Лекция

Нейронные сети Хопфилда

До сих пор рассматривались лишь сети без обратных связей. Такие сети выдают ответ сразу после прохождения через них входного сигнала. Каждый нейрон при этом срабатывает лишь один раз, поэтому многоэтапная обработка данных подразумевает наличие многих слоев, что усложняет обучение сетей.

Естественным обобщением таких однопроходных схем служат рекуррентные сети, выходы которых возвращаются обратно на их входы. Тем самым, информация пропускается через одну и ту же сеть многократно. Принципиально иной архитектурой НС по сравнению с рассмотренными выше являются *НС **с обратными связями*** **или же** ***рекуррентными НС (РНС).*** В отличие от НС ПР, которые можно назвать статическими, РНС реализуют динамический отклик сети. Суть этого состоит в том, что сигналы некоторых нейронов подаются на входы нейронов, располагающихся в этом же или предыдущих слоях. Традиционной сферой применения РНС являются задачи распознавания образов.

Рассмотрим один из типов РНС, называемых *НС Хопфилда*. Такие сети состоят из одного слоя нейронов (не считая входного), причем выходы всех нейронов единственного слоя подаются на свои входы. В НС Хопфилда считается, что элементы входного вектора могут принимать значение -1 и 1. Такая сеть функционирует потактово.

Американский физик Джон Хопфилд представил первую ассоциативную сеть (на основе автоассоциативной памяти) в 1982 г. в Национальной Академии Наук. В честь Хопфилда и нового подхода к моделированию, эта сетевая парадигма упоминается как сеть Хопфилда. Одной из наиболее известных моделей такого рода, которая оказала важнейшее воздействие на возрождение интереса к нейронным сетям в восьмидесятые годы, является сеть Хопфилда.

Автоассоциативной памятью называют память, которая может завершить или исправить образ, но не может ассоциировать полученный образ с другим образом. Данный факт является результатом одноуровневой структуры ассоциативной памяти, в которой вектор (образец некоторого класса) появляется на выходе тех же нейронов, на которые поступает входной вектор.

Нейронная сеть Хопфилда реализует важное свойство ассоциативной памяти – восстановление по искаженному (зашумленному) сигналу (образу) по ближайшему к нему эталонному. Входной вектор при обучении сети используется как начальное состояние сети, в дальнейшем сеть функционирует согласно выбранной динамике. При работе сети любой обучающий пример, находящийся в области эталона образа используется в качестве указателя для его восстановления. Полученный выходной образ определяется в том случае, если есть достигает равновесного состояния.

Структура сети Хопфилда. Принципиально иной архитектурой НС по сравнению с рассмотренными выше являются нейронные сети с обратными связями, или рекуррентными нейронными сетями (РНС).

Сеть Хопфилда состоит из трех слоев: входной, слой Хопфилда и выходной слой. Каждый слой имеет одинаковое количество нейронов. Входы слоя Хопфилда подсоединены к выходам соответствующих нейронов входного слоя через изменяющиеся веса соединений. Выходы слоя Хопфилда подсоединяются ко входам всех нейронов слоя Хопфилда, за исключением самого себя, а также к соответствующим элементам в выходном слое.

То есть выходы нейронов единственного слоя Хопфилда подаются на свои входы. Выход каждого нейрона соединен с входами всех остальных нейронов.

Сети прямого распространения можно отнести к статистическим. РНС образуют динамический отклик сети, при котором сигналы некоторый нейронов подаются на входы нейронов, расположенным в текущем при предыдущем слоях.

В режиме функционирования, сеть направляет данные из входного слоя через фиксированные веса соединений к слою Хопфилда. Структурная схема сети Хопфилда представлена на рис. 1.

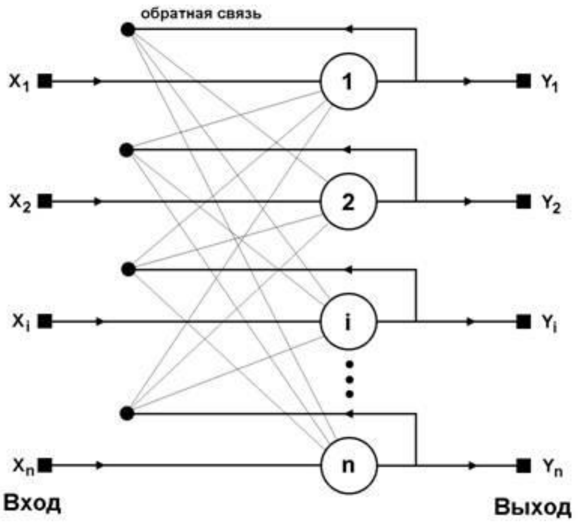


Рис.1. Структура сети Хопфилда

Пример конфигурации сети Хопфилда с тремя нейронами приведен на рис.2 (связи с одинаковым весом обозначены одинаковыми линиями).

