbf{v}\| \equiv \sqrt{\sum_{i=0}^d v_i^2}$, where $\mathbf{v} \equiv [v_0, \dots, v_d]^T$.

III. PROPERTIES OF AN INVARIANT SET OF THE WEIGHTS OF THE PERCEPTRON

In this section, an invariant set of the weights of the perceptron is defined and its properties are studied. However, before we study properties of an invariant set of the weights of the perceptron, further discussions on exploring the relationship between $\tilde{\mathfrak{Z}}_k^F: \mathfrak{R}^{d+1} \rightarrow \mathfrak{R}^{d+1}$ and $\tilde{\mathfrak{Z}}_k^B: \mathfrak{R}^{d+1} \rightarrow \mathfrak{R}^{d+1} \quad \forall k \in Z$, finding the values of $y'(j)$ for $j = 0, 1, \dots, k-1$, and characterizing the injective property of $\tilde{\mathfrak{Z}}_k^F: \mathfrak{R}^{d+1} \rightarrow \mathfrak{R}^{d+1}$ for $k \in Z$ are presented in Lemma 1, Lemma 2 and Lemma 3, respectively. These lemmas are useful for deriving properties of an invariant set of the weights of the perceptron.

Lemma 1

$$\tilde{\mathfrak{Z}}_{k-1}^F(\tilde{\mathfrak{Z}}_k^B(\mathbf{w}(k))) = \mathbf{w}(k) \quad \forall k \in Z.$$

Proof:

$$\begin{aligned} & \tilde{\mathfrak{Z}}_{k-1}^F(\tilde{\mathfrak{Z}}_k^B(\mathbf{w}(k))) \\ &= \tilde{\mathfrak{Z}}_{k-1}^F\left(\mathbf{w}(k) - \frac{t(k-1) - Q(\mathbf{w}^T(k)\mathbf{x}(k-1))}{2} \mathbf{x}(k-1)\right) \\ &= \begin{cases} \tilde{\mathfrak{Z}}_{k-1}^F(\mathbf{w}(k)) & t(k-1) = Q(\mathbf{w}^T(k)\mathbf{x}(k-1)) \\ \tilde{\mathfrak{Z}}_{k-1}^F(\mathbf{w}(k) - \mathbf{x}(k-1)) & t(k-1) = 1 \text{ and } Q(\mathbf{w}^T(k)\mathbf{x}(k-1)) = -1 \\ \tilde{\mathfrak{Z}}_{k-1}^F(\mathbf{w}(k) + \mathbf{x}(k-1)) & t(k-1) = -1 \text{ and } Q(\mathbf{w}^T(k)\mathbf{x}(k-1)) = 1 \end{cases} \\ &= \begin{cases} \mathbf{w}(k) + \frac{t(k-1) - Q(\mathbf{w}^T(k)\mathbf{x}(k-1))}{2} \mathbf{x}(k-1) & t(k-1) = Q(\mathbf{w}^T(k)\mathbf{x}(k-1)) \\ \mathbf{w}(k) - \mathbf{x}(k-1) + \frac{t(k-1) - Q(\mathbf{w}^T(k)\mathbf{x}(k-1)) - \|\mathbf{x}(k-1)\|^2}{2} \mathbf{x}(k-1) & t(k-1) = 1 \text{ and } Q(\mathbf{w}^T(k)\mathbf{x}(k-1)) = -1 \\ \mathbf{w}(k) + \mathbf{x}(k-1) + \frac{t(k-1) - Q(\mathbf{w}^T(k)\mathbf{x}(k-1)) + \|\mathbf{x}(k-1)\|^2}{2} \mathbf{x}(k-1) & t(k-1) = -1 \text{ and } Q(\mathbf{w}^T(k)\mathbf{x}(k-1)) = 1 \end{cases} \\ &= \begin{cases} \mathbf{w}(k) & t(k-1) = Q(\mathbf{w}^T(k)\mathbf{x}(k-1)) \\ \mathbf{w}(k) - \mathbf{x}(k-1) + \mathbf{x}(k-1) & t(k-1) = 1 \text{ and } Q(\mathbf{w}^T(k)\mathbf{x}(k-1)) = -1 \\ \mathbf{w}(k) + \mathbf{x}(k-1) - \mathbf{x}(k-1) & t(k-1) = -1 \text{ and } Q(\mathbf{w}^T(k)\mathbf{x}(k-1)) = 1 \end{cases} \\ &= \mathbf{w}(k) \end{aligned}$$

$\forall k \in Z$. This completes the proof.

