657 Assignment 2

Zhen, Gong
20673670
z33gong@uwaterloo.ca

Wang, Eugene
20657282
yj9wang@uwaterloo.ca

June 20, 2022

Q1

$$
\begin{aligned}
\frac{\delta J(n)}{\delta w_{(n)}} &= \frac{\delta}{\delta w_{(n)}}\frac{1}{2}[y_d(n) - \sum_{k=1}^{N} w_k(n) exp(-\frac{||\mathbf{x}(n)-\mathbf{c}_k(n)||^2}{\sigma_k^2(n)})]^2 \\
&= (y_d(n) - \sum_{k=1}^{N} w_k(n) exp(-\frac{||\mathbf{x}(n)-\mathbf{c}_k(n)||^2}{\sigma_k^2(n)})(-\sum_{k=1}^{N} exp(-\frac{||\mathbf{x}(n)-\mathbf{c}_k(n)||^2}{\sigma_k^2(n)}) \\
&= -(y_d(n) - y(n))(\sum_{k=1}^{N} \phi\{\mathbf{x}(n), \mathbf{c}_k(n), \sigma_k(n)\}) \\
&= -e(n)\psi(n)
\end{aligned}
$$

$$
\begin{aligned}
&Therefore, \\
\mathbf{w}(n+1) &= \mathbf{w}(n) - \mu_w \frac{\delta J(n)}{\delta w(n)} \\
&= \mathbf{w}(n) + \mu_w e(n)\psi(n)
\end{aligned}
$$

$$
\begin{aligned}
\frac{\delta J(n)}{\delta \mathbf{c}_k(n)} &= \frac{\delta}{\delta \mathbf{c}_k(n)}\frac{1}{2}[y_d(n) - \sum_{k=1}^{N} w_k(n) exp(-\frac{||\mathbf{x}(n)-\mathbf{c}_k(n)||^2}{\sigma_k^2(n)})]^2 \\
&= (y_d(n) - \sum_{k=1}^{N} w_k(n) exp(-\frac{||\mathbf{x}(n)-\mathbf{c}_k(n)||^2}{\sigma_k^2(n)})(-w_k(n) exp(-\frac{||\mathbf{x}(n)-\mathbf{c}_k(n)||^2}{\sigma_k^2(n)})(-\frac{2}{\sigma_k^2(n)}(x(n)-\mathbf{c}_k(n))(-1)) \\
&= -(y_d(n) - y(n))(\frac{2w_k(n)}{\sigma_k^2(n)})\phi\{\mathbf{x}(n), \mathbf{c}_k(n), \sigma_k(n)\}(x(n)-\mathbf{c}_k(n)) \\
&= -(\frac{e(n)2w_k(n)}{\sigma_k^2(n)})\phi\{\mathbf{x}(n), \mathbf{c}_k(n), \sigma_k(n)\}(x(n)-\mathbf{c}_k(n))
\end{aligned}
$$

$$
\begin{aligned}
&Therefore, \\
\mathbf{c}_k(n+1) &= \mathbf{c}_k(n) - \mu_c \frac{\delta J(n)}{\delta \mathbf{c}_k(n)} \\
&= \mathbf{c}_k(n) + \mu_c(\frac{e(n)2w_k(n)}{\sigma_k^2(n)})\phi\{\mathbf{x}(n), \mathbf{c}_k(n), \sigma_k(n)\}[\mathbf{x}(n)-\mathbf{c}_k(n)]
\end{aligned}
$$

