

Kisitli Boltzmann Makinalari (Restricted Boltzmann Machines -RBM-)

Bir RBM icinde ikisel (binary) degerler tasiyan gizli (hidden) h degiskenler, ve yine ikisel gorunen (visible) degiskenler v vardir. Z aynen once gordugumuz Boltzman Makinalarinda (BM) oldugu gibi normalizasyon sabitidir.

$$p(x, h; W) = \exp(-E(x, h))/Z$$

E tanimina “enerji” olarak ta atif yapilabiliyor.

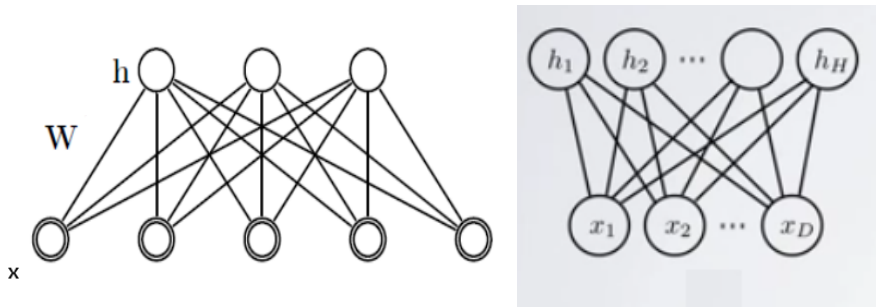
$$E(x, h) = -h^T W x - c^T x - b^T h$$

BM’lerden farkli olarak RBM taniminda c, b degiskenleri var. Bu degiskenler yanlilik (bias) icin, yani veri icindeki genel egilimi saptamaları icin modele konulmustur. Ayrica $h^T W x$ terimi var, bu BM’deki $x^T W x$ biraz farkli, gizli degiskenler, h üzerinden x ’ler arasinda baglanti yapiyor. Bir baska ilginc farklilik BM ile tum x ogeleri birbirine baglanabiliyordu, RBM ile daha az (ya da fazla) olabilecek h katmaninda baglantilar paylasiliyor. Ozellikle azaltma durumunda RBM ozellik alanini azaltarak bir tur genellemeyi gerceklestirebiliyor.

Formul alttaki gibi de acilabilir,

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

RBM’lerin alttaki gibi resmedildigini gorebilirsiniz.



h, x degiskenleri olasilik teorisinden bilinen rasgele degiskenlerdir, yani hem x ’e hem de h ’e “zar attirabiliriz” / bu degiskenler uzerinden orneklem toplayabiliriz.

Ayrica, RBM’ler aynen BM’ler gibi bir olasilik yogunluk fonksiyonu uzerinden tanimlanirlar, onceki formilde gordugumuz gibi, tum mumkun degerleri uzerinden entegralleri (ya da toplamlari) alinca sonuc 1 olur, vs.

RBM’lerin “kisitli” olarak tanimlanmalarinin sebebi gizli degiskenlerin kendi aralarinda, ve gorunen degiskenlerin kendi aralarinda direk baglantiya izin verilmemis olmasidir, bu bakimdan “kisitlanmislardir”. Baglantilara, W uzerinden

sadece gizli ve gorunen degiskenler (tabakalar) arasinda izin verilmistir. Bu tabii ki matematiksel olarak bazi kolayliklar sagliyor.

Devam edelim, ana formulden hareketle cebirsel olarak sunlar da dogrudur,

$$\begin{aligned}
 p(x, h; W) &= \exp(-E(x, h))/Z \\
 &= \exp(h^T W x + c^T x + b^T h)/Z \\
 &= \exp(h^T W x) \exp(c^T x) \exp(b^T h)/Z
 \end{aligned} \tag{2}$$

cunku bir toplam uzerindeki exp, ayri ayri exp'lerin carpimi olur. Ayni mantikla, eger ana formulu matris / vektor yerine ayri degiskenler olarak gormek istersek,

$$p(x, h; W) = \frac{1}{Z} \prod_j \prod_k \exp(W_{jk} h_j x_k) \prod_k \exp(c_k x_k) \prod_j \exp(b_j h_j)$$

Notasyonu kolaylastirmak amaciyla b, c terimlerini W icine absorbe edebiliriz, $x_0 = 1$ ve $h_0 = 1$ degerlerini mecbur tutarsak ve $w_{0,:} = c$ ve $w_{:,0} = b$ dersek, yani W 'nin sifirinci satirinin tamamini c 'ye set edip, sifirinci kolonunun tamamini b 'ye set edersek, RBM ana formulunu tekrar elde etmis oluruz, fakat artik

$$\begin{aligned}
 E(x, h) &= -h^T W x \\
 &= - \sum_j \sum_k W_{j,k} h_j x_k
 \end{aligned}$$

ve

$$p(x, h; W) = \exp(h^T W x)/Z$$

yeterli olacaktir. Bir diger kolaylik x, h yerine tek degisken kullanmak,

Eger $y \equiv (x, h)$ olarak alirsak,

$$P(x, h; W) = \frac{1}{Z(W)} \exp \left[\frac{1}{2} y^T W y \right]$$

Aslinda acik konusmak gerekirse “enerji” gibi kavramlarla ugrasmak, ya da icinde eksi terimler iceren bir grup degiskenin tekrar eksisini almak ve eksilerin etkisinin notralize etmis olmak yerine bastan (2)'deki ifadeyle yola cikmak daha

kisa. Giristeki aciklamalari literaturde gorulebilecek bazi anlatimlari aciklamak icin yaptik.