It is trivial to see that $\tilde{\mathfrak{Z}}_k^F: \mathfrak{R}^{d+1} \rightarrow \mathfrak{R}^{d+1} \quad \forall k \in Z$ governs the forward dynamics of the weights of the perceptron. On the other hand, Lemma 1 reveals that one of the possible backward dynamics of the weights of the perceptron could be governed by $\tilde{\mathfrak{Z}}_k^B: \mathfrak{R}^{d+1} \rightarrow \mathfrak{R}^{d+1} \quad \forall k \in Z$ even though the time index of the weights in the activation function in $\tilde{\mathfrak{Z}}_k^B: \mathfrak{R}^{d+1} \rightarrow \mathfrak{R}^{d+1} \quad \forall k \in Z$ is not equal to $k-1$. Although $\tilde{\mathfrak{Z}}_{k-1}^F(\tilde{\mathfrak{Z}}_k^B(\mathbf{w}(k))) = \mathbf{w}(k) \quad \forall k \in Z$, it is worth noting that the inverse of $\tilde{\mathfrak{Z}}_k^F: \mathfrak{R}^{d+1} \rightarrow \mathfrak{R}^{d+1} \quad \forall k \in Z$ may not exist. The

details will be explained in Lemma 3.

Lemma 2

$$y'(0) = Q\left(\mathbf{w}'^T(0)\mathbf{x}(0) + \frac{1}{2} \begin{bmatrix} Q(\mathbf{w}^T(0)\mathbf{x}(0)) \\ \vdots \\ Q(\mathbf{w}^T(k-1)\mathbf{x}(k-1)) \end{bmatrix} - \begin{bmatrix} Q(\mathbf{w}^T(0)\mathbf{x}(0)) \\ \vdots \\ Q(\mathbf{w}^T(k-1)\mathbf{x}(k-1)) \end{bmatrix}\right) \begin{bmatrix} \mathbf{x}^T(0)\mathbf{x}(0) \\ \vdots \\ \mathbf{x}^T(k-1)\mathbf{x}(0) \end{bmatrix}$$

and

$$y'(j) = Q\left(\mathbf{w}'^T(0)\mathbf{x}(j) + \frac{1}{2} \begin{bmatrix} t(0) \\ \vdots \\ t(j-1) \\ Q(\mathbf{w}^T(j)\mathbf{x}(j)) \\ \vdots \\ Q(\mathbf{w}^T(k-1)\mathbf{x}(k-1)) \end{bmatrix} - \begin{bmatrix} Q(\mathbf{w}^T(0)\mathbf{x}(0)) \\ \vdots \\ Q(\mathbf{w}^T(k-1)\mathbf{x}(k-1)) \end{bmatrix}\right) \begin{bmatrix} \mathbf{x}^T(0)\mathbf{x}(j) \\ \vdots \\ \mathbf{x}^T(k-1)\mathbf{x}(j) \end{bmatrix}$$

for $j = 1, 2, \dots, k-1$.

Proof:

Since $y'(j) \equiv Q(\mathbf{w}'^T(j)\mathbf{x}(j))$ and $\mathbf{w}'(j+1) \equiv \tilde{\mathfrak{Z}}_j^F(\mathbf{w}'(j)) \quad \forall j \in Z$ as well as $\exists k \in Z$ such that $\mathbf{w}'(k) = \mathbf{w}(k)$, we have

$$\begin{aligned} & \mathbf{w}'(0) + \frac{1}{2} [\mathbf{x}(0), \dots, \mathbf{x}(k-1)] \begin{bmatrix} t(0) - Q(\mathbf{w}^T(0)\mathbf{x}(0)) \\ \vdots \\ t(k-1) - Q(\mathbf{w}^T(k-1)\mathbf{x}(k-1)) \end{bmatrix} \\ &= \mathbf{w}(0) + \frac{1}{2} [\mathbf{x}(0), \dots, \mathbf{x}(k-1)] \begin{bmatrix} t(0) - Q(\mathbf{w}^T(0)\mathbf{x}(0)) \\ \vdots \\ t(k-1) - Q(\mathbf{w}^T(k-1)\mathbf{x}(k-1)) \end{bmatrix} \end{aligned}$$

This further implies that

$$\mathbf{w}'(0) = \mathbf{w}(0) + \frac{1}{2} [\mathbf{x}(0), \dots, \mathbf{x}(k-1)] \begin{bmatrix} Q(\mathbf{w}^T(0)\mathbf{x}(0)) - Q(\mathbf{w}^T(0)\mathbf{x}(0)) \\ \vdots \\ Q(\mathbf{w}^T(k-1)\mathbf{x}(k-1)) - Q(\mathbf{w}^T(k-1)\mathbf{x}(k-1)) \end{bmatrix}$$

and

$$\mathbf{w}'(j) = \mathbf{w}(0) + \frac{1}{2} [\mathbf{x}(0), \dots, \mathbf{x}(k-1)] \begin{bmatrix} t(0) \\ \vdots \\ t(j-1) \\ Q(\mathbf{w}^T(j)\mathbf{x}(j)) \\ \vdots \\ Q(\mathbf{w}^T(k-1)\mathbf{x}(k-1)) \end{bmatrix} - \begin{bmatrix} Q(\mathbf{w}^T(0)\mathbf{x}(0)) \\ \vdots \\ Q(\mathbf{w}^T(k-1)\mathbf{x}(k-1)) \end{bmatrix}$$

for $j = 1, 2, \dots, k-1$. As $y'(j) \equiv Q(\mathbf{w}'^T(j)\mathbf{x}(j)) \quad \forall j \in Z$, the result follows directly.

