Neural Learning Can Form Structures From Computational Geometry

Thomas Martinetz

Abstract: It is shown that a competitive Hebbian learning rule forms so-called Delaunay triangulations, which play an important role in computational geometry for efficiently solving proximity problems. Given a set of neural units i, $i = 1, ..., N$, the synaptic weights of which can be interpreted as points $\mathbf{w}_i$ in $\Re^D$, by sequentially presenting input patterns $\mathbf{v} \in \Re^D$ the competitive Hebbian learning rule leads to a connectivity structure between the units i which corresponds to the Delaunay triangulation of the points $\mathbf{w}_i$, $i = 1, ..., N$. Such competitive Hebbian rule develops connections ($C_{ij} > 0$) between neural units i, j with neighboring receptive fields (Voronoi polygons) V_i, V_j, whereas between all other units i, j no connections evolve ($C_{ij} = 0$).

1. Introduction

Information processing tasks often require to solve repetitively so-called *proximity problems*. The most prominent examples of such proximity problems are the k-nearest-neighbor search, the construction of the Euclidean minimum spanning tree and the triangulation problem. These proximity problems occur in applications ranging from speech- and image processing over network design to efficient storage and transfer of data [1-5]. A powerful structure from computational geometry which, after having been constructed in a preprocessing stage, allows one to solve these proximity problems with significantly reduced computational effort[1], is the Delaunay triangulation and its dual, the Voronoi diagram [6, 7].

The Voronoi diagram $\mathcal{V}_S$ of a set $S = \{\mathbf{w}_1, \ldots, \mathbf{w}_N\}$ of points $\mathbf{w}_i \in \Re^D$ is given by N D-dimensional polyhedra, the Voronoi polyhedra V_i. The Voronoi polyhedron V_i of a point $\mathbf{w}_i \in S$ is given by the set of points $\mathbf{v} \in \Re^D$ which are closer to $\mathbf{w}_i$ than to any other $\mathbf{w}_j \in S$:

$$V_i = \{\mathbf{v} \in \Re^D \,\Big|\, \|\mathbf{v} - \mathbf{w}_i\| \leq \|\mathbf{v} - \mathbf{w}_j\| \ \forall j\}.$$

The Delaunay triangulation $\mathcal{D}_S$ of a set $S = \{\mathbf{w}_1, \ldots, \mathbf{w}_N\}$ of points $\mathbf{w}_i \in \Re^D$ is defined by the graph whose vertices are the $\mathbf{w}_i$ and whose *adjacency matrix* $\mathbf{A}$, $A_{ij} \in \{0, 1\}$, $i, j = 1, ..., N$ carries the value one if and only if $V_i \cap V_j \neq \emptyset$. Two vertices $\mathbf{w}_i, \mathbf{w}_j$ are connected by an edge if and only if their Voronoi polyhedra V_i, V_j are adjacent. An illustration of the Voronoi diagram and the Delaunay triangulation of a set of points in a plane is given in Fig. 1.

[1]at most linearly increasing in the number of vertices.

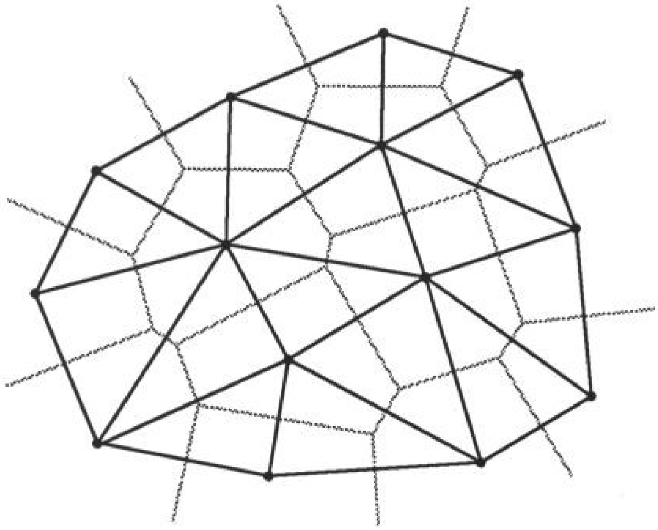

Figure 1: The Voronoi diagram and the Delaunay triangulation (dark lines) of a set of points.