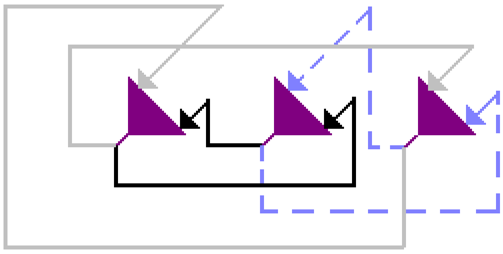


Рис.2. Конфигурация сети Хопфилда

В отличие от многослойных сетей типа персептрона, в которых входные и выходные нейроны пространственно разделены. В модели Хопфилда все нейроны одновременно являются и входными, и скрытыми, и выходными. Роль входа в таких сетях выполняет начальная конфигурация активностей (весовых коэффициентов) нейронов, а роль выхода – конечная стационарная конфигурация их активностей.

В отличие от многих нейронных сетей, функционирующих до получения ответа через определённое количество тактов, сети Хопфилда функционируют до достижения равновесия, когда следующее состояние сети в точности равно предыдущему: начальное состояние является входным образом, а при равновесии получают выходной образ.

Проблема устойчивости сети Хопфилда была решена после того, как Кохоненом и Гроссбергом была доказана теорема, определяющая достаточное условие устойчивости сетей с обратными связями, а именно сеть с обратными связями является устойчивой, если матрица весов симметрична и имеет нули на главной диагонали.

В сети Хопфилда на веса связей в выражении $s_j = w_{ij}\ x_i + w_{0j}$ следующие условия:

1) все элементы связаны со всеми;

2) w ij = w ji – прямые и обратные связи симметричны;

3) wii = 0 – диагональные элементы матрицы связей равны нулю.

Последнее условие обычно добавляется, чтобы исключить непосредственную обратную связь с выхода нейрона на собственный вход.

Каждый эталон является точкой из конечного множества равновесных точек, определяющих минимум энергии сети (функция Ляпунова).

Нейронная сеть Хопфилда — полносвязная нейронная сеть с симметричной матрицей связей. В процессе работы динамика таких сетей сходится к одному из положений равновесия. Эти положения равновесия являются локальными минимумами функционала, называемого энергией сети (в простейшем случае — локальными минимумами отрицательно определённой квадратичной формы на n-мерном кубе).

Области применения сетей Хопфилда – это распознавание образов, сжатие массивов данных. Пусть сеть состоит из N нейронов, то граница ёмкости памяти для сети (то есть количество образов, которое она может запомнить) составляет приблизительно 15% от числа нейронов в слое Хопфилда (N*0,15). При этом запоминаемые образы не должны быть сильно коррелированны.

Размерности входных и выходных сигналов в сети ограничены при программной реализации только возможностями вычислительной системы, на которой моделируется нейронная сеть, при аппаратной реализации —

технологическими возможностями. Размерности входных и выходных сигналов совпадают.

Динамическое изменение состояний сети может быть выполнено двумя способами: синхронно и асинхронно. В первом случае все элементы изменяются синхронно на каждом цикле обучения, а во втором случае – в каждый момент времени изменяется и подвергается обработке один элемент. Этот элемент выдирается случайно.

Рассмотрим синхронную бинарную сеть Хопфилда, представляющую собой пример сети с дискретными состояниями и дискретным временем. В качестве функции активации используется пороговая, при которой выходы нейронов принимают значения о или 1, при превышении взвешенной суммы значений входов некоторого порогового уровня.

Каждый нейрон системы может принимать одно из двух состояний (что аналогично выходу нейрона с пороговой функцией активации):

Благодаря своей биполярной природе нейроны сети Хопфилда иногда называют спинами. Взаимодействие спинов сети описывается выражением («энергетической» функцией, которая уменьшается в процессе функционирования сети):

$$E = -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij} x_i x_j$$

где wij элемент матрицы взаимодействий W, которая состоит из весовых коэффициентов связей между нейронами. В эту матрицу в процессе обучения записывается M «образов» — N-мерных бинарных векторов: Sm = (sm1,sm2,...,smN).

Предварительно матрицей весовых коэффициентов задан набор эталонов.

$$E = -\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij} y_i y_j - \sum_{j=1}^{n} x_j y_j + \sum_{j=1}^{n} \theta_j y_j ,$$

где E – искусственная энергия сети; w_{ij} – вес от выхода i-го ко входу j-го нейрона; x_j, y_j – вход и выход j-го нейрона; θ_j – порог j-го нейрона.

Главное свойство энергетической функции состоит в том, что в процессе эволюции состояний нейронной сети согласно уравнению она уменьшается и достигает локального минимума (аттрактора), в котором сохраняет постоянную энергию. Это позволяет решать задачи комбинаторной оптимизации, если они могут быть сформулированы как задачи минимизации энергии.