To evaluate $y'(k-1)$, as

$$y'(k-1) = Q\left(\mathbf{w}'^T(0)\mathbf{x}(k-2) + \frac{1}{2} \begin{bmatrix} t(0) \\ \vdots \\ t(k-2) \\ Q(\mathbf{w}^T(k-1)\mathbf{x}(k-1)) \end{bmatrix} - \begin{bmatrix} Q(\mathbf{w}^T(0)\mathbf{x}(0)) \\ \vdots \\ Q(\mathbf{w}^T(k-1)\mathbf{x}(k-1)) \end{bmatrix}\right) \begin{bmatrix} \mathbf{x}^T(0)\mathbf{x}(k-2) \\ \vdots \\ \mathbf{x}^T(k-1)\mathbf{x}(k-2) \end{bmatrix}$$

it can be seen easily that the above equation is satisfied if $y'(k-1) = y(k-1)$. However, the above equation may also be satisfied when $y'(k-1) = -y(k-1)$. Once all the possible values of $y'(k-1)$ are determined, then $y'(k-2)$ can also be determined as follows. As

$$y'(k-2) = Q\left(\mathbf{w}'^T(0)\mathbf{x}(k-2) + \frac{1}{2} \begin{bmatrix} t(0) \\ \vdots \\ t(k-3) \\ Q(\mathbf{w}^T(k-2)\mathbf{x}(k-2)) \\ Q(\mathbf{w}^T(k-1)\mathbf{x}(k-1)) \end{bmatrix} - \begin{bmatrix} Q(\mathbf{w}^T(0)\mathbf{x}(0)) \\ \vdots \\ Q(\mathbf{w}^T(k-1)\mathbf{x}(k-1)) \end{bmatrix}\right) \begin{bmatrix} \mathbf{x}^T(0)\mathbf{x}(k-2) \\ \vdots \\ \mathbf{x}^T(k-1)\mathbf{x}(k-2) \end{bmatrix}$$

all possible values of $y'(k-1)$ have already been determined and $y'(k-2) \in \{-1, 1\}$, all possible values of $y'(k-2)$ could be determined accordingly. Similarly, all possible values of $y'(j)$ for $j = 0, 1, \dots, k-1$ could be determined accordingly.

Lemma 3

Assume that $\|\mathbf{x}(k)\|^2 \neq |\mathbf{w}^T(k)\mathbf{x}(k)|$. Then $\tilde{\mathfrak{Z}}_k^F$ is not injective if and only if $\|\mathbf{x}(k)\|^2 > |\mathbf{w}^T(k)\mathbf{x}(k)|$.

Proof:

For the necessity, $\|\mathbf{x}(k)\|^2 > |\mathbf{w}^T(k)\mathbf{x}(k)|$ implies that $\mathbf{x}^T(k)\mathbf{x}(k) > \mathbf{w}^T(k)\mathbf{x}(k)$ for $\mathbf{w}^T(k)\mathbf{x}(k) \geq 0$ and $\mathbf{x}^T(k)\mathbf{x}(k) > -\mathbf{w}^T(k)\mathbf{x}(k)$ for $\mathbf{w}^T(k)\mathbf{x}(k) < 0$. This implies that $Q(\mathbf{w}^T(k)\mathbf{x}(k)) = -Q(\mathbf{w}^T(k)\mathbf{x}(k) - \mathbf{x}^T(k)\mathbf{x}(k))$ for $\mathbf{w}^T(k)\mathbf{x}(k) \geq 0$ and

$Q(\mathbf{w}^T(k)\mathbf{x}(k)) = -Q(\mathbf{w}^T(k)\mathbf{x}(k) + \mathbf{x}^T(k)\mathbf{x}(k))$ for $\mathbf{w}^T(k)\mathbf{x}(k) < 0$.

This further implies that

$$Q(\mathbf{w}^T(k)\mathbf{x}(k)) = -Q(\mathbf{w}^T(k)\mathbf{x}(k) - \mathbf{x}^T(k)Q(\mathbf{w}^T(k)\mathbf{x}(k))\mathbf{x}(k)).$$

As $y(k) = Q(\mathbf{w}^T(k)\mathbf{x}(k))$, we have

$$y(k) = -Q(\mathbf{w}^T(k)\mathbf{x}(k) - \mathbf{x}^T(k)y(k)\mathbf{x}(k)) = -Q((\mathbf{w}(k) - y(k)\mathbf{x}(k))^T \mathbf{x}(k)).$$

Define $\mathbf{w}''(k) \equiv \mathbf{w}(k) - y(k)\mathbf{x}(k)$ and $y''(k) \equiv Q(\mathbf{w}''^T(k)\mathbf{x}(k))$.

