AN ALGORITHM FOR COMPUTATION OF INTER-PATTERN INTERFERENCE NOISE IN BAM

Dilip Sarkar

Abstract: Standard Bidirectional Associative Memory (BAM) stores sum-of-the-correlation-matrices of the pairs of patterns. When a pattern of an encoded pair is presented, the other is expected to be recalled. It has been shown that standard BAM cannot correctly recall a pattern pair if it is not at local minima of the energy function. To overcome this problem, novel methods for encoding have been proposed. The efficient novel-encoding methods require knowledge of the interference noise in the standard BAM. In this paper, we propose an algorithm for computing the exact amount of interference noise in standard encoding of BAM. The computational complexity of the algorithm is the same as that of computing the correlation matrix for the standard BAM.

Key words: Bidirectional associative memory (BAM), coding strategies, interference noise, multiple training

Received: April 30, 2001
Revised and accepted: July 31, 2001

1. Introduction

Associative information recalling process of humans has motivated researchers to work in the area of associative memory. In an associative memory system, related information can be recalled by giving only a part of it as input. For an introduction to the subject one can refer to Kohonen’s book [2]. Kosko has added a new dimension to the associative memory by introducing bidirectional associative memory (BAM) [4], [3], [5]. In a BAM system, a set of pattern-pairs \((A_i, B_i)\) are stored. To recall \(B_i\), \(A_i\) is supplied as input to the system. For recalling \(A_i\), system takes \(B_i\) as input. A BAM is usually implemented using two-layer nonlinear neurons. A brief introduction of the standard BAM is presented next.

*Dilip Sarkar
Department of Computer Science, University of Miami, Coral Gables, FL 33124, e-mail: sarkar@cal.cs.miami.edu

©ICS AS CR 2001
1.1 BAM: A Brief Review

Let \( \{(A_1, B_1), (A_2, B_2), ..., (A_p, B_p)\} \) be \( p \) training pairs, where \( A_i = [a_{i1}, a_{i2}, ..., a_{in}] \), and \( B_i = [b_{i1}, b_{i2}, ..., b_{im}] \). In binary mode \( a_{ij}, b_{ij} \) are 0 or 1, and in bipolar mode \( a_{ij}, b_{ij} \) are \(-1\) and \(1\). Since the conversion from a binary to a bipolar mode is an easy task, and the bipolar mode encoding enhances recalling performance [3], without any loss of generality, in the sequel it is assumed that \( a_{ij}, b_{ij} \) are in a bipolar mode. For encoding these training pairs a correlation matrix is constructed as \( M = \sum_{i=1}^{p} A_i^T B_i \), where \( A_i^T \) is the transpose of \( A_i \). This correlation matrix \( M \) defines the weights of a BAM. A set of nonlinear activation functions \( F(.) \) is used to transform the input of neurons to their outputs. The activation functions used here are defined by, \( F(V) = [f(v_1), f(v_2), ..., f(v_n)] \), where \( V \) is a \( n \)-dimension vector \( (v_1, v_2, ..., v_n) \), and

\[
  f(v_i) = \begin{cases} 
  1 & \text{if } v_i > 0 \\
  \text{unchanged} & \text{if } v_i = 0 \\
  -1 & \text{if } v_i < 0 
  \end{cases}
\]

In such a BAM structure, if \( X_i^0 \) is applied as input, a bidirectional process proceeds as follows:

\[
  \begin{align*}
  F(X_i^0) & = Y_i^0 \\
  \rightarrow F(X_i^1) & = Y_i^1 \rightarrow F(Y_i^1 M^T) = X_i^2 \\
  & \vdots \\
  \rightarrow F(X_i^j) & = Y_i^j \rightarrow F(Y_i^j M^T) = X_i^{j+1} \\
  & \vdots \\
  \rightarrow F(X_i^f) & = Y_i^f \rightarrow F(Y_i^f M^T) = X_i^f 
  \end{align*}
\]

The final pair \( (X_f, Y_f) \) is one of the training pairs \((A_i, B_i)\) if the encoded pairs satisfy some constraints. These constraints were characterized in [4], [3], [5], [8]. An important property of the final state is that it has lower energy value (defined next) than any other neighboring state. Following Kosko [3], the energy function for a pattern pair \((A_i, B_i)\) is defined as, \( E_i = -A_i MB_i^T \). Thus for correct recalling of a training pair \((A_i, B_i)\), it is required to have a local minimum of the energy function at that point.

