

# **METODY SZTUCZNEJ INTELIGENCJI**

## **Laboratorium nr 3**

**Temat: Uczenie neuronów liniowych**

## Uczenie jednowarstwowych sieci neuronowych

### 1. Cel ćwiczenia

Celem ćwiczenia jest zapoznanie studentów z gradientowymi metodami uczenia adaptacyjnego i wsadowego neuronów, z zastosowaniem funkcji aktywacji typu: funkcje liniowe, funkcje sigmoidalne bipolarne.

Zajmiemy się przypadkiem, gdy SN zbudowana jest tylko z jednego neuronu o **liniowej** funkcji aktywacji .

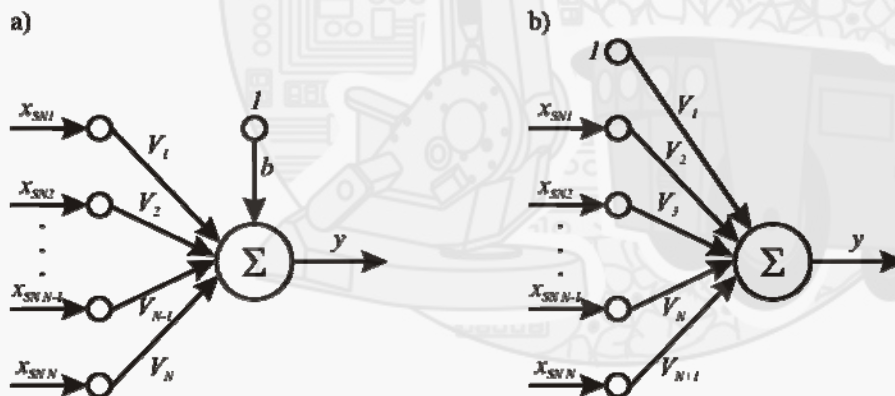
#### 1.1. Sieć neuronowa z liniowymi funkcjami aktywacji neuronów

Opiszmy liniową funkcję aktywacji neuronu przy pomocy zależności:

$$y = f(x) = \mathbf{w}x + \mathbf{b} , \quad (1)$$

gdzie  $w, b$  - to wagi.

Neuron liniowy o  $N$  wejściach schematycznie przedstawiono na rys. 1.a), gdzie  $V$  to wektor wag warstwy wejściowej do sieci, którym zastąpiliśmy wagę  $w$  występującą w równaniu (1) odpowiadającą wejściu skalarnemu, natomiast  $b$  to wolna waga, tzw. **bias**. Schemat ten jest równorzędny ze schematem neuronu liniowego pokazanym na rys. 1.b), gdzie dla zwięzłego zapisu wolną wagę  $b$  włączono do wektora wag wejściowych  $V$ , przez co rozmiar wektora  $V$  wzrósł do  $N+1$ , oraz poszerzono wektor wejściowy sieci (neuronu) do postaci  $\mathbf{x}_v = [1, x_{SN}]^T$ , gdzie  $x_{SN}$  to wektor wejściowy  $x$  .



Rys. 1. Schemat neuronu liniowego a) z wolną wagą  $b$ , b) z poszerzonym wektorem wejściowym  $x_{SN}$

**W procesie uczenia neuronu, uczeniu podlegać będą wagi warstwy wejściowej do sieci  $V$ .**

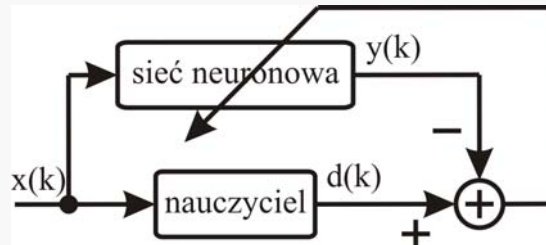
#### Uczenie sztucznych sieci neuronowych

Celem procesu uczenia jest taki dobór wag, który umożliwi odwzorowania sygnałów wejściowych w odpowiednie sygnały wyjściowe. Wyróżnia się dwa sposoby uczenia:

- **sposób przyrostowy (iteracyjny, adaptacyjny)** – aktualizacja wag następuje bezpośrednio po podaniu każdej pary uczącej; funkcja błędu zmniejsza się w każdym kolejnym kroku iteracyjnym,
- **sposób grupowy (wsadowy)** – korekta wag następuje po podaniu całego zestawu uczącego, wszystkich par uczących .