Then $y(k) = -Q(\mathbf{w}''^T(k)\mathbf{x}(k)) = -y''(k)$ and

$$\begin{aligned} & \tilde{\mathfrak{Z}}_k^F(\mathbf{w}''(k)) \\ &= \mathbf{w}''(k) + \frac{t(k) - Q(\mathbf{w}''^T(k)\mathbf{x}(k))}{2} \mathbf{x}(k) \\ &= \mathbf{w}''(k) + \frac{t(k) - y''(k)}{2} \mathbf{x}(k) \\ &= \mathbf{w}(k) - y(k)\mathbf{x}(k) + \frac{t(k) + y(k)}{2} \mathbf{x}(k) \\ &= \mathbf{w}(k) + \frac{t(k) - y(k)}{2} \mathbf{x}(k) \\ &= \mathbf{w}(k) + \frac{t(k) - Q(\mathbf{w}^T(k)\mathbf{x}(k))}{2} \mathbf{x}(k) \\ &= \tilde{\mathfrak{Z}}_k^F(\mathbf{w}(k)) \end{aligned}$$

Obviously, $\mathbf{w}''(k) \neq \mathbf{w}(k)$ because $y(k) \neq 0$ and $\mathbf{x}(k) \neq \mathbf{0}$.

Hence, $\tilde{\mathfrak{Z}}_k^F$ is not injective. This proves the necessity.

To prove the sufficiency, if $\tilde{\mathfrak{Z}}_k^F$ is not injective, then there exists $\mathbf{w}(k), \mathbf{w}''(k) \in \mathfrak{R}^{d+1}$ such that $\mathbf{w}(k) \neq \mathbf{w}''(k)$ and

$$\mathbf{w}(k) + \frac{t(k) - Q(\mathbf{w}^T(k)\mathbf{x}(k))}{2} \mathbf{x}(k) = \mathbf{w}''(k) + \frac{t(k) - Q(\mathbf{w}''^T(k)\mathbf{x}(k))}{2} \mathbf{x}(k).$$

This implies that

$$\mathbf{w}''(k) - \mathbf{w}(k) = \frac{Q(\mathbf{w}''^T(k)\mathbf{x}(k)) - Q(\mathbf{w}^T(k)\mathbf{x}(k))}{2} \mathbf{x}(k).$$

This further implies that

$$Q(\mathbf{w}''^T(k)\mathbf{x}(k)) = -Q(\mathbf{w}^T(k)\mathbf{x}(k))$$

and

$$\mathbf{w}''(k) = \mathbf{w}(k) - Q(\mathbf{w}^T(k)\mathbf{x}(k))\mathbf{x}(k).$$

Consequently, we have

$$\begin{aligned} Q(\mathbf{w}''^T(k)\mathbf{x}(k)) &= Q((\mathbf{w}(k) - Q(\mathbf{w}^T(k)\mathbf{x}(k))\mathbf{x}(k))^T \mathbf{x}(k)) \\ &= Q(\mathbf{w}^T(k)\mathbf{x}(k) - \mathbf{x}^T(k)Q(\mathbf{w}^T(k)\mathbf{x}(k))\mathbf{x}(k)) \end{aligned}$$

As $Q(\mathbf{w}''^T(k)\mathbf{x}(k)) = -Q(\mathbf{w}^T(k)\mathbf{x}(k))$, if $\mathbf{w}^T(k)\mathbf{x}(k) > 0$, then $\mathbf{w}^T(k)\mathbf{x}(k) - \mathbf{x}^T(k)\mathbf{x}(k) < 0$. This implies that

$\|\mathbf{x}(k)\|^2 > \mathbf{w}^T(k)\mathbf{x}(k)$. This further implies that

$\|\mathbf{x}(k)\|^2 > |\mathbf{w}^T(k)\mathbf{x}(k)|$. If $\mathbf{w}^T(k)\mathbf{x}(k) < 0$, then

$\mathbf{w}^T(k)\mathbf{x}(k) + \mathbf{x}^T(k)\mathbf{x}(k) \geq 0$. This implies that

$\|\mathbf{x}(k)\|^2 \geq -\mathbf{w}^T(k)\mathbf{x}(k)$. This further implies that $\|\mathbf{x}(k)\|^2 \geq |\mathbf{w}^T(k)\mathbf{x}(k)|$. Since $\|\mathbf{x}(k)\|^2 \neq |\mathbf{w}^T(k)\mathbf{x}(k)|$, we have $\|\mathbf{x}(k)\|^2 > |\mathbf{w}^T(k)\mathbf{x}(k)|$. If $\mathbf{w}^T(k)\mathbf{x}(k) = 0$, since we assume that $\mathbf{x}(k) \neq \mathbf{0}$, then $\|\mathbf{x}(k)\|^2 > |\mathbf{w}^T(k)\mathbf{x}(k)|$. Hence, this proves the sufficiency and it completes the proof.

Corollary 1

Assume that $\|\mathbf{x}(k-1)\|^2 \neq |\mathbf{w}^T(k-1)\mathbf{x}(k-1)|$. Then

$$\|\mathbf{x}(k-1)\|^2 > |\mathbf{w}^T(k-1)\mathbf{x}(k-1)|,$$

$$Q(\mathbf{w}^T(k-1)\mathbf{x}(k-1)) = -Q(\mathbf{w}^T(k-1)\mathbf{x}(k-1))$$

and

$$\mathbf{w}'(k-1) = \mathbf{w}(k-1) - Q(\mathbf{w}^T(k-1)\mathbf{x}(k-1))\mathbf{x}(k-1).$$

Proof:

The result follows directly from Lemma 3, so the proof is omitted here.

