Рециркуляционные нейронные сети

Сиднев Андрей Андреевич
Бакалавр НИУ МИЭТ, Россия, г. Москва, г. Зеленоград
E-mail: sidnev38@gmail.com
«НИУ МИЭТ»

Рециркуляционные нейронные сети.

Рециркуляционные нейронные сети представляют собой многослойные нейронные сети с обратным распространение информации. При этом обратное распространение информации происходит по двунаправленным связям, которые имеют в различных направлениях разные весовые коэффициенты. При обратном распространении сигналов, в таких сетях осуществляется преобразование их с целью восстановления входного образа. В случае прямого распространения сигналов происходит сжатие входных данных. Обучение рециркуляционных сетей производится без учителя.

Рециркуляционные сети характеризуются как прямым Y=f(X), так и обратным X=f(Y) преобразованием информации. Задачей такого преобразования является достижение наилучшего автопрогноза или самовоспроизводимости вектора X. Рециркуляционные нейронные сети применяются для сжатия (прямое преобразование) и восстановления исходной (обратное преобразование) информации. Такие сети являются самоорганизующимися в процессе работы. Они были предложены в 1988 году. Теоретической основой рециркуляционных нейронных сетей является анализ главных компонент.

Метод главных компонент

Метод главных компонент применяется в статистике для сжатия информации без существенных потерь ее информативности. Он состоит в линейном ортогональном преобразовании входного вектора X размерности n в выходной вектор Y размерности p, где p<n. При этом компоненты вектора Y являются некоррелированными и общая дисперсия после преобразования остается неизменной. Совокупность входных паттернов представим в виде матрицы:

$$X = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \dots & \dots & \dots & \dots \\ x_{L1} & x_{L2} & \dots & x_{Ln} \end{bmatrix}$$

где

$x^k = (x_{k1}, x_{k2}, \dots, x_{kn})$ соответствует k-му входному образу, L — общее количество образов.

Будем считать, что матрица X является центрированной, то есть вектор математических

ожиданий μ=0. Этого добиваются при помощи следующих преобразований:

$$x_{ij} = x_{ij} - \mu_j$$

$$\mu_j = \sum_{i=1}^{L} \frac{x_{ij}}{L}$$

Матрица ковариаций входных данных X определяется как

$$K = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \dots & \sigma_{2n} \\ \dots & \dots & \dots & \dots \\ \sigma_{n1} & \sigma_{n2} & \dots & \sigma_{nn} \end{bmatrix}$$

где σij — ковариация между i-ой и j-ой компонентой входных образов.

Элементы матрицы ковариаций можно вычислить следующим образом:

$$\sigma_{ij} = \frac{1}{L} \sum_{k=1}^{L} \left(x_i^k - \mu^i\right) \times \left(x_j^k - \mu^j\right)$$

где i,j = 1,…,n.

Метод главных компонент состоит в нахождении таких линейных комбинаций исходных переменных

$$y_1 = w_{11}x_1 + w_{21}x_2 + \ldots + w_{n1}x_n$$
$$y_2 = w_{12}x_1 + w_{22}x_2 + \ldots + w_{n2}x_n$$
$$\dots$$
$$y_p = w_{1p}x_1 + w_{2p}x_2 + \ldots + w_{np}x_n$$

что

$$\sigma(y_i, y_j) = 0; i, j = \overline{1, n}$$
$$\sigma(y_1) \geq \sigma(y_2) \geq \ldots \geq \sigma(y_p)$$
$$\sum_{i}^{n} \sigma_{ii} = \sum_{i}^{n} \sigma(y_i)$$

Из последних выражений следует, что переменные yi некоррелированы, упорядочены по возрастанию дисперсии и сумма дисперсий входных образов остается без изменений. Тогда подмножество первых p переменных y характеризует большую часть общей дисперсии. В результате получается представление входной информации.

Переменные y, i = 1,…,p называются главными компонентами. В матричной форме преобразование главных компонент можно представить как