Обозначим вектор, описывающий k-й эталон, через $\mathbf{X}^k = \{x_i^k\}$, $k = 1...K$, K – число эталонов.

На вход сети подается произвольный вектор $\mathbf{X} = \{x_i\}$.

В результате серии итераций сеть должна выделить эталон, соответствующий входному вектору, или дать заключение о том, что входные данные не соответствуют ни одному из эталонов.

После отдельной итерации общее изменение энергии сети, вызванное изменением состояний всех нейронов, составит:

$$\Delta E = -\sum_{j=1}^{n}\left[\sum_{i \neq j}(w_{ij} y_i) + x_j - \theta_j\right]\Delta y_j ,$$

где Δy_j – изменение выхода j-го нейрона после итерации.

Анализ этого выражения показывает, что любое изменение состояния нейронов либо уменьшит значение E, либо оставит его без изменения. Второй случай указывает на достижение сетью устойчивого состояния и выделение ею эталона, наилучшим образом сочетающимся с входным вектором.

При распознании входного вектора (частично представленного или искаженного) выходы сети будут содержать соответствующий эталон, т. е. $\mathbf{Y} = \mathbf{X}^k$, где $\mathbf{Y} = \{y_i\}$ – выходной вектор. В противном случае, выходной вектор не совпадет ни с одним эталоном.

Для безошибочной работы сети Хопфилда число запоминаемых эталонов N не должно превышать $0{,}15n$.

Кроме того, в случае высокой степени корреляции нескольких эталонов возможно возникновение перекрестных ассоциаций при их предъявлении на входах сети. Требование достаточного (но не необходимого) условия слабой коррелируемости образов можно представить как выполнение следующего неравенства:

$$\sum_{k \neq j}^{N} \left|\left(x^k, x^j\right)\right| < n, \quad j = 1...n,$$

или в виде более сильного условия:

65

Поведение системы в пространстве состояний напоминает движение шарика, который стремится скатиться в точку минимума некоторого потенциального рельефа со множеством локальных минимумов (рис.). Эти минимумы будут устойчивыми состоянию памяти, а окружающие точки на склонах – переходными состояниями. Начальное состояние шарика соответствует вектору, содержащему неполную информацию об образе памяти, которому отвечает дно лунки.

Характер рельефа определяется видом целевой функции E и формируется в процессе обучения сети. Обучение производится путем демонстрации эталонных образов, которые сеть должна запомнить, хранить и потом воспроизводить (узнавать). Алгоритм обучения (формирование весовых коэффициентов w ij) основывается на правиле Хебба.

Одно из достоинств симметричной квадратной матрицы связей, характерной для сети Хопфилда, состоит в том, что поведение сети можно описать через стремление к минимуму простой целевой функции

$$E = w_{ij}\, x_i x_j = \min, \ i\ j$$

Обычно E интерпретируется как некоторая «обобщенная энергия». Такая интерпретация берет начало от известной физической модели Изинга, в которой совокупность взаимодействующих магнитных диполей (спинов) стремится занять такую конфигурацию, в которой суммарная энергия будет минимальна. Модель Хопфилда обобщает модель Изинга по двум причинам:

- коэффициенты связей могут принимать любые значения, как положительные, так и отрицательные;

- эти значения не являются константами, а меняются в **процессе обучения, с помощью чего вводится динамика изменений состояний нейронов.**

Расстояние между состояниями сети можно измерять в метрике Хэмминга [3]. Если два вектора и бинарные, то Хэммингово расстояние между ними определяется как количество различающихся компонент. Так, если векторы имеют вид b1=(1,0,0,0,1) и b2=(1,1,0,0,0), то Хэммингово расстояние между ними будет равно двум, поскольку в точности две компоненты этих векторов (вторая и пятая) имеют различные значения. Формально, Хэммингово расстояние для таких (Булевых) векторов может быть определено как

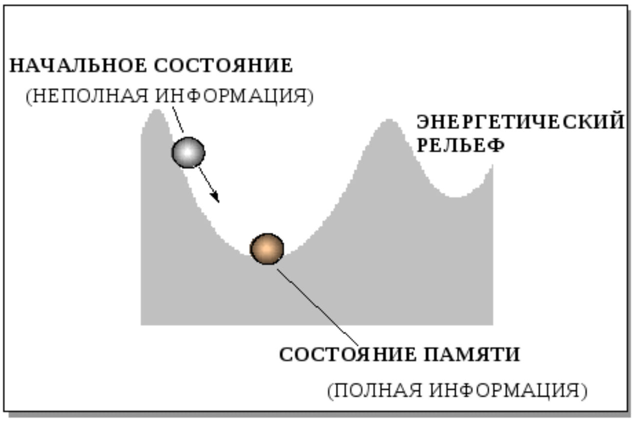


Рис.

