

Equivalence and differences in recall and storage dynamics of associative memory

Hiro-Fumi Yanai

Department of Information and Communication Engineering

Faculty of Engineering, Tamagawa University

Machida, Tokyo 194, Japan

E-mail: yanai@eng.tamagawa.ac.jp

Abstract

I introduce two-stage-dynamics neuron, and discuss the relationship between recall and storage dynamics of associative memory. In conventional studies of neural nets, recall and storage dynamics have been treated separately. But recent studies on two-stage neuron revealed their close relationship. I discuss equivalence and differences of those dynamics that might be interesting to be noted from experimental as well as theoretical points of view. Although this paper focuses on associative memory, the kind of equivalence as is shown here could also be seen in other types of neural nets.

*Proceedings of
the 1996 International Conference
on Neural Information Processing
(ICONIP'96)
pp. 581-586*

1 Recall and storage dynamics

Neural nets process information by dynamically updating states of neurons as well as modifying weights of synapses. I shall call those two phases of dynamics as recall and storage dynamics, respectively. Recall is often called retrieval, and storage dynamics is usually called learning. For studies of various aspects of recall dynamics see [2, 3, 5, 6, 8, 9, 10, 11, 12, 13, 14, 15], and for the combined discussions of recall and storage dynamics see [1, 4, 7].

1.1 Fundamental models

The simplest model of recall dynamics is

$$x_i(t+1) = \text{sgn} \left(\sum_{j=1}^n w_{ij} x_j \right), \quad (1)$$

where $x_i(t) = \pm 1$ is the output of the i th neuron at discrete time t , w_{ij} is the weight of synapse which connects j th and i th neurons, and $\text{sgn}(u) = 1(u > 0)$ and $-1(u \leq 0)$. In matrix form, Eq. (1) is

$$\mathbf{x}_{t+1} = \text{sgn}(W \mathbf{x}_t), \quad (2)$$

where \mathbf{x}_t is a column vector whose i th component is $x_i(t)$, and W , called synaptic weight matrix, is a $n \times n$ matrix whose (i, j) component is w_{ij} .

Widely analyzed type of synaptic weight matrix is correlation one. In the case of auto-associative memory, given m column vectors $\boldsymbol{\xi}^1, \boldsymbol{\xi}^2, \dots, \boldsymbol{\xi}^m$ to be stored (I call these vectors *embedded vectors*), synaptic weight is given by

$$w_{ij}^c = \frac{1}{n} \sum_{\mu=1}^m \xi_i^\mu \xi_j^\mu. \quad (3)$$

By defining the $n \times n$ matrix $S = (\boldsymbol{\xi}^1, \boldsymbol{\xi}^2, \dots, \boldsymbol{\xi}^m)$, the above synaptic weight is expressed as matrix form

$$W^c = \frac{1}{n} S S' \quad (4)$$

$$= \frac{1}{n} \sum_{\mu=1}^m \boldsymbol{\xi}^\mu (\boldsymbol{\xi}^\mu)', \quad (5)$$

where $'$ denotes transposition. The matrix W^c is the auto-correlation matrix.

The above correlation-type synapses are obtained by the storage dynamics of

$$W(s+1) = (1 - \gamma)W(s) + \beta \boldsymbol{\xi}^{\mu(s)} (\boldsymbol{\xi}^{\mu(s)})', \quad (6)$$

where $\mu(s) \in \{1, 2, \dots, m\}$, and γ and β are positive constants. Apart from fluctuations, if conditions are appropriate, the matrix $W(s)$ converges to W^c [1].

Dynamics of Eq. (1) with correlation-type synapses (cross-correlation as well as auto-correlation) are reviewed and analyzed in [3].

1.2 Generalized models

In [4] Caianiello has made a unified treatment on the generalized model of the above two phases of dynamics, where they are described by *neuronal equation* and *mnemonic equation*, respectively. The neuronal equation has the form

$$\mathbf{x}_{t+1} = \text{sgn} \left(\sum_{k=0}^{\infty} W^{(k)} \mathbf{x}_{t-k} \right). \quad (7)$$

The mnemonic equation is an extension of the correlation synapse for multiple delays. By adding threshold and external force terms, the neuronal equation includes most models that have ever been analyzed as associative memories.

Another direction of generalizing recall dynamics is the *two-stage model*.

$$\begin{cases} \mathbf{u}_t = W \mathbf{x}_t, \\ \tilde{\mathbf{u}}_t = F(\mathbf{u}_t, W), \\ \mathbf{x}_{t+1} = \text{sgn}(\tilde{\mathbf{u}}_t), \end{cases} \quad (8)$$

where $F(\mathbf{u}, W)$ is, in general, a nonlinear function of \mathbf{u} and W .