Z punktu widzenia systemu uczącego sieci neuronowe, można wyróżnić między innymi uczenie

- **pod nadzorem (z nauczycielem)**, co pokazano na rys. 2 – dla każdego wektora ze zbioru uczącego znana jest poprawna odpowiedź; korekcja wag przeprowadzana jest na podstawie różnicy pomiędzy odpowiedzią rzeczywistą sieci i odpowiedzią zadaną,



Rys. 2. Schemat uczenia pod nadzorem

## 1.2 Adaptacyjne uczenie neuronu liniowego

**Przykład. 1.** Uczenie wag  $V$  neuronu liniowego

Przeprowadźmy proces uczenia SN z liniowymi funkcjami aktywacji neuronów, zbudowanej z jednego neuronu liniowego, której schemat przedstawiono na rys.3.



Rys.3. Schemat SN zbudowanej z pojedynczego neuronu liniowego

Zadaniem SN jest aproksymacja nieznanego odwzorowania na podstawie dyskretnych danych pomiarowych zgromadzonych w czasie eksperymentu (np. pomiar rezystancji potencjometru w zależności od wysunięcia suwaka), przedstawionych w tab. 1. Wartości wejścia  $x$  oraz nieznanego odwzorowania  $f(x)$  podane w tab. 1 będą stanowiły zbiór wzorców uczących (danych treningowych) w procedurze adaptacji wag warstwy wejściowej  $V$  SN z neuronem liniowym.

Tab. 1. Dane pomiarowe z eksperymentu

Nr pomiaru	$x$ (np. wysunięcie suwaka potencjometru [cm])	$f(x)$ (np. rezystancja potencjometru [ $\Omega$ ])
1	0	17.74
2	.1	23.98
3	.2	27.73
4	.3	34.51
5	.4	38.3
6	.5	45.29
7	.6	48.32
8	.7	54.91
9	.8	62.39
10	.9	65.62

Zastosowanie metody gradientowej (metody najszybszego spadku), powszechnie stosowanej w teorii optymalizacji, do adaptacji wag  $V$  warstwy wejściowej SN z liniowymi funkcjami aktywacji neuronów, pozwala na iteracyjne wyznaczanie kolejnych przybliżeń optymalnych wag sieci (w rozważanym przypadku neuronu) w kierunku ujemnego gradientu funkcji błędu.

Rozwiązanie zadania polega na doborze wag ( $V_2=w, V_1=b$ ) wg kryterium błędu kwadratowego (kryterium błędu) wg równania:

$$J = \frac{1}{2} \sum_{i=1}^N e_i^2, \quad (2)$$

gdzie  $y_i - (b + wx_i) = y_i - \hat{y}_i = e_i$  jest błędem dopasowania,  $N$  jest liczbą obserwacji.