The importance of Lemma 3 and Corollary 1 is to reveal that $\tilde{\mathfrak{Z}}_k^F$ may not be invertible because $\tilde{\mathfrak{Z}}_k^F$ is not injective when $\|\mathbf{x}(k)\|^2 > |\mathbf{w}^T(k)\mathbf{x}(k)|$.

Now we can proceed to derive properties of an invariant set of the weights of the perceptron. Define $\wp \equiv \{\mathbf{w}(qN) \mid \forall q \in Z \text{ such that } \forall j \in Z \text{ and } \forall n \in Z\}$

$\mathbf{w}(jN) \neq \mathbf{w}(nN) + \sum_{p=0}^{N-1} \frac{Q((\mathbf{w}(p+jN))^T \mathbf{x}(p)) - Q((\mathbf{w}(p+nN))^T \mathbf{x}(p))}{2} \mathbf{x}(p)$.

Define $\mathfrak{Z}^F : \wp \rightarrow \wp$ such that

$$\mathfrak{Z}^F(\mathbf{w}(qN)) \equiv \tilde{\mathfrak{Z}}_{N-1}^F \circ \dots \circ \tilde{\mathfrak{Z}}_0^F(\mathbf{w}(qN)) \quad \forall \mathbf{w}(qN) \in \wp.$$

Theorem 1

\mathfrak{Z}^F is bijective and \wp is an invariant set under the system map \mathfrak{Z}^F .

Proof:

As $\forall \mathbf{w}(qN) \in \wp$,

$$\mathfrak{Z}^F(\mathbf{w}(qN)) = \tilde{\mathfrak{Z}}_{N-1}^F \circ \dots \circ \tilde{\mathfrak{Z}}_0^F(\mathbf{w}(qN)) = \mathbf{w}((q+1)N) \in \wp \quad \forall q \in Z.$$

Hence, $\mathfrak{Z}^F(\wp) \subseteq \wp$. $\forall \mathbf{w}(qN) \in \wp$, since $\mathbf{w}((q-1)N) \in \wp$ and

$$\mathfrak{Z}^F(\mathbf{w}((q-1)N)) = \tilde{\mathfrak{Z}}_{N-1}^F \circ \dots \circ \tilde{\mathfrak{Z}}_0^F(\mathbf{w}((q-1)N)) = \mathbf{w}(qN) \in \wp.$$

Hence, $\mathfrak{Z}^F(\wp) \supseteq \wp$ and \mathfrak{Z}^F is surjective. Consequently,

$\mathfrak{Z}^F(\wp) = \wp$ and \wp is an invariant set under the system map \mathfrak{Z}^F .

Assume that $\mathbf{w}(jN) \neq \mathbf{w}(nN)$ such that $\mathfrak{Z}^F(\mathbf{w}(nN)) = \mathfrak{Z}^F(\mathbf{w}(jN))$. This implies that

$$\begin{aligned} & \mathbf{w}(jN) + \sum_{p=0}^{N-1} \frac{t(p) - Q((\mathbf{w}(p+jN))^T \mathbf{x}(p))}{2} \mathbf{x}(p) \\ &= \mathbf{w}(nN) + \sum_{p=0}^{N-1} \frac{t(p) - Q((\mathbf{w}(p+nN))^T \mathbf{x}(p))}{2} \mathbf{x}(p) \end{aligned}$$

This further implies that

$$\mathbf{w}(jN) = \mathbf{w}(nN) + \sum_{p=0}^{N-1} \frac{Q((\mathbf{w}(p+jN))^T \mathbf{x}(p)) - Q((\mathbf{w}(p+nN))^T \mathbf{x}(p))}{2} \mathbf{x}(p).$$

However, there is a contradiction. Consequently, \mathfrak{Z}^F is injective. As a result, \mathfrak{Z}^F is bijective and this completes the proof.

The importance of Theorem 1 is to reveal that the weights

of the perceptron will be trapped inside \wp if the initial weight is inside \wp .

Lemma 4

$\forall jN < k$ and $\forall nN < k$,

$$\mathbf{w}(nN) + \sum_{p=0}^{k-1-nN} \frac{t(p) - Q(\mathbf{w}^T(p+nN)\mathbf{x}(p))}{2} \mathbf{x}(p) - \sum_{p=0}^{k-1-jN} \frac{t(p) - Q(\mathbf{w}^T(p+jN)\mathbf{x}(p))}{2} \mathbf{x}(p) \notin \wp$$

Proof:

Since $\exists k \in Z$ such that $\mathbf{w}'(k) = \mathbf{w}(k)$. This implies that $\forall jN < k$ and $\forall nN < k$

$$\mathbf{w}'(jN) = \mathbf{w}(nN) + \sum_{p=0}^{k-1-nN} \frac{t(p) - Q(\mathbf{w}^T(p+nN)\mathbf{x}(p))}{2} \mathbf{x}(p) - \sum_{p=0}^{k-1-jN} \frac{t(p) - Q(\mathbf{w}^T(p+jN)\mathbf{x}(p))}{2} \mathbf{x}(p)$$

Suppose that $\mathbf{w}'(jN) \in \wp$. Then $\exists p, q \in Z^+ \cup \{0\}$ and $\exists m \in \{0, 1, \dots, N-1\}$ such that $\mathbf{w}'((j+p)N) \in \wp$, $\mathbf{w}((n+q)N) \in \wp$ and $\mathbf{w}'((j+p)N+m) = \mathbf{w}((n+q)N+m) = \mathbf{w}(k)$. This implies that $\mathbf{w}'((j+p+1)N) = \mathbf{w}((n+q+1)N) \in \wp$. However, it contradicts to Theorem 1. Hence, $\mathbf{w}'(jN) \notin \wp$ and this completes the proof.

