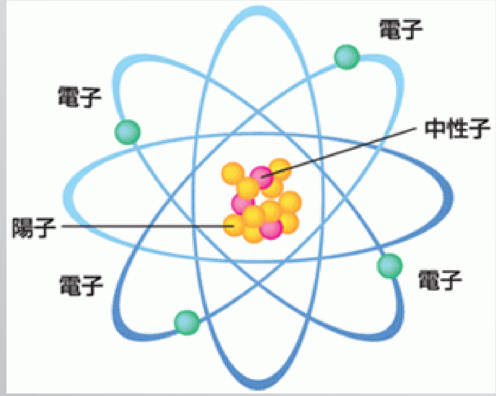
統計力学と機械学習

統計力学/イジング模型の概要と

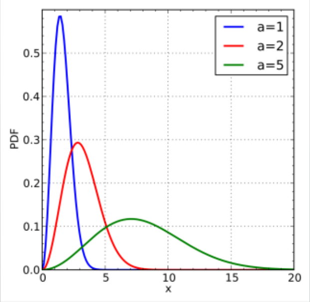
関連した機械学習(ボルツマンマシン)について

統計力学とは？
＝統計的な視点から系の熱力学を記述

ミクロな世界

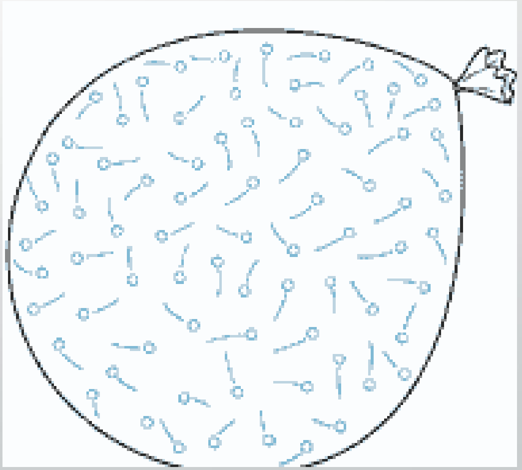


統計的視点



熱力学の世界

気体/磁石/金属など



統計力学の概要

- 温度Tにおいてi番目のエネルギー E_i の状態が実現する確率 p_i
- 熱力学量は確率分布の期待値で与えられる(k_B :ボルツマン定数)

$$p_i = \frac{\exp(-E_i/k_B T)}{Z} \qquad < E > = \sum_i E_i p_i$$

- 確率分布の規格化定数 Z ：分配関数
- 自由エネルギーFは分配関数の対数

$$Z = \sum_i \exp(-E_i/k_B T) \qquad F = k_B T \ln Z$$

簡単な例：２状態系

- σ_1,σ_2=±１の値をとる２つの自由度があり、同じ値をとるとエネルギーEが下がる

$$\sigma_1,\sigma_2=\pm 1 \qquad E=-J\sigma_1\sigma_2\ (J>0)$$

- ４通りの状態({++,+-,-+,--})について、それぞれ確率を求める

簡単な例：２状態系

- σ1,σ2=±１の値をとる２つの自由度があり、同じ値をとるとエネルギーEが下がる

$$\sigma_1, \sigma_2 = \pm 1 \qquad E = -J\sigma_1\sigma_2 \; (J > 0)$$

- ４通りの状態({++,+-,-+,--})について、それぞれ確率を求める

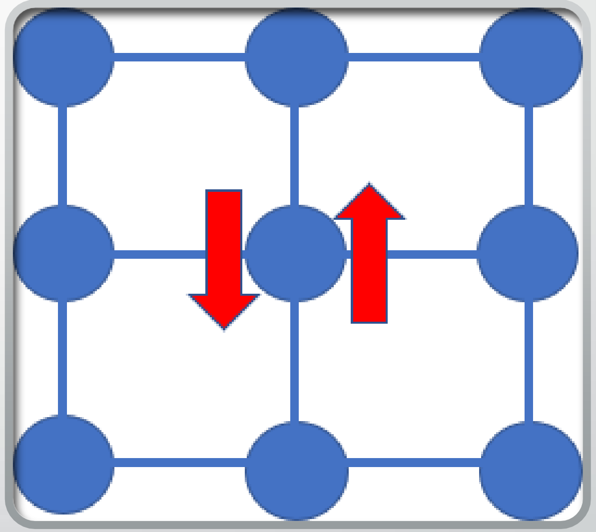
$$K \equiv \frac{J}{k_B T} \qquad Z = \exp(K) + \exp(-K) + \exp(-K) + \exp(K) = 4\cosh(K)$$