Динамика состояний определяет главное свойство сети Хопфилда - способность восстанавливать возмущенное состояние равновесия – «вспоминать» искаженные или потерянные биты информации. Восстановление полной информации по какой-либо ее части - вспоминание по ассоциации - наделяет модель Хопфилда свойством ассоциативной памяти.

Другими словами, основное свойство такой сети, несколько напоминающее голограмму, состоит в том, что одна и та же сеть с одними и теми же весами связей может хранить и воспроизводить несколько различных эталонов. Каждый эталон является аттрактором (стационарным оптимальным состоянием), вокруг которого существует **область притяжения (от англ. attract – притягивать**) Любая система с несколькими аттракторами, к которым она тяготеет, может рассматриваться как содержательно-адресуемая память, т.е. память, из которой информация об эталоне извлекается путем задания нескольких признаков эталона. Если системе задается некоторое начальное состояние, отличное от эталонного, то это равносильно заданию частичной информации об эталоне. Если начальное состояние достаточно близко к эталону и попадает в область его притяжения, то система начинает двигаться к этому эталону - "вспоминает" его. Это выглядит как восстановление неверно заданных или отсутствующих признаков эталонного образа, отыскание полной информации о нем. Если одним из признаков, предъявлявшихся при обучении, является имя класса, то его восстановление будет равносильно отнесению образа к определенному классу, т.е распознаванию.

Обратим внимание на следующий факт. Если взять одну из точек минимума энергии E в пространстве признаков X и поменять значения всех признаков на противоположные, то величина E, как видно из выражения E = wij x ixj = min , не изменится, т.е. останется минимальной. Следовательно, "негатив" эталона является таким же аттрактором, как и сам эталон. Другими словами образы, хранящиеся в памяти нейросети, обладают инвариантностью по отношению к позитив-негативным преобразованиям. В естественных условиях обитания это свойство вряд ли могло принести какую-то пользу, поскольку в природе таких преобразований не бывает. Однако в искусственном мире человеческой цивилизации оно нашло себе применение: мы одинаково хорошо узнаем знаки, написанные чернилами на белой бумаге и мелом на черной доске, черные и белые контурные рисунки и т.п., поскольку важнейшие отношения между элементами образа при таком преобразовании сохраняются.

Далее, если какие-то два фрагмента эталона независимы, то один из них можно поменять на негативный и такая комбинация негатива и позитива снова будет точкой минимума E, т.е. аттрактором. В сетях с большим количеством элементов всегда много достаточно независимых фрагментов. Позитив-негативные комбинации таких фрагментов могут порождать ложные эталоны, "призраки", которые никогда не предъявлялись при обучении, но тем не менее являются аттракторами и способны притягивать к себе близкие изображения, т.е. "узнаваться". Проектировщикам нейросетей эти призраки только мешают, но для психолога они представляют определенный интерес, порождая известный феномен "ложного узнавания" [4].

Сети Хопфилда отличаются от ранее рассмотренных типов нейронных сетей следующими признаками:

- наличие обратных связей, идущих с выходов сетей на их входы по принципу «со всех на все»;

- расчет весовых коэффициентов осуществляется на основе исходной информации только перед началом работы сети;

- при предъявлении входного вектора, сеть «сходится» к одному из запомненных в сети эталонов, представляющих множество равновесных точек, которые являются локальными минимумами функции энергии, содержащей в себе всю структуру взаимосвязей сети.

Алгоритм функционирования и режимы работы сети Хопфилда

1. На стадии инициализации сети весовые коэффициенты синапсов (связей, по которым входные сигналы одних нейронов поступают на входы других) устанавливаются

$$w_{ij} = \begin{cases} \sum_{k=0}^{m} x_i^k x_j^k, i \neq j \\ 0, i \neq j \end{cases}$$

где i, j – индексы предсинаптического (входного) и постсинаптического (выходного) сигналов нейронов; w_{ij} – i-й синаптический вес j- го нейрона; x_i^k, x_j^k – i, j элементы k-го образа.

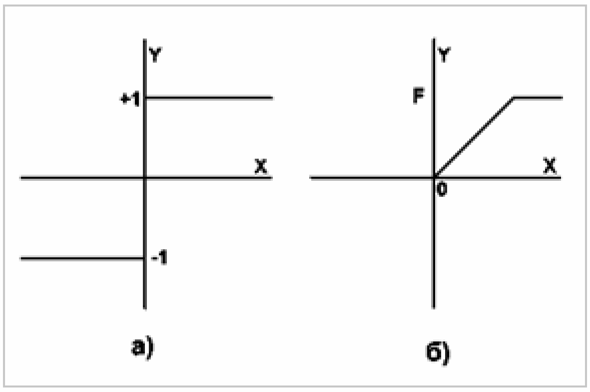


Рис.1. Функции активации: а) жесткая пороговая (передаточная) функция сети Хопфилда; б) активационная функция сети Хопфилда

2.На входы сети подается неизвестный сигнал, вычисляется выход сети на нулевом шаге итерации, фактически его ввод осуществляется непосредственной установкой значения

$$y_i(0) = x_j, j = 1....n-1$$

где y_i – аксон, т.е. выход для j-го нейрона. Нуль в скобках означает нулевую итерацию.