$$
\begin{aligned}
\frac{\delta J(n)}{\delta \sigma_k(n)} &= \frac{\delta}{\delta \sigma_k(n)}\frac{1}{2}[y_d(n) - \sum_{k=1}^{N} w_k(n) exp(-\frac{||\mathbf{x}(n)-\mathbf{c}_k(n)||^2}{\sigma_k^2(n)})]^2 \\
&= (y_d(n) - \sum_{k=1}^{N} w_k(n) exp(-\frac{||\mathbf{x}(n)-\mathbf{c}_k(n)||^2}{\sigma_k^2(n)})(-w_k(n) exp(-\frac{||\mathbf{x}(n)-\mathbf{c}_k(n)||^2}{\sigma_k^2(n)})(\frac{-||\mathbf{x}(n)-\mathbf{c}_k(n)||^2}{\sigma_k^3(n)}(-2)) \\
&= -(y_d(n) - y(n))(\frac{2w_k(n)}{\sigma_k^3(n)})\phi\{\mathbf{x}(n), \mathbf{c}_k(n), \sigma_k(n)\}||x(n)-\mathbf{c}_k(n)||^2 \\
&= -(\frac{e(n)2w_k(n)}{\sigma_k^3(n)})\phi\{\mathbf{x}(n), \mathbf{c}_k(n), \sigma_k(n)\}||x(n)-\mathbf{c}_k(n)||^2
\end{aligned}
$$

$$
\begin{aligned}
&Therefore, \\
\sigma_k(n+1) &= \sigma_k(n) - \mu_\sigma \frac{\delta J(n)}{\delta \sigma_k(n)} \\
&= \sigma_k(n) + \mu_\sigma(\frac{e(n)2w_k(n)}{\sigma_k^3(n)})\phi\{\mathbf{x}(n), \mathbf{c}_k(n), \sigma_k(n)\}||x(n)-\mathbf{c}_k(n)||^2
\end{aligned}
$$

The 3 proofs were finished, and the coefficient 2 is just a scalar which will not affect the optimization.

Q2

Question: The paper titled "The Capacity of the Hopfield Associative Memory" can be accessed through UWaterloo Library. It is also uploaded in the assignment folder on LEARN. In your words, and in no longer than 2 pages, Summarize the paper thoroughly heightening the subject of research, the major contributions, the conclusions, and your own commentary.

1 answer

It is always fill my scholarly heart with a sense of joy to review great works that establish an entirely new unprecedented field. The Robert McEliece et al's "Copacity of the Hopefield Associative Memory" is one such work that exemplifies such unprecedented novoelty; all the while providing the rigorous proofs that supports its points.

1.1 Subject of Study

The author of the paper investigate associete memory proposed by Hopfield. However the author noticed that to efficiently utilized connectiveness and storage efficiency one can tap into the structure of that of an encoder. For example the encoder that encodes the bits in m fundamental memories as codwords, and will "have very low rates and hence find limited and specialized use for channel coding". From here the author noticed that he is able to find weights that connect two states in the network. In this case Hamming distance is benchmarked as the natural similarity measures (n) between the two states in binary space. Anything less than n/2 will work.

1.2 Contribution

This is where the author starts to make his contribution towards the scientific community and for the benefits of human tech advancement overall.

In the introduction, the author discussed two ways of state change of a neuron: 1. synchronous operation, each of the n neurons evaluates and updates its state in parallel. 2. In asynchronous operation, the components of the current state vector x are updated sequentially to produce a new state vector.

1.2.1 Sum-of-outer Products Connection matrix

A Hopfield connection matrix for a set of m memories x1 to xm can be calculated as the follow

$$T = \sum_{\alpha=1}^{m} (x^{\alpha})((x^{\alpha})^{T} - I_n).$$

In a sum of outer products connection matrix once a connection matrix T has been calculated all other x^{α} will be forgotten; thus the importance to learn. In the paper a proof was provided that for number of memories m small enough comprared to the number of components n, the memory is stable. Then the author further elaborate on method that compared subpar to this sum of matrix methodology.

1.2.2 Stability of Memory

As alluded to earlier when discussing stability of memories in sum of product matrix, the author described several error correction method that would remediate had a unknown component error. The author described 3 possibilities of convergence as: sphere of radius pn directly attacted to its fundamental memory, a random step in the right direction, and components change back and forth during sojourn but get better on average.

1.2.3 Encoding Method

The author talked about the capacity of the encoding method, and attributed it as a growth rate. The capacity when direct attacttion is desired and fundamental memories recallable correctly is

$$\frac{(1-2p)^2}{4} n/logn \tag{1}$$

for $0 < p \leq 1$ maximum number of memories stored in Hopfield is gotten through Hamming distance n/(2 log n).