$$p_{++} = \frac{\exp(-K)}{Z}(= p_{--}) \qquad p_{+-} = \frac{\exp(K)}{Z}(= p_{-+})$$

$$< E >= -J * p_{++} + J * p_{+-} + J * p_{-+} - J * p_{--} = -J\tanh(K)$$

統計力学の代表例：イジング模型

- 格子上の点に変数（スピン）$\sigma=\pm1$を与える
  - 磁石のモデル（前のスライドの系を複雑にしたもの）
- 系全体のエネルギー
  - J_{ij}：結合定数, h_k：外部磁場
  - i, j, k はサイトのindex



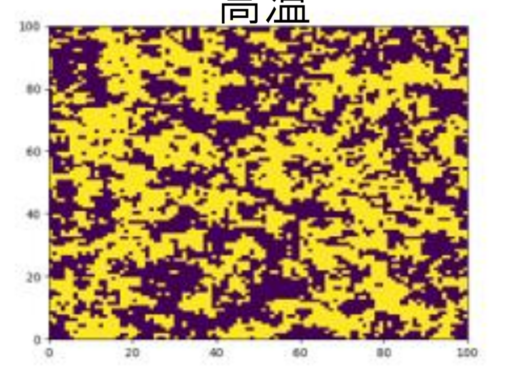
$$E = -\sum_{ij} J_{ij}\sigma_i\sigma_j - \sum_k h_k\sigma_k$$

統計力学の代表例：イジング模型

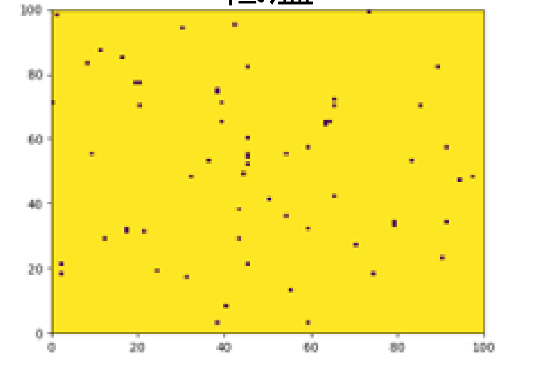
$$E = -\sum_{ij} J_{ij}\sigma_i\sigma_j - \sum_k h_k\sigma_k$$