3.Рассчитывают новое состояние нейронов и новое состояние аксонов

$$s_j(t+1) = \sum_{i=0}^{n-1} w_{ij} y_i(t)$$

$$y_j(t+1) = f(s_j(t+1)$$

t - номер конкретной итерации

4. Проверка изменения состояния аксонов за последнюю итерацию, если да, то осуществляется переход к пункту 2, если нет – по значения стабилизировались и конец работы алгоритма. При этом выходной вектор представляет собой образец, наилучшим образом сочетающийся с входными данными.

Алгоритм обучения сети Хопфилда имеет существенные отличия в сравнении с классическими алгоритмами обучения персептрона методом коррекции ошибок или метод обратного распространения ошибки. Отличие заключается в том, что вместо последовательного приближения к нужному состоянию вместо последовательного приближения к нужному состоянию с вычислением ошибок, все коэффициенты матрицы рассчитываются по одной формуле, за один цикл обучения, после чего сеть сразу готова к работе. Вычисление весовых коэффициентов базируется на правиле: для всех запомненных сетью образов Xi, матрица связи должна удовлетворять условию:

$$X_i^T = WX_i$$

Если выполняется условие по формуле (), состояние сети будет устойчивым.

Некоторые авторы относят сеть Хопфилда к обучению без учителя. Но это неверно, т.к. обучение без учителя предполагает отсутствие информации о том, к каким классам нужно относить тот или иной образ. Для сети Хопфилда без этой информации нельзя настроить весовые коэффициенты, поэтому здесь можно говорить лишь о том, что такую сеть можно отнести к классу оптимизирующих сетей (фильтров). Отличительной особенностью фильтров является то, что матрица весовых коэффициентов настраивается детерминированным алгоритмом раз и навсегда, и затем весовые коэффициенты больше не изменяются.

В сети Хопфилда есть обратные связи и поэтому необходимо решать проблему устойчивости. Веса между нейронами в сети Хопфилда могут рассматриваться в виде матрицы взаимодействий W. Сеть с обратными связями является устойчивой, если её матрица симметрична и имеет нули на главной диагонали. В случае сети Хопфилда условие симметричности является необходимым, но не достаточным, в том смысле, что на достижение устойчивого состояния влияет ещё и режим работы сети. Ниже будет показано, что только асинхронный режим работы сети гарантирует достижение устойчивого состояния сети, в синхронном случае возможно бесконечное переключение между двумя разными состояниями (такая ситуация *называется динамическим аттрактором*, в то время как устойчивое состояние принято называть статическим аттрактором).

Расчёт весовых коэффициентов (обучением сети) происходит по следующей формуле:

$$w_{ij} = \frac{1}{N} \sum_{d=1..m} X_{id} X_{jd}$$

d- номер запоминаемого выходного вектора

Xij – i-компонента запоминаемого выходного –jго вектора

Весовая матрица может быть найдена через вычисление внешнего произведения каждого запоминаемого вектора с самим собой и суммированием матриц, полученных таким образом

$$W = \frac{1}{N} \sum_{i=1}^{n} X_i X_i^T$$

где Xi запоминаемый вектор-строка

Когда веса заданы, сеть может быть использована для получения запомненного выходного вектора по данному входному вектору, который может быть частично неправильным или неполным. Для этого выходам сети сначала придают значения этого начального вектора. Затем сеть последовательно меняет свои состояния согласно формулам 3-его этапа алгоритма функционирования сети. Полученное устойчивое состояние yj (статический аттрактор), или, возможно, в синхронном случае пара {yj, yj+1} (динамический аттрактор), является ответом сети на данный входной образ.

Нейродинамика в модели Хопфилда

Рассмотрим сеть из n формальных нейронов, в которой степень возбуждения каждого из нейронов s_i, i=1..n, может принимать только два значения {-1, +1}. Любой нейрон имеет связь со всеми остальными нейронами s_j, которые в свою очередь связаны с ним. Силу связи от i-го к j-му нейрону обозначим как w_{ij}.