Models of storage dynamics other than correlation type include *pseudoinverse-matrix-type* [1, 7] and *perceptron-type*. For the case of auto-associative memory, pseudoinverse-matrix-type storage dynamics is

$$W(s+1) = (1 - \gamma)W(s) + \beta \left[\boldsymbol{\xi}^{\mu(s)} - W(s)\boldsymbol{\xi}^{\mu(s)} \right] (\boldsymbol{\xi}^{\mu(s)})', \quad (9)$$

which converges to the orthogonal-projection matrix

$$P = SS^+, \quad (10)$$

where S^+ is the pseudoinverse of S , and is expressed as $S^+ = (S'S)^{-1}S'$. By these synaptic weights the neural net can store linearly independent vectors as fixed points of the recall dynamics of Eq. (1). Perceptron-type storage dynamics is

$$W(s+1) = (1 - \gamma)W(s) + \beta \left[\boldsymbol{\xi}^{\mu(s)} - \text{sgn}(W(s)\boldsymbol{\xi}^{\mu(s)}) \right] (\boldsymbol{\xi}^{\mu(s)})', \quad (11)$$

by which the net can store linearly separable vectors.

2 Recall dynamics of associative memory

When the embedded vectors are random, recall dynamics of associative memory are characterized by *capacity* and *basin of attraction*. Capacity is the largest number of embedded vectors that can be stably stored as fixed points (auto-association) or sequences (sequence-association). There are two definitions of capacity, *absolute capacity* and *relative capacity* [2]. Allowing no error, we have absolute capacity, and allowing errors, we have relative capacity¹. For example, auto- and sequence-associative memory with correlation-type synapses (and no selfcouplings, i.e. $w_{ii} = 0$) have the absolute capacity of $n/(2 \log n)$ with the number of neurons n tending to infinity. For finite n , the absolute capacity is monotonically decreasing with n .

¹The relative capacity is defined by the phase diagram of the equilibrium states without ambiguity.

The relative capacity of the above auto-associative memory is well known to be *near* $0.15n$, exact value being still unclear. For sequence-associative memory, it is quite close to the corresponding value $0.27n$ of layered net [5].

The basin of attraction of an embedded vector is the domain of state space $\{-1, 1\}^n$ which is attracted to (absolute case) or attracted close to (relative case) the embedded vector by the recall dynamics. To discuss the property, let me define *distance* and *overlap*. Distance of the output state \mathbf{x}_t from the embedded vector $\boldsymbol{\xi}$ in concern is

$$d_t = \frac{1}{2n} \sum_{i=1}^n |\xi_i - x_i(t)|, \quad (12)$$

and the overlap is the direction cosine of \mathbf{x}_t and $\boldsymbol{\xi}$

$$l_t = \frac{1}{n} \sum_{i=1}^n \xi_i x_i(t), \quad (13)$$

$$= 1 - 2d_t \quad (14)$$

When the number of embedded vectors is less than the absolute capacity, that is $m = \kappa n / (2 \log n)$ with $\kappa < 1$, for initial distance satisfying $d_0 < (1 - \sqrt{\kappa})/2$, the embedded vector in concern is perfectly recalled by one step, i.e. $d_1 = 0$. In this case the radius of basin of attraction is $(1 - \sqrt{\kappa})/2$. When the number of embedded vectors is larger, recall dynamics is characterized by the *loading rate* $r = m/n$. The phase diagram for correlation-type sequence-associative memory can be seen in [5]. The phase diagram for correlation-type auto-associative memory is qualitatively the same according to the theory with several approximations [2, 3, 14].

3 Relationship between two-stage model and pseudoinverse matrix

3.1 Their correspondence

Two-stage neuron model of Eq. (8) is an extension of the discrete non-monotonic neuron model proposed in [8]². For the case of correlation-type auto-associative memory without selfcouplings, recall dynamics in the regime of absolute capacity are studied for various choices of the function F in Eq. (8) [11, 12]. A good choice of the function F is

$$F(u, W) = u + W_0^c f(u) \quad (15)$$

with

$$f(u) = -au + (2a - 1)\text{sgn}(u), \quad a > 0, \quad (16)$$

where $W_0^c = W^c - \frac{m}{n}I$ (I is a unit matrix). With the above function, the absolute capacity is, independently of a ,

$$\frac{n}{\sqrt{2 \log n}}. \quad (17)$$

And the relative capacity is greater than $0.3n$.