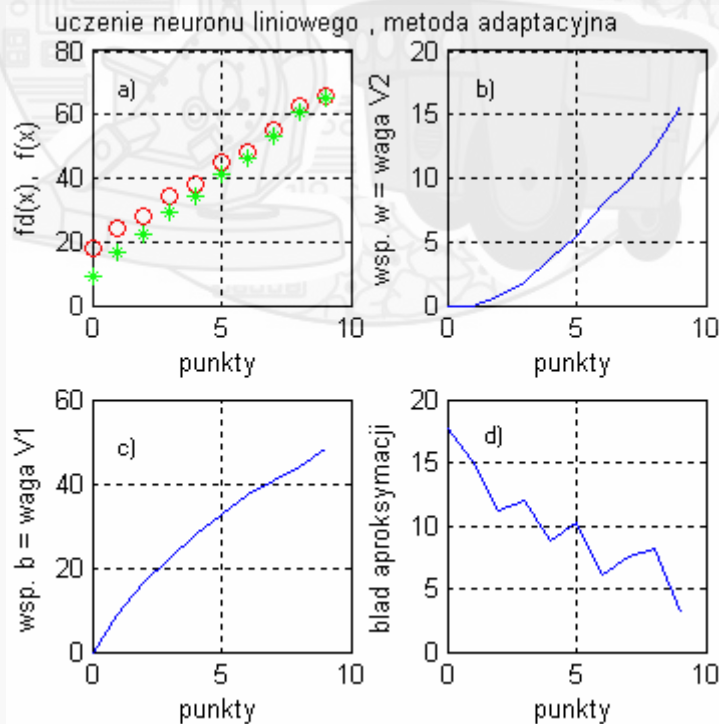
Wyznaczenie wag  $w$  i  $b$  przeprowadzono wg metody gradientowej (najszybszego spadku). W tej metodzie wagi  $w$  i  $b$  oblicza się następująco:

$$\begin{aligned} w(k+1) &= w(k) + \eta e(k)x(k) \\ b(k+1) &= b(k) + \eta e(k) \end{aligned} \quad (3)$$

Proces adaptacji wag wejściowych neuronu liniowego przeprowadzono w pakiecie Matlab, jak podano na wykładzie.

Przeprowadzono procedurę adaptacji wag wejściowych  $V$  neuronu liniowego przy pomocy metody uczenia nadzorowanego z zastosowaniem zależności (3), prezentując sekwencyjnie SN wszystkie wzorce uczące ze zbioru treningowego (tab. 1).

Wybrano wagi warstwy wejściowej  $V$  równe zero w procesie inicjalizacji SN, oraz współczynnik wzmocnienia uczenia wag  $\alpha=0.5$ . Dla prezentacji wzorców uczących ze zbioru treningowego otrzymano przebiegi pokazane na rys. 4.



Rys.4. Wyniki neuronowego odwzorowania

Na rys. 4.a przedstawiono zadane dyskretne wartości funkcji  $y=fd(x)$  (kółka) oraz wartości wyjścia z neuronu liniowego  $\hat{y}=f(x)$  (gwiazdki) otrzymane podczas procesu aproksymacji. Rys.4.b i 4c

przedstawiają odpowiednio wagi  $V_2$  i  $V_1$  neuronu liniowego. Wagi  $V=[V_1, V_2]^T$  początkowo mają wartości równe zero, następnie są adaptowane zgodnie z zależnością (3). Na rys. 4.d) przedstawiono wartości błędu  $e_i$  dla dyskretnego kroku.

**Do ilościowej oceny jakości estymacji** zadanego odwzorowania zastosowano wskaźniki jakości:  
- suma kwadratów błędu (Sum of Squared Error, SSE)

$$SSE = \sum_{i=1}^n e_i^2 \quad (4)$$

- pierwiastek błędu średniokwadratowego (*Root Mean Square Error, RMSE*),

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n e_i^2} \quad (5)$$

Wartości poszczególnych wskaźników jakości dla wybranych etapów procesu prezentacji wzorców uczących przedstawiono w tabelicy 2.

Tab.2.

SSE	RMSE
1.1619e+003	10.7791

Zaprezentowane w przykładzie rozwiązania nie są jedynymi. Możliwe są również inne rozwiązania. Należy podkreślić, że metoda gradientowa zastosowana w uczeniu neuronu gwarantuje osiągnięcie **jedynie minimum lokalnego**.

