Contents

1 Notation

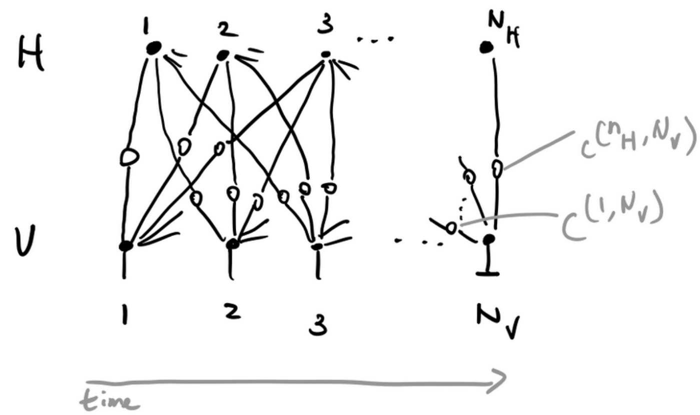
Firstly, define N_H and N_V to be the number of hidden and visible units respectively. H_h is a vector of weights for the hidden units, where $h = 1, 2, \ldots, N_H$. Similarly, V_v is a vector of weights for the visible units, where $v = 1, 2, \ldots, N_V$. W_{hv} is a matrix of weights. $C^{(hv)}_{kl}$ is the coupling matrix between the hth hidden unit and the vth visible unit. For simplicity, we will assume that C is a $d \times d$ matrix.

1.0.1 Note on the coordination number

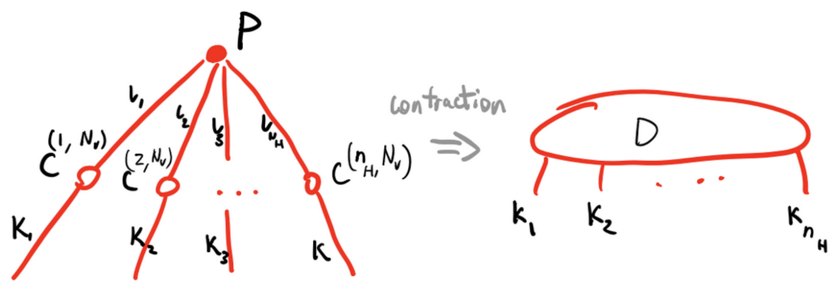
$n^{(h)}_V$ is the coordination number for visible units of the h^{th} hidden unit. This is the number of connected visible units. Similarly, $n^{(v)}_H$ is the coordination number for hidden units of the v^{th} visible unit.

Throughout this section, we are operating on the last visible unit (the ${N_V}^{\text{th}}$ unit), which is connected to $n^{(N_V)}_H$ hidden units. Since it is clear which visible unit we are referring to, the superscript will be omitted.

In the general case, $n_H = N_H$ as the hidden and visible units form a complete bipartite graph. In the Snake RBM case, each visible unit is connected to exactly one hidden unit so $n_H = 1$.



2 Imposing the causality condition on RBMs

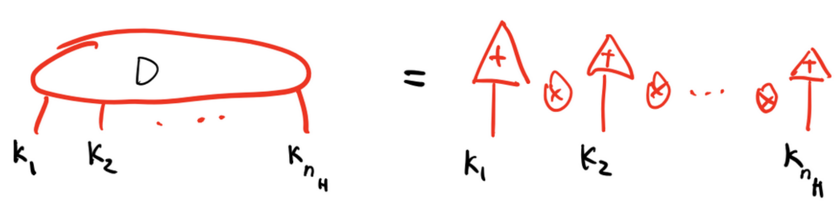


Let $D_{k_1 k_2 \ldots k_{n_H}}$ be the contraction of the red network, and $|+\rangle$ to be a vector of 1s.

If the causality condition were to hold, then,

$$D = \bigotimes_{n_H} |+\rangle \,. \tag{1}$$

This is represented visually as



The RHS is then a tensor of 1s. This amounts to an a set of equations where each element of D must equal 1.

2.1 Preliminary steps

Let P be the contraction of the vectorised identity with the copy tensor, and let $k = \sqrt{d}$. The index at which a one appears in the vectorised identity is given by a sequence $S = (s_1, s_2, \ldots, s_k)$ where $s_i = k(i-1)+i$ for $1 \leq i \leq k$.

$$P_{l_1 l_2 \ldots l_{n_H}} = \begin{cases} 1 & l_1 = l_2 = \cdots = l_{n_H} \text{ and } l_1 \in S \\ 0 & \text{otherwise} \end{cases}$$

2.2 Defining the coupling matrices

$C^{(hv)}_{kl}$ is a $d \times d$ coupling matrix between the hth hidden and vth visible unit.