Define $\tilde{\mathfrak{Z}}^F : \mathfrak{R}^{d+1} \rightarrow \mathfrak{R}^{d+1}$ such that

$$\tilde{\mathfrak{Z}}^F(\mathbf{w}''(qN)) \equiv \tilde{\mathfrak{Z}}_{N-1}^F \circ \dots \circ \tilde{\mathfrak{Z}}_0^F(\mathbf{w}''(qN)) \quad \forall \mathbf{w}''(qN) \in \mathfrak{R}^{d+1}.$$

Theorem 2

$\tilde{\mathfrak{Z}}^F$ is not injective.

Proof:

As $\mathbf{w}'(k) = \mathbf{w}(k)$, then $\exists m \in \{0, 1, \dots, N-1\}$, $\exists \mathbf{w}'(jN) \notin \wp$ and $\exists \mathbf{w}(nN) \in \wp$ such that $\mathbf{w}'(jN+m) = \mathbf{w}(nN+m) = \mathbf{w}(k)$. Obviously, $\mathbf{w}'(jN) \neq \mathbf{w}(nN)$ and

$$\tilde{\mathfrak{Z}}^F(\mathbf{w}'(jN)) = \mathbf{w}'((j+1)N) = \mathbf{w}((n+1)N) = \tilde{\mathfrak{Z}}^F(\mathbf{w}(nN)).$$

Hence, $\tilde{\mathfrak{Z}}^F$ is not injective and this completes the proof.

The importance of Lemma 4 and Theorem 2 is to reveal that there exist some initial weights that are not in \wp , but the weights will be eventually moved to \wp . That implies that \wp is attracting.

Lemma 5

$$w_0(n) - w_0(0) \in Z \quad \forall n \in Z.$$

Proof:

Since $\mathbf{w}(n) = \mathbf{w}(0) + \sum_{k=0}^{n-1} \frac{t(k) - y(k)}{2} \mathbf{x}(k) \quad \forall n \in Z$, the first element of $\mathbf{x}(n)$ is 1 $\forall n \in Z$ and $\frac{t(n) - y(n)}{2} \in \{1, 0, -1\} \quad \forall n \in Z$,

the result follows directly and this completes the proof.

The implication of Lemma 5 is that $\mathbf{w}(n)$ occurs only in certain discrete hyperplanes, which depends on $w_0(0)$.

Theorem 3

$\tilde{\mathfrak{Z}}^F$ is surjective.

Proof:

$\forall \mathbf{w} \in \mathfrak{R}^{d+1}$, define $\mathbf{v} = \tilde{\mathfrak{Z}}_1^B \circ \dots \circ \tilde{\mathfrak{Z}}_N^B(\mathbf{w})$. Obviously,

$\mathbf{v} \in \mathfrak{R}^{d+1}$. By Lemma 1, we have $\tilde{\mathfrak{Z}}^F(\mathbf{v}) = \tilde{\mathfrak{Z}}_{N-1}^F \circ \dots \circ \tilde{\mathfrak{Z}}_0^F \circ \tilde{\mathfrak{Z}}_1^B \circ \dots \circ \tilde{\mathfrak{Z}}_N^B(\mathbf{w}) = \mathbf{w}$. Hence, $\tilde{\mathfrak{Z}}^F$ is surjective and this completes the proof.

IV. COMPUTER NUMERICAL SIMULATION RESULTS

Consider the following set of the training feature vectors

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix} \right\} \text{ as shown in Figure 1a. Assume that}$$

the corresponding desirable outputs are $\{1, -1, 1, -1\}$, respectively. Also, assume that $\mathbf{w}(0) = [0, 2, 0]^T$. The set of the weights only contains a single point, that is $\mathbf{w}(k) = [0, 2, 0]^T \quad \forall k \in Z$ as shown in Figure 2. Hence, the dynamics of the weights exhibits a fixed point behavior. This result agrees with the perceptron convergence theorem because the set of the training feature vectors is linearly separable. Consequently, $\wp = \{[0, 2, 0]^T\}$. It can be checked easily that the map $\mathfrak{Z}^F : \{[0, 2, 0]^T\} \rightarrow \{[0, 2, 0]^T\}$ is bijective and the map $\tilde{\mathfrak{Z}}^F : \mathfrak{R}^3 \rightarrow \mathfrak{R}^3$ is not injective because $\|\mathbf{x}(k)\|^2 > |\mathbf{w}^T(k)\mathbf{x}(k)| \quad \forall k \in Z$. Since some points in the three dimensional real space would converge to this fixed point, \wp is attracting.

