

Чи завжди можна побудувати нейромережу, що виконує перетворення, задане будь навчальною вибіркою, і яким вимогам ця нейромережа повинна задовольняти?

Це питання зводиться до наступного: чи можна будь-яку функцію багатьох змінних представити у вигляді суперпозиції функцій меншої кількості змінних. Воно цікавило математиків протягом декількох останніх століть (відоме як 13-та проблема Гільберта, на початку XX ст. таке представлення вважалося неможливим). Проте, надалі А. Н. Колмогоровим і В. І. Арнольдом були отримані фундаментальні теоретичні результати, що свідчать про принципову можливість подання неперервних функцій декількох змінних у вигляді суперпозиції функцій меншого числа змінних. Потім Хехт-Нільсеном ці результати були перероблені стосовно нейронних мереж. Зокрема, була доведена принципова **можливість побудови нейромережі, яка виконує перетворення, задане будь навчальною вибіркою попарно різних між собою прикладів, і встановлено, що такою універсальною нейромережею є двошаровий персептрон (тобто персептрон з одним прихованим шаром) зі скінченним числом нейронів і сигмоїдними передавальними функціями.**

Для визначення необхідного числа нейронів в прихованих шарах персептрона була запропонована формула, яка є наслідком теорем **Арнольда-Колмогорова-Хехт-Нільсена**:

$$\frac{N_y Q}{1 + \log_2 Q} \leq N_w \leq N_y \left(\frac{Q}{N_x} + 1 \right) (N_x + N_y + 1) + N_y,$$

де N_y - розмірність вихідного сигналу; Q - кількість елементів навчальної вибірки; N_w - необхідне число синаптичних ваг; N_x - розмірність вхідного сигналу.

Оцінивши за допомогою цієї формули необхідне число синаптичних ваг, можна розрахувати число нейронів в прихованих шарах. Наприклад, число нейронів прихованого шару двошарового персептрона

$$N = \frac{N_w}{N_x + N_y}.$$

Теоретично для побудови нейромережевої моделі будь-якого як завгодно складного об'єкта досить використовувати персептрон з одним прихованим шаром сигмоїдних нейронів, число яких визначається формулами вище. Однак в практичних реалізаціях персептронів як кількість шарів, так і число нейронів в кожному з них часто відрізняються від теоретичних. Число прихованих шарів персептрона згідно теорем Арнольда-Колмогорова-Хехт-Нільсена має бути не менше одного. Число нейронів в прихованих шарах може бути наближено оцінено за вищезгаданими формулами, однак його бажано оптимізувати для кожного конкретного завдання.

Існують два способи оптимізації числа нейронів в прихованих шарах - **деструктивний** і **конструктивний**.

Деструктивний спосіб полягає в тому, що спочатку будується мережа із свідомо надлишковим числом ступенів свободи, а потім в процесі навчання з неї поступово виключаються зайві синаптичні зв'язки і нейрони. Існує досить велике різноманіття алгоритмів виключення надлишкових ступенів свободи нейромережі. Проте, загальним недоліком деструктивних алгоритмів є значна тривалість їх роботи, оскільки початкові обчислення проводяться в мережах, що мають надмірну кількість нейронів.

Цього недоліку позбавлені альтернативні *конструктивні алгоритми*, які передбачають поступове додавання нейронів до мережі, в якій їх свідомо недостатньо. Нові нейрони додаються кожен раз після певного числа епох навчання, а синаптичним вагам і порогам нових нейронів присвоюються випадкові числа. Цікаво відзначити, що навички, набуті нейромережею в процесі попереднього навчання, не втрачаються повністю при додаванні в неї нового нейрона.

Вибір алгоритму навчання

Винахід алгоритму зворотного поширення помилки (який фактично реалізовує відомий метод *найшвидшого спуску*) відкрив шлях широкого практичного застосування багат шарового персептрона. Разом з тим з розширенням фронту наукових досліджень виявилися і недоліки цього алгоритму. Є дуже багато методів, які використовують то й же принцип (**градієнтні**). Проте, якщо поверхня помилок має складну будову і локальні мінімуми, то такі методи досить часто "заходять" в ці локальні мінімуми.

Є ще група методів, що називаються **евристичними**. Вони є дуже вузькоспеціалізованими, не мають строгого теоретичного підґрунтя, і в них відображається особистий досвід роботи авторів із неймережами.