- J, h, Tのパラメーターから、確率分布が定まる
  - ※実際はモンテカルロ法などで近似的に計算
- 確率分布に従って系の配位（≒２値画像）が生成

高温



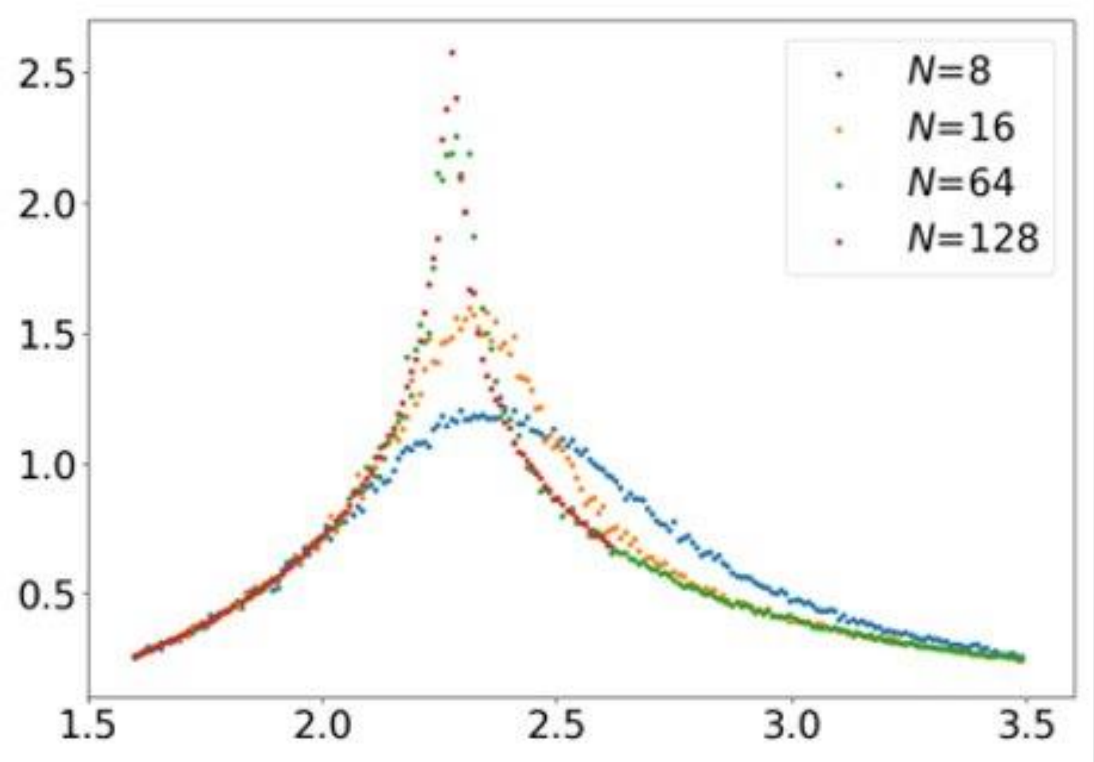
低温



例)相互作用が一様かつ正(J_ij=J>0)なら、高温ではスピン変数はバラバラで、低温では値が揃う

計算例：温度vs比熱
(マルコフ連鎖モンテカルロ法で計算：卒業研究)

比熱

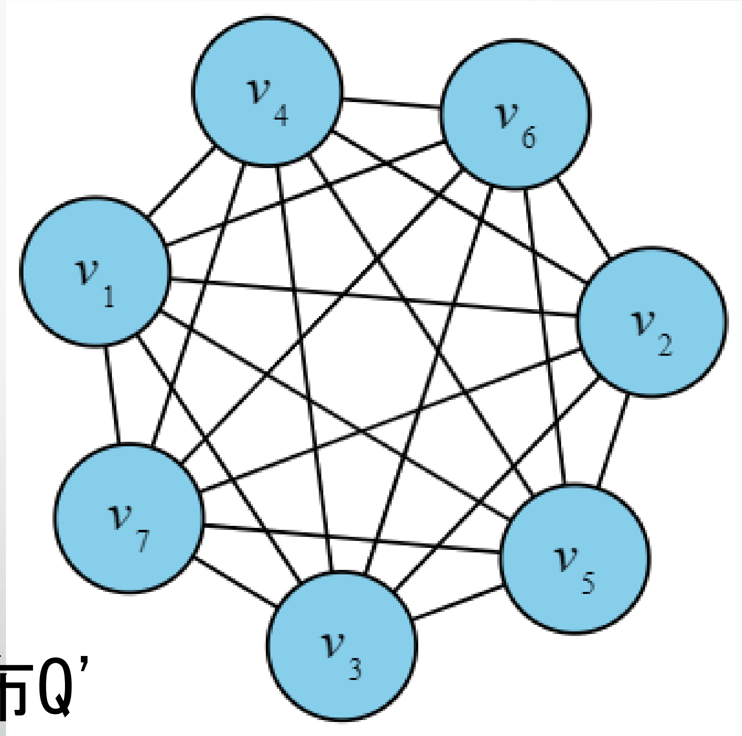


温度

N：系の一辺の大きさ
(J_{ij}=J> 0 , h_k = 0)