1.3 Conclusion

The paper concluded that for capacity m of a Hopfield associetive memory of length n with m of random independent +-1 probability at most pn away from fundamental memory is 1/2 eq(1) if no fundamental memory can be exponential then m is eq(1). If some moves permitted a small fraction of fundamental memorys then n/(2logn) else n/(4 log n).

1.4 Commentary

Indeed, this is a spectacular work, a very indepth one that examine the nature of Hopfield network and how itse capacity can be linked with that of its stability and its learning property through its connectiveness. Going forward we should definitely integrate this in future auto association and optimization tasks. I see it as useful in computing the mechanical action of humans and structures in an autonomous vehicle which is my field of research.

Q3

Write a python code that creates an RBF Network to approximate the mapping defined by ...

Please see A2Q3.ipynb write up paragraphs.

Q4

(a) Please see the figures below

(b) Conclusion: We can observe that when σ_0 is small, the range of each cluster is small and the boundaries are sharp and clear. As σ_0 gets larger, the range of each cluster gets bigger and the boundaries turns smooth and the picture becomes vague. The output is better as the number the number of epochs is growing, and a model with smaller σ_0 converges faster.

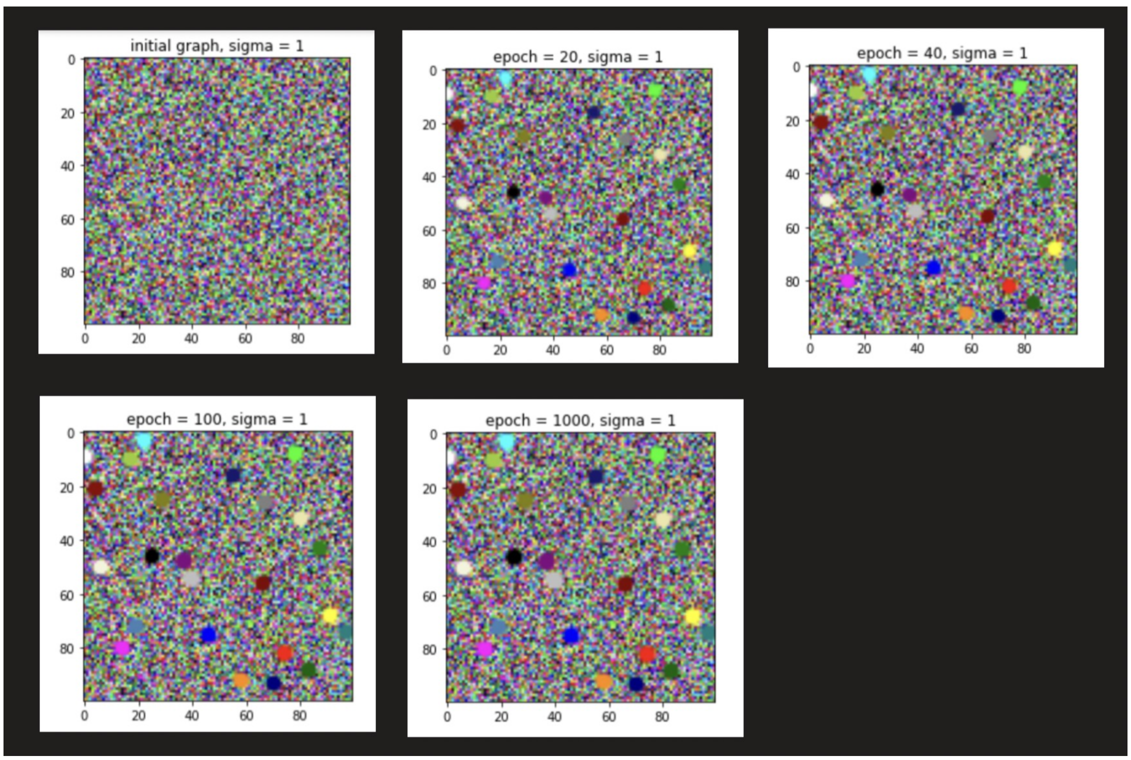

Figure 4.1 Figures of SOM, given $\sigma_0 = 1$

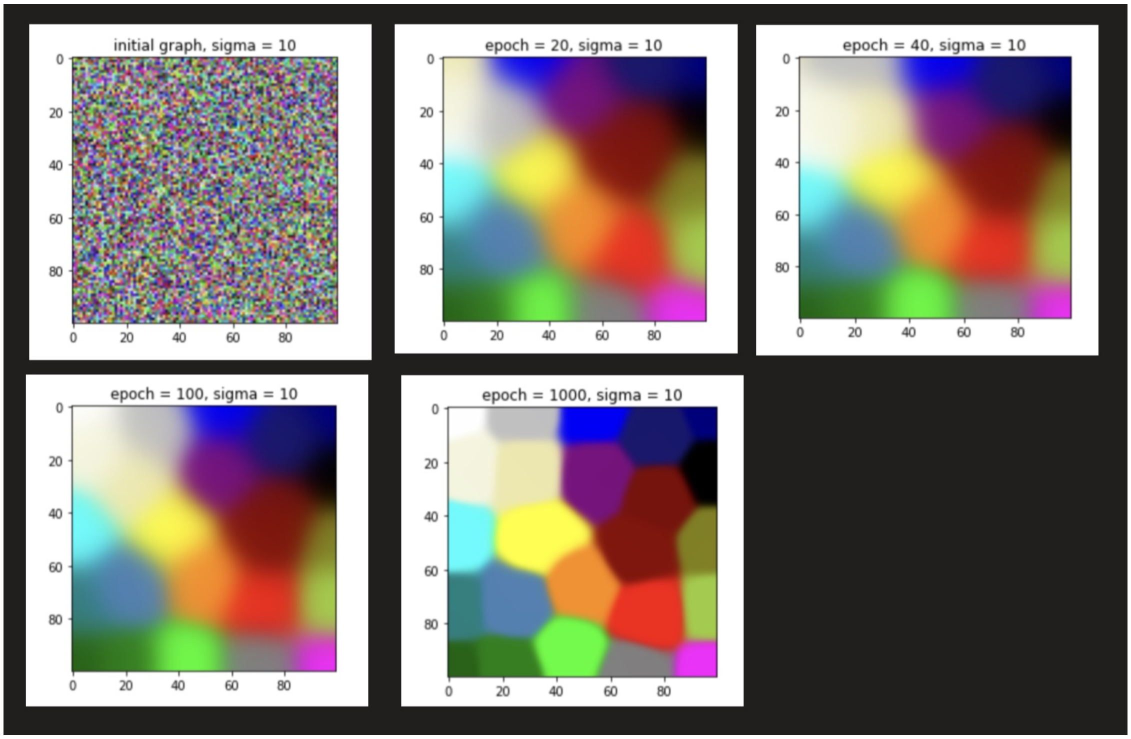

Figure 4.2 Figures of SOM, given $\sigma_0 = 10$

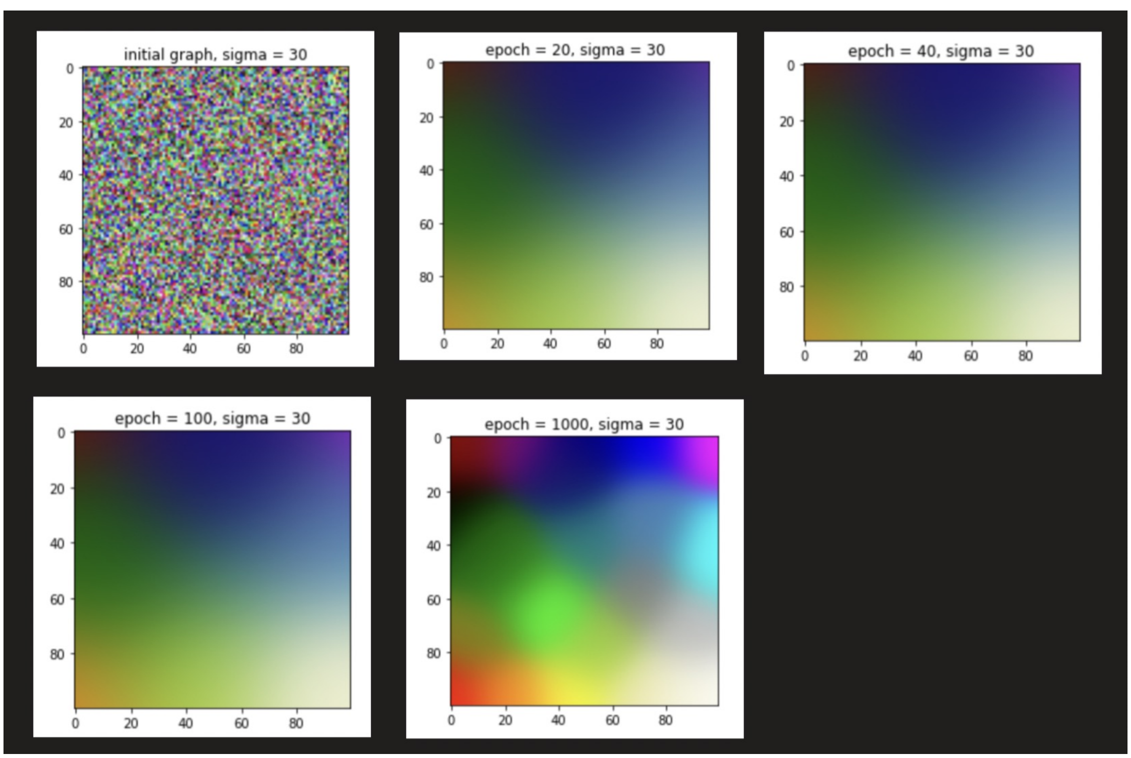

Figure 4.3 Figures of SOM, given $\sigma_0 = 30$

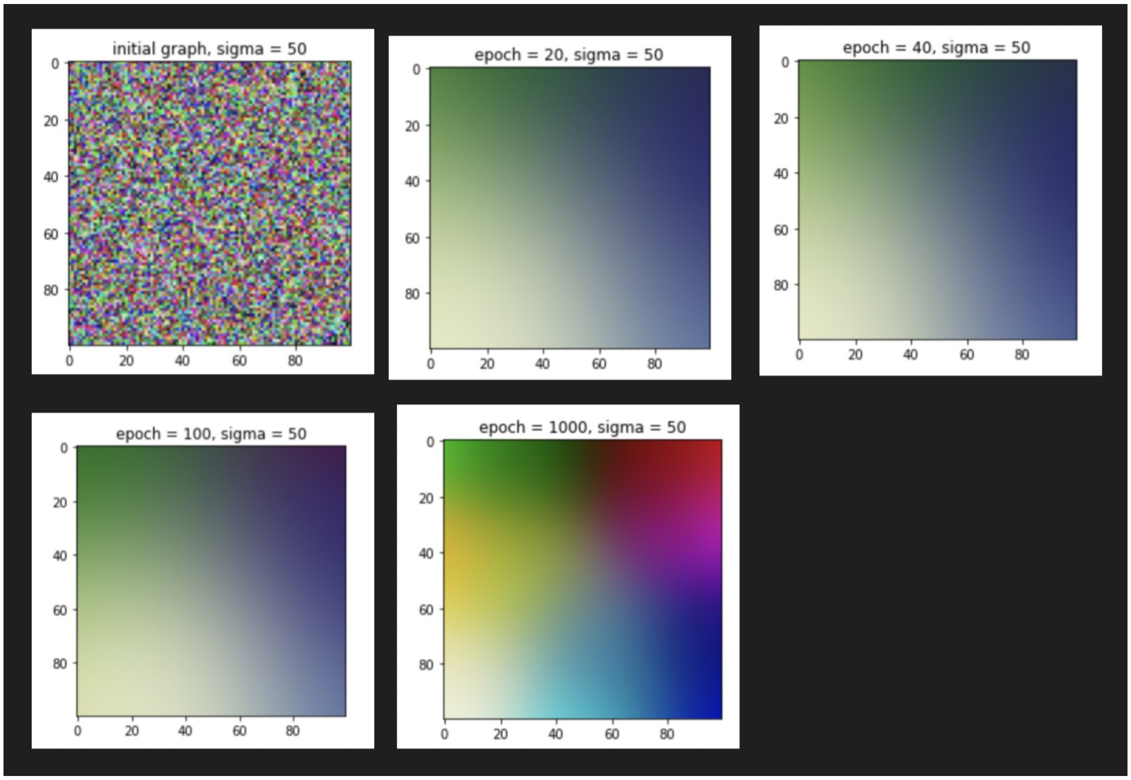

Figure 4.4 Figures of SOM, given $\sigma_0 = 50$

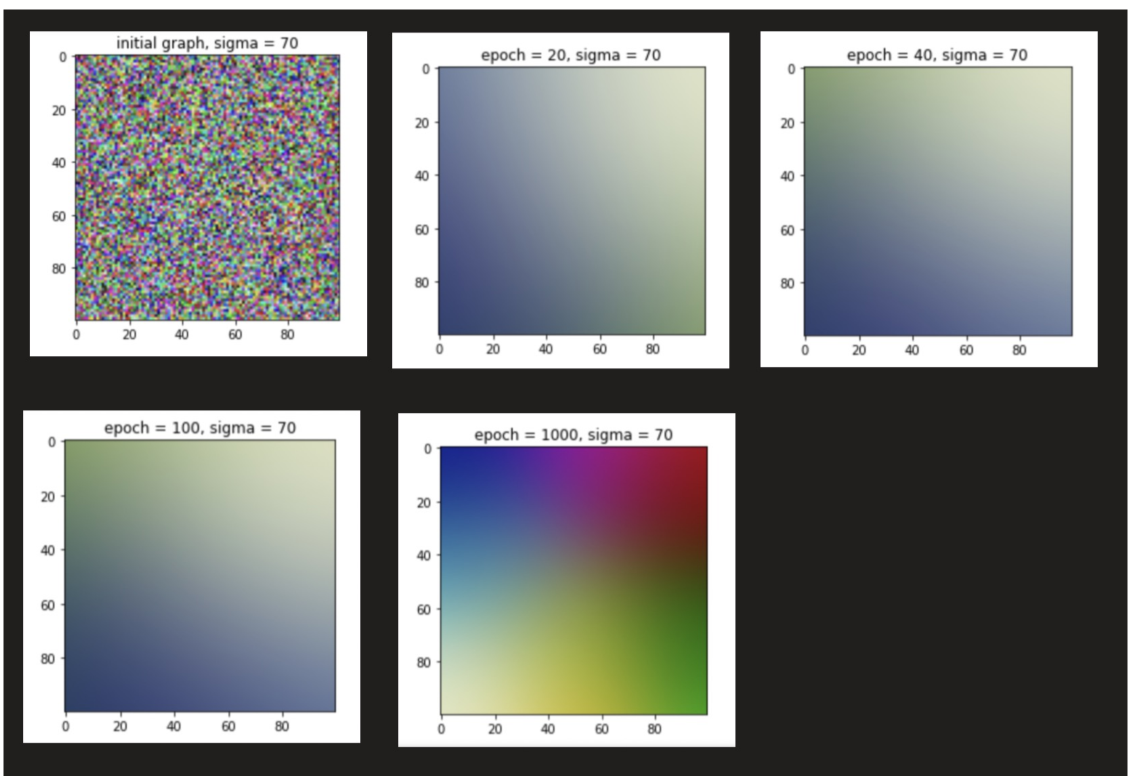

Figure 4.5 Figures of SOM, given $\sigma_0 = 70$