The simplest form of Eq. (16) is the linear function obtained by putting $a = 1/2$. To write down the recall dynamics of Eq. (8) explicitly for $a = 1/2$, we have

$$\mathbf{x}_{t+1} = \text{sgn} \left[\left(I - \frac{1}{2} W_0^c \right) W_0^c \mathbf{x}_t \right]. \quad (18)$$

²For theoretical analyses of continuous non-monotonic models, see [10, 15]

In fact, the matrix $(I - \frac{1}{2}W_0^c)W_0^c$ corresponds to the first-degree approximation of the orthogonal-projection matrix P that realizes error-less autoassociation. That is, since the von Neumann expansion of the pseudoinverse S^+ of S is³

$$S^+ = \frac{\alpha}{n} \sum_{k=0}^{\infty} \left(I - \frac{\alpha}{n} S' S \right)^k S', \quad (19)$$

the corresponding orthogonal-projection matrix P is

$$\begin{aligned} P &= S S^+ \\ &= \sum_{k=0}^{\infty} (I - \alpha W^c)^k \alpha W^c. \end{aligned} \quad (20)$$

The first two terms of the expansion is

$$\sum_{k=0}^1 (I - \alpha W^c)^k \alpha W^c = 2\alpha \left(I - \frac{\alpha}{2} W^c \right) W^c. \quad (21)$$

This proves the correspondence.

If we focus on the linear two-stage model, since we know the correspondence between recall and storage dynamics through Eq. (20), it is straightforward to extend the linear two-stage model to higher-degree ones which better approximate the storage dynamics that yield P . Let me define a linear two-stage model of degree τ as

$$\text{Stage 1 : } \quad \mathbf{u}_t = W^c \mathbf{x}_t, \quad (22)$$

$$\text{Stage 2 : } \quad \tilde{\mathbf{u}}_t^{(\tau)} = \sum_{k=0}^{\tau} (I - \alpha W^c)^k \alpha \mathbf{u}_t, \quad (23)$$

$$\text{Output stage : } \quad \mathbf{x}_{t+1} = \text{sgn} \left(\tilde{\mathbf{u}}_t^{(\tau)} \right), \quad (24)$$

where t is the time. Stages 1, 2 and output stage are applied cyclically. Initial condition is given by \mathbf{x}_0 and \mathbf{u}_t is calculated (This is the stage 1), and according to Eq. (23) $\mathbf{u}_1^{(\tau)}$ is calculated (This is the stage 2), then $\mathbf{u}_1^{(\tau)}$ is substituted into Eq. (24) (This is the output stage), that determines \mathbf{x}_1 . In this way, the output state vectors $\mathbf{x}_t (t = 2, 3, \dots)$ are determined. Stage 2 dynamics, Eq. (23), can be realized by recurrent rules which are plausible as neural net models. One example is

$$\tilde{\mathbf{u}}_t^{(k+1)} = \tilde{\mathbf{u}}_t^{(k)} + (I - \alpha W^c)(\tilde{\mathbf{u}}_t^{(k)} - \tilde{\mathbf{u}}_t^{(k-1)}) \quad \text{with} \quad \tilde{\mathbf{u}}_t^{(0)} = \alpha W^c \mathbf{x}_t \quad \text{and} \quad \tilde{\mathbf{u}}_t^{(-1)} = 0, \quad (25)$$

and another is

$$\begin{cases} \mathbf{v}_t^{(k+1)} = (I - \alpha W^c) \mathbf{v}_t^{(k)} \\ \tilde{\mathbf{u}}_t^{(k+1)} = \tilde{\mathbf{u}}_t^{(k)} + \mathbf{v}_t^{(k+1)} \end{cases} \quad \text{with} \quad \mathbf{v}_t^{(0)} = \alpha \mathbf{u}_t \quad \text{and} \quad \tilde{\mathbf{u}}_t^{(0)} = 0. \quad (26)$$

Equivalence of recall and storage dynamics is exact when $\tau \rightarrow \infty$. That is, the two-stage dynamics with correlation matrix is equivalent to the conventional (threshold) dynamics with pseudoinverse matrix. Thus the two-stage model with $\tau \rightarrow \infty$ is fully understood by the analysis of pseudoinverse matrix associative memories, e.g. Kohonen [7] and Kanter & Sompolinsky [6].

³Convergence of the series, Eq. (19), is guaranteed if $0 < \alpha < 2/\lambda$ is satisfied, where λ is the maximum eigenvalue of $(1/n)S'S$.