Часто процес навчання набуває характеру тривалого експериментування, в ході якого пробуються різні оптимізаційні алгоритми з різними параметрами. В результаті такого успіху застосування неймережових технологій ставиться в залежність від досвіду та інтуїції фахівця, числа різних оптимізаційних алгоритмів, що є в його розпорядженні. У зв'язку з цим актуальним є розвиток методів **глобальної оптимізації**, тобто таких, які дозволяють знайти глобальний мінімум багатоекстремальної цільової функції.

Серед безлічі можливих підходів найбільш успішним визнається ідея **генетичних алгоритмів**. Ця ідея, вперше запропонована Дж.Холландом в 1970-х рр., полягає в імітації природних оптимізаційних процесів, що відбуваються при еволюції живих організмів. Відповідно до еволюційної теорії природа оптимізує все живе завдяки двом біологічним механізмам - *природному відбору* і *генетичній спадковості*. При навчанні персептронів зазвичай застосовують різні варіації генетичних алгоритмів, пов'язані з вибором параметрів і способів основних процесів алгоритму (*селекції, схрещування і мутацій*).

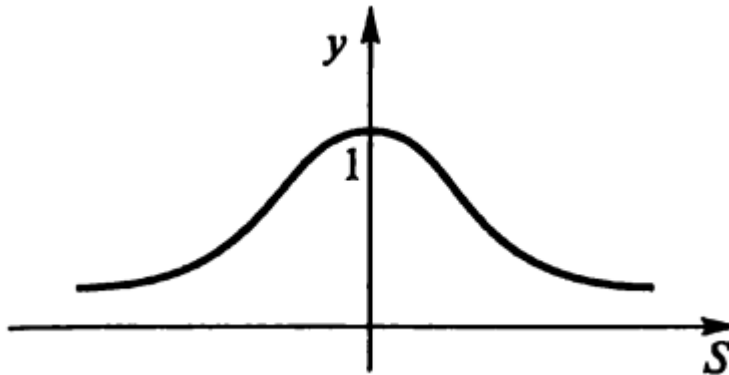
Ефективним виявилось поєднання генетичних алгоритмів з локальними алгоритмами оптимізації: на початкових стадіях працює генетичний алгоритм, а потім особина-переможець або група найбільш досконалих особин продовжують спуск до мінімуму, наприклад одним із градієнтних методів.

Радіально-базисні мережі

Останнім часом набули поширення нейромережі, нейрони яких мають активаційні функції в формі функції Гаусса

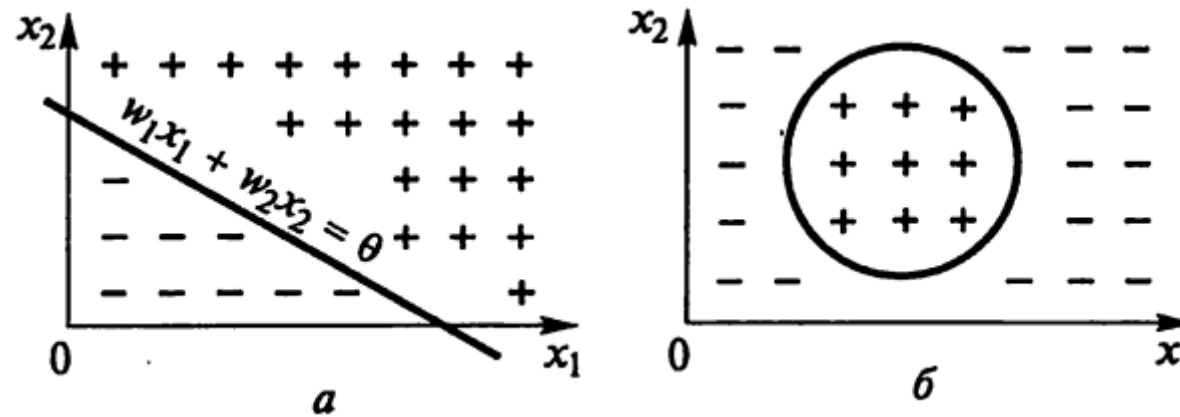
$$y = e^{-\frac{s^2}{2\sigma^2}},$$

де S – евклідова відстань між вхідним вектором X і деяким C центром активаційної функції, σ - параметр гаусової кривої.



