Neural networks

Restricted Boltzmann machine - extensions
Topics: RBM, visible layer, hidden layer, energy function

Energy function:

\[ E(x, h) = -h^T \mathbf{W} x - c^T x - b^T h \]

\[ = - \sum_j \sum_k W_{j,k} h_j x_k - \sum_k c_k x_k - \sum_j b_j h_j \]

Distribution:

\[ p(x, h) = \exp(-E(x, h))/Z \]

partition function (intractable)
GAUSSIAN-BERNOULLI RBM

**Topics:** Gaussian-Bernoulli RBM

- Inputs \( \mathbf{x} \) are unbounded reals
  - add a quadratic term to the energy function
    \[
    E(\mathbf{x}, \mathbf{h}) = -\mathbf{h}^T \mathbf{W} \mathbf{x} - \mathbf{c}^T \mathbf{x} - \mathbf{b}^T \mathbf{h} + \frac{1}{2} \mathbf{x}^T \mathbf{x}
    \]
  - only thing that changes is that \( p(\mathbf{x}|\mathbf{h}) \) is now a Gaussian distribution with mean \( \mu = \mathbf{c} + \mathbf{W}^T \mathbf{h} \) and identity covariance matrix
  - recommended to normalize the training set by
    - subtracting the mean of each input
    - dividing each input \( x_k \) by the training set standard deviation
  - should use a smaller learning rate than in the regular RBM
FILTERS

(LAROCHELLE ET AL., JMLR2009)
OTHER TYPES OF OBSERVATIONS

**Topics:** extensions to other observations

- Extensions support other types:
  - real-valued: Gaussian-Bernoulli RBM
  - Binomial observations:
    - Rate-coded Restricted Boltzmann Machines for Face Recognition.
      Yee Whye Teh and Geoffrey Hinton, 2001
  - Multinomial observations:
    - Replicated Softmax: an Undirected Topic Model.
      Ruslan Salakhutdinov and Geoffrey Hinton, 2009
    - Training Restricted Boltzmann Machines on Word Observations.
      George Dahl, Ryan Adam and Hugo Larochelle, 2012
  - and more (see course website)
**Topics:** Boltzmann machine

- The original Boltzmann machine has lateral connections in each layer

\[
E(x, h) = -h^\top W x - c^\top x - b^\top h - \frac{1}{2} x^\top V x - \frac{1}{2} h^\top U h
\]

- when only one layer has lateral connection, it’s a semi-restricted Boltzmann machine
BOLTZMANN MACHINE

**Topics:** Boltzmann machine

- The original Boltzmann machine has lateral connections in each layer

\[
E(x, h) = -x^T \mathbf{W} x - c^T x - b^T h - \frac{1}{2} x^T \mathbf{V} x - \frac{1}{2} h^T \mathbf{U} h
\]

- when only one layer has lateral connection, it’s a semi-restricted Boltzmann machine
**BOLTZMANN MACHINE**

**Topics:** Boltzmann machine

- The original Boltzmann machine has lateral connections in each layer

\[
E(x, h) = -h^\top W x - c^\top x - b^\top h - \frac{1}{2}x^\top V x - \frac{1}{2}h^\top U h
\]

- when only one layer has lateral connection, it's a semi-restricted Boltzmann machine