Implemented Equations

Keyan Ghazi-Zahedi [*1]

[1]Max Planck Institute for Mathematics in the Sciences, Leipzig, Germany

December 1, 2016

1 Binary CRBM

$$V \in \mathbb{R}^{m \times k} \tag{1}$$
$$W \in \mathbb{R}^{m \times n} \tag{2}$$
$$b \in \mathbb{R}^{n} \tag{3}$$
$$c \in \mathbb{R}^{m} \tag{4}$$
$$y \in \mathbb{R}^{k} \tag{5}$$
$$x \in \mathbb{R}^{n} \tag{6}$$
$$z \in \mathbb{R}^{m} \tag{7}$$

Binary up:

$$z = c + V \cdot y + W \cdot x \tag{8}$$
$$z = \text{sigm}(z) \tag{9}$$
$$z = \text{set randomly to true false, based on } \mathit{uniform}() > z_i \tag{10}$$

Algorithm 1 Binary up

Input: $x \in \mathbb{R}^n$
Input: $y \in \mathbb{R}^k$
Output: $z \in \mathbb{R}^m$

1: $z \leftarrow c + V \cdot y + W \cdot x$
2: $z \leftarrow \text{sigm}(z)$
3: $z \leftarrow$ set randomly to true false, based on $\mathit{uniform}() > z_i$

*zahedi@mis.mpg.de

Algorithm 2 Binary down

Input: $z \in \mathbb{R}^m$
Output: $x \in \mathbb{R}^n$

$x \leftarrow b + W^T \cdot z$
$x \leftarrow \text{sigm}(x)$
$x \leftarrow$ set randomly to true false, based on $\mathit{uniform}() > x_i$

Algorithm 3 Control update

Input: $x \in \mathbb{R}^n$
Input: $y \in \mathbb{R}^k$
Input: $u \in \mathbb{N}$
Output: $x \in \mathbb{R}^n$
Output: $z \in \mathbb{R}^m$

$z \leftarrow \text{binary\_up}(y, x)$
for i from 1 to u **do**
    $x \leftarrow \text{binary\_down}(z)$
    $z \leftarrow \text{binary\_up}(y, x)$
end for

Algorithm 4 Learn update

Input: $x \in \mathbb{R}^n$
Input: $y \in \mathbb{R}^k$
Input: $u \in \mathbb{N}$
Output: $x \in \mathbb{R}^n$
Output: $z \in \mathbb{R}^m$

$z \leftarrow \text{binary\_up}(y, x)$
for i from 1 to $u - 1$ **do**
    $x \leftarrow \text{binary\_down}(z)$
    $z \leftarrow \text{binary\_up}(y, x)$
end for
$x \leftarrow \text{binary\_down}(z)$

Algorithm 5 Initialise b

Input: $n \in \mathbb{N}$ ▷ Nr. of input units
Input: $t \in \mathbb{N}$ ▷ Nr. of rows in the data

for i from 1 to n **do**
    $s \leftarrow \text{columnsum}(\hat{S}(i))$
    if $s = t$ **then** ▷ Only 1's in the column
        $p \leftarrow {}^{t-1}/_{t}$
    else if $s == 0$ **then** ▷ Only 0's in the column
        $p \leftarrow {}^{1}/_{t}$
    else
        $p \leftarrow {}^{s}/_{t}$
    end if
end for

Algorithm 6 Training

Input: CRBM
Input: $S \in \mathbb{R}^{t \times |S|}$
Input: $A \in \mathbb{R}^{t \times |A|}$
Input: bins $\in \mathbb{N}$
Input: batchsize $\in \mathbb{N}$
Input: numepochs $\in \mathbb{N}$
Output: $x \in \mathbb{R}^n$
Output: $z \in \mathbb{R}^m$
Output: $m \in \mathbb{R}$ ▷ Momentum parameter
Output: $\gamma \in \mathbb{R}$ ▷ Weight cost

$N \leftarrow \lceil \log_2(\text{bins}) \rceil$
$\hat{S} \leftarrow \text{binarise}(S)$
$\hat{A} \leftarrow \text{binarise}(A)$
$rbm.W \leftarrow \mathcal{N}(0, 0.01)$
$rbm.V \leftarrow \mathcal{N}(0, 0.01)$
$rbm.c \leftarrow 0$
$rbm.b \leftarrow$ initialise based on data (see Alg. 5
for i in 1 to numepochs **do**
    $r \leftarrow \text{rand}(t - \text{batchsize})$
    $\text{indices} \leftarrow [r, r + \text{batchsize}]$
    $\tilde{S} \leftarrow \hat{S}(\text{rows} = \text{indices})$
    $\tilde{A} \leftarrow \hat{A}(\text{rows} = \text{indices})$
    $z \leftarrow \text{binary\_up}(\tilde{S}, \tilde{A})$
    $\bar{A} \leftarrow$ random action matrix $\in [0, 1]^{|\tilde{A}|}$
    $(\bar{A}, Z) \leftarrow \text{learn}(s, \bar{A})$
    $\mathbb{E}\{b\} \leftarrow \text{avg}(\tilde{A}) - \text{avg}(\bar{A})$
    $\mathbb{E}\{c\} \leftarrow \text{avg}(\tilde{z}) - \text{avg}(\bar{Z})$
    $\mathbb{E}\{W\} \leftarrow {}^{z^T a - Z^T \bar{A}}/_{\text{batchsize}}$
    $\mathbb{E}\{V\} \leftarrow {}^{z^T s - Z^T \bar{s}}/_{\text{batchsize}}$
    $b \leftarrow b + \alpha \mathbb{E}\{b\}$
    $c \leftarrow c + \alpha \mathbb{E}\{c\}$
    $W \leftarrow W + \alpha \mathbb{E}\{W\}$
    $V \leftarrow V + \alpha \mathbb{E}\{V\}$
    if $m > 0$ **then**
        $b \leftarrow b + \alpha \cdot m \cdot b^m$
        $c \leftarrow c + \alpha \cdot m \cdot c^m$
        $W \leftarrow W + \alpha \cdot m \cdot W^m$
        $V \leftarrow V + \alpha \cdot m \cdot V^m$
    end if
    if $\gamma > 0$ **then**
        $W \leftarrow (1 - \gamma) W$
        $V \leftarrow (1 - \gamma) V$
    end if
    $b^m \leftarrow \mathbb{E}(b)$
    $c^m \leftarrow \mathbb{E}(c)$
    $W^m \leftarrow \mathbb{E}(W)$
    $V^m \leftarrow \mathbb{E}(V)$
end for