Raport Badawczy

Research Report

RB/57/2010

Model neuronu jako nomogram

B. Fiksak

Instytut Badań Systemowych
Polska Akademia Nauk

Systems Research Institute
Polish Academy of Sciences

MODEL NEURONU JAKO NOMOGRAM

Bogumił Fiksak
Studia Doktoranckie IBS PAN

In this paper it is considered the 13th Hilbert's Problem and its solution given by young Russian mathematician Vladimir Arnold. It seems this solution can be successfully used to construct nomograms. Next it shown that construction of nomograms has many similarities with neural networks. As a simply example a model of an artificial neuron was compare to a proper nomogram.

Słowa kluczowe: *13 Problem Hilberta, sieci neuronowe, nomogramy.*

Wprowadzenie

W 1900 r. na Międzynarodowym Kongresie Matematyków w Paryżu David Hilbert przedstawił 23 nierozwiązane zagadnienia, zwane później *23 Problemami Hilberta*. Jednym z tych problemów był *13 Problem Hilberta* stawiający pytanie czy można równanie 7-go stopnia z trzema zmiennymi przedstawić w postaci skończonej liczby funkcji dwóch zmiennych.

Oryginalne postawienie 13-tego Problemu Hilberta ma postać [1]:

Rozwiązania $x = f(a,b,c)$ równania

$$x^7 + a\,x^3 + bx^2 + cx + 1 = 0 \tag{1}$$

nie mogą być przedstawione jako superpozycje ciągłych funkcji dwóch zmiennych.

Problem ten został rozwiązany przez Vladimira Arnolda i Andreja Kołmogorova w 1957 roku, którzy udowodnili następujące twierdzenie.

Każda ciągła funkcja *n* zmiennych $f(x_1, x_2, \ldots, x_n)$ określona w obszarze $[0,1]^n$ może być przedstawiona w postaci:

$$f(x_1, x_2, \ldots, x_n) = \sum_{i=0}^{2n} g_i\left(\sum_{j=1}^{n} h_{ij}(x_j)\right) \tag{2}$$

gdzie funkcje g_i i h_{ij} są ciągłymi funkcjami jednej zmiennej.

Zależność (2) początkowo dotyczyła możliwości konstruowana nomogramów, ale rozwiązanie tego problemu zostało także efektywnie wykorzystane do konstruowania sztucznych sieci neuronowych. Istnieją zatem pewne wspólne teoretyczne podstawy nomogramów i sztucznych sieci neuronowych.

W pracy zostanie przedstawiona procedura konstruowania nomogramów jako wielowarstwowych sieci neuronowych.

1. Model neuronu

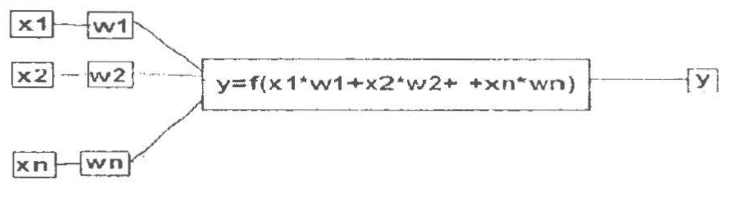
W 1943 roku uproszczony model neuronu został zaproponowany przez McCullock'a i Pitts'a. Taki model oblicza sumę ważoną sygnałów wejściowych transmitowanych przez dendryty i daje na wyjściu aksonu sygnał w postaci funkcji Heaviside'a.

Od tamtego czasu zaproponowano wiele różnorodnych funkcji przejścia odmiennych od funkcji Heaviside'a, takich jak funkcja sigmoidalna, tangens hiperboliczny i inne. W oparciu o takie modele neuronów zbudowano wiele poprawnie działających przykładów sztucznych sieci neuronowych.

Nieliniowy element w sieciach neuronowych może być opisany funkcją aktywacji neuronu:

$$y = f(x_1 w_1 + x_2 w_2 + ... + x_n w_n) \qquad (3)$$

gdzie $w_1, w_2, ..., w_n$ oznaczają wagi sygnałów wejściowych; $x_1, x_2, ..., x_n$ oznaczają sygnały wejściowe, Rys.1.



Rys. 1. Model neuronu

Uczenie pojedynczego neuronu polega na takim doborze wag $w_1, w_2, \ldots, w_n$, aby jak najlepiej dopasować funkcję $f(x_1, x_2, \ldots, x_n, y) = 0$ do funkcji aproksymowanej.