$$Y = W^T X$$

где строки матрицы WT должны удовлетворять условию ортогональности, т.е

$$W_i W_j^T = 1, \forall i = j$$

$$W_i W_j^T = 0, \forall i \neq j$$

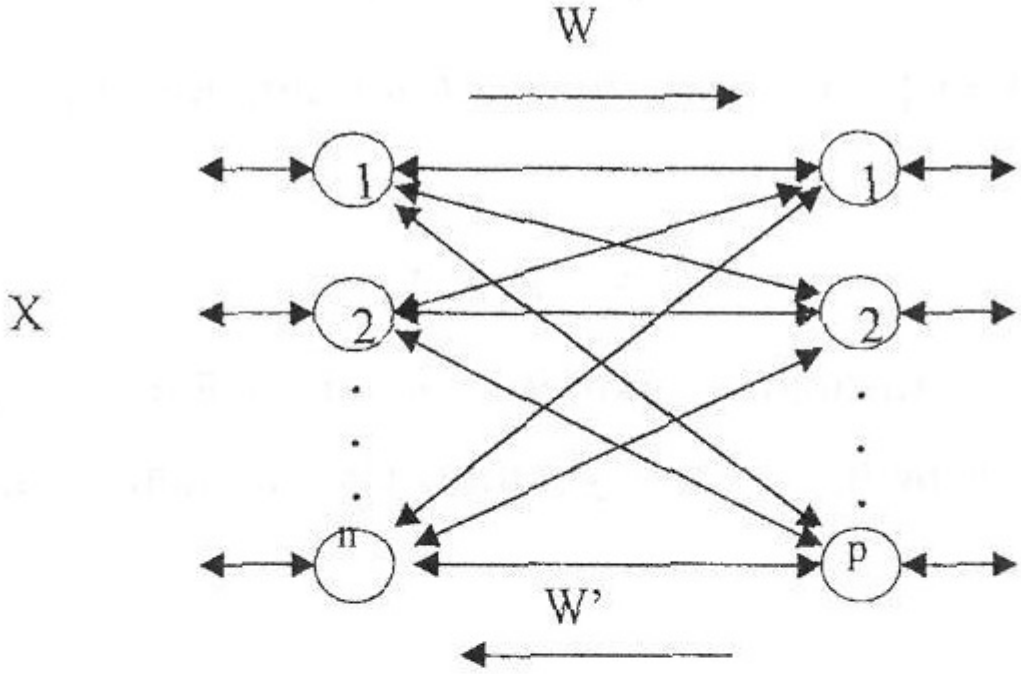
при этом вектор Wi определяется как

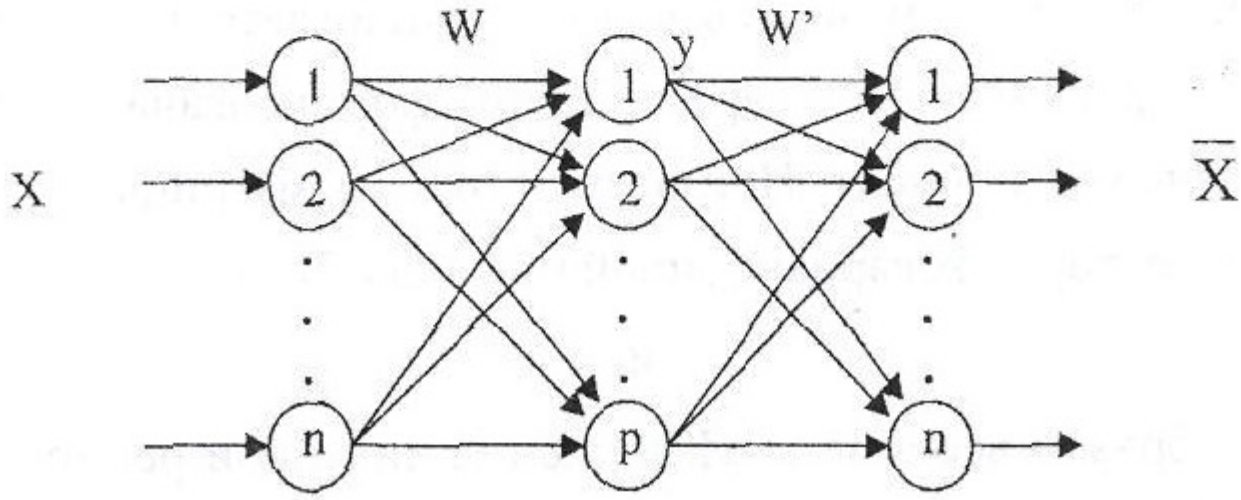
$$W_i = (w_{1i}, w_{2i}, \ldots, w_{ni})$$

Для определения главных компонент необходимо определить весовые коэффициенты Wi,j = 1, …,p.

Архитектура рециркуляционной нейронной сети

Рециркуляционная нейронная сеть представляет собой совокупность двух слоев нейронных элементов, которые соединены между собой двунаправленными связями.





Каждый из слоев нейронных элементов может использоваться в качестве входного или выходного. Если слой нейронных элементов служит в качестве входного, то он выполняет распределительные функции.

В противном случае нейронные элементы слоя являются обрабатывающими. Весовые коэффициенты соответствующие прямым и обратным связям характеризуются матрицей весовых коэффициентов W и W'. Для наглядности, рециркуляционную сеть можно представить в развернутом виде.

Такое представление сети является эквивалентным и характеризует полный цикл преобразования информации. При этом промежуточный слой нейронных элементов производит кодирование (сжатие) входных данных X, а последний слой осуществляет восстановление сжатой информации Y. Назовем слой нейронной сети, соответствующий матрице связи W прямым, а соответствующий матрице связей W' — обратным.

Рециркуляционная сеть предназначена как для сжатия данных, так и для восстановления сжатой информации. Сжатие данных осуществляется при прямом преобразовании информации в соответствие с выражением: $Y = F(W^T X)$

Восстановление или реконструкция данных происходит при обратном преобразовании информации:

$$X = F(WX)$$

В качестве функции активации нейронных элементов F может использоваться как линейная, так и нелинейная функции. При использовании линейной функции активации:

$$Y = W^T X$$
$$\overline{X} = WY$$

Линейные рециркуляционные сети, в которых весовые коэффициенты определяются в соответствии с методом главных компонент называются PCA сетями.

Обработка изображений

Рециркуляционные нейронные сети можно применять для сжатия и восстановления изображений. Изображение делится на блоки. Блок называется окном, которому в соответствие ставится рециркуляционная нейронная сеть. Количество нейронов первого слоя сети соответствует размерности окна (количеству пикселей; иногда каждый цвет отдельно). Сканируя изображение при помощи окна и подавая его на нейронную сеть, можно сжать входное изображение. Сжатое изображение можно восстановить при помощи обратного распространения информации.

Пример работы нейронной сети можно увидеть ниже:



Слева исходное изображение, справа — восстановленное после сжатия. Использовалось окно размером 3 на 3 пикселя, количество нейронов на втором слое — 21, максимальная допустимая ошибка — 50. Коэффициент сжатия составил 0,77. Для обучения нейронной сети

потребовалось 129 итераций.