В модели Хопфилда предполагается условие симметричности связей $w_{ij} = w_{ji}$, с нулевыми диагональными элементами $w_{ii}=0$. К сожалению, это условие имеет весьма отдаленное отношение к известным свойствам биологических сетей, в которых, наоборот, если один нейрон передает возбуждение другому, то тот, в большинстве случаев, непосредственно не связан с первым. Однако именно симметричность связей, как будет ясно из дальнейшего, существенно влияет на устойчивость динамики.

Изменение состояния каждого нейрона s_j в модели Хопфилда происходит по известному правилу для формальных нейронов МакКаллока и Питтса. Поступающие на его входы сигналы s_i в момент t взвешиваются с весами матрицы связей w_{ij} и суммируются, определяя уровень силы входного сигнала:

$$h_j = \sum_{i \neq j} w_{ij} S_i$$

Далее в момент времени t+1 нейрон изменяет свое состояние в зависимости от уровня сигнала h и индивидуального порога T для каждого нейрона:

$$\begin{cases} S_j(t+1) = -1, h_j(t) < T_j \\ S_j(t+1) = +1, h_j(t) > T_j \\ S_j(t+1) = S_j(t), h_j(t) = T_j \end{cases}$$

Изменение состояний возбуждения всех нейронов может происходить одновременно, в этом случае говорят о параллельной нейродинамике. Рассматривается также и последовательная нейродинамика, при которой в данный момент времени происходит изменение состояния только одного нейрона.

Многочисленные исследования показали, что свойства памяти нейронной сети практически не зависят от типа динамики.

При моделировании нейросети на обычном компьютере удобнее последовательная смена состояний нейронов. В аппаратных реализациях нейросетей Хопфилда применятся параллельная динамика.

Совокупность значений возбуждения всех нейронов Si в некоторый момент времени образует вектор состояния S сети. Нейродинамика приводит к изменению вектора состояния S(t). Вектор состояния описывает траекторию в пространстве состояний нейросети. Это пространство для сети с двумя уровнями возбуждения каждого нейрона, очевидно, представляет собой множество вершин гиперкуба размерности, равной числу нейронов N. Возможные наборы значений координат вершин гиперкуба () и определяют возможные значения вектора состояния.

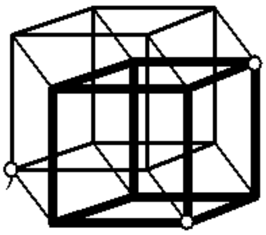


Рис.. Нейродинамика сети Хопфилда

Проекция 4-х мерного гиперкуба на плоскость. Указанные на рисунке три точки служат примерами возможных состояний нейронной сети из 4-х нейронов.

Рассмотрим теперь проблему устойчивости динамики изменения состояний. Поскольку на каждом временном шаге некоторый нейрон i изменяет свое состояние в соответствии со знаком величины (hi-Ti), то приведенное ниже соотношение всегда неположительно:

Введенная таким образом величина E является функцией состояния E=E(S) и называется энергетической функцией (энергией) нейронной сети Хопфилда. Поскольку она обладает свойством невозрастания при динамике сети, то одновременно является для нее функцией Ляпунова (А.М. Ляпунов, 1892). Поведение такой динамической системы устойчиво при любом исходном векторе состояния S(t=0) и при любой симметричной матрице связей W с нулевыми диагональными элементами. Динамика при этом заканчивается в одном из минимумов функции Ляпунова, причем активности всех нейронов будут совпадать по знаку с входными сигналами h.

$$\Delta E_i = -(S_i(t+1) - S_i(t))(h_i - T_i) \le 0$$

$$E = -\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij} x_i x_j$$

$$E = -\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij} S_i S_j + \sum_{i}^{n} S_i T_i$$

$$\Delta E_i = -(S_i(t+1) - S_i(t)) \cdot (h_i - T_i) \le 0$$

Таким образом, оответствующая величина E, являющаяся суммой отдельных значений E_i, может только убывать, либо сохранять свое значение в процессе нейродинамики.

$$E = -1/2 \sum_{i} \sum_{j \ne i} W_{ij} S_i S_j + \sum_{i} S_i T_i$$

Поверхность энергии E(S) в пространстве состояний имеет весьма сложную форму с большим количеством локальных минимумов, образно напоминая стеганое одеяло. **Стационарные состояния, отвечающие минимумам, могут интерпретироваться, как образы памяти нейронной сети.** Эволюция к такому образу соответствует процессу извлечения из памяти. При произвольной матрице связей W образы также произвольны. Для записи в память сети какой-либо осмысленной информации требуется определенное значение весов W, которое может получаться в процессе обучения.

Сети Хопфилда используются для того, чтобы научить компьютер ассоциативно мыслить. Нейронная сеть Хопфилда – это полный граф. Нейроны в сети линейные. Сеть с обратной связью между нейронами, поэтому сеть может функционировать в синхронном и асинхронном режимах:

В синхронном режиме все весовые коэффициенты считывают свой результат и изменяются одновременно.

