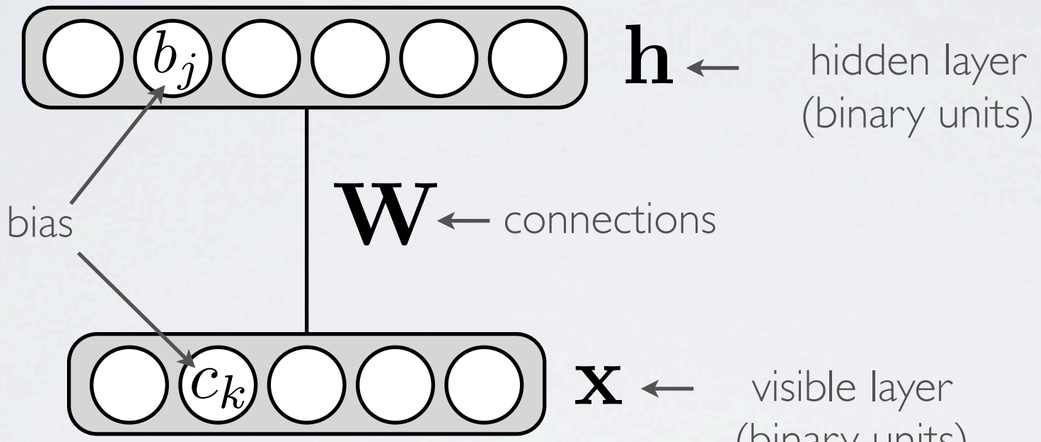
Neural networks

Restricted Boltzmann machine - extensions

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)

GAUSSIAN-BERNOULLI RBM

Topics: Gaussian-Bernoulli RBM

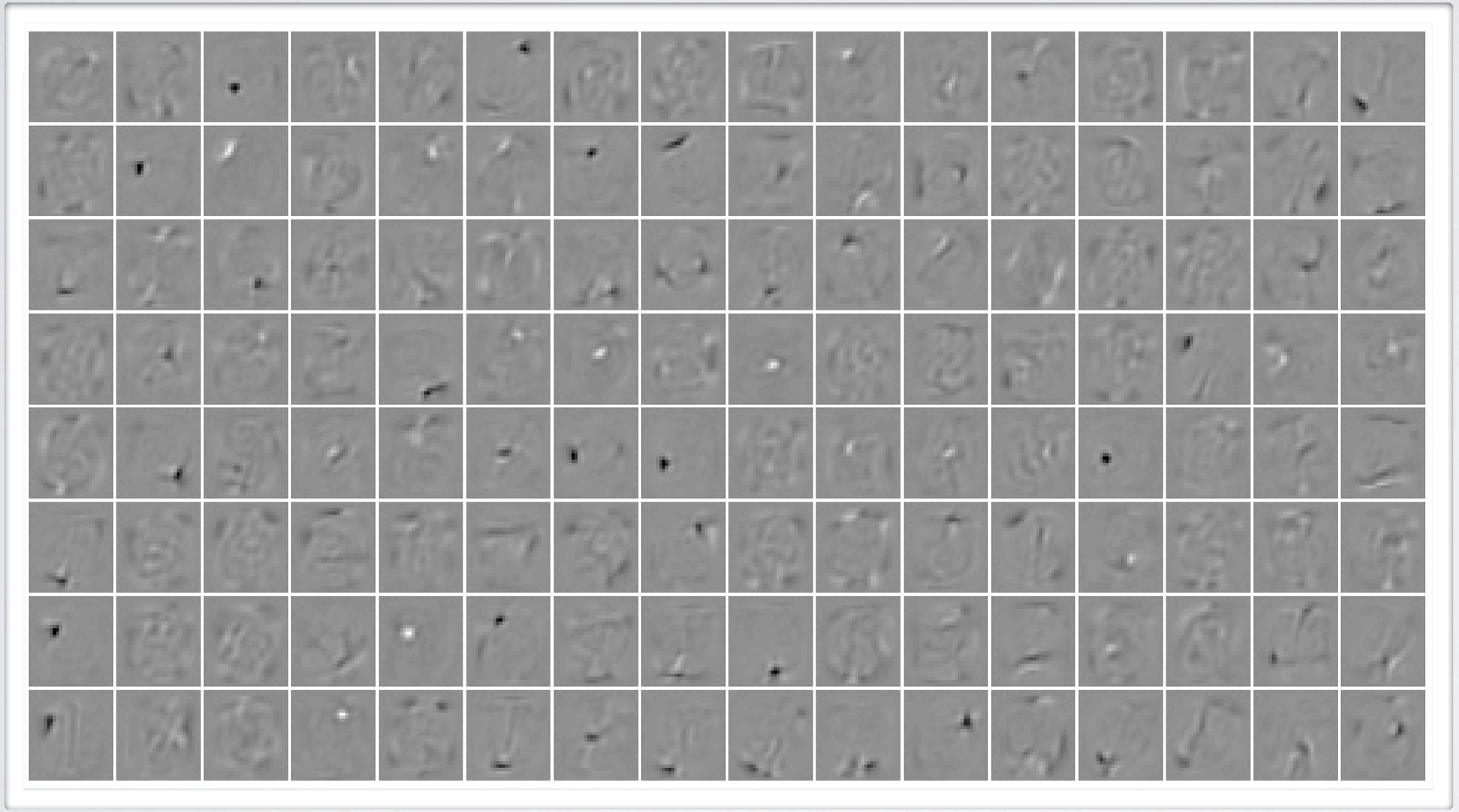
- Inputs $\mathbf{x}$ are unbounded reals
  - add a quadratic term to the energy function

$$E(\mathbf{x}, \mathbf{h}) = -\mathbf{h}^\top \mathbf{W}\mathbf{x} - \mathbf{c}^\top \mathbf{x} - \mathbf{b}^\top \mathbf{h} + \tfrac{1}{2}\mathbf{x}^\top \mathbf{x}$$