機械学習での対応物：ボルツマンマシン

互いに結合したノードから構成
ノードは変数 v=0 or 1を持つ
ノード間にバイアスbと重みW が定義

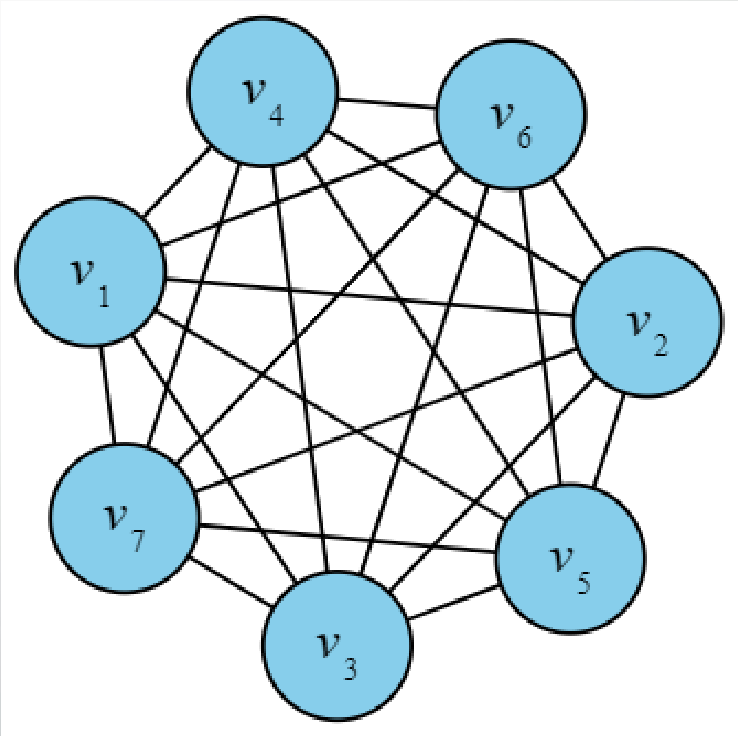


概要：

●未知の確率分布Qから出力されたデータ列
X={X_1, X_2, . . . }を入力として、Qを推定。
●バイアス/重みを調節して、Qに近い確率分布Q'を作る

機械学習での対応物：ボルツマンマシン

- 互いに結合したノードから構成
- ノードは変数 v=0 or 1を持つ
- ノード間にバイアスbと重みW が定義



”エネルギー”: Φ

$$\Phi[\boldsymbol{v}] = -\sum_{ij} w_{ij} v_i v_j - \sum_k b_k v_k$$

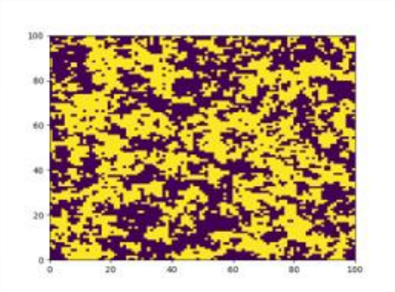
$$p(\boldsymbol{v}) = \frac{1}{Z}\exp(-\Phi[\boldsymbol{v}]) \qquad Z = \sum_{\boldsymbol{v}} \exp(-\Phi[\boldsymbol{v}])$$

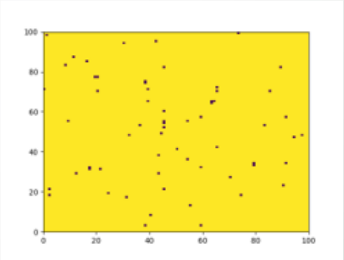
イジング模型 ⇔ ボルツマンマシン

$$\{J_{ij}, h_k\}$$

既知

分布{p}



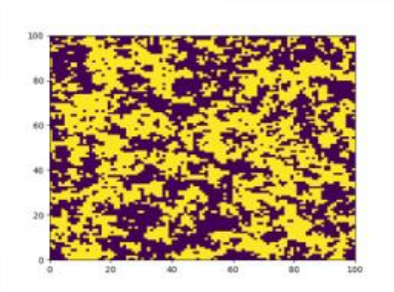


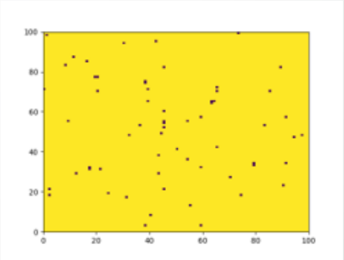
生成物

- 統計力学にインスパイアされた学習の手法も多い
(ex,平均場近似,モンテカルロ法)