В асинхронном режиме изменение весов производится для каждого нейрона.

Суть обучения сети заключается в том, чтобы сеть сходилась к заранее заданному набору воспоминаний $\{x_i\}$

Воспоминание – это множество значений каждого веса $\{x_i\}$

Для запоминания набора воспоминаний выполняется условия

$$X_i^T = W X_i$$

$$w_{ij} = \frac{1}{N} \sum_{d=1..m} X_{id} X_{jd}$$

$$W = \frac{1}{N} \sum_{i} X_i^T X_i$$

N – число воспоминаний

Существуют 3 режима работы сети Хопфилда:

1) режим фильтрации;

2) синхронный режим;

3) асинхронный режим.

Рассмотрим коротко работу сети в режиме фильтрации (восстановление повреждённых образов). Как только веса заданы, сеть может быть использована для получения запомненного выходного вектора по данному входному вектору, который может быть частично неправильным или неполным. Для этого выходам сети сначала придают значения этого начального вектора. Затем сеть последовательно меняет свои состояния согласно формуле нахождения общего алгоритма. Этот процесс называется конвергенцией сети.

Это же можно описать так называемым локальным полем ai действующим на нейрон x_i со стороны всех остальных нейронов сети:

$$a_i(t) = \sum_{j=1, j\neq i}^{N} w_{ij} x_j(t-1)$$

После расчёта локального поля нейрона ai(t) это значение используется для расчёта значения выхода через функцию активации, которая в данном случае является пороговой (с нулевым порогом). Соответственно, значение выхода xi(t) нейрона i в текущий момент времени t рассчитывается по формуле:

$$x_i(t) = sign(\sum_{j=1, j\neq i}^{N} w_{ij} x_j(t-1))$$

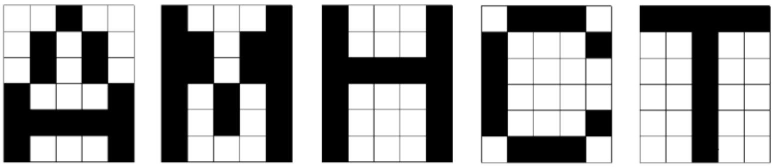
Обычно ответом является такое устойчивое состояние, которое совпадает с одним из запомненных при обучении векторов, однако при некоторых условиях (в частности, при слишком большом количестве запомненных образов) результатом работы может стать так называемый **ложный аттрактор ("химера"),** состоящий из нескольких частей разных запомненных образов, а также в синхронном режиме сеть может прийти к динамическому аттрактору. Обе эти ситуации в общем случае являются нежелательными, поскольку не соответствуют ни одному запомненному вектору - а соответственно, не определяют класс, к которому сеть отнесла входной образ.

Рассмотрим коротко работу сети в синхронном режиме работы. Если работа сети моделируется на одном процессоре, то при данном режиме последовательно просматриваются нейроны, однако их состояния запоминаются отдельно и не меняются до тех пор, пока не будут пройдены все нейроны сети. Когда все нейроны просмотрены, их состояния одновременно (т.е. синхронно, отсюда и название) меняются на новые. Таким образом, достигается моделирование параллельной работы последовательным алгоритмом. При реально параллельном моделировании, этот режим фактически означает, что время передачи ?ij для каждой связи между элементами ui и uj одинаковое для каждой связи, что приводит к параллельной работе всех связей, они одновременно меняют свои состояния, основываясь только на предыдущем моменте времени. Наличие таких синхронных тактов, которые можно легко выделить и приводит к пониманию синхронного режима. При синхронном режиме возможно (хотя и далеко не всегда наблюдается) бесконечное чередование двух состояний с разной энергией (динамический аттрактор). Поэтому синхронный режим практически для сети Хопфилда не используется, и рассматривается лишь как основа для понимания более сложного асинхронного режима.