Neyse, h uzerinden marjinalize edersek,

$$\begin{aligned} P(x; W) &= \sum_h \frac{1}{Z(W)} \exp \left[\frac{1}{2} y^T W y \right] \\ P(x; W) &= \frac{1}{Z(W)} \sum_h \exp \left[\frac{1}{2} y^T W y \right] \end{aligned} \quad (1)$$

Ve $Z(W)$

$$Z(W) = \sum_{h,x} \exp \left[\frac{1}{2} y^T W y \right]$$

(1) denkleminde bolumunden sonraki kisma $Z_x(W)$ dersek, sanki ayni exp denkleminin "farkli bir sekilde marjinalize edilmiş hali" olarak gosteremis oluruz onu, ve boylece daha kısa bir formül kullanabiliriz,

$$P(x; W) = \frac{1}{Z(W)} \underbrace{\sum_h \exp \left[\frac{1}{2} y^T W y \right]}_{Z_x(W)}$$

O zaman

$$P(x; W) = \frac{Z_x(W)}{Z(W)}$$

elde ederiz. Veri uzerinden maksimum olurluk icin, yine log uzerinden bir hesap yapariz, BM icin yapmistik bunu,

$$\begin{aligned} \mathcal{L} &= \ln \left(\prod_{n=1}^N P(x^n; W) \right) = \sum_{n=1}^N \ln P(x^n; W) \\ &= \sum_{n=1}^N \ln \frac{Z_{x^{(n)}}(W)}{Z(W)} = \sum_{n=1}^N (\ln Z_{x^{(n)}} - \ln Z) \\ \frac{\partial \mathcal{L}}{\partial w_{ij}} &= \sum_{n=1}^N \left(\frac{\partial \ln Z_{x^{(n)}}}{\partial w_{ij}} - \frac{\partial \ln Z}{\partial w_{ij}} \right) \end{aligned} \quad (3)$$

Parantez icindeki 1. turevi alalim,

$$\begin{aligned}
\frac{\partial \ln Z_{x^{(n)}}}{\partial w_{ij}} &= \frac{\partial}{\partial w_{ij}} \ln \left[\sum_h \exp \left(\frac{1}{2} y^{n\top} W y^n \right) \right] \\
&= \frac{1}{Z_{x^{(n)}}} \left[\sum_h \frac{\partial}{\partial w_{ij}} \exp \left(\frac{1}{2} y^{n\top} W y^n \right) \right] \\
&= \frac{1}{Z_{x^{(n)}}} \left[\sum_h \exp \left(\frac{1}{2} y^{n\top} W y^n \right) \frac{\partial}{\partial w_{ij}} y^{n\top} W y^n \right] \\
&= \frac{1}{Z_{x^{(n)}}} \sum_h \exp \left(\frac{1}{2} y^{n\top} W y^n \right) y_i y_j \\
&= \sum_h \frac{1}{Z_{x^{(n)}}} \exp \left(\frac{1}{2} y^{n\top} W y^n \right) y_i y_j
\end{aligned}$$

$Z_{x^{(n)}}$ 'nin ne olduğunu hatırlarsak, exp ifadesinin h üzerinden marjinalize edilmiş hali,

$$= \sum_h \frac{\exp \left(\frac{1}{2} y^{n\top} W y^n \right)}{\sum_h \exp \left(\frac{1}{2} y^{n\top} W y^n \right)} y_i y_j$$

Eğer bolumun üstünü ve altını Z ile bösek,

$$= \sum_h \frac{\exp \left(\frac{1}{2} y^{n\top} W y^n \right) / Z}{\sum_h \exp \left(\frac{1}{2} y^{n\top} W y^n \right) / Z} y_i y_j$$

Üst kısım $P(y; W)$ yani $P(x, h; W)$ alt kısım $P(x; W)$ olmaz mı? Evet! Ve,

$$P(h|x^n; W) = \frac{P(x^n, h; W)}{P(x^n; W)}$$

olduguna göre,

$$= \sum_h P(h|x^n; W) y_i y_j$$

elde ederiz. Bunu da $\langle y_i y_j \rangle_{P(h|x^n; W)}$ olarak yazabiliriz.