イジング模型 ⇔ ボルツマンマシン

$$\{w_{ij}, b_k\}$$





分布{p}

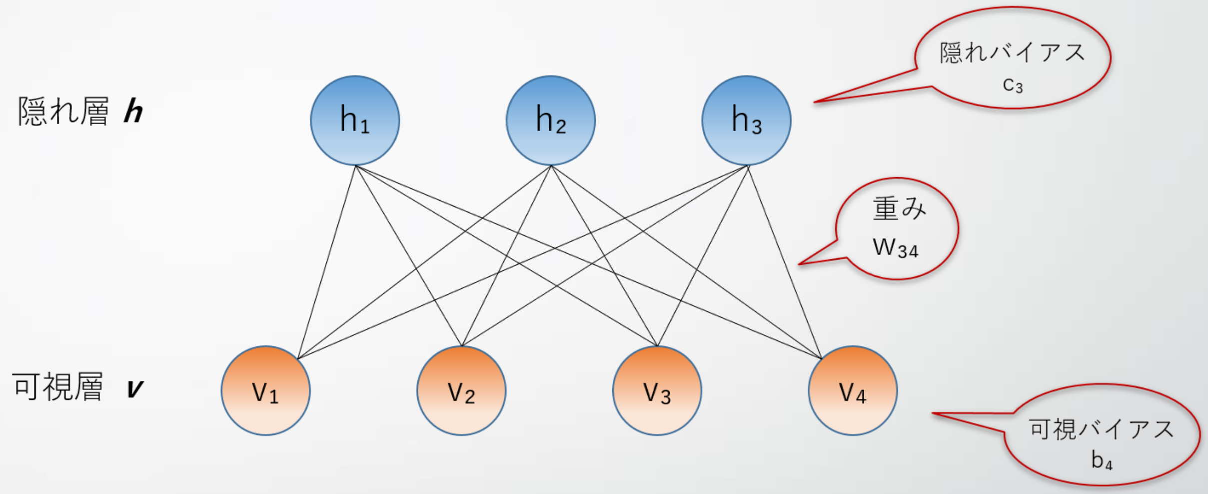
推定

入力データ

- 統計力学にインスパイアされた学習の手法も多い
(ex,平均場近似,モンテカルロ法)

応用例：制限付きボルツマンマシン(RBM)

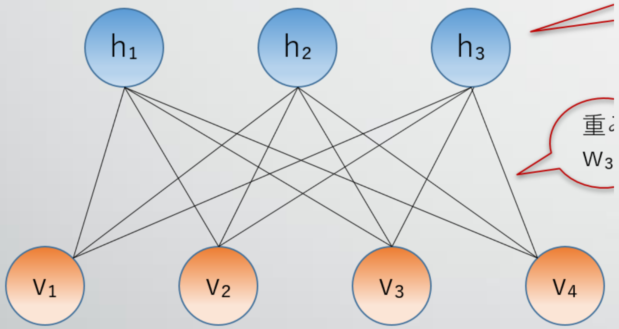
- 隠れ層と可視層に分け、可視層と隠れ層の間でのみ結合



$$\Phi[\boldsymbol{v}, \boldsymbol{h}] = -\sum_{ij} w_{ij} v_i h_j - \sum_k b_k v_k - \sum_k c_k h_k$$

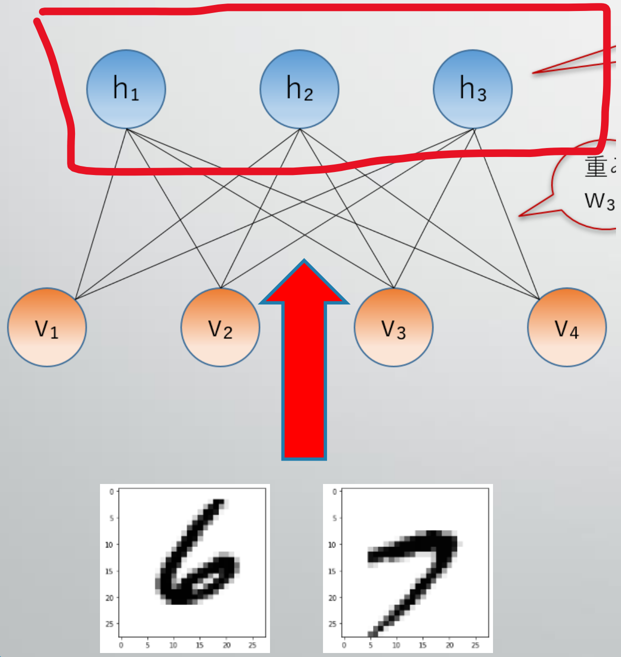
計算例：MNISTを用いた自己符号化

- １. MNISTの画像を10000個ランダムにRBMの可視層に入力し、学習
  - CD法(contrastive divergence)



