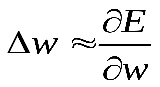
**Паралич сети**

Если один из весов при обучении получает слишком большое значение, то при обычных значениях этого входа выход нейрона окажется в насыщении, т.е. будет близок к предельному значению функции активации. Выход нейрона будет мало зависеть от *w*, и поэтому производная ∂*E/∂w =* 0. Обучение по этому весу будет очень медленным, т.к. изменение веса пропорционально производной. Выходной сигнал нейрона будет мало зависеть не только от веса, но и от входного сигнала *x* данного нейрона, а производная по *x* участвует в обратном распространении ошибки. Следовательно, предшествующие нейроны тоже будут обучаться медленно. Такое замедление обучения называется параличом сети.

Чтобы избежать паралича при обучении, можно предпринять следующие действия:

1. Уменьшить размер шага по *W* и Θ. При этом увеличится время обучения.

2. В области больших весов отказаться от зависимости , т.е. считать, что длина шага связана с модулем градиента.

3. Применять эвристические правила для ограничения роста весов. Эти правила пока не систематизированы, нет выкладок, оправдывающих тот или иной выбор ограничений.

<https://studfile.net/preview/1672047/page:7/>

# Математические основы радиальных сетей

Математическую основу функционирования радиальных сетей составляет теорема Т. Ковера о распознаваемости образов, в соответствии с которой нелинейные проекции образов в некоторое многомерное пространство могут быть линейно разделены с большей вероятностью, чем при их проекции в пространство с меньшей размерностью.

Если вектор радиальных функций в https://studfile.net/html/2706/521/html_ntuXmQJfBe.PnxW/img-UuBeLY.png- мерном входном пространстве обозначить https://studfile.net/html/2706/521/html_ntuXmQJfBe.PnxW/img-lFBdMU.png, то это пространство является нелинейно https://studfile.net/html/2706/521/html_ntuXmQJfBe.PnxW/img-lM2J_5.png- разделяемым на два пространственных класса https://studfile.net/html/2706/521/html_ntuXmQJfBe.PnxW/img-4aDIyG.pngи https://studfile.net/html/2706/521/html_ntuXmQJfBe.PnxW/img-Ckki_E.pngтогда, когда существует такой вектор весов https://studfile.net/html/2706/521/html_ntuXmQJfBe.PnxW/img-KAth5m.png, что

https://studfile.net/html/2706/521/html_ntuXmQJfBe.PnxW/img-RsYsoL.png

Граница между этими классами определяется уравнением https://studfile.net/html/2706/521/html_ntuXmQJfBe.PnxW/img-KEbB96.png.

Доказано, что каждое множество образов, случайным образом размещенных в многомерном пространстве, является https://studfile.net/html/2706/521/html_ntuXmQJfBe.PnxW/img-nePfEQ.png- разделяемым с вероятностью 1 при условии соответственно большой размерности этого пространства. На практике это означает, что применение достаточно большого количества скрытых нейронов, реализующих радиальные функции https://studfile.net/html/2706/521/html_ntuXmQJfBe.PnxW/img-XcHSJB.png, гарантирует решение задачи классификации при построении всего лишь двухслойной сети: скрытый слой должен реализовать вектор https://studfile.net/html/2706/521/html_ntuXmQJfBe.PnxW/img-u9SLI6.png, а выходной слой может состоять из единственного линейного нейрона, который выполняет суммирование выходных сигналов от скрытых нейронов с весовыми коэффициентами, заданными вектором https://studfile.net/html/2706/521/html_ntuXmQJfBe.PnxW/img-1H8ds2.png.

Простейшая нейронная сеть радиального типа функционирует по принципу многомерной интерполяции, состоящей в отображении https://studfile.net/html/2706/521/html_ntuXmQJfBe.PnxW/img-kFs7tu.pngразличных входных векторов https://studfile.net/html/2706/521/html_ntuXmQJfBe.PnxW/img-tu3edJ.pngиз входного https://studfile.net/html/2706/521/html_ntuXmQJfBe.PnxW/img-HYnMhN.png-мерного пространства во множество из p чисел https://studfile.net/html/2706/521/html_ntuXmQJfBe.PnxW/img-L7RYfL.png. Для реализации этого процесса необходимо использовать https://studfile.net/html/2706/521/html_ntuXmQJfBe.PnxW/img-rCw3A_.pngскрытых нейронов радиального типа и задать такую функцию отображения https://studfile.net/html/2706/521/html_ntuXmQJfBe.PnxW/img-yiJWdV.png, для которой выполняется условие интерполяции