In this paper we show that a competitive version of a well-known neural learning rule, the Hebbian adaptation rule, leads to interneural connections within a network of neural units i corresponding to the edges of the Delaunay triangulation of the synaptic weight vectors $\mathbf{w}_i \in \Re^D$. We show that by sequentially presenting input patterns $\mathbf{v}$ drawn from a manifold $M = \Re^D$ the competitive version of the Hebbian adaptation rule forms the complete Delaunay triangulation $\mathcal{D}_S$.

2. Competitive Hebbian learning rule

In the following we assume a set of neural units i, $i = 1, ..., N$ which can develop *lateral connections* between each other. A neural unit connects itself with another unit by developing a *synaptic link* to this unit. The lateral connections are described by a connection strength matrix $\mathbf{C}$ with elements $C_{ij} \in \Re_0^+$. The larger a matrix element C_{ij}, the stronger is the connection from unit i to unit j. Only if $C_{ij} > 0$, we regard neural unit i as being connected with unit j. If $C_{ij} = 0$, neural unit i is *not* connected with unit j. Negative values for C_{ij} do not arise.

The basic principle which governs the change of interneural connection strength has first been formulated by Hebb [8]. According to Hebb's postulate a presynaptic unit i increases the strength of its synaptic link to a postsynaptic unit j if both units are concurrently active, i.e., if both activities do correlate. A variety of quantitative formulations of this conjunctive mechanism have been proposed, e.g., for modeling Pavlovian conditioning [9], motor learning [10], or associative memory [11]. In its simplest mathematical formulation Hebb's rule is described by the equation

$$\Delta C_{ij} \propto y_i \cdot y_j, \tag{1}$$

in which the change of the strength C_{ij} of the connection from unit i to unit j is linearly proportional to the presynaptic activity y_i and the postsynaptic activity y_j. The quantities y_i, $i = 1, \ldots, N$ denote the output activities of the neural units.

In the following we assume that to each neural unit i a weight vector $\mathbf{w}_i \in \Re^D$ is assigned. Further, we assume that each neural unit i, $i = 1, \ldots, N$ receives the same external input pattern $\mathbf{v} \in \Re^D$. The weight vector $\mathbf{w}_i$ determines the center of the receptive field of unit i in the sense that with the reception of an input pattern $\mathbf{v}$ the output activity y_i of unit i is the larger the closer its $\mathbf{w}_i$ is to $\mathbf{v}$. In mathematical terms, we assume that $y_i = R(\|\mathbf{v} - \mathbf{w}_i\|)$ is valid, with $R(.)$ being a positive and continuously monotonically decreasing function, e.g., a Gaussian.

Employing the Hebb rule in the simple form as given in eq. (1) yields the rather trivial result that each neural unit i develops connections to all the other units $j \neq i$, with lateral connection strengths C_{ij} each of which is simply proportional to the overlap of receptive field $R(\|\mathbf{v} - \mathbf{w}_i\|)$ with receptive field $R(\|\mathbf{v} - \mathbf{w}_j\|)$:

$$\Delta C_{ij}(t \to \infty) \propto \int_{\Re^D} R(\|\mathbf{v} - \mathbf{w}_i\|) \cdot R(\|\mathbf{v} - \mathbf{w}_j\|) d\mathbf{v}.$$

However, as in many systems governed by self-organizing processes, the connectivity pattern which evolves on the set of neural units becomes significantly more structured if we introduce competition. Analog to the competition among the units in a winner-take-all network we introduce competition among the connections $i - j$. Instead of being based on the output activities of the neural units itself as in a winner-take-all network, the competition among the connections $i - j$ is determined by the *correlated* output activities $Y_{ij} = y_i \cdot y_j$. Keeping the analogy to winner-take-all networks, with the presentation of an input pattern $\mathbf{v}$ only the connection $i - j$ whose "activity" $Y_{ij} = y_i \cdot y_j$ is highest is modified. Instead of changing the connection strengths according to (1), the *competitive* Hebbian learning rule as a winner-take-all or competitive version of (1) changes the connection strengths according to

$$\Delta C_{ij} \propto \begin{cases} y_i \cdot y_j & \text{if } y_i \cdot y_j \geq y_k \cdot y_l \quad \forall\, k, l \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

With an input pattern $\mathbf{v}$ only the connection between the unit with the largest and the unit with the second largest output is modified, all the other connections remain unchanged.