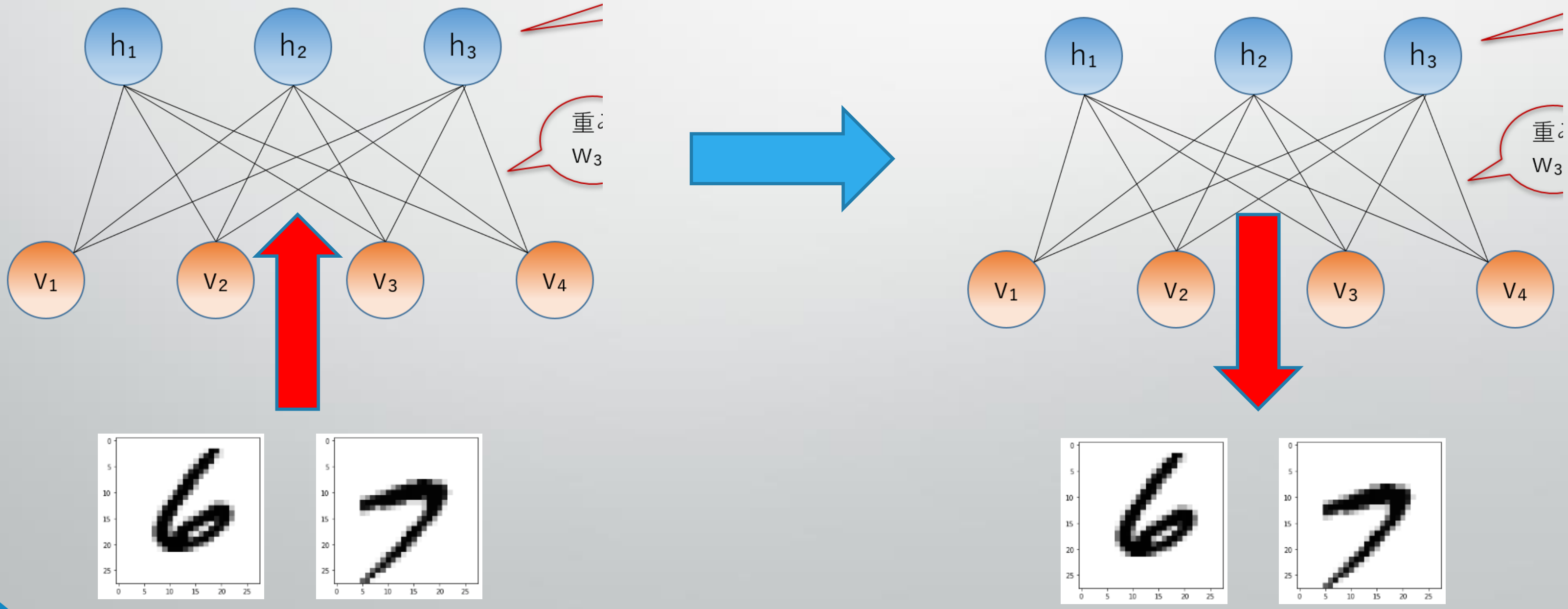
計算例：MNISTを用いた自己符号化

- １. MNISTの画像を10000個ランダムにRBMの可視層に入力し、学習
- ２. 学習後、可視層に画像を入力して隠れ層の出力を得る



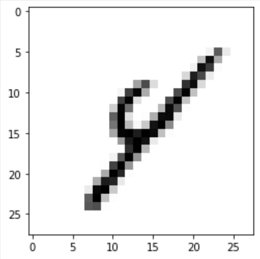
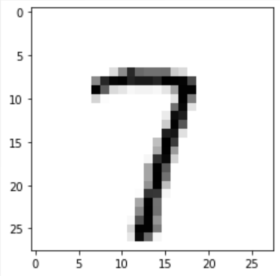
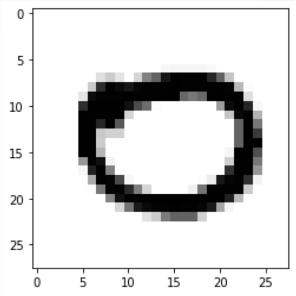
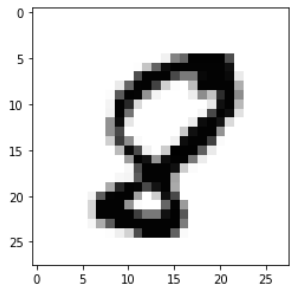
計算例：MNISTを用いた自己符号化

- １. MNISTの画像を10000個ランダムにRBMの可視層に入力し、学習
- ２. 学習後、可視層に画像を入力して隠れ層の出力を得る
- ３. 得られた隠れ層の出力で、再度可視層の値を出力



計算結果：MNISTを用いた自己符号化

MNIST

生成された画像