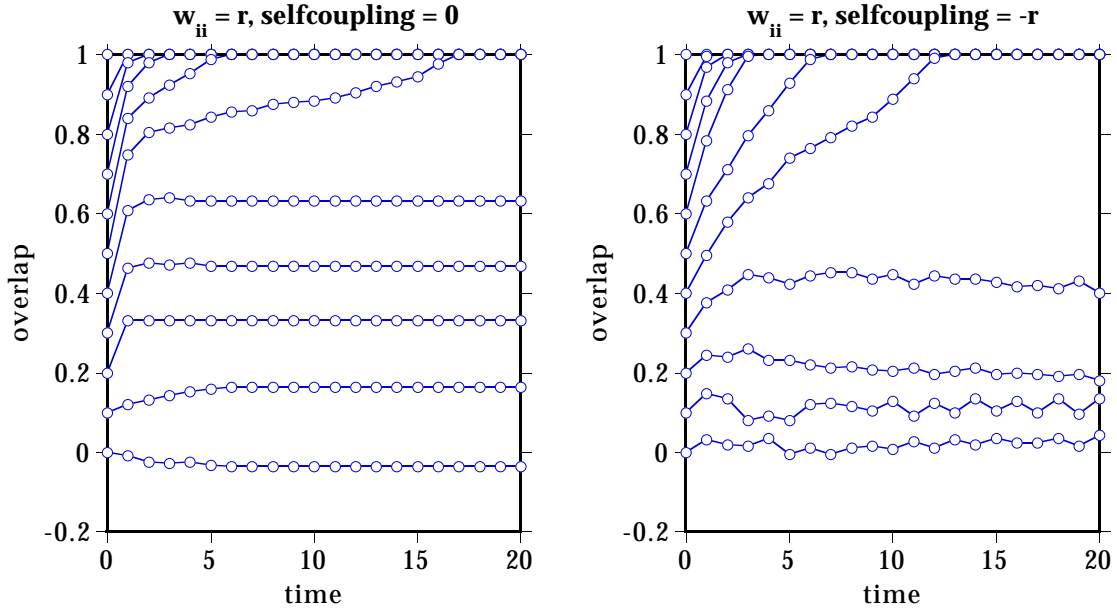


Figure 1: Numerically simulated recall processes of associative memory with two-stage-neurons. $n = 500, r = 0.3, \tau = 5, \alpha = 0.5$. w_{ii} represents the diagonal elements of the correlation matrix, so that $w_{ii} = r$ means that W^c is used. Left: corresponds to fundamental orthogonal projection matrix model. Right: corresponds to orthogonal projection matrix model with zero diagonal elements.

3.2 Examples from numerical simulations

Some numerically simulated results for recall of a memory vector from noisy initial conditions are presented in Figures 1 and 2. In the simulations the output stage is slightly modified to demonstrate interesting properties of the model. That is, the selfcoupling term is added as

$$\mathbf{x}_{t+1} = \text{sgn} \left(\hat{\mathbf{u}}_t^{(\tau)} + \text{selfcoupling } \mathbf{x}_t \right). \quad (27)$$

According to [6], if the every diagonal elements of the orthogonal projection matrix are set to zero, the basin of attraction of memory is enlarged (compare two figures in Figure 1). In the present model, since we are concerned with the orthogonal projection matrix only implicitly, we cannot manipulate those diagonal elements directly. To achieve the equivalent situation, the selfcoupling term helps — if we put $\text{selfcoupling} = -r = -m/n$, the present model is equivalent to the orthogonal projection matrix model with zero diagonal elements. But note that, in the exact sense, the equivalence here is in a statistical sense, where the fluctuation is of the order $1/\sqrt{n}$.

Next I show an interesting behavior in the recall dynamics that is specific to the two-stage-model.

The dynamics of the stage 2 is written using W^c in order to be compatible with Eq. (20). But what happens if we replace W^c with W_0^c as is often done in the studies of associative memory?⁴ The result is shown in Figure 2. The model discriminates the memory vector itself and noisy vectors (The exactness of the discrimination differs depending on the various parameters

⁴If the dynamics is simple and described by Eq. (2), there is no essential difference by the replacement.