### 1.3. Uczenie wsadowe neuronu liniowego

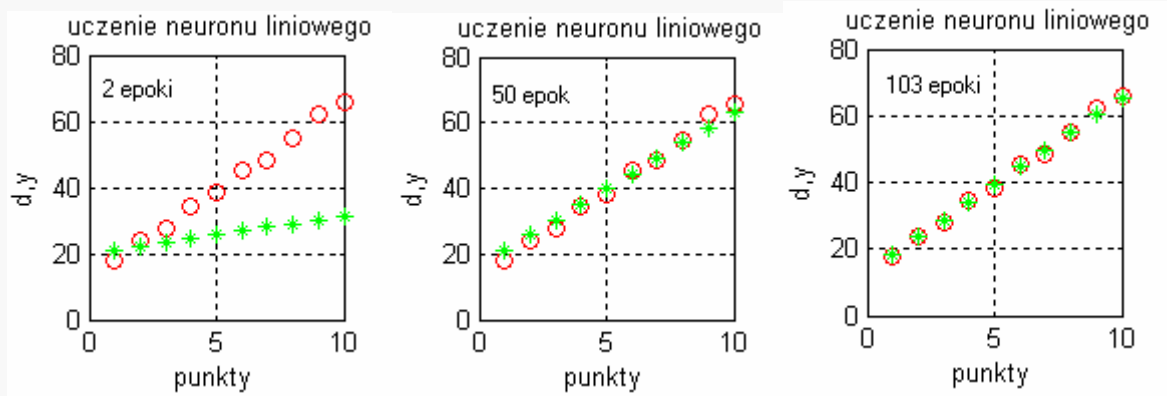
W uczeniu wsadowym do sieci podaje się pełny ciąg uczący (wszystkie pary  $(x^p, y^p)$  dla  $p = 1, 2, \dots, P$ ) w danym kroku iteracyjnym. Ten krok iteracyjny nazywa się **epoką**. Błąd sieci jest kumulowany, aż do momentu gdy wszystkie pary danych  $(x^p, y^p)$  zostaną zaprezentowane sieci. Na końcu tej procedury uaktualnia się wagi sieci wg zależności

$$\begin{aligned} w(k+1) &= w(k) + \eta \sum_{p=1}^P e^p x^p \\ b(k+1) &= b(k) + \eta \sum_{p=1}^P e^p \end{aligned} \quad (6)$$

W uczeniu wsadowym indeks iteracji  $k$  odpowiada liczbie prezentacji całego zbioru danych o zawartości  $P$ , to znaczy  $k$  odpowiada liczbie epok.

Stosując zależności (6) przeprowadzono aproksymację nieznanego odwzorowania na podstawie dyskretnych danych podanych w tab.1.

Uczenie wsadowe przeprowadzono według programu podanego na wykładzie. Na rys 5 zamieszczono uzyskane wyniki .



Rys.5. Wyniki neuronowego odwzorowania po 2 50 i 103 epokach uczenia

Przebiegi wyjścia z SN dla wag  $V$  warstwy wejściowej neuronu liniowego, otrzymanych po wybranych etapach procesu adaptacji wag przedstawiono na rys. 5. Zauważamy, iż przebieg wygenerowany przez sieć po 2 epokach przebiegu procesu adaptacji wag w znacznym stopniu odbiega od wartości zadanych, jednakże kontynuacja procesu adaptacji wag zbliża przebieg generowany przez SN do wartości zadanych. Po 103 epokach procesu prezentacji wzorców uczących, oraz adaptacji wag, wyjście SN generuje wartości bardzo bliskie wartościom zadanim. W procesie uczenia zatrzymanie procedury nastąpiło po spełnieniu kryterium stopu  $SSE(k) < 10$ .

**Wartości poszczególnych wskaźników jakości** (4), (5) dla wybranych liczby epok prezentacji wzorców uczących przedstawiono w tablicy 3.

**Tab.3.**

Epoki	SSE	RMSE
2	3.9271e+003	19.8170
50	46.9575	2.1670
103	9.9800	0.9990

Jak wynika z analizy wyników umieszczonych w tab.3. kryterium stopu jest spełnione po zrealizowaniu 103 epok uczenia.