3. Competitive Hebbian learning forms Delaunay triangulations

We show that, instead of connecting each unit with all the other units, the competitive Hebbian learning rule (2) forms a connectivity structure among the neural units i, $i = 1, \ldots, N$ which corresponds to the Delaunay triangulation of the weight vectors $\mathbf{w}_1, \ldots, \mathbf{w}_N$. More precisely, we show that if we present sequentially input patterns $\mathbf{v}$ with a distribution $P(\mathbf{v})$ which has support (is nonzero) everywhere on $\Re^D$, then the elements C_{ij} of the connection strength matrix $\mathbf{C}$ obey asymptotically

$$\theta(C_{ij}(t \to \infty)) = A_{ij} \qquad i, j = 1, \ldots, N$$

with $\theta(.)$ as the Heavyside step function and A_{ij} as the elements of the adjacency matrix $\mathbf{A}$ of the Delaunay triangulation of the points $\mathbf{w}_1, \ldots, \mathbf{w}_N$, for which

$$A_{ij} = \begin{cases} 1 & \text{if } V_i \cap V_j \neq \emptyset \quad \text{(adjacent)} \\ 0 & \text{if } V_i \cap V_j = \emptyset \quad \text{(not adjacent)} \end{cases}$$

is valid. V_i, V_j again denote the Voronoi polyhedra of $\mathbf{w}_i$, $\mathbf{w}_j$.

To prove that the adjacency matrix $A_{ij} = \theta(C_{ij})$ defined by the connectivity structure between the neural units becomes equivalent to the adjacency matrix of the Delaunay triangulation $\mathcal{D}_S$ of the set of points $S = (\mathbf{w}_1, \ldots, \mathbf{w}_N)$, we introduce the *second order Voronoi polyhedra* V_{ij}, $i, j = 1, ..., N$. The second order Voronoi polyhedron V_{ij} is given by all the $\mathbf{v} \in \Re^D$ for which $\mathbf{w}_i$ and $\mathbf{w}_j$ are the two closest points of S; i.e., V_{ij} is defined by

$$V_{ij} = \{ \quad \mathbf{v} \in \Re^D \;\Big|\; \|\mathbf{v} - \mathbf{w}_i\| \leq \|\mathbf{v} - \mathbf{w}_k\| \\ \wedge \|\mathbf{v} - \mathbf{w}_j\| \leq \|\mathbf{v} - \mathbf{w}_k\| \; \forall k \neq i, j\}.$$

As the Voronoi polyhedron of first order V_i, also V_{ij} forms a convex polyhedron.

It is obvious that two units i, j can become connected by the competitive Hebbian rule if and only if $V_{ij} \neq \emptyset$ is valid. Only then there is an input pattern $\mathbf{v}$ for which unit i (or j) has the largest output $y_i = R(\|\mathbf{v} - \mathbf{w}_i\|)$ and unit j (or i, respectively) has the second largest output $y_j = R(\|\mathbf{v} - \mathbf{w}_j\|)$ and, hence, for which the correlated output activity $Y_{ij} = y_i \cdot y_j$ of the units i, j wins. We prove that $V_{ij} \neq \emptyset$ is valid if and only if the corresponding first order Voronoi polyhedra V_i, V_j are adjacent, i.e., if and only if $V_i \cap V_j \neq \emptyset$ is valid. Then, in case $\int_{V_{ij}} P(\mathbf{v}) d\mathbf{v} \neq 0$ holds for each $V_{ij} \neq \emptyset$, the connections formed by he competitive Hebbian rule correspond to the edges of the Delaunay triangulation of the points $\mathbf{w}_1, \ldots, \mathbf{w}_N$ [2].

Theorem 1 *For a set $S = \{\mathbf{w}_1, \ldots, \mathbf{w}_N\}$ of points $\mathbf{w}_i \in \Re^D$ the relation*

$$V_i \cap V_j \neq \emptyset \;\Leftrightarrow\; V_{ij} \neq \emptyset$$

is valid for each pair i, j. V_i denotes the first order Voronoi polyhedron of point $\mathbf{w}_i$, and V_{ij} denotes the second order Voronoi polyhedron of the points $\mathbf{w}_i$, $\mathbf{w}_j$.

Proof: If $V_i \cap V_j \neq \emptyset$ is valid, there is a $\mathbf{v} \in \Re^D$ with $\mathbf{v} \in V_i$ and $\mathbf{v} \in V_j$. Then we obtain $\|\mathbf{v} - \mathbf{w}_i\| = \|\mathbf{v} - \mathbf{w}_j\| \leq \|\mathbf{v} - \mathbf{w}_k\|$ for all $\mathbf{w}_k \in S$, and, therefore, $\mathbf{v} \in V_{ij}$, i.e., $V_{ij} \neq \emptyset$, is valid.