Şimdi parantez içindeki 2. türevi alalım, yani $\frac{\partial \ln Z}{\partial w_{ij}}$,

$$\frac{\partial \ln Z}{\partial w_{ij}} = \sum_{h,x} \frac{1}{Z} \exp \left(\frac{1}{2} y^{n\top} W y^n \right) y_i y_j = \sum_{h,x} P(y^n; W) y_i y_j$$

ki bu son ifadeyi de $\langle y_i y_j \rangle_{P(y^n; W)}$ olarak yazabiliriz. Tamamini, yani (3) ifadesini, artık şöyle yazabiliriz,

$$\sum_{n=1}^N \left(\frac{\partial \ln Z_{x^{(n)}}}{\partial w_{ij}} - \frac{\partial \ln Z}{\partial w_{ij}} \right) = \sum_{n=1}^N \langle y_i y_j \rangle_{P(h|x^n; W)} - \langle y_i y_j \rangle_{P(y^n; W)}$$

```
import numpy as np
import itertools
```

```
class RBM:
```

```
    def __init__(self, num_visible, num_hidden, learning_rate = 0.1, \
                  max_epochs = 1000):
        self.num_hidden = num_hidden
        self.num_visible = num_visible
        self.learning_rate = learning_rate
        self.norm_dict = {}
        self.weights = 0.1 * np.random.randn(self.num_visible, self.num_hidden)
        self.weights = np.insert(self.weights, 0, 0, axis = 0)
        self.weights = np.insert(self.weights, 0, 0, axis = 1)
        self.max_epochs = max_epochs
```

```
    def fit(self, data):
```

```
        num_examples = data.shape[0]
```

```
        data = np.insert(data, 0, 1, axis = 1)
```

```
        for epoch in range(self.max_epochs):
```

```
            pos_hidden_activations = np.dot(data, self.weights)
```

```
            pos_hidden_probs = self._logistic(pos_hidden_activations)
```

```
            pos_hidden_states = pos_hidden_probs > \
                np.random.rand(num_examples, self.num_hidden + 1)
```

```
            tmp = np.array(pos_hidden_states).astype(float)
```

```
            pos_visible_states = self.run_hidden(tmp[:,1:])
```

```
            for h,v in itertools.izip(pos_hidden_states.astype(float),
                                      pos_visible_states):
```

```
                v = np.insert(v, 0, 1)
```

```
                self.norm_dict[(tuple(h),tuple(v))] = 1
```

```
            pos_associations = np.dot(data.T, pos_hidden_probs)
```

```
            neg_visible_activations = np.dot(pos_hidden_states, self.weights.T)
```

```
            neg_visible_probs = self._logistic(neg_visible_activations)
```

```
            neg_visible_probs[:,0] = 1 # Fix the bias unit.
```

```
            neg_hidden_activations = np.dot(neg_visible_probs, self.weights)
```

```
            neg_hidden_probs = self._logistic(neg_hidden_activations)
```

```
            neg_associations = np.dot(neg_visible_probs.T, neg_hidden_probs)
```

```
            self.weights += self.learning_rate * \
```

```
                ((pos_associations - neg_associations) / num_examples)
```

```

        error = np.sum((data - neg_visible_probs) ** 2)

    print 'dict', len(self.norm_dict)
    self.norm_c = self.norm_constant()

def run_hidden(self, data):

    num_examples = data.shape[0]

    visible_states = np.ones((num_examples, self.num_visible + 1))

    data = np.insert(data, 0, 1, axis = 1)

    visible_activations = np.dot(data, self.weights.T)
    visible_probs = self._logistic(visible_activations)
    visible_states[:, :] = visible_probs > \
        np.random.rand(num_examples, self.num_visible + 1)

    visible_states = visible_states[:, 1:]
    return visible_states

def run_visible(self, data):
    num_examples = data.shape[0]

    hidden_states = np.ones((num_examples, self.num_hidden + 1))

    data = np.insert(data, 0, 1, axis = 1)

    hidden_activations = np.dot(data, self.weights)
    hidden_probs = self._logistic(hidden_activations)
    hidden_states[:, :] = hidden_probs > \
        np.random.rand(num_examples, self.num_hidden + 1)
    hidden_states = hidden_states[:, 1:]
    return hidden_states

def norm_constant(self):
    sum = 0
    for h,v in self.norm_dict:
        h = np.array(h); v = np.array(v)
        sum += np.dot(np.dot(h.T, self.weights.T), v)
    return sum

def predict_proba(self, X):
    hs = self.run_visible(X)
    hs = np.insert(hs, 0, 1, axis=1)
    res = []
    for i in range(len(X)):
        tmp = np.dot(hs[i], self.weights.T)
        res.append(np.dot(tmp.T, np.insert(X[i], 0, 1)))
    return np.array(res) / self.norm_c

def _logistic(self, x):
    return 1.0 / (1 + np.exp(-x))

```