

## 11. Aproksymacja metodą najmniejszych kwadratów

W naukach przyrodniczych wykonujemy często eksperymenty polegające na pomiarach par wielkości, które, jak przypuszczamy, są ze sobą powiązane jakąś zależnością funkcyjną  $y=f(x)$ , np. wydłużenie sprężyny w zależności od wiszącego na niej ciężaru, stężenie reagenta w zależności od czasu itp. Sensownym posunięciem jest znalezienie takiej krzywej, która możliwie najlepiej przybliży te punkty doświadczalne. Znajdowanie takich krzywych jest celem teorii **aproksymacji**.

W teorii aproksymacji zajmujemy się generalnie rzecz biorąc dwoma typami problemów. Pierwszym z nich jest wspomniane wyżej dopasowywanie do punktów eksperymentalnych krzywych, które w najlepszy sposób mogą reprezentować dane. Drugim przypadkiem, kiedy stosujemy aproksymację jest sytuacja, gdy funkcja jest znana, ale chcemy znaleźć prostszą funkcję, taką jak np. wielomian, której można użyć do określenia przybliżonych wartości funkcji danej.

Oba problemy były już zarysowane w rozdziale traktującym o interpolacji. Podano w nim sposób znajdowania wielomianu przybliżającego funkcję daną w postaci punktów lub tablicowaną w tym celu, by uzyskać te punkty (węzłowe). Wiemy już, że interpolacja może nam służyć do aproksymacji wartości funkcji między punktami węzłowymi, tyle że tych punktów nie może być zbyt wiele (efekt Rungego). Jak zatem postąpić w sytuacji, gdy mamy, powiedzmy, 20 punktów układających się w przybliżeniu na prostej. Możemy z tego wywnioskować, że nieznana funkcja  $y = f(x)$  musi wyrażać zależność liniową i użycie do jej przybliżenia wielomianu 19 stopnia jest raczej bez sensu. A co w sytuacji, w końcu nie tak rzadkiej, kiedy tych punktów eksperymentalnych są setki, jeśli nie tysiące?

Trzeba się po prostu zabrać do tego inaczej. Po pierwsze należy zrezygnować z warunku, że funkcja przybliżająca punkty doświadczalne (lub punkty uzyskane po stabilizowaniu skomplikowanej funkcji, którą chcemy wyrażać za pomocą funkcji prostszej, rezygnując przy tym z absolutnej dokładności) musi przez nie przechodzić. Wystarczy zażądać, by funkcja aproksymująca przebiegała jak najbliżej wszystkich punktów, była jak najlepiej do nich „dopasowana”. Jak to osiągnąć?

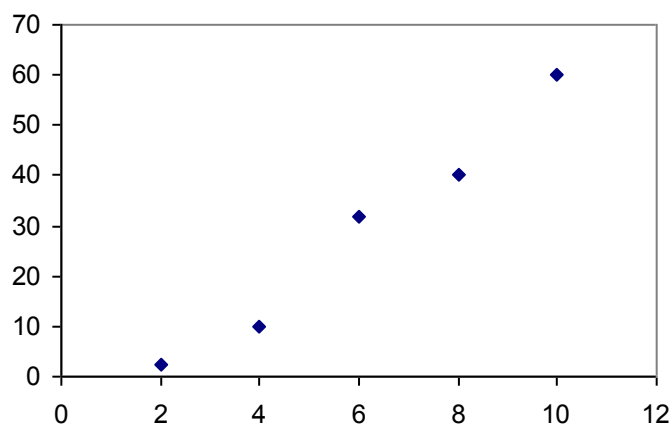
## Regresja liniowa.

Załóżmy, że mamy pięć punktów doświadczalnych danych w tabeli:

Tabela 11.1

$i$	$x_i$	$y_i$
1	2	2,5
2	4	10
3	6	32
4	8	40
5	10	60

Jeśli wykreślimy te punkty, otrzymamy Rysunek 11.1.



Rysunek 11.1

Widać, że chociaż punkty są nieco porzrucane na skutek, powiedzmy, błędów pomiarowych, to jednak wyraźnie układają się wzdłuż prostej. Świadczy to o tym, że mamy tu do czynienia z zależnością liniową i wszelkie próby aproksymowania tych punktów wielomianem interpolacyjnym ( czwartego stopnia), który musiałby przez nie przechodzić, nie mają większego sensu.

Równanie prostej ma postać następującą:

$$y = ax + b,$$

gdzie  $a$  i  $b$  to współczynniki, których chwilowo nie znamy.