## 2. Zadania do wykonania

Zrealizować dla tych samych danych uczenie **adaptacyjne i wsadowe**.

**2.1.** Przeprowadzić proces uczeni SN zbudowanej z jednego neuronu typu liniowego, metodą **adaptacyjna i wsadową** dla odwzorowania zadanego w **tab. 4** dla poszczególnych zespołów.

**Zespoły: 1, 5** - pomiar rezystancji potencjometru w zależności od wysunięcia suwaka.

**Zespoły: 2, 6** - pomiar mocy rozwijanej przez elektrownię wiatrową w zależności od bieżącej prędkości wiatru.

**Zespoły: 3, 7** - pomiar wydatku przepływu paliwa w funkcji czasu otwarcia wtryskiwacza przy stałej prędkości obrotowej silnika spalinowego z zapłonem iskrowym.

**Zespoły: 4, 8** - pomiar wartości przepływu czynnika chłodzącego w zależności od poziomu otwarcia zaworu.

**2.2.** Zbadać wpływ zmiany współczynnika wzmocnienia uczenia  $\alpha$  na jakość aproksymacji zadanej funkcji oraz szybkość zbieżności procesu adaptacji wag  $V$ .



**Tab. 4. Parametry SN dla poszczególnych zespołów, dane pomiarowe**

nr	typ f. aktywacji	$\alpha$	Zadane nieznane odwzorowanie										
			[cm]	0	2	4	6	8	10	12	14	16	18
1.	liniowa	0.5	[ $\Omega$ ]	137	195	261	307	338	398	434	484	544	611
			[m/s]	4	5	6	7	8	9	10	11	12	13
2.	liniowa	0.4	[MW]	0.03	0.32	0.46	0.68	0.74	0.89	0.86	0.95	1.03	1.01
			[ms]	1	2	3	4	5	6	7	8	9	10
3.	liniowa	0.3	[g/s]	1.14	2.88	4.11	5.86	6.7	8.01	9.87	11.59	12.56	14.4
			[%]	10	20	30	40	50	60	70	80	90	100
4.	liniowa	0.8	[hl/min]	0.16	0.31	0.59	0.62	0.72	0.83	0.93	0.91	0.91	1.02
			[cm]	0	1	2	3	4	5	6	7	8	9
5.	liniowa	0.7	[k $\Omega$ ]	8.97	8.59	7.95	7.63	6.85	6.3	6.03	5.69	4.88	4.5
			[m/s]	5.0	6.2	7.4	8.6	9.8	11.0	12.2	13.4	14.6	15.8
6.	liniowa	0.6	[MW]	0.09	0.23	0.53	0.54	0.65	0.78	0.9	0.87	0.88	1.02
			[ms]	2	4	6	8	10	12	14	16	18	20
7.	liniowa	0.5	[g/s]	3.15	6.21	9.8	14.36	16.85	20.54	24.41	28.67	32.56	37.01
			[%]	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
8.	liniowa	0.4	[hl/min]	0.2	0.39	0.68	0.72	0.81	0.9	0.98	0.95	0.95	1.0

**Sprawozdanie powinno zawierać:**

1. Wstęp teoretyczny
  - podstawowe wiadomości na temat metod uczenia neuronu:
    - o liniowego,
2. Przebieg ćwiczenia
  - a) SN z liniowymi funkcjami aktywacji neuronów, **uczenie adaptacyjne i wsadowe** :
    - listing kodu Matlab-a zrealizowany w postaci *m*-pliku, służący do uczenia SN z neuronem liniowym dla zadanego odwzorowania,
    - wykresy aproksymacji funkcji, wag, błędu oceny SSE, (wzór rys.4)
    - tabelę z wyznaczonymi wartościami wskaźników jakości (analogicznie do tab. 2 lub tab. 3),
3. Wnioski

Uwaga. Każdy realizowany podpunkt sprawozdania powinien być odpowiednio skomentowany.