Current Model

To solve the blind signal problem, the current model I am exploring is a two layered neural network.

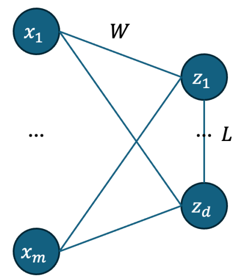

Figure 1 Two-layer neural network with lateral connections

Input layer $\boldsymbol{x} \in \mathbb{R}^m$, output layer $\boldsymbol{z} \in \mathbb{R}^d$, weight matrix $W \in \mathbb{R}^{d\times m}$.

We have $\boldsymbol{v} = W\boldsymbol{x}$. This is the initial value of the second layer activations, $\boldsymbol{z}$. Afterwards, we take gradient steps to minimise an energy function, modifying $\boldsymbol{z}$.

The energy function for the ith node in the second layer is given by

$$E_i = \frac{1}{2}\left[A(z_i - v_i)^2 + \frac{B}{d-1}\sum_{j\neq i}\langle x_i^2 x_j^2\rangle^2 + C\big(1 - var(x_i)\big)^2\right]$$

where A, B, C are hyperparameters that weight the importance of each term. The first term is the usual squared error, the second term aims to minimise the covariances and the final term aims to make the variance of each dimension close to 1.

To keep track of the variances and covariances we use an exponential moving average. These estimates are stored as weights in matrices V and L.

$$V_i \leftarrow \alpha z_i^2 + (1-\alpha)V_i \; for \; i \in \{1, \dots, d\}$$

$$L_{ij} \leftarrow \alpha z_i z_j + (1-\alpha)L_{ij} \; for \; i \neq j$$

Then,

$$\frac{\partial E_i}{\partial z_i} = A(z_i - v_i) + \frac{B}{d-1}\sum_{j\neq i} L_{ij}(\alpha)(z_j) + C(1 - V_i)(-2\alpha z_i)]$$

assuming $\langle x_i^2 x_j^2\rangle \approx L_{ij}, \frac{\partial L_{ij}}{\partial z_i} = \alpha z_j, var(z_i) \approx V_i, \frac{\partial V_i}{\partial z_i} = 2\alpha z_i$.

We take multiple gradient steps updating $\boldsymbol{z}$. Optionally, we can also introduce activation function in the second layer that we apply after relaxation. Then, we take a single gradient step updating W.

$$\frac{\partial E_i}{\partial W_{ij}} = A(z_i - v_i)(-x_j)$$

This is performed for every input datapoint and the process is repeated.