  - only thing that changes is that $p(\mathbf{x}|\mathbf{h})$ is now a Gaussian distribution with mean $\boldsymbol{\mu} = \mathbf{c} + \mathbf{W}^\top \mathbf{h}$ 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

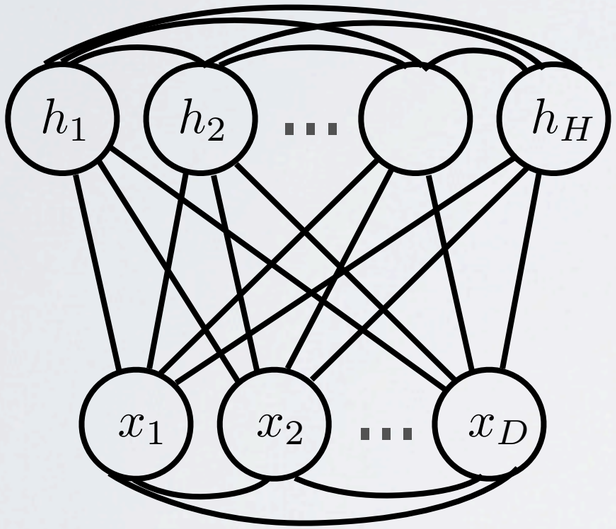
Topics: extensions to other observations

- 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)

BOLTZMANN MACHINE

Topics: Boltzmann machine

- The original Boltzmann machine has lateral connections in each layer



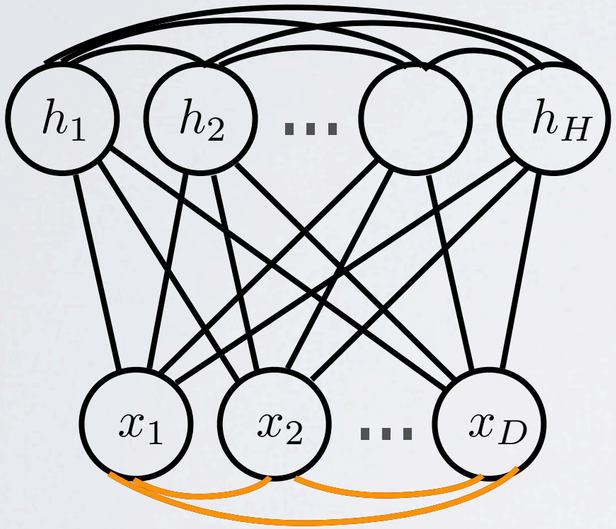
$$E(\mathbf{x}, \mathbf{h}) = -\mathbf{h}^\top \mathbf{W}\mathbf{x} - \mathbf{c}^\top \mathbf{x} - \mathbf{b}^\top \mathbf{h} - \frac{1}{2}\mathbf{x}^\top \mathbf{V}\mathbf{x} - \frac{1}{2}\mathbf{h}^\top \mathbf{U}\mathbf{h}$$

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

BOLTZMANN MACHINE

Topics: Boltzmann machine

- The original Boltzmann machine has lateral connections in each layer



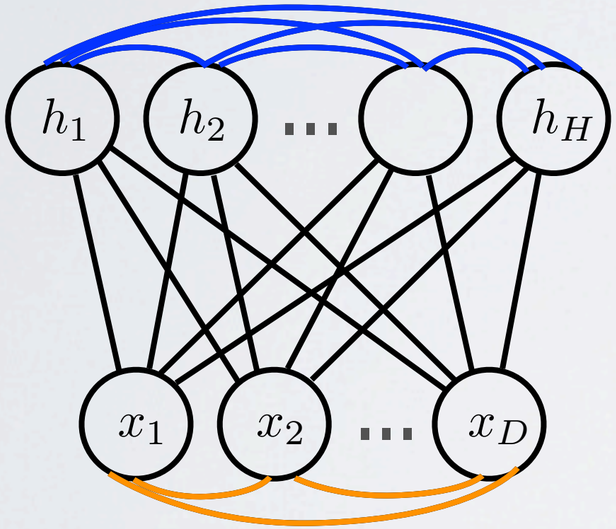
$$E(\mathbf{x}, \mathbf{h}) = -\mathbf{h}^\top \mathbf{W}\mathbf{x} - \mathbf{c}^\top \mathbf{x} - \mathbf{b}^\top \mathbf{h} - \frac{1}{2}\mathbf{x}^\top \mathbf{V}\mathbf{x} - \frac{1}{2}\mathbf{h}^\top \mathbf{U}\mathbf{h}$$

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

BOLTZMANN MACHINE

Topics: Boltzmann machine

- The original Boltzmann machine has lateral connections in each layer



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

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