Neural networks

Restricted Boltzmann machine - persistent CD

CD-K: PSEUDOCODE

Topics: contrastive divergence

1. For each training example $\mathbf{x}^{(t)}$
   i. generate a negative sample $\tilde{\mathbf{x}}$ using k steps of Gibbs sampling, starting at $\mathbf{x}^{(t)}$
   ii. update parameters

$$\mathbf{W} \Longleftarrow \mathbf{W} + \alpha \left(\mathbf{h}(\mathbf{x}^{(t)})\, \mathbf{x}^{(t)\top} - \mathbf{h}(\tilde{\mathbf{x}})\, \tilde{\mathbf{x}}^{\top} \right)$$

$$\mathbf{b} \Longleftarrow \mathbf{b} + \alpha \left(\mathbf{h}(\mathbf{x}^{(t)}) - \mathbf{h}(\tilde{\mathbf{x}}) \right)$$

$$\mathbf{c} \Longleftarrow \mathbf{c} + \alpha \left(\mathbf{x}^{(t)} - \tilde{\mathbf{x}} \right)$$

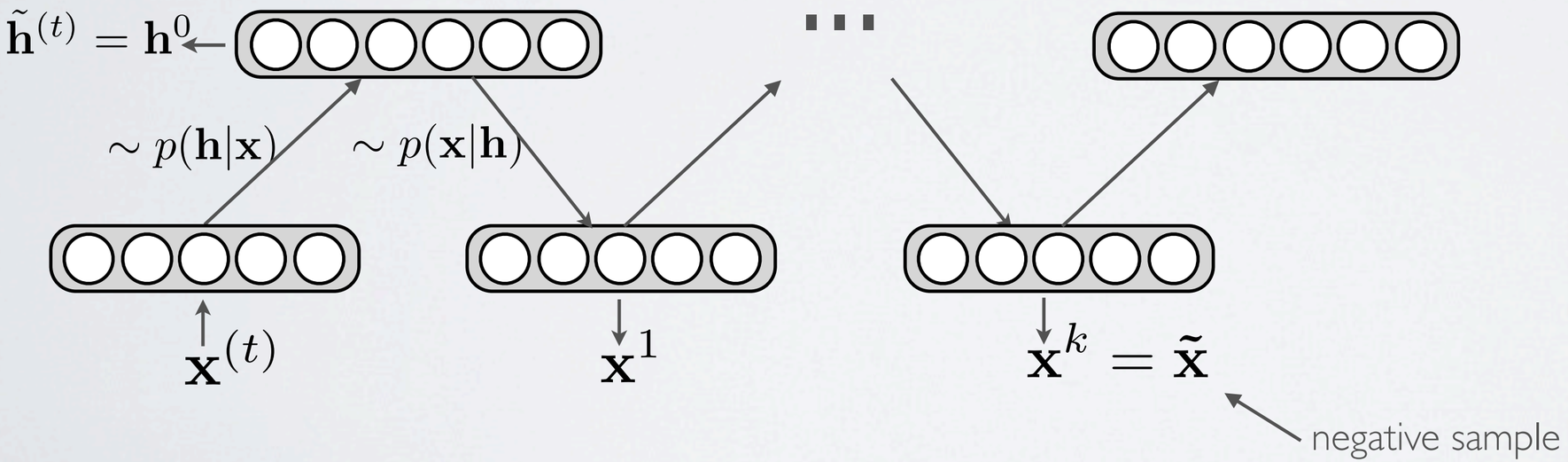
2. Go back to 1 until stopping criteria

PERSISTENT CD (PCD)

(TIELEMAN, ICML 2008)

Topics: persistent contrastive divergence

- Idea: instead of initializing the chain to $\mathbf{x}^{(t)}$, initialize the chain to the negative sample of the last iteration

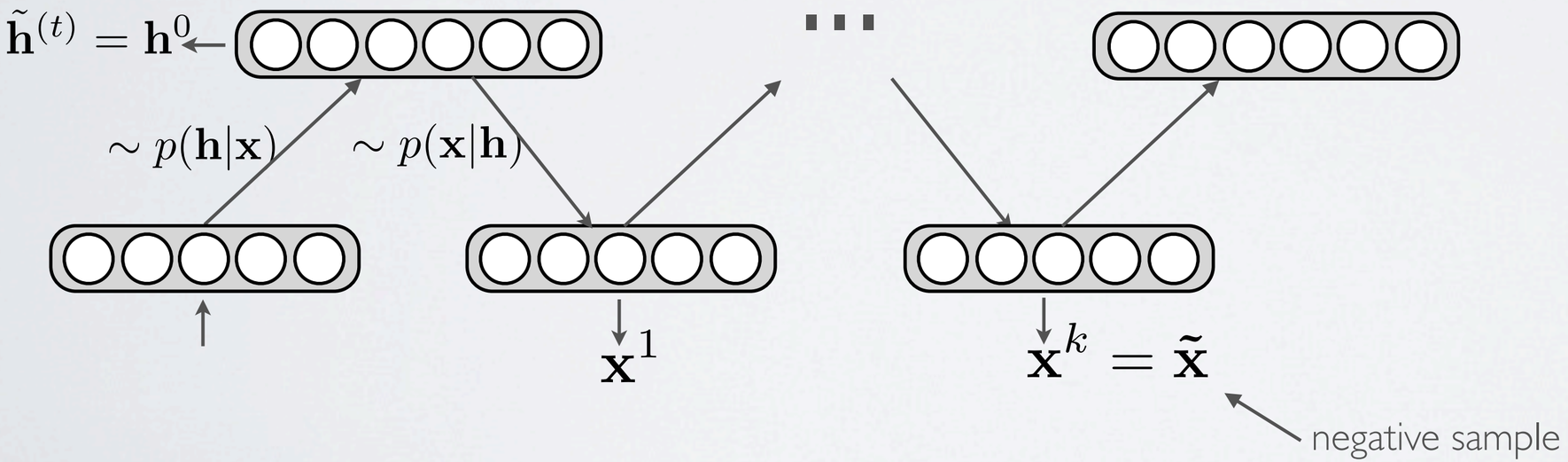


PERSISTENT CD (PCD)

(TIELEMAN, ICML 2008)

Topics: persistent contrastive divergence

- Idea: instead of initializing the chain to $\mathbf{x}^{(t)}$, initialize the chain to the negative sample of the last iteration



PERSISTENT CD (PCD)

(TIELEMAN, ICML 2008)

Topics: persistent contrastive divergence

- Idea: instead of initializing the chain to $\mathbf{x}^{(t)}$, initialize the chain to the negative sample of the last iteration