Нейрон з радіально-базисною функцією також ділить простір вхідних параметрів на дві частини, проте розділяючою поверхнею тут є гіперсфера. Для точок простору, що лежать всередині гіперсфери,

вихід нейрона додатній (позитивний), а для точок, що лежать зовні гіперсфери, він дорівнює нулю (негативний).

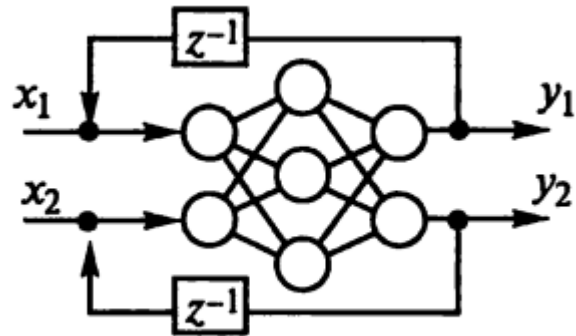


Радіально-базисні нейрони мають перевагу, що полягає в тому, що з їх допомогою легше побудувати поверхню, що забезпечує поділ вхідних параметрів на класи. У зв'язку з цим в радіально-базисних мережах відсутня необхідність використання великої кількості прихованих шарів. Типова радіально-базисна мережа має тільки один прихований шар, причому синаптичні ваги нейронів прихованого шару дорівнюють одиниці, а нейрони вхідного і вихідного шарів мають лінійні активаційні функції.

Рекурентні мережі

Як показали нейрофізіологічні дослідження, мозок людини має набагато складнішу структуру і механізми взаємодії між нейронами, ніж ті, які реалізовані в розглянутих вище штучних нейронних мережах. Зокрема, між біологічними нейронами виявлено велику кількість не тільки прямих, але і зворотних зв'язків. У зв'язку з цим були зроблені спроби доповнити штучні нейронні мережі зворотними зв'язками.

На рисунку наведено приклад персептрона, у якого вихідні сигнали y_1 і y_2 через елементи одиничних затримок z^{-1} подаються назад на входи персептрона. Таким чином, під впливом вхідних сигналів x_1 і x_2 на виході мережі в момент часу t виробляються сигнали $y_1(t)$ і $y_2(t)$, а в наступний момент часу під впливом цих сигналів, що подаються на вхід, виробляються нові вихідні сигнали $y_1(t + 1)$ і $y_2(t + 1)$.



Показано, що для навчання таких мереж також можна застосовувати алгоритм зворотнього розповсюдження помилки.

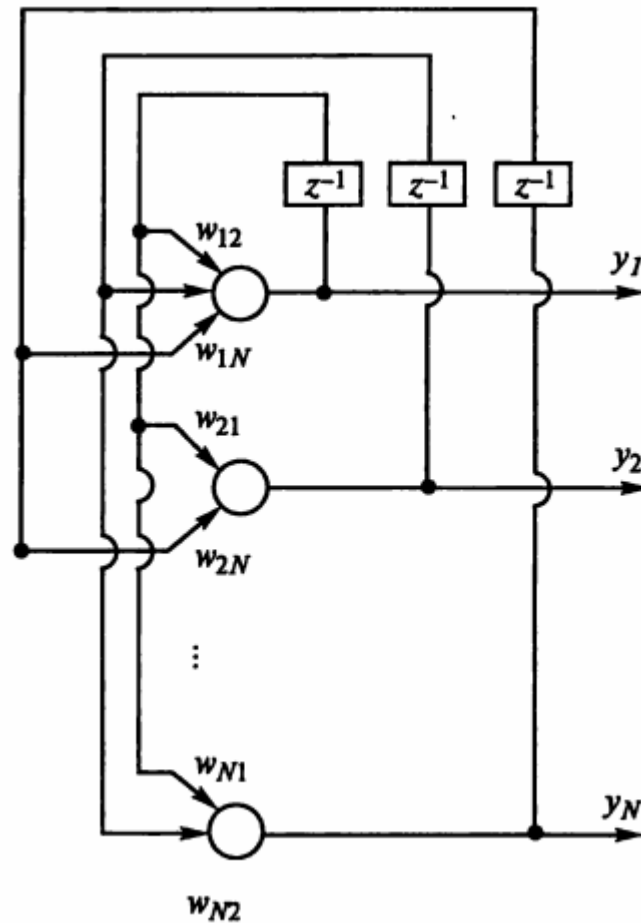
Рекурентні мережі розглянутого типу широко застосовуються (зокрема) для математичного моделювання динамічних об'єктів.