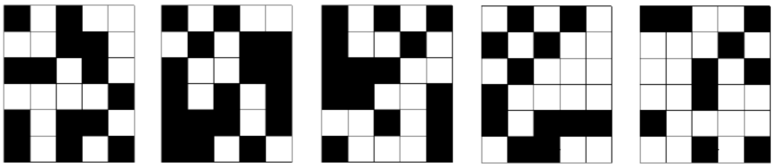
Рассмотрим коротко работу сети в асинхронном режиме работы сети. Если моделировать работу сети как последовательный алгоритм, то в данном режиме работы состояния нейронов в следующий момент времени меняются последовательно: вычисляется локальное поле для первого нейрона в момент t, определяется его реакция, и нейрон устанавливается в новое состояние (которое соответствует его выходу в момент t+1), потом вычисляется локальное поле для второго нейрона с учётом нового состояния первого, меняется состояние второго нейрона, и так далее - состояние каждого следующего нейрона вычисляется с учетом всех изменений состояний рассмотренных ранее нейронов. По сути при последовательной реализации сети Хопфилда явно не видно в чём заключается асинхронность, но это видно если сеть Хопфилда реализовать с параллельными вычислениями. В этом случае асинхронный режим сети Хопфилда упрощён, и носит частный случай по сравнению с общим видом асинхронных сетей, где время передачи ?ij для каждой связи между элементами ui и uj свое, но постоянное. Чтобы рассмотреть работу сети при параллельной реализации, необходимо ввести понятие такта - как минимальное время за которое происходит передача сигнала по связи, т.е. при ?ij = 1. Тогда за промежуток времени между t и (t+1) происходит определённое количество тактов N. И именно в пределах времени из N тактов происходит асинхронность протекания сигналов и выполнения расчётов. То есть, например, когда нужно рассчитать состояние нейрона №3 необходимо рассчитать состояния нейрона №1 и состояния нейрона №2 и умножить это на соответствующие веса (рис.1). Но оказывается, для того чтобы рассчитать состояние нейрона №2 нам нужно знать обновленное состояние нейрона №1 и старое состояние нейрона №3, умножить их на соответствующие веса. Понятно, что физически невозможно рассчитать состояние нейрона №1 и состояние нейрона №2 за одно и то же время, т.к. состояние нейрона №2 зависит от состояния нейрона №1. Поэтому связь между нейроном №1 и нейроном №3 имеете время передачи ?ij = 2, и достигает нейрона №3 за два такта. Именное такое разное время передачи ?ij и позволяет говорить о сети Хопфилда как о сети с асинхронным режимом.

В асинхронном режиме невозможен динамический аттрактор - вне зависимости от количества запомненных образов и начального состояния, сеть непременно придет к устойчивому состоянию (статическому аттрактору).

3. Области применения сети



а)



б)

Сеть Хопфилда может быть использована как ассоциативная память, для решения некоторых задач оптимизации, а также как фильтр (задачи распознавания образов).

Чтобы организовать устойчивую автоассоциативную память с помощью данной сети с обратными связями, веса должны выбираться так, чтобы образовывать энергетические минимумы в нужных вершинах единичного гиперкуба.

На каждом итерации алгоритма функционирования сети понижается значение энергии нейронной сети. Это позволяет решать комбинаторные задачи оптимизации, если они могут быть сформулированы как задачи минимизации энергии.

Рассмотрим пример восстановления повреждённого изображения.

Если во время обучения сформировать матрицу весовых коэффициентов на основании эталонных бинарных векторов, то нейронная сеть в процессе работы будет менять состояния нейронов до тех пор, пока не перейдет к одному из устойчивых состояний.

Пусть имеется нейронная сеть размерностью N=100 , в матрицу связей записан набор чёрно-белых картинок (-1 — чёрный цвет, +1 — белый), среди которых есть изображение собачки (рис. 3б). Если установить начальное состояние сети близким к этому вектору (рис. 3а), то в ходе динамики нейронная сеть восстановит исходное изображение (рис. 3б). В этом смысле можно говорить о том, что сеть Хопфилда решает задачу распознавания образов (хотя строго говоря, полученное эталонное изображение ещё нужно превратить в номер класса, что в некоторых случаях может быть весьма вычислительно ёмкой задачей).

Достоинства, недостатки и модификации Сети Хопфилда. Достоинством сети Хопфилда является то, что она имеет огромное историческое значение. С этой модели началось возрождение интереса к нейронным сетям в середине 80-х годов. Также имеющиеся модификации применимы к решению современных задач области применения данной сети.

2)

К сожалению, у нейронной сети Хопфилда есть ряд недостатков:

Относительно небольшой объём памяти, величину которого можно оценить выражением:

Попытка записи большего числа образов приводит к тому, что нейронная сеть перестаёт их распознавать.

Достижение устойчивого состояния не гарантирует правильный ответ сети. Это происходит из-за того, что сеть может сойтись к так называемым ложным аттракторам, иногда называемым "химерой" (как правило, химеры склеены из фрагментов различных образов).

При использовании коррелированных векторов-образцов возможно зацикливание сети в процессе функционирования.

Наряду с запомненными образами в сети хранятся и их негативы.

сети Хопфилда существуют модификации. Одна из них предназначена для решения задач оптимизации, в частности задачи распределения работ между исполнителями.

Существует модель сети Хопфилда с бинарными входными сигналами.

Для увеличения ёмкости сети и повышения качества распознавания образов используют мультипликативные нейроны. Сети, состоящие из таких нейронов, называются сетями высших порядков.

Разработаны к настоящему моменту многослойные сети Хопфилда, которые обладают определёнными преимуществами по сравнению с первоначальной моделью.