The standard BAM encoding fails to recall a pattern correctly if other patterns have too much interference with it. Examples of this phenomenon can be found in [8]. To overcome this problem, several novel-encoding methods have been proposed [8], [7], [9], [6], [1]. All but one of these novel-encoding methods eliminate the interference noise by augmenting the patterns or the correlation matrices. Thus computation of an exact interference noise is very useful in encoding the BAMs. In the next Section, we present a necessary condition for a correct recall and use it for obtaining an algorithm for computation of interference noise in recalling each pattern.
2. Inter-Pattern Interference Noise

The basic principle behind the recalling process can be explained as follows. When an element $A_i$ is premultiplied with the correlation matrix, we obtain:

$$A_iM = A_i \sum_{j=1}^{p} A_j^T B_j = A_i A_i^T B_i + \sum_{j \neq i}^{p} A_i A_j^T B_j$$

$$= n B_i + \sum_{j \neq i}^{p} A_i A_j^T B_j.$$

The last equation can be viewed as a sum of two $m$-dimension vectors and rewritten as,

$$A_i M = n B_i + C_i = D_i$$

where $C_i = [c_{i1}, c_{i2}, \ldots, c_{im}]$, such that $c_{ik} = \sum_{j \neq i}^{p} A_i A_j^T b_{jk}$, and $D_i = [d_{i1}, d_{i2}, \ldots, d_{im}]$, such that $d_{ik} = n b_{ik} + c_{ik}$. If $F(D_i) = B_i$, a correct recall is obtained. If we look more closely, we find that when the signs of $c_{ik}$ and $b_{ik}$ are the same, $f(d_{ik}) = b_{ik}$, which means that, $b_{ik}$ is correctly recalled. Now, what remains to be examined are situations where the signs of $b_{ik}$ and $c_{ik}$ are different. In this situation, two cases are possible and are discussed next:

2.1 Case 1: $b_{ik} = 1$, and $c_{ik} < 0$

In this case, if $|c_{ik}| \leq n$, we have $d_{ik} = (nb_{ik} + c_{ik}) \geq 0$. Thus $f(d_{ik}) = 1 = b_{ik}$, and a correct recall is obtained. On the other hand, if $|c_{ik}| > n$, we have $d_{ik} = (nb_{ik} + c_{ik}) < 0$. Hence, $f(d_{ik}) = -1 = -b_{ik}$, and during a recall this bit has an opposite sign. Therefore, during the encoding process a corrective measure is necessary to ensure a correct recall of this bit.

Let $d_{ik}' = d_{ik} + q_{ik} b_{ik} = d_{ik} + q_{ik}$ (since $b_{ik} = 1 > 0$, or $q_{ik} > -d_{ik} = -c_{ik} - nb_{ik} = |c_{ik}| - n$). Thus the bit $b_{ik}$ is correctly recalled if we choose $q_{ik} \geq (|c_{ik}| - n + 1)$ and add $q_{ik}$ with $d_{ik}$. Let $q_{ik}' = (|c_{ik}| - n + 1)$, that is, the smallest integer value of $q_{ik}$ that ensures a correct recall of $b_{ik}$. Next we discuss the other case.

2.2 Case 2: $b_{ik} = -1$, $c_{ik} \geq 0$

In this case, if $c_{ik} \leq n$, we have $d_{ik} = (nb_{ik} + c_{ik}) \leq 0$. Thus $f(d_{ik}) = -1 = b_{ik}$, and a correct recall is obtained. On the other hand, if $c_{ik} > n$, we have $d_{ik} = (nb_{ik} + c_{ik}) > 0$. Hence, $f(d_{ik}) = 1 = -b_{ik}$, and during a recall this bit has opposite sign. Hence, during the encoding process some corrective measure is necessary to obtain correct recall of this bit.