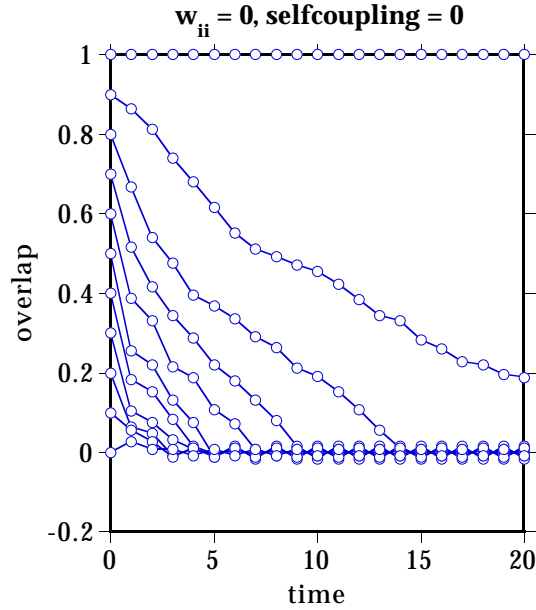


Figure 2: Effect of using W_0^c instead of W^c for associative memory with two-stage-neurons. $n = 500, r = 0.3, \tau = 5, \alpha = 0.5$.

involved, e.g. r, τ, α). This might be an important property as a subsystem of an autonomous memory system.

4 Interesting differences — a summary

Finally I summarize interesting differences between two-stage model and pseudoinverse-matrix model.

- Since the matrix involved in two-stage model is correlation one, addition and removal of vectors are easily manipulated by a simple procedure.
- Although two-stage neurons are more complex than threshold neurons, the entire net could be less complex than the net of threshold neurons with pseudoinverse or similar types of synapses. This is because, in general, the number of synapses in a neural net with n neurons is $O(n^2)$.
- Pseudoinverse-type associative memory is more sensitive than correlation-type to noise in storage processes. In two-stage model the sensitivity could be adjusted by choosing suitable degree τ .
- By replacing W^c in Eq. (22) and (23) with W_0^c , and taking τ large enough, the associative memory performs as a “match detector”. That is, if the initial condition is one of the embedded vectors itself, the state of the net remain unchanged, but if it is noisy, the state goes far away from it converging to the state that is correlated to neither of embedded vectors.

References

- [1] Amari, S. (1977): Neural theory of association and concept-formation, *Biological Cybernetics*, **26**, 175–185
- [2] Amari, S. and Maginu, K. (1988): Statistical neurodynamics of associative memory, *Neural Networks*, **1**, 63–73
- [3] Amari, S. and Yanai, H.-F. (1993): Statistical neurodynamics of various types of associative nets, *Associative Neural Memories — Theory and Implementation*, M. H. Hassoun ed. , Oxford University Press: New York
- [4] Caianiello, E. (1961): Outline of a theory of thought-processes and thinking machines, *J. Theoret. Biol.*, **2**, 204–235
- [5] Domany, E., Kinzel, W. and Meir, R. (1989): Layered neural networks, *J. Phys. A: Math. Gen.*, **22**, 2081–2102
- [6] Kanter, I. and Sompolinsky, H. (1987): Associative recall of memory without errors, *Phys. Rev. A*, **35**, 380–392
- [7] Kohonen, T. (1989): *Self-Organization and Associative Memory*, Springer-Verlag
- [8] Morita, M. (1993): Associative memory with nonmonotone dynamics, *Neural Networks*, **6**, 115–126
- [9] Okada, M. (1995): A hierarchy of macrodynamical equations for associative memory, *Neural Networks*, **8**, 833–838
- [10] Shiino, M. and Fukai, T. (1993): Self-consistent signal-to-noise analysis of the statistical behavior of analog neural networks and enhancement of the storage capacity, *Physical Review E*, **48**, 867–897
- [11] Yanai, H.-F. and Amari, S. (1993): A theory on a neural net with non-monotone neurons, *Proceedings of 1993 IEEE Conference on Neural Networks*, 1385–1390
- [12] Yanai, H.-F. and Amari, S. (1996): Auto-associative memory with two-stage dynamics of non-monotonic neurons, *IEEE Trans. Neural Networks*, **7**, Issue 3, in press
- [13] Yanai, H.-F. and Sawada, Y. (1990): Integrator neurons for analogue neural networks, *IEEE Trans. on Circuits and Systems*, **37**, 854–856
- [14] Yanai, H.-F., Sawada, Y. and Yoshizawa, S. (1991): Dynamics of an auto-associative neural network model with arbitrary connectivity and noise in the threshold, *Network: Computation in Neural Systems*, **2**, 295–314
- [15] Yoshizawa, S., Morita, M. and Amari, S. (1993): Capacity of associative memory using a non-monotonic neuron model, *Neural Networks*, **6**, 167–176