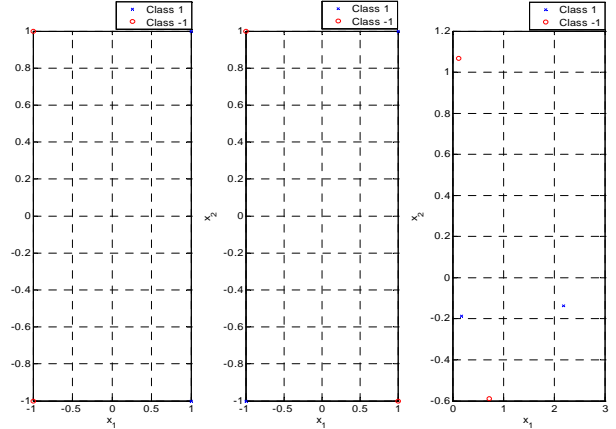


Fig. 1. The plot of x_1 against x_2 when the set of the training feature

$$\text{vectors is (a) } \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix} \right\}, \text{ (b)}$$

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \right\} \text{ and (c)}$$

$$\left\{ \begin{bmatrix} 1 \\ 0.1746 \\ -0.1867 \end{bmatrix}, \begin{bmatrix} 1 \\ 0.7258 \\ -0.5883 \end{bmatrix}, \begin{bmatrix} 1 \\ 2.1832 \\ -0.1364 \end{bmatrix}, \begin{bmatrix} 1 \\ 0.1139 \\ 1.0668 \end{bmatrix} \right\}.$$

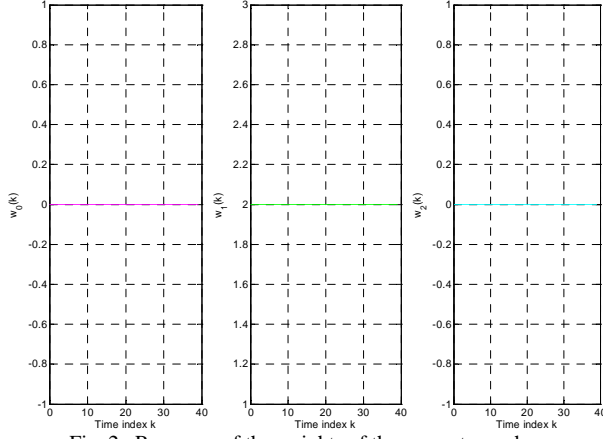


Fig. 2. Response of the weights of the perceptron when

$$\mathbf{x}(k) \in \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix} \right\}, \mathbf{w}(0) = [0, 2, 0]^T \text{ and}$$

$$t(k) \in \{1, -1, 1, -1\}. \text{ (a) } w_0(k). \text{ (b) } w_1(k). \text{ (c) } w_2(k).$$

Consider the following set of the training feature vectors

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \right\} \text{ as shown in Figure 1b. Assume that}$$

the corresponding desirable outputs are $\{1, -1, 1, -1\}$, respectively. Also, assume that $\mathbf{w}(0) = [-1, -1, -1]^T$. The

set of weights is $\left\{ \begin{bmatrix} -1 \\ -1 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ -2 \\ 0 \end{bmatrix} \right\}$. That is

$$\mathbf{w}(4k) = [-1, -1, -1]^T, \quad \mathbf{w}(4k+1) = [0, 0, 0]^T,$$

$$\mathbf{w}(4k+2) = [-1, -1, 1]^T \quad \text{and} \quad \mathbf{w}(4k+3) = [0, -2, 0]^T$$

$$\forall k \in \mathbb{Z}. \text{ Hence, the dynamics of the weights exhibits a limit cycle behavior with period 4 as shown in Figure 3.}$$

Consequently, $\wp = \left\{ \begin{bmatrix} -1 \\ -1 \\ -1 \end{bmatrix} \right\}$. It can be checked easily that the

map $\mathfrak{F}^F : \left\{ \begin{bmatrix} -1 \\ -1 \\ -1 \end{bmatrix} \right\} \rightarrow \left\{ \begin{bmatrix} -1 \\ -1 \\ -1 \end{bmatrix} \right\}$ is bijective and the map

$\tilde{\mathfrak{F}}^F : \mathfrak{R}^3 \rightarrow \mathfrak{R}^3$ is not injective because $\|\mathbf{x}(k)\|^2 \geq \|\mathbf{w}^T(k)\mathbf{x}(k)\|$ $\forall k \in \mathbb{Z}$. Since some points in the three dimensional real space would converge to this set of fixed points, \wp is attracting.

