

Суханов А.Я., Суханов Д.Я.

saysandr@mail.ru

Томский Университет Систем Управления и Радиоэлектроники

Томский Государственный Университет

АЛГОРИТМ НАСТРОЙКИ НЕЙРОННОЙ СЕТИ ДЛЯ РЕШЕНИЯ ОБРАТНОЙ ЗАДАЧИ ЛИДАРНОГО ЗОНДИРОВАНИЯ С ПОМОЩЬЮ ПСЕВДО-ОБРАТНЫХ МАТРИЦ

В настоящее время нейронные сети находят широкое применение в различных задачах и областях, особенно целесообразно использовать их там, где недостаточно формализованы или неизвестны данные для строгого решения задачи. Обратная задача лидарного зондирования является некорректной задачей, поэтому предлагается для ее решения использовать нейронные сети. Решение задачи восстановления концентрации газа осуществляется из решения интегрального уравнения:

$$\int_0^H K_{abs}^{gas}(I, h)n(h)dh + t_m(I, h) + t_a(I, h) = t(I, h),$$

где $t_m(h)$ - оптическая толщина как вклад молекулярного поглощения;

$t_a(h)$ - оптическая толщина как вклад аэрозольного поглощения;

$n(h)$ - концентрация газа;

$K_{abs}^{gas}(I, h)$ - коэффициент поглощения;

$t(I, h)$ - оптическая толщина.

Нейронная сеть строится таким образом, что в качестве входа служат значения оптической толщины на каждой высоте, а выходом служат значения концентрации. При этом удобно использовать сеть типа многослойный персептрон.

В данной работе предлагается алгоритм настройки многослойной нейронной сети, основанный на матрично-векторных преобразованиях, для решения задачи лидарного зондирования. Если вход и выход нейронной сети представить в виде векторов, а весовые коэффициенты в виде матрицы, то связь входа нейронной сети с ее выходом можно представить в виде следующего выражения:

$$Y = F^K \left(W^K \cdot F^{K-1} \left(\mathbf{K} F^2 \left(W^2 \cdot F^1 \left(W^1 \cdot X \right) \right) \right) \right),$$

где, Y - матрица выходных обучающих примеров;

F^K - вектор активационных функций k - го слоя;

W^K - матрица весовых коэффициентов k - го слоя;

X - матрица входных обучающих примеров.

Во входных и выходных матрицах обучающие примеры расположены в столбцах матрицы, каждая строка матрицы весовых коэффициентов принадлежит одному нейрону. Алгоритм заключается в выражении матриц коэффициентов через входную, выходную матрицы и активационные функции. Для однослойной сети данные выражения можно получить следующим образом:

$Y = F(W \cdot X)$, $F^{-1}(Y) = W \cdot X$, где F^{-1} - функция обратная к активационной функции.

Например, для активационной функции $f(g) = \arctg(g)$ активационная функция это

$$g(f) = \operatorname{tg}(f), \text{ а для } f(g) = \frac{1}{1+e^{-g}} \text{ это } g(f) = -\ln\left(\frac{1}{f} - 1\right).$$

Далее умножаем справа на транспонированную матрицу входных примеров, $F^{-1}(Y) \cdot X^T = W \cdot X \cdot X^T$,

$$F^{-1}(Y) \cdot X^T \cdot (X \cdot X^T)^{-1} = W \cdot X \cdot X^T \cdot (X \cdot X^T)^{-1},$$

$$F^{-1}(Y) \cdot X^T \cdot (X \cdot X^T)^{-1} = W.$$

Алгоритм вычисления весовых матриц для многослойных сетей можно определить следующим образом:

$$(F^K)^{-1}(Y) = W^K \cdot F^{K-1}(\mathbf{K} F^2(W^2 \cdot F^1(W^1 \cdot X))),$$

$$W^K = F^{-1}(Y) \cdot FF^{K-1,T} \cdot (FF^{K-1} \cdot FF^{K-1,T})^{-1},$$

где $FF^K = F^K(\mathbf{K} F^2(W^2 \cdot F^1(W^1 \cdot X)))$;

$$(W^K)^T \cdot (F^K)^{-1}(Y) = (W^K)^T \cdot W^K \cdot FF^{K-1};$$

$$((W^K)^T \cdot W^K)^{-1} \cdot (W^K)^T \cdot (F^K)^{-1}(Y) = ((W^K)^T \cdot W^K)^{-1} \cdot (W^K)^T \cdot W^K \cdot FF^{K-1};$$

$$((W^K)^T \cdot W^K)^{-1} \cdot (W^K)^T \cdot (F^K)^{-1}(Y) = F^{K-1}(W^{k-1} F^{k-2}[\mathbf{K} F^2(W^2 \cdot F^1(W^1 \cdot X))]);$$

$$(F^{K-1})^{-1} [((W^K)^T \cdot W^K)^{-1} \cdot (W^K)^T \cdot (F^K)^{-1}(Y)] \cdot FF^{K-2,T} \cdot (FF^{K-2} \cdot FF^{K-2,T})^{-1} = W^{k-1}.$$

Обозначим псевдо-обратные матрицы в виде:

$$WI^K = ((W^K)^T \cdot W^K)^{-1} \cdot (W^K)^T,$$

$$FI^K = FF^{T,K} \cdot (FF^K \cdot FF^{T,K})^{-1}.$$

Используя эти псевдо-обратные матрицы, получаем итерационные формулы (1). Элементы матрицы весовых коэффициентов сначала задаются случайным образом. Затем на

основе входной, выходной матриц, а также на основе $1, \dots, K-1$ весовых матриц вычисляется матрица W^K . Затем, используя новую матрицу W^K , вычисляется матрица W^{K-1} и т.д. для остальных слоев нейронной сети:

$$\begin{aligned}
 W^K &= (F^K)^{-1}(Y) \cdot FI^{K-1}, \\
 W^{K-1} &= (F^{K-1})^{-1}(WI^K \cdot (F^K)^{-1}(Y)) \cdot FI^{K-2}, \\
 W^1 &= (F^1)^{-1}(WI^2 \cdot (F^2)^{-1}(\dots WI^K \cdot (F^K)^{-1}(Y))) \cdot X^T \cdot (X \cdot X^T)^{-1}.
 \end{aligned} \tag{1}$$

Уточняя последовательно весовые матрицы, можно получить исходное решение. Было проведено тестирование метода, для двухслойной и однослойной нейронных сетей. Точность восстановления параметров для однослойной нейронной сети составляет 2-3 %, для двухслойной 1-2 %.

Достоинство алгоритма в том, что он обладает большей по сравнению с другими алгоритмами скоростью, но при этом возможны сложности с вычислением псевдо-обратных матриц, в этом случае необходимо прибегнуть к различным методам регуляризации. Также можно в обучающие матрицы ввести слабый шум. Для обучения сети данным методом рекомендуется выбирать линейно-независимые обучающие примеры во входных и выходных выборках. Также для нейронных сетей с большим числом слоев необходимо увеличивать регуляризирующий параметр. Начальные значения матриц весовых коэффициентов рекомендуется задать случайным образом с распределением, соответствующим обратной активационной функции.