Let $d_{ik}' = d_{ik} + q_{ik} b_{ik} = d_{ik} + q_{ik}$ (since $b_{ik} = -1 < 0$, or $q_{ik} < -d_{ik}$, or $q_{ik} > c_{ik} + nb_{ik} = |c_{ik}| - n$). Thus the bit $b_{ik}$ is correctly recalled if we have $q_{ik} \geq (|c_{ik}| - n + 1)$ and add $q_{ik} b_{ik}$ with $d_{ik}$. Like Case 1, we define $q_{ik}'$ as the smallest integer value of $q_{ik}$ that assures a correct recall of $b_{ik}$.

It is interesting to note that if there is an incorrect recall of a bit $b_{ik}$ of $B_i$, we need to add $q_{ik}' b_{ik}$ with corresponding $d_{ik}$, where $q_{ik}'$ is a positive quantity in both the cases discussed. Let us assume that $q_{ik}' = 0$ if the corresponding $b_{ik}$ is correctly
recalled. Thus one can now find a positive quantity $q_i^*$ such that if $D_i' = D_i + q_i^* B_i$, pattern $B_i$ is correctly recalled. In fact, $q_i^* = \max_{1 \leq k \leq m} \{ q_{ik}^* \}$ assures a correct recall of $B_i$. Let $q_i^*$ be called a *noise correction coefficient* (NCC) of $B_i$. Although the preceding discussion and the analysis look a little complicated, one can use it to design a simple algorithm for computing NCCs. We present such an algorithm next.

3. **An Algorithm for Computation of NCCs**

First, we start with Equation (1), then we will introduce a few definitions and notations, and finally present a pseudo code of an algorithm for computation of NCCs. An example illustrating snapshots of major intermediate computation steps of the algorithm is presented in the next subsection.

From Equation (1) we write:

\[
\begin{align*}
A_1 M &= n B_1 + C_1 \\
A_2 M &= n B_2 + C_2 \\
&\vdots \quad \vdots \quad \vdots \\
A_i M &= n B_i + C_i \\
&\vdots \quad \vdots \quad \vdots \\
A_p M &= n B_p + C_p \\
\end{align*}
\]

(2)

In the matrix notation, these equations can be rewritten as,

\[
M = n \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_i \\ \vdots \\ A_p \end{bmatrix} = \begin{bmatrix} B_1 \\ B_2 \\ \vdots \\ B_i \\ \vdots \\ B_p \end{bmatrix} + \begin{bmatrix} C_1 \\ C_2 \\ \vdots \\ C_i \\ \vdots \\ C_p \end{bmatrix}
\]

Or,

\[AM = nB + C,
\]

where $A = [A_1 A_2 \ldots A_i \ldots A_p]^T$, $B = [B_1 B_2 \ldots B_i \ldots B_p]^T$, and $C = [C_1 C_2 \ldots C_i \ldots C_p]^T$.

Thus, one can compute $C$ using the following equation:

\[C = AM - nB.
\]

(3)

Matrix $C$ is a good starting point for computation of NCCs. It can be recalled from the previous section that if the signs of $c_{ij}$ and $b_{ij}$ are the same, $b_{ij}$ is recalled correctly. Thus to aid in the identification of these elements of $C$, a matrix $E$ is computed. In $E$, an entry $e_{ij}$ is set to $c_{ij} \times b_{ij}$. Thus if both $c_{ij}$ and $b_{ij}$ have the same sign, $e_{ij}$ is positive. From $E$ we remove the elements that are correctly recalled, and obtain a matrix $CC$. In other words, $cc_{ij} = 0$, if $e_{ij}$ is nonnegative, and $cc_{ij} = c_{ij}$ otherwise. Only non-zero entries of $CC$ are those that require
further closer examination as they correspond to possible bits that may not be recalled correctly. The two cases we discussed earlier can now be used to eliminate the entries of CC that are correctly recalled. To be more precise, if the magnitude of \( cc_{ij} \) is greater than \( n \), the corresponding bit \( b_{ij} \) is incorrectly recalled, and the minimum noise correction value \( q_{ij}^* = |cc_{ij}| - n + 1 \). Otherwise, the bit \( b_{ij} \) is correctly recalled. We store all the noise bits in a matrix \( Q^* \). A pseudo code for computation of matrix \( Q^* \) and \( q_i^* \)'s is presented next. It is assumed that the pattern pairs are available in the matrix form as denoted in Equation (3).