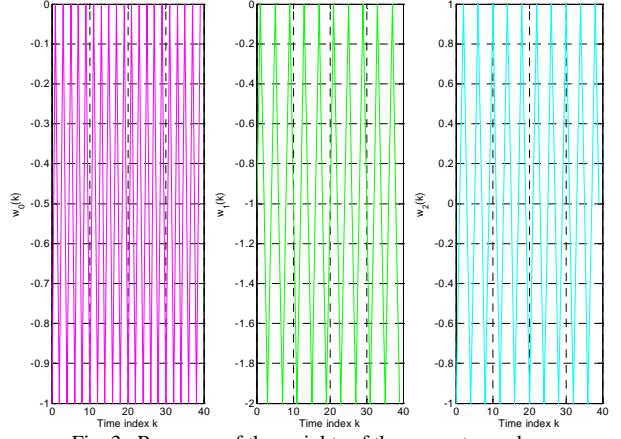


Fig. 3. Response of the weights of the perceptron when

$$\mathbf{x}(k) \in \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \right\}, \mathbf{w}(0) = [-1, -1, -1]^T \text{ and}$$

$$t(k) \in \{1, -1, 1, -1\}. \text{ (a) } w_0(k). \text{ (b) } w_1(k). \text{ (c) } w_2(k).$$

Consider the following set of the training feature vectors

$$\left\{ \begin{bmatrix} 1 \\ 0.1746 \\ -0.1867 \end{bmatrix}, \begin{bmatrix} 1 \\ 0.7258 \\ -0.5883 \end{bmatrix}, \begin{bmatrix} 1 \\ 2.1832 \\ -0.1364 \end{bmatrix}, \begin{bmatrix} 1 \\ 0.1139 \\ 1.0668 \end{bmatrix} \right\} \text{ as shown in}$$

Figure 1c. Assume that the corresponding desirable outputs are $\{1, -1, 1, -1\}$, respectively. Also, assume that

$\mathbf{w}(0) = [-1, 0.7923, -0.2133]^T$. The weights consist of three hyperplanes and exhibit chaotic behaviors within these three hyperplanes as shown in Figure 4. Hence, the set of the weights and an invariant set of the weights consist of three hyperplanes. It can be checked easily that the map from \wp to itself is bijective and the map $\tilde{\mathfrak{F}}^F : \mathfrak{R}^3 \rightarrow \mathfrak{R}^3$ is not injective because $\exists k \in \mathbb{Z}$ such that $\|\mathbf{x}(k)\|^2 > \|\mathbf{w}^T(k)\mathbf{x}(k)\|$.

Since some points in the three dimensional real space would converge to these three hyperplanes, \wp is attracting.

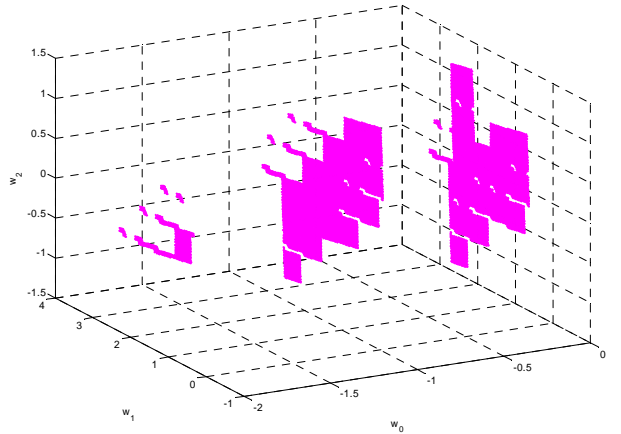


Fig. 4. Phase diagram of the weights of the perceptron when

$$\mathbf{x}(k) \in \left\{ \begin{bmatrix} 1 \\ 0.1746 \\ -0.1867 \end{bmatrix}, \begin{bmatrix} 1 \\ 0.7258 \\ -0.5883 \end{bmatrix}, \begin{bmatrix} 1 \\ 2.1832 \\ -0.1364 \end{bmatrix}, \begin{bmatrix} 1 \\ 0.1139 \\ 1.0668 \end{bmatrix} \right\},$$

$$\mathbf{w}(0) = [-1, 0.7923, -0.2133]^T \text{ and } t(k) \in \{1, -1, 1, -1\}.$$

V. CONCLUSIONS

In this paper, an invariant set that results to a bijective invariant map is characterized. Also, it is found that some weights outside the invariant set will be moved to the invariant set. Hence, the invariant set is attracting.

ACKNOWLEDGMENT

The work obtained in this paper was supported by a research grant from Queen Mary College, University of London.

REFERENCES

- [1] Sankar K. Pal and Sushmita Mitra, "Multilayer perceptron, fuzzy sets, and classification," *IEEE Transactions on Neural Networks*, vol. 3, pp. no. 5, pp. 683-697, 1992.
- [2] Mitra Basu and Quan Liang, "The fractional correction rule: a new perspective," *Neural Networks*, vol. 11, pp. 1027-1039, 1998.
- [3] Uri Erez, Simon Litsyn and Ram Zamir, "Lattices which are good for (almost) everything," *IEEE Transactions on Information Theory*, vol. 51, no. 10, pp. 3401-3416, 2005.
- [4] Tie Liu, Pierre Moulin and Ralf Koetter, "On error exponents of modulo lattice additive noise channels," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 454-471, 2006.
- [5] Charlotte Yuk-Fan Ho, Bingo Wing-Kuen Ling, Hak-Keung Lam and Muhammad H U Nasir, "Global convergence and limit cycle behavior of weights of perceptron," to appear in *IEEE Transactions on Neural Networks*, June, 2008.