Альтернативная схема вывода правил обучения ограниченной машины Больцмана

Крощенко Александр Александрович

УО «Брестский государственный университет имени А.С. Пушкина»

Рассматривая ограниченную машину Больцмана (RBM) как автоэнкодер и выполняя минимизацию ошибок восстановления образов на видимом и скрытом слое (с использованием итераций сэмплирования Гиббса), можно получить классические правила обучения более простым путем, не прибегая к рассмотрению понятия энергии RBM.
Рассмотрим ограниченную машину Больцмана, которую будем представлять в виде трех слоев нейронных элементов [1]: видимый, скрытый и видимый (рисунок 1).

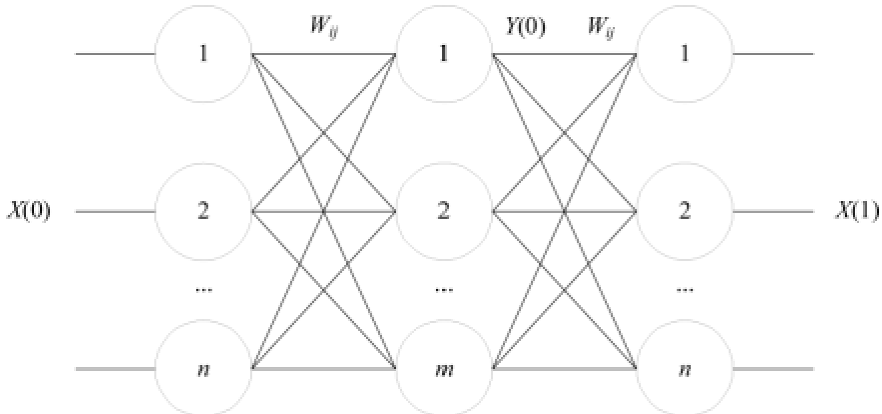

Рисунок 1 – Развернутое представление RBM

Представим сэмплирование Гиббса, используя развернутое представление RBM (рисунок 2)

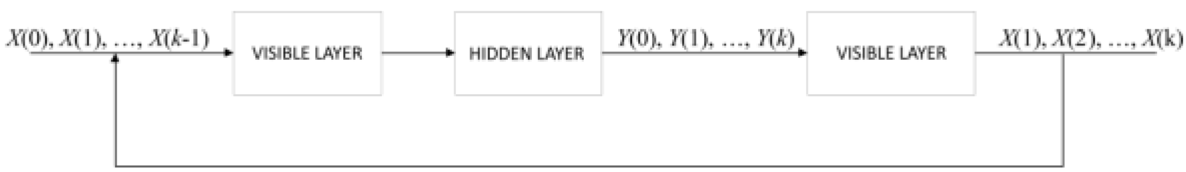

Рисунок 2 – Сэмплирование Гиббса

Сэмплирование Гиббса заключается в следующей процедуре. Пусть $x(0)$ – входной вектор, который поступает на видимый слой в момент времени 0. Тогда выходные значения нейронов скрытого слоя:

$$y_j(0) = F(S_j(0)), \tag{1}$$

$$S_j(0) = \sum_i \omega_{ij} x_i(0) + T_j \ . \tag{2}$$

Инверсный (последний) слой реконструирует входной вектор на основе данных со скрытого слоя. В результате получается восстановленный вектор $x(1)$ в момент времени 1:

$$x_i(1) = F(S_i(1)), \tag{3}$$

$$S_i(1)=\sum_j \omega_{ij} y_j(0)+T_i \,. \tag{4}$$

Затем вектор $x(1)$ поступает на видимый слой, и вычисляются выходные значения нейронов скрытого слоя:

$$y_j(1)=F(S_j(1)), \tag{5}$$

$$S_j(1)=\sum_i \omega_{ij} x_i(1)+T_j \,. \tag{6}$$

Продолжая данный процесс, можно получить на шаге k

$$x_i(k)=F(S_i(k)),$$
$$S_i(k)=\sum_j \omega_{ij} y_j(k-1)+T_i.$$
$$y_j(k)=F(S_j(k)),$$
$$S_j(k)=\sum_i \omega_{ij} x_i(k)+T_j.$$

Дж. Хинтоном была предложена энергетическая модель, базирующаяся на идее максимизации функции правдоподобия распределения входных данных $P(x)$ [2]. Нами было предложено использование двух разных критериев для обучения RBM [1]. Первый критерий основывается на минимизации среднеквадратичной ошибки (MSE). Второй – на минимизации кросс-энтропийной функции ошибки.