**PROCEDURE** \((A, B, M, n)\);

```
BEGIN

\( C := AM - nB; \) \% \( n \) is the dimension of \( A_i s \)

FOR \( i := 1 \) TO \( p \) DO \% \( p \) is the number of the input pairs

FOR \( j := 1 \) TO \( m \) DO \% \( m \) is the dimension of \( B_i s \)

\( e(i, j) := c(i, j) * b(i, j); \)

IF \( e(i, j) \geq 0 \) THEN \( cc(i, j) := 0; \)
ELSE \( cc(i, j) := c(i, j); \)

FOR \( i := 1 \) TO \( p \) DO

FOR \( j := 1 \) TO \( m \) DO

IF \( |cc(i, j)| > n \) THEN \( q^*(i, j) := |cc(i, j)| - n + 1; \)
ELSE \( q^*(i, j) := 0; \)

FOR \( i := 1 \) TO \( p \) DO

\( q_i^* := 0; \)

FOR \( j := 1 \) TO \( m \) DO

IF \( q^*(i, j) > q_i^* \) THEN \( q_i^* := q^*(i, j); \)

END;
```

### 3.1 An Example

In this section we illustrate an application of the algorithm for computation of NCCs. We take the six pattern pairs \((A_i, B_i)\) for \( i = 1, 2, ..., 6 \) from Wang and Lee [6], convert them in a bipolar mode, and show them as row vectors.

\[
A_1 = \begin{bmatrix} 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \end{bmatrix} \\
A_2 = \begin{bmatrix} -1 & 1 & -1 & 1 & 1 & 1 & 1 & -1 \end{bmatrix} \\
A_3 = \begin{bmatrix} 1 & 1 & -1 & -1 & 1 & 1 & -1 & 1 \end{bmatrix} \\
A_4 = \begin{bmatrix} -1 & -1 & -1 & 1 & -1 & 1 & 1 & 1 \end{bmatrix} \\
A_5 = \begin{bmatrix} 1 & -1 & -1 & -1 & 1 & 1 & -1 & 1 \end{bmatrix} \\
A_6 = \begin{bmatrix} -1 & -1 & 1 & 1 & 1 & -1 & -1 & 1 \end{bmatrix}
\]
\[ B_1 = \begin{bmatrix} 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \end{bmatrix} \]
\[ B_2 = \begin{bmatrix} -1 & 1 & -1 & 1 & -1 & 1 & 1 \end{bmatrix} \]
\[ B_3 = \begin{bmatrix} -1 & -1 & 1 & -1 & -1 & 1 & -1 \end{bmatrix} \]
\[ B_4 = \begin{bmatrix} 1 & -1 & -1 & 1 & -1 & 1 & 1 \end{bmatrix} \]
\[ B_5 = \begin{bmatrix} -1 & 1 & 1 & -1 & 1 & -1 & 1 \end{bmatrix} \]
\[ B_6 = \begin{bmatrix} 1 & -1 & -1 & -1 & 1 & -1 & -1 \end{bmatrix} \]

The corresponding correlation matrix \( M \) is given next.

