概率图模型 第7次作业

1.

(1) Gibbs distribution

$$P(H_1,H_2,V_1,V_2,V_3)=\frac{1}{Z}e^{\alpha(H_1+H_2+0.2(V_1+V_2+V_3))-bH_1(V_1+V_2+V_3)+bH_2(V_1+V_2+V_3)}$$

(2)
$$P(H_1=1|V_1,V_2,V_3)=\frac{e^{\alpha(1+0.2(V_1+V_2+V_3))-b(V_1+V_2+V_3)}}{e^{0.2\alpha(V_1+V_2+V_3)}+e^{\alpha(1+0.2(V_1+V_2+V_3))-b(V_1+V_2+V_3)}}$$

$$=\frac{1}{1+e^{-\alpha+b(V_1+V_2+V_3)}}=\text{sigmoid}(\alpha-b(V_1+V_2+V_3))$$

$$P(H_2=1|V_1,V_2,V_3)=\text{sigmoid}(\alpha+b(V_1+V_2+V_3))$$

$$P(V_1=1|H_1,H_2)=\text{sigmoid}(0.2\alpha-b(H_1-H_2))$$

$$P(V_2=1|H_1,H_2)=\text{sigmoid}(0.2\alpha-b(H_1-H_2))$$

$$P(V_3=1|H_1,H_2)=\text{sigmoid}(0.2\alpha-b(H_1-H_2))$$

① input $V=(V_1,V_2,V_3)$

calculate $P(H_1|V)$ $P(H_2|V)$

sample $H'=(H_1',H_2')$

caculate $P(V_1|H')$ $P(V_2|H')$ $P(V_3|H')$

sample $V'=(V_1',V_2',V_3')$

$$l(\theta|X)=\frac{1}{M}\left(\sum_{i=1}^{M}\left(\alpha(H_{1m}'+H_{2m}'+0.2(V_1^m+V_2^m+V_3^m))-bH_1(V_1+V_2+V_3)+bH_2(V_1+V_2+V_3)\right)\right)-\ln Z$$

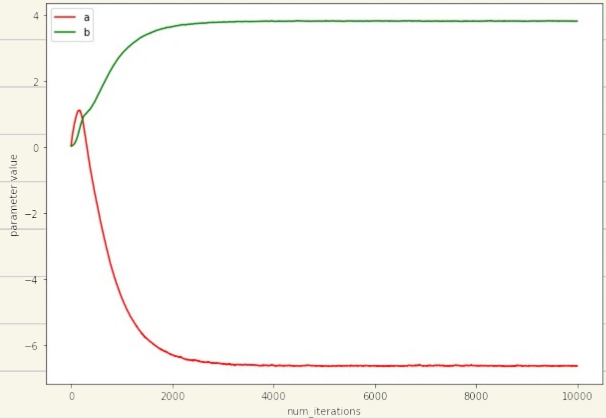
$$\frac{\partial l}{\partial a}=\frac{1}{M}\sum_{i=1}^{M}(H_{1m}'+H_{2m}'+0.2(V_{1m}+V_{2m}+V_{3m}))-E_{a,b}(H_1+H_2+0.2(V_1+V_2+V_3))$$

$$\frac{\partial l}{\partial b}=\frac{1}{M}\sum_{i=1}^{M}\left((V_{1m}+V_{2m}+V_{3m})(H_{2m}'-H_{1m}')\right)-E_{ab}((V_1+V_2+V_3)(H_2-H_1))$$

故 $a\_new=a+\lambda\cdot\frac{\partial l}{\partial a}$ $\quad b\text{-}new=b+\lambda\frac{\partial l}{\partial b}$

① 假定固定住矩阵自由度，只有a, b 2个参数

学习过程如下图：



得 $a = -6.631945$

$b = 3.806938$

训练结果如下：

```
a = -6.631945, b = 3.806938
rbm.bias_visi =  [-1.32638905 -1.32638905 -1.32638905]
rbm.bias_hidd =  [-6.63194526 -6.63194526]
rbm.weights =
[[-3.80693811 -3.80693811 -3.80693811]
 [ 3.80693811  3.80693811  3.80693811]]
Real prob(V1) = 0.588940, prob(V2) = 0.589840, prob(V3) = 0.589150
gibbs sampling prob(V1) = 0.589270, prob(V2) = 0.587330, prob(V3) = 0.586600
The inferred hidden nodes of P(H=1) =
[[2.92710190e-05 5.59877008e-02]
 [6.50306446e-07 7.27491191e-01]
 [1.44472811e-08 9.91746818e-01]
 ...
 [6.50306446e-07 7.27491191e-01]
 [1.44472811e-08 9.91746818e-01]
 [1.44472811e-08 9.91746818e-01]]
```

接着取消固定a, b，使得参数量从2变为11，训练结果如下

```
rbm2.bias_visi =  [-3.1746378 -3.1679378 -3.0950878]
rbm2.bias_hidd =  [-1.4651286   1.51243861]
rbm2.weights =
[[-4.74279287 -5.0625912  -5.06562189]
 [ 3.60546391  3.60425608  3.53895355]]
Real prob(V1) = 0.588940, prob(V2) = 0.589840, prob(V3) = 0.589150
RBM2 gibbs sampling prob(V1) = 0.592520, prob(V2) = 0.593360, prob(V3) = 0.595740
The inferred hidden nodes of P(H=1) =
[[1.45578888e-03 9.93640288e-01]
 [1.27430389e-05 9.99837091e-01]
 [8.04094095e-08 9.99995268e-01]
 ...
 [1.27430389e-05 9.99837091e-01]
 [8.04094095e-08 9.99995268e-01]
 [8.04094095e-08 9.99995268e-01]]
```

可以看到表现稍差，

最后导入 sklearn 的 RBM 包，进行对比实验

```
sklearn_rbm.intercept_visible_ =  [0.133  0.1175 0.1401]
sklearn_rbm.intercept_hidden_ =  [-0.00925213 -0.00945231]
sklearn_rbm.components_ =
[[0.19398407 0.19731675 0.18424678]
 [0.19344732 0.20447702 0.19260772]]
Real prob(V1) = 0.588940, prob(V2) = 0.589840, prob(V3) = 0.589150
gibbs sampling prob(V1) = 0.586770, prob(V2) = 0.586350, prob(V3) = 0.591770
The inferred hidden nodes of P(H=1) =
[[0.54363736 0.54566128]
 [0.59436713 0.59591482]
 [0.63790793 0.64131582]
 ...
 [0.59436713 0.59591482]
 [0.63790793 0.64131582]
 [0.63790793 0.64131582]]
```

可以看到结果截然不同，考虑到只能调用API，难以验证是否收敛，因此无法与我自己写的代码比较.