In our notation, we extend k and l to range over $[1, d]$, so a more useful definition could be as follows.

$$C^{(hv)}_{kl} = e^{W_{h,v}(k-1)(l-1) + \frac{H_h(k-1)}{N_H} + \frac{V_v(l-1)}{N_V}} \tag{2}$$

2.3 Equation for elements of the contracted tensor

In fact, we can solve for specific elements of the D tensor.

$$\begin{aligned}
D_{k_1 k_2 \ldots k_{n_H}} &= \sum_{l_1 l_2 \ldots l_{n_H}} P_{l_1 l_2 \ldots l_{n_H}} C^{(1,N_V)}_{k_1 l_1} C^{(2,N_V)}_{k_2 l_2} \cdots C^{(n_H,N_V)}_{k_{n_H} l_{n_H}} && (3) \\
&= \sum_{s \in S} P_{ss\ldots s} C^{(1,N_V)}_{k_1 s} C^{(2,N_V)}_{k_2 s} \cdots C^{(n_H,N_V)}_{k_{n_H} s} && (4) \\
&= \sum_{s \in S} C^{(1,N_V)}_{k_1 s} C^{(2,N_V)}_{k_2 s} \cdots C^{(n_H,N_V)}_{k_{n_H} s} && (5) \\
&= \sum_{s \in S} \prod_{h=1}^{n_H} C^{(h,N_V)}_{k_h,s} && (6) \\
&= \sum_{s \in S} \prod_{h=1}^{n_H} e^{W_{h,N_V}(k_h-1)(s-1) + \frac{V_{N_V}(s-1)}{N_V} + \frac{H_h(k_h-1)}{N_H}} && (7)
\end{aligned}$$

2.3.1 A special case with no solutions

Using the general equation (7), we can solve for the very first element of the tensor, $D_{1,1,\ldots,1}$. (Here, let $k_1 = k_2 = \cdots = k_{n_H} = 1$.)

$$D_{1,1,\ldots,1} = \sum_{s\in S} \prod_{h=1}^{n_H} e^{(s-1)\frac{V_{N_V}}{N_V}} \tag{8}$$

$$= \sum_{s\in S} e^{(s-1)\frac{n_H V_{N_V}}{N_V}} \tag{9}$$

Suppose that $d = 4$ (the case that $C^{(hv)}$ is a 2×2 matrix), then $S = (1, 4)$.

$$D_{1,1,\ldots,1} = \sum_{s\in\{1,4\}} e^{(s-1)\frac{n_H V_{N_V}}{N_V}} \tag{10}$$

$$= e^0 + e^{\frac{3n_H V_{N_V}}{N_V}} \tag{11}$$

But requiring $D_{1,1,\ldots,1} = 1$ implies $\exp(\frac{3n_H V_{N_V}}{N_V}) = 0$, which has no solution.

In fact, the number of terms in the summation in (7) is equal to the length of $S = \sqrt{d}$.

2.4 Interpreting results for Uniform RBMs

Up until now, we have been considering the general case of fully-connected RBMs. Consider the special subset of RBMs where all coupling matrices are identical. That is, for all $1 \le h_1, h_2 \le n_H$, and $1 \le v_1, v_2 \le N_V$,

$$C^{(h_1 v_1)}_{kl} = C^{(h_2 v_2)}_{kl} \tag{12}$$

It immediately follows from the definition of the coupling matrices, that all entries of W, H, and V are also identical. So the subscripts for the indices can be ommitted. This means that (6) can be simplified.

$$D_{k_1 k_2 \ldots k_{n_H}} = \sum_{s\in S} \prod_{h=1}^{n_H} C^{(h,N_V)}_{k_h,s} \tag{6 revisited}$$

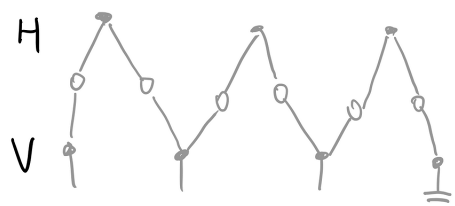
$$= \sum_{s\in S} \prod_{h=1}^{n_H} e^{W(k_h-1)(s-1)+\frac{V(s-1)}{N_V}+\frac{H(k_h-1)}{N_H}} \tag{13}$$

2.5 Interpreting results for minimally-connected RBMs

Consider a subset of RBMs that are connected in a minimal way…

We will call this the *Snake RBM*.

"Snake RBM" has this form.



Solving the Snake RBM amounts to solving the following simplified equation.