If $V_{ij} \neq \emptyset$ is valid, there is a $\mathbf{v} \in \Re^D$ for which the points $\mathbf{w}_i$ and $\mathbf{w}_j$ are the two nearest neighbors. Without loss of generality we assume that $\mathbf{w}_i$ is the nearest neighbor. Since for each $\mathbf{u} \in \overline{\mathbf{v}\mathbf{w}_j}$ the point $\mathbf{w}_j$ is either the nearest or the second nearest neighbor of $\mathbf{u}$, and since for $\mathbf{u} = \mathbf{v}$ the point $\mathbf{w}_i$ is closest and for $\mathbf{u} = \mathbf{w}_j$ the point $\mathbf{w}_j$ is closest to $\mathbf{u}$, there is a $\mathbf{u}^* \in \overline{\mathbf{v}\mathbf{w}_j}$ for which $\|\mathbf{u}^* - \mathbf{w}_i\| = \|\mathbf{u}^* - \mathbf{w}_j\|$ is valid. Hence, we obtain $\mathbf{u}^* \in V_i$ and $\mathbf{u}^* \in V_j$, and, therefore, $\mathbf{u}^* \in V_i \cap V_j$, i.e., $V_i \cap V_j \neq \emptyset$, is valid.

[2] The following theorem was formulated together with Philippe Dalger and Bennoit Noël [12].

4. Summary

We showed that formal neural units i form connectivity structures corresponding to Delaunay triangulations, if the Hebb rule together with competition among the connections is employed. Each neural unit i has to have a localized receptive field within the feature space M. By sequentially presenting patterns $\mathbf{v} \in M$ and each time connecting those two units i, j which have the highest correlated output activity $y_i \cdot y_j$, the Delaunay triangulation $\mathcal{D}_S$ of the receptive field centers $\mathbf{w}_i$ evolves. Delaunay triangulations play an important role in a variety of information processing tasks. These tasks range from data compression over discrete optimization to pattern recognition and function approximation [1, 5, 13]. We demonstrated that the Delaunay triangulation with its significant role for information processing can be established by a self-organizing neural network: a set of formal neural units with lateral connections formed by an input driven, Hebbian learning rule.

Acknowledgement

I am indebted to Klaus Schulten for many fruitful discussions. This work has been supported by the BMFT under Grant No. 01IN102A7.

References

[1] Duda RO, Hart PE (1973) Pattern Classification and Scene Analysis. Wiley, New York.

[2] Makhoul J, Roucos S, Gish H (1985) Vector quantization in speech coding. Proceedings of the IEEE 73:1551–1588.

[3] Nasrabadi NM, King RA (1988) Image coding using vector quantization: A review. IEEE Trans Comm 36(8):957-971.

[4] Zahn CT (1971) Graph-theoretical methods for detecting and describing gestalt clusters. IEEE Trans. Comp. C-20(1), pp 68–86.

[5] Strang G, Fix G (1973) An analysis of the finite element Method. Prentice-Hall, Englewood Cliffs, NJ, 1973.

[6] Preparata FP, Shamos MI (1988) Computational Geometry: An Introduction. Springer-Verlag, New York.

[7] Delaunay B (1934) Sur la Spère Vide. Bull. Acad. Sci. USSR (VII), Classe Sci. Mat. Nat., pp 793–800.

[8] Hebb D (1949) Organization of Behavior. Wiley, New York.

[9] Grossberg S (1974) Classical and instrumental learning by neural networks. In: Prog. Theor. Biol., vol III, Rosen E., Snell F. (eds.), Academic Press, New York, pp. 51-141.

[10] Marr D (1969) A Theory of Cerebellar Cortex. J. Physiol., 202:437–470.

[11] Palm G (1982) Neural Assemblies: An Alternative Approach to Artificial Intelligence, Springer, New York.

[12] Dalger P, Noël B, Martinetz T, Schulten K (1992) Delaunay triangulation and Voronoi polyhedra of second order. Beckman Institute Technical Report TB 98-01, University of Illinois at Urbana-Champaign.

[13] Kruskal JB (1956) On the shortest spanning subtree of a graph and the traveling salesman problem. Proc. AMS 7, pp 48–50.

Author

Dr. Thomas Martinetz
Siemens AG
Corporate Research & Development
Neural Networks Group
81730 München, Germany