Neural networks

Restricted Boltzmann machine - definition

UNSUPERVISED LEARNING

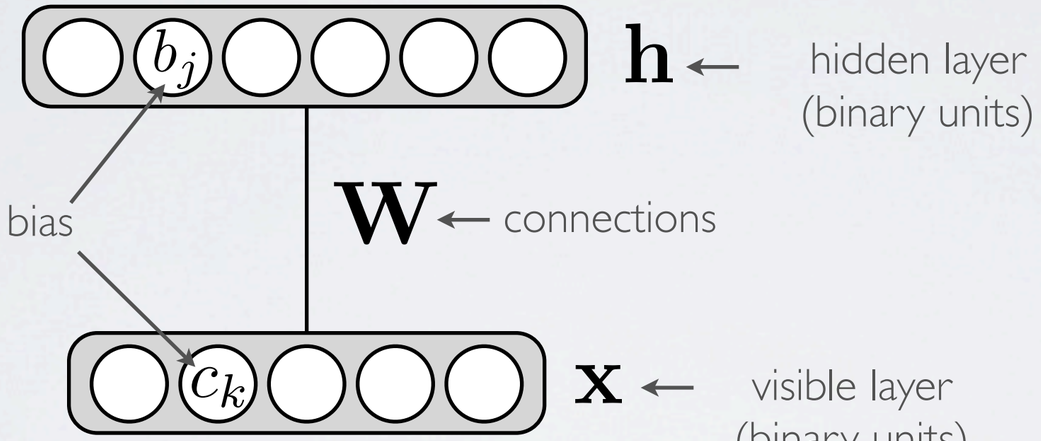
Topics: unsupervised learning

- Unsupervised learning: only use the inputs $\mathbf{x}^{(t)}$ for learning
  - automatically extract meaningful features for your data
  - leverage the availability of unlabeled data
  - add a data-dependent regularizer to training ($-\log p(\mathbf{x}^{(t)})$)

- We will see 3 neural networks for unsupervised learning
  - **restricted Boltzmann machines**
  - autoencoders
  - sparse coding model

RESTRICTED BOLTZMANN MACHINE

Topics: RBM, visible layer, hidden layer, energy function



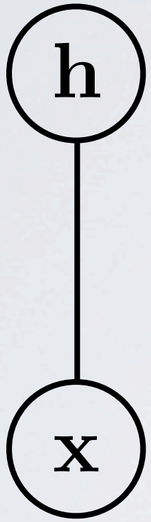
Energy function:
$$E(\mathbf{x}, \mathbf{h}) = -\mathbf{h}^\top \mathbf{W}\mathbf{x} - \mathbf{c}^\top \mathbf{x} - \mathbf{b}^\top \mathbf{h}$$
$$= -\sum_j \sum_k W_{j,k} h_j x_k - \sum_k c_k x_k - \sum_j b_j h_j$$

Distribution:
$$p(\mathbf{x}, \mathbf{h}) = \exp(-E(\mathbf{x}, \mathbf{h}))/Z$$

partition function
(intractable)

MARKOV NETWORK VIEW

Topics: Markov network (with vector nodes)

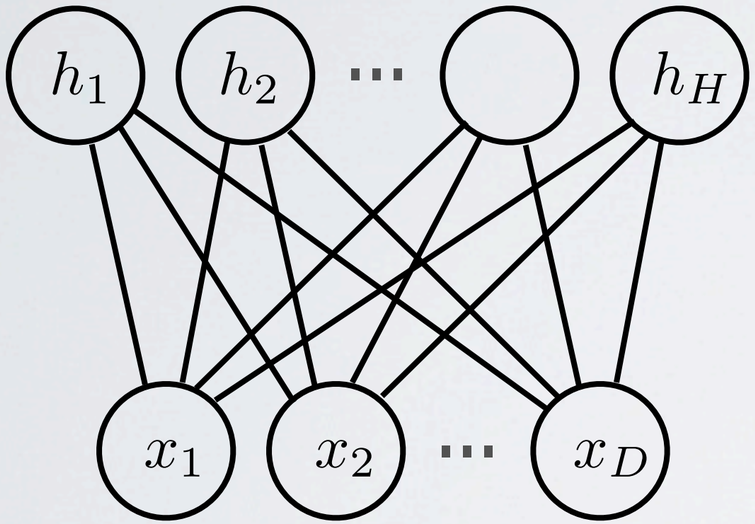


$$
\begin{aligned}
p(\mathbf{x}, \mathbf{h}) &= \exp(-E(\mathbf{x}, \mathbf{h}))/Z \\
&= \exp(\mathbf{h}^\top \mathbf{W}\mathbf{x} + \mathbf{c}^\top \mathbf{x} + \mathbf{b}^\top \mathbf{h})/Z \\
&= \underbrace{\exp(\mathbf{h}^\top \mathbf{W}\mathbf{x}) \exp(\mathbf{c}^\top \mathbf{x}) \exp(\mathbf{b}^\top \mathbf{h})}_{\text{factors}}/Z
\end{aligned}
$$

- The notation based on an energy function is simply an alternative to the representation as the product of factors

MARKOV NETWORK VIEW

Topics: Markov network (with scalar nodes)



$$p(\mathbf{x}, \mathbf{h}) = \frac{1}{Z} \overbrace{\prod_j \prod_k \exp(W_{j,k} h_j x_k)}^{\text{pair-wise factors}} \left.\begin{aligned} &\prod_k \exp(c_k x_k) \\ &\prod_j \exp(b_j h_j) \end{aligned}\right\} \text{unary factors}$$

- The scalar visualization is more informative of the structure within the vectors

FACTOR GRAPH VIEW

Topics: factor graph of an RBM