https://studfile.net/html/2706/521/html_ntuXmQJfBe.PnxW/img-LwcVvr.png

Использование https://studfile.net/html/2706/521/html_ntuXmQJfBe.PnxW/img-wQUkTF.pngскрытых нейронов, соединяемых связями с весами с выходными линейными нейронами, означает формирование выходных сигналов сети путем суммирования взвешенных значений соответствующих базисных функций. Рассмотрим радиальную сеть с одним выходом и https://studfile.net/html/2706/521/html_ntuXmQJfBe.PnxW/img-HqGXWr.pngобучающими парами https://studfile.net/html/2706/521/html_ntuXmQJfBe.PnxW/img-pyWq9W.png. Примем, что координаты каждого из https://studfile.net/html/2706/521/html_ntuXmQJfBe.PnxW/img-xyZl8t.pngцентров узлов сети определяются одним из векторов https://studfile.net/html/2706/521/html_ntuXmQJfBe.PnxW/img-k1EzlC.png, т.е. https://studfile.net/html/2706/521/html_ntuXmQJfBe.PnxW/img-P8KnCj.png. В этом случае взаимосвязь между входными и выходными сигналами сети может быть определена системой уравнений, линейных относительно весов, которая в матричной форме имеет вид:

https://studfile.net/html/2706/521/html_ntuXmQJfBe.PnxW/img-ZGS2tR.png(1)

где https://studfile.net/html/2706/521/html_ntuXmQJfBe.PnxW/img-cMrAG5.pngопределяет радиальную функцию с центром в точке https://studfile.net/html/2706/521/html_ntuXmQJfBe.PnxW/img-W6KNsf.pngс вынужденным вектором https://studfile.net/html/2706/521/html_ntuXmQJfBe.PnxW/img-2IwqUG.png, https://studfile.net/html/2706/521/html_ntuXmQJfBe.PnxW/img-UPz9zP.pngи https://studfile.net/html/2706/521/html_ntuXmQJfBe.PnxW/img-vNFK7N.png.

Доказано, что для ряда радиальных функций в случае

https://studfile.net/html/2706/521/html_ntuXmQJfBe.PnxW/img-jTwLIy.png

квадратная интерполяционная матрица https://studfile.net/html/2706/521/html_ntuXmQJfBe.PnxW/img-ka46jE.pngявляется невырожденной и при этом неотрицательно определенной. Поэтому существует решение уравнения (1) в виде

https://studfile.net/html/2706/521/html_ntuXmQJfBe.PnxW/img-1BfLcf.png(2)

что позволяет получить вектор весов выходного нейрона сети.

Теоретическое решение проблемы, представленное выражением (2), не может считаться абсолютно истинным по причине серьезного ограничения общих свойств сети, вытекающих из сделанных вначале допущений. При очень большом количестве обучающих выборок и равном ему количестве радиальных функций проблема с математической точки зрения становится бесконечной (плохо структурированной), поскольку количество уравнений начинает превышать число степеней свободы физического процесса, моделируемого уравнением (1). Это означает, что результатом такого чрезмерного количества весовых коэффициентов станет адаптация модели к разного рода шумам или нерегулярностям, сопровождающим обучающие выборки. Как следствие, интерполирующая эти данные гиперповерхность не будет гладкой, а обобщающие возможности останутся очень слабыми.

Чтобы их усилить, следует уменьшить количество радиальных функций и получить из избыточного объема данных дополнительную информацию для регуляризации задачи и улучшения ее обусловленности.

**Тесты по ИНС**

[**https://studfile.net/preview/1672044/**](https://studfile.net/preview/1672044/)

<https://fundamental-research.ru/ru/article/view?id=21239&ysclid=ls7fd0u5j5287212167>

### НЕЙРОННЫЕ СЕТИ КОХОНЕНА И НЕЧЕТКИЕ НЕЙРОННЫЕ СЕТИ В ИНТЕЛЛЕКТУАЛЬНОМ АНАЛИЗЕ ДАННЫХ