\[
M = \begin{bmatrix}
-2 & 0 & 6 & -4 & 4 & -4 & 2 & 2 & -2 & -2 \\
-2 & 0 & 2 & 0 & 0 & -4 & 6 & -2 & 2 & -2 \\
4 & -2 & 0 & -2 & 2 & 2 & 0 & 0 & 0 & -4 \\
4 & -2 & -4 & 2 & -2 & 2 & 0 & -4 & 4 & 0 \\
-2 & 0 & 2 & -4 & 0 & -4 & 2 & -2 & -2 & -2 \\
-2 & 4 & -2 & 4 & 0 & 0 & -2 & 2 & 2 & 6 \\
2 & 0 & -2 & 4 & 0 & 4 & -2 & 2 & 2 & 2 \\
-2 & 0 & -2 & 4 & -4 & 0 & 2 & -2 & 2 & 2 \\
0 & 2 & -4 & 2 & -2 & 2 & -4 & 0 & 0 & 4 \\
0 & -2 & 0 & 2 & -2 & 2 & 0 & 0 & 0 & 0 \\
\end{bmatrix}
\]

A 6 by 10 matrix \( A \) (respectively \( B \)) is formed by using vectors \( A_4 \) (respectively \( B_4 \)). Using \( A, B, M \), and the dimension of \( A_{18} \), in Equation (3) matrix \( C \) is computed and displayed next.

\[
C = \begin{bmatrix}
-6 & 2 & 6 & -14 & 2 & -6 & 6 & 2 & -14 & -14 \\
-6 & 6 & 10 & -2 & 10 & -10 & 10 & 6 & 2 & -2 \\
-6 & 14 & -14 & 18 & -6 & 10 & -6 & 10 & 2 & -2 \\
2 & 6 & 2 & -2 & 10 & -6 & 2 & 2 & 6 & 2 \\
\end{bmatrix}
\]

The first two nested FOR LOOPS compute matrices \( E \) and \( CC \). Bits corresponding to the nonnegative entries of \( E \) are correctly recalled, and hence set to zero for obtaining matrix \( CC \). These two intermediate matrices are given next.

\[
E = \begin{bmatrix}
-6 & -2 & 6 & 14 & 2 & 6 & 6 & -2 & -14 & 14 \\
6 & -6 & 10 & 2 & -10 & 10 & 10 & -6 & -2 & 2 \\
-6 & 14 & 14 & 18 & 6 & 10 & 6 & -10 & 2 & 14 \\
2 & -6 & -2 & 2 & -10 & -6 & -2 & -2 & -6 & -2 \\
\end{bmatrix}
\]

\[
CC = \begin{bmatrix}
-6 & 2 & 0 & 0 & 0 & 0 & 0 & 2 & -14 & 0 \\
2 & -2 & 0 & -2 & 0 & 2 & -6 & 0 & -2 & 0 \\
0 & 6 & 0 & 0 & 10 & 0 & 0 & 6 & 2 & 0 \\
-6 & 14 & 0 & 0 & 0 & 0 & 0 & 10 & 0 & 0 \\
0 & 0 & -2 & 0 & -2 & 0 & 2 & -2 & 0 & -2 \\
0 & 6 & 2 & 0 & 10 & -6 & 2 & 2 & 6 & 2 \\
\end{bmatrix}
\]

Only the nonzero entries of matrix \( CC \) need further evaluation using justifications given in Cases 1 and 2. These cases require absolute values of the elements.
of $CC$. Each element of $CC$ whose absolute value is greater than $n$ (the dimension of $A_i$s) causes a recalling error of the corresponding bit of $B_i$s. The middle two nested FOR LOOPS use absolute values of the entries of $CC$ and compute $Q$.

$$Q = \begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 5 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}$$

Finally, the last two nested FOR LOOPS compute $q_1^* = 5$, $q_2^* = 0$, $q_3^* = 0$, $q_4^* = 5$, $q_5^* = 0$, and $q_6^* = 0$. It is clear that only two bits $b_{19}$ and $b_{42}$ are incorrectly recalled.

4. Conclusion

It is known that the standard BAM cannot correctly recall a pattern pair if it is not at the local minimum of the energy function. To overcome this problem, several novel methods for encoding BAM have been proposed. Most of these novel-encoding methods require the value of interference noise in the standard BAM for efficient encoding. Here we have presented an algorithm for computing inter-pattern interference noise and corresponding NCCs (noise correction coefficients). The computational complexity of the algorithm is the same as that of computing the correlation matrix for the standard BAM.

References