Podstawowy algorytm uczenia został wprowadzony przez Widrowa i Hoffa i nazwany jest regułą DELTA. Na wejście neuronu jest podawany sygnał x, a sygnał wyjściowy y jest obliczany jako iloczyn wag i sygnału wejściowego. Różnica pomiędzy odpowiedzią neuronu y, a zadanym sygnałem wejściowym jest podstawą do korekcji wag. Matematyczny dowód poprawności uczenia pojedynczego neuronu został przedstawiony w pracy Ryszarda Tadeusiewicza, Sieci Neuronowe. Zmiana wag zależy od iloczynu sygnału wejściowego, błędu pomiędzy sygnałem wyjściowym zadanym, a obliczonym i specjalnego współczynnika liczbowego decydującego o szybkości uczenia.

2. Nomogramy

Nomogramem nazywamy wykres umożliwiający szybkie, proste i przybliżone wyznaczenie - bez obliczania - wartości dowolnej zmiennej z równania:

$$f(x_1, x_2, \ldots, x_n, y) = 0 \tag{4}$$

gdy znane są wartości pozostałych zmiennych.

Na nomogramie znajdują się najczęściej dwie osie liczbowe, na których zaznaczono dane wejściowe. Wartość wyjściową odczytuje się z nomogramu. Przykład dla odczytywania kwadratu liczb z przedziału od -10 do 10 przedstawiono poniżej na Rys 2.

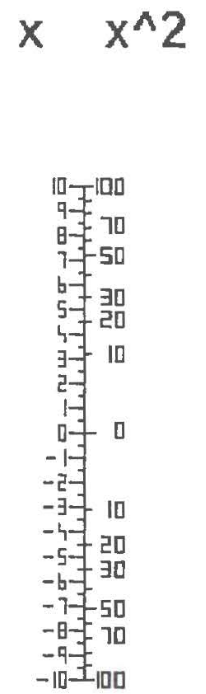
W celu ułatwienia wizualizacji przebiegu zmienności funkcji konstruujemy wykresy takich funkcji. Jest to potężne narzędzie stosowane w wielu dziedzinach nauki i techniki. Szczególnie w technice i zagadnieniach inżynierskich stosowane są specjalne wykresy funkcji zwane *nomogramami kolineacyjnymi*. Słowo *kolineacyjny* oznacza, że dwa punkty odpowiadające danym wartościom pewnych zmiennych leżą na wspólnej linii prostej.

Zaletami nomogramów kolineacyjnych są:

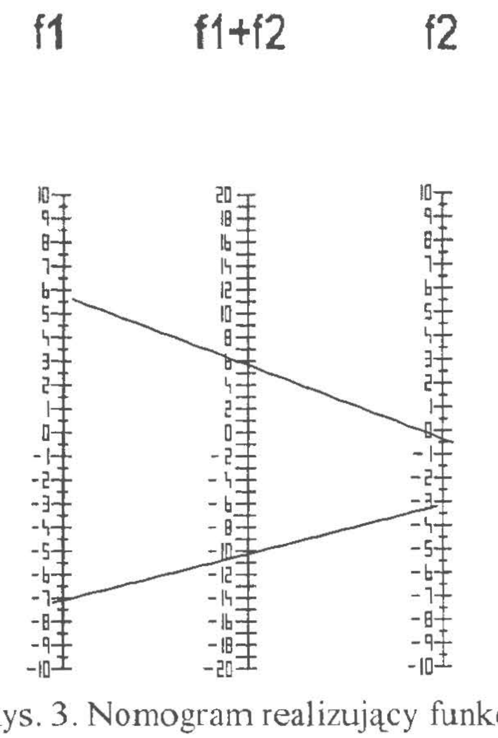
- prostota
- całościowy zakres zmienności dziedziny i przeciwdziedziny
- łatwość operowania
- możliwość wyznaczania granic zmienności.

Nomogramy, obok wielu zalet, posiadają również istotne wady takie jak:

- ograniczona dokładność wyznaczanych wielkości
- trudności w konstruowaniu.



Rys. 2. Prosty nomogram realizujący funkcję $y = x^2$



Rys. 3. Nomogram realizujący funkcję $y = f_1 + f_2$

3. Model neuronu a nomogram

Wyjście neuronu opisuje następująca zależność

$$y = f\left(x_1 w_1 + x_2 w_2 + \ldots + x_n w_n\right) \tag{5}$$

Wprowadźmy następujące oznaczenia:

$$\begin{aligned} f_1(x_1 w_1) &= x_1 w_1 \\ f_2(x_2 w_2) &= x_2 w_2 \\ &\vdots \\ f_n(x_n w_n) &= x_n w_n . \end{aligned} \tag{6}$$