Розглянемо ще один приклад рекурентних мереж.

Мережа Хопфілда

Хопфілд звернув увагу на те, що динамічний процес, що виникає в замкнутій самій на себе рекуррентній мережі може привести до деякого стійкого стану, що відрізняється від початкового. Іншими словами, ітераційний процес рекуррентної мережі може вивести на стаціонарний режим, при якому стан мережі перестане змінюватися. Причому цей кінцевий стаціонарний стан мережі залежить як від її початкового стану, так і від значень елементів матриці синаптичних ваг.

Мережа Хопфілда в класичному варіанті виконання наведена на рисунку. Вважається, що вона не має вхідних елементів, а вхідний вектор задає початкову активність нейронів, яка потім змінюється в ході ітераційного процесу, обумовленого наявністю зворотних зв'язків.



Мережа Хопфілда поводитья подібно пам'яті, що зберігає деякий заданий заздалегідь набір образів, яка намагається згадати один з них, якщо їй пред'являється будь-якої з цих образів, спотворений перешкодами.

Самонавчальні та гібридні мережі

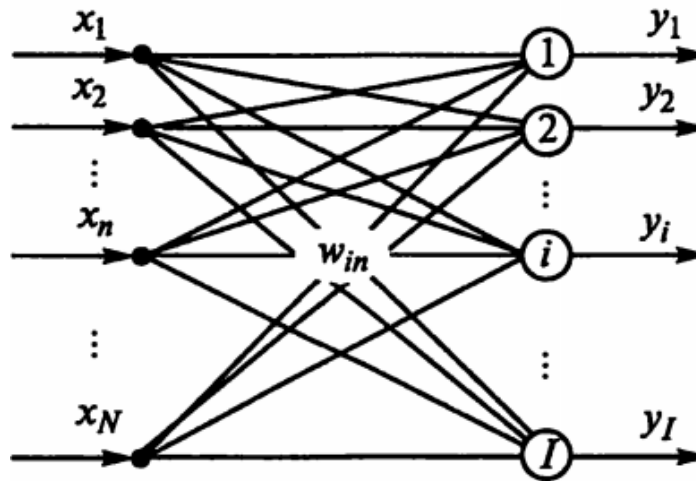
Розглянуті вище нейронні мережі персептронного типу навчалися шляхом тренування на прикладах. Для навчання надавалася первинна інформація про предметну область у вигляді набору вхідних векторів X і відповідних їм вихідних векторів D - свого роду підказок, з використанням яких мережа навчалася давати правильні відповіді на запитання. Тому такий спосіб називають *навчанням з учителем*.

Механізм навчання з учителем безсумнівно властивий мозку будь-якої живої істоти. Однак мозок людини володіє і іншими механізмами навчання. Ми в змозі вирішувати багато інтелектуальних завдань і без попереднього навчання.

Нейронна мережа, яка без допомоги вчителя автоматично налаштовує свої синаптичні ваги, вирішуючи завдання кластеризації (розбиття на підмножини-класи) вхідних векторів, називається **мережею Кохонена**, на честь вченого, який запропонував алгоритм самонавчання мережі. Така мережа має один шар (так зв. *шар Кохонена*) з I нейронами і N входами. Алгоритм навчання полягає в наступному.

- для кожного нейрона датчиком випадкових чисел задаються початкові вектори ваг, розмірність яких співпадає з розмірністю вхідних векторів.

- знаходимо евклідову відстань між кожним вектором ваг і вхідним сигналом нейрона X . нейрон, у якого вектор синаптичних ваг є "найближчим" до вхідного вектора оголошується "переможцем".
- ваги коригуються або по принципу "переможець забирає все" (коригується тільки переможець) або по принципу "переможець забирає більше" (переможець корегується в значній мірі, інші - в меншій).



Варто відмітити, що згідно останніх даних, мозок людини є каскадним об'єднанням біологічних нейронних мереж різних функціональних призначень. Тому *гібридна мережа*, яка об'єднує в собі шари різних нейромережєвих парадигм, щонайбільшою мірою відповідає сучасним уявленням про структуру і спосіб функціонування мозку, а також є більш ефективною.