В случае CD-k кросс-энтропийная функция ошибки для видимого слоя определяется следующим образом:

$$CE_v(k)=-\sum_{l=1}^{L}\left[\sum_{p=1}^{k}\ \sum_{i=1}^{n}\left(x_i^l(p-1)\log(x_i^l(p))+(1-x_i^l(p-1))\log(1-x_i^l(p))\right)\right]. \tag{7}$$

где L – размер обучающей выборки.

Аналогично для скрытого слоя:

$$CE_h(k)=-\sum_{l=1}^{L}\left[\sum_{p=1}^{k}\ \sum_{j=1}^{m}\left(y_j^l(p-1)\log(y_j^l(p))+(1-y_j^l(p-1))\log(1-y_j^l(p))\right)\right]. \tag{8}$$

Общая функция определяется как сумма кросс-энтропийных функций ошибки видимого и скрытого слоев:

$$CE_s(k)=CE_h(k)+CE_v(k)\,. \tag{9}$$

Докажем следующую теорему.

Максимизация логарифма функции распределения входных данных $P(x)$ в пространстве синаптических весов ограниченной машины Больцмана эквивалентно минимизации кросс-энтропийной целевой функции CE*s*(1).

Доказательство. В случае CD-1 кросс-энтропийная функция для одного обучающего примера будет иметь следующий вид:

$$CE(1) = -\sum_{i=1}^{n}\left(x_i(0)\log(x_i(1)) + (1-x_i(0))\log(1-x_i(1))\right) - \sum_{j=1}^{m}\left(y_j(0)\log(y_j(1)) + (1-y_j(0))\log(1-y_j(1))\right).$$

Найдем частные производные кросс-энтропийной функции по весовым элементам. Получим

$$\frac{\partial CE(1)}{\partial \omega_{ij}} = \frac{\partial CE_v(1)}{\partial \omega_{ij}} + \frac{\partial CE_h(1)}{\partial \omega_{ij}}$$

Тогда

$$\frac{\partial CE_v(1)}{\partial \omega_{ij}} = -\frac{x_i(0)}{x_i(1)} x_i(1)(1-x_i(1))y_j(0) + \frac{1-x_i(0)}{1-x_i(1)} x_i(1)(1-x_i(1))y_j(0) = -x_i(0)(1-x_i(1))y_j(0) +$$
$$+(1-x_i(0))x_i(1)y_j(0) = -x_i(0)y_j(0) + x_i(0)x_i(1)y_j(0) + x_i(1)y_j(0) - x_i(0)x_i(1)y_j(0) = -x_i(0)y_j(0) +$$
$$x_i(1)y_j(0)$$

и

$$\frac{\partial CE_h(1)}{\partial \omega_{ij}} = -y_j(0)(1-y_j(1))x_i(1) + (1-y_j(0))y_j(1)x_i(1) = -y_j(0)x_i(1) + y_j(0)y_j(1)x_i(1) + y_j(1)x_i(1) -$$
$$y_j(0)y_j(1)x_i(1) = -y_j(0)x_i(1) + y_j(1)x_i(1).$$

Окончательно получим

$$\frac{\partial CE(1)}{\partial \omega_{ij}} = \frac{\partial CE_v(1)}{\partial \omega_{ij}} + \frac{\partial CE_h(1)}{\partial \omega_{ij}} = -x_i(0)y_j(0) + x_i(1)y_j(0) - y_j(0)x_i(1) + y_j(1)x_i(1) = x_i(1)y_j(1) - x_i(0)y_j(0).$$

Рассуждая аналогичным образом, можно получить для пороговых элементов:

$$\frac{\partial CE(1)}{\partial T_i} = x_i(1) - x_i(0),$$

$$\frac{\partial CE(1)}{\partial T_j} = y_j(1) - y_j(0).$$

Теорема доказана.

Данная теорема может быть обобщена на случай CD-k ([1]).

Как следует из теоремы, классические правила обучения RBM сети могут быть получены более простым путем по сравнению с конвенциальным методом, основанным на энергии. Таким образом, используя минимизацию функции кросс-энтропии и сэмплирование Гиббса, можно получить классические выражения для обучения RBM.

Использованные источники

1. Golovko, V. The Nature of Unsupervised Learning in Deep Neural Networks: A New Understanding and Novel Approach / Vladimir Golovko, Aliaksandr Kroshchanka, Douglas Treadwell // Optical Memory And Neural Networks (Springer Link). – 2016. – Vol. 25, № 3. – P. 127–141.
2. Hinton, G. A fast learning algorithm for deep belief nets / G. Hinton, S. Osindero, Y. Teh // Neural Computation. – 2006. – Vol. 18. – P. 1527–1554.