Po podstawieniu (6) do wyrażenia (5) otrzymujemy

$$y = f\left(f_1 + f_2 + \ldots + f_n\right) \tag{7}$$

Funkcja odwrotna funkcji $f(.)$ ma następującą postać

$$f^{-1}(y) = g(y) = f_1 + f_2 + \ldots + f_n \tag{8}$$

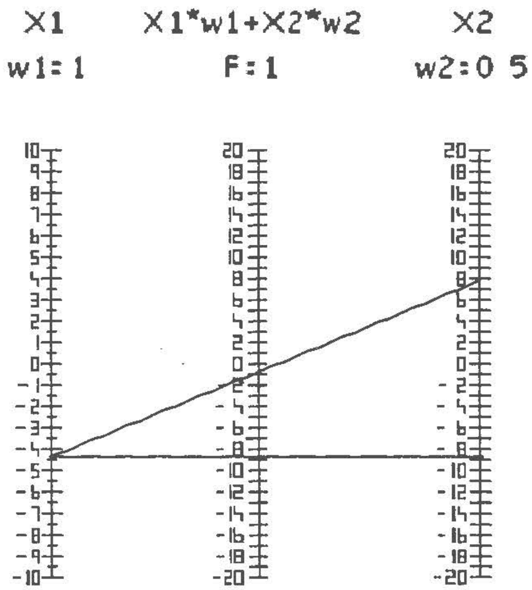
Wyznaczoną powyżej postać funkcji można przedstawić w postaci kolineacyjnej. Jeżeli sumujemy tylko dwie funkcje to wówczas neuron możemy przedstawić jako nomogram składający się z trzech osi liczbowych. Jeżeli wejść neuronu jest więcej to musimy skorzystać z metody geometrycznego łączenia nomogramów przedstawionych w pracy Edwarda Otto *Nomografia.*

Przykład:

Najprostszy model neuronu składa się z dwóch wejść x_1 i x_2, sygnałom wejściowym przyporządkowane są wagi w_1 i w_2. Ważona suma sygnałów wejściowych $x_1 w_1 + x_2 w_2$ jest argumentem funkcji przejścia $f(.)$, wyjście neuronu ma zatem następującą postać

$$y = f\left(x_1 w_1 + x_2 w_2\right) \tag{9}$$

Taki model można przedstawić w postaci nomogramu składającego się z trzech prostych równoległych położonych w równych odstępach.



Rys. 4. Nomogram realizujący funkcję $f(x_1, w_1 + x_2 w_2)$

Na osiach skrajnych występują dwie skale liniowe x_1, x_2 przemnożone przez wagi w_1 i w_2, a na osi środkowej jedna skala wyznaczająca wyjście neuronu w postaci funkcji $f(x_1, w_1 + x_2 w_2)$.

Funkcja f w zasadzie może być dowolnie stosowaną w modelach liniowych i nieliniowych neuronów. Jednak warunkiem dla zachowania ścisłości jednak jest, aby posiadała funkcję odwrotną.

W podanym przykładzie f jest funkcją liniową i równą 1. Model więc przedstawia neuron liniowy.

Przeprowadzając prostą przecinającą wszystkie trzy osie możemy odczytać wartości liczbowe na tych osiach x_1, x_2 i y.

Sam proces analizy jest szybki i praktycznie sprowadza się do przyłożenia linijki i odczytu na trzech osiach.

W przeciągu krótkiego czasu możemy w sposób graficzny przeanalizować sposób działania neuronu.

Wnioski

Praca przybliża analizę działania pojedynczego neuronu o nieliniowej funkcji przejścia. Działanie to można prześledzić na prostym kolineacyjnym nomogramie. W znaczącym stopniu ułatwia to dydaktykę, tym bardziej, że nomogram w małym stopniu zależy od rodzaju funkcji przejścia. Na jednym nomogramie można analizować wiele typów neuronów o różnych funkcjach przejścia.

Analiza tych funkcji w przestrzeniach wielowymiarowych pozwala porównać intuicyjne wywody z geometryczną ścisłością. Nasuwa się następujące pytanie: jeżeli pojedynczy neuron można przedstawić w postaci nomogramu kolineacyjnego to czy można przedstawić w takiej postaci całą sieć?

Odpowiedź na to pytanie jest twierdząca ponieważ, należałoby dołączyć kolejny nomogram reprezentujący neuron kolejnej warstwy.

Niech jednak nie zmyli nas prostota takiego modelu ponieważ nomogram reprezentujący sieć złożona już z trzech takich neuronów reprezentuje przekształcenie liniowe płaszczyzny w samą siebie.

Literatura

[1] V. I. Arnold (1959): Representation of continuous functions of three variables by the superposition of continuous functions of two variables. *Mat. Sb.* 48 (90).

[2] N. Kolmogorov (1957): On the representation of continuous functions of several variables by superpositioins of continuous functions of one variable, and addition, *Dokl. Akad. Nauk ZSSR* 114, 953-957.

[3] R. Tadeusiewicz (1993): *Sieci Neuronowe*, Akademicka Oficyna Wydawnicza, Warszawa.

[4] E. Otto (1964): *Nomografia*, PWN, Warszawa.