Poszukiwanie parametrów takiej prostej, która by przechodziła możliwie najbliżej wszystkich punktów doświadczalnych  $(x_i, y_i)$ , polega na minimalizacji sumy:

$$S(a, b) = \sum_{i=1}^n [y_i - y(x_i)]^2 = \sum_{i=1}^n [y_i - (ax_i + b)]^2 \quad (11.1)$$

gdzie  $y(x_i)$  to wartości współrzędnej  $y$  obliczonej z równania prostej dla danych  $x_i$ . Różnice między dokładnymi wartościami  $y_i$  oraz wartościami obliczonymi z równania prostej są podniesione do kwadratu, aby uniknąć możliwości, że będą się nawzajem znosiły na skutek różnicy znaków. Z tego też względu przedstawiona w tym rozdziale metoda postępowania nosi nazwę **metody najmniejszych kwadratów**.

Dla danych z Tabeli 11.1 wielkość  $S$  będzie równa:

$$S(a,b) = [(2,5 - (2a+b))^2 + [10 - (4a+b)]^2 + [32 - (6a+b)]^2 + [40 - (8a+b)]^2 + [60 - (10a+b)]^2$$

Formalnie rzecz biorąc jest to funkcja dwóch zmiennych  $a$  i  $b$ . Interesują nas takie wartości tych zmiennych, dla których  $S(a,b)$  jest minimalna. Wiadomo, że funkcja wielu zmiennych ma minimum w punkcie, dla którego pochodne cząstkowe tej funkcji po wszystkich zmiennych są równe zeru, a zatem w tym przypadku muszą być spełnione warunki:

$$\begin{aligned} \frac{\partial S(a,b)}{\partial a} &= 0 \\ \frac{\partial S(a,b)}{\partial b} &= 0 \end{aligned} \quad (11.2)$$

czyli:

$$2[2,5 - 2a - b](-2) + 2[10 - 4a - b](-4) + 2[32 - 6a - b](-6) + 2[40 - 8a - b](-8) + 2[60 - 10a - b](-10) = 0$$

oraz

$$2[2,5 - 2a - b](-1) + 2[10 - 4a - b](-1) + 2[32 - 6a - b](-1) + 2[40 - 8a - b](-1) + 2[60 - 10a - b](-1) = 0$$

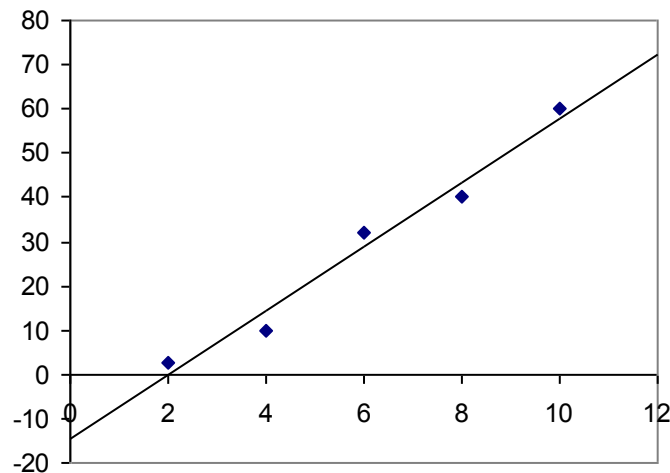
Po uproszczeniu otrzymujemy:

$$\begin{aligned} 440a + 60b &= 2314 \\ 60a + 10b &= 289 \end{aligned} \quad (11.3)$$

Rozwiązaniem tego układu równań liniowych są wielkości:

$$a = 7,25 \text{ oraz } b = -14,6$$

Na Rysunku 11.2 przedstawiono punkty z Tabeli 11.1 oraz linię prostą określoną przez równanie :  $y = 7,25 * x - 14,6$ .



Rysunek 11.2

Aproksymacja danych doświadczalnych krzywymi nosi często miano nazwę **regresji**. W przypadku, gdy do tych danych dopasowujemy prostą, mówimy o *regresji liniowej*.

Wyprowadźmy teraz wzory dla regresji liniowej w sposób ogólny. Równanie linii zapiszmy teraz nieco inaczej niż poprzednio jako:

$$p_1(x) = a_0 + a_1x = y(x)$$

Funkcja  $S(a_0, a_1)$  dla danych  $\{(x_i, y_i) [i = 1, 2, \dots, n]\}$  ma postać:

$$S(a_0, a_1) = \sum_{i=1}^n (y_i - a_0 - a_1x_i)^2$$

gdzie  $n$  oznacza liczbę punktów.

Chcąc znaleźć minimum tej funkcji, musimy rozwiązać układ równań, który nosi nazwę układu *równań normalnych* (równania (11.2) też są normalne).

$$\frac{\partial S(a_0, a_1)}{\partial a_0} = 2 \sum_{i=1}^n (y_i - a_0 - a_1x_i)(-1) = 0 \quad (11.4)$$

$$\frac{\partial S(a_0, a_1)}{\partial a_1} = 2 \sum_{i=1}^n (y_i - a_0 - a_1x_i)(-x_i) = 0$$

Po uporządkowaniu, otrzymamy układ równań:

$$a_0n + a_1 \sum_{i=1}^n x_i = \sum_{i=1}^n y_i$$

$$a_0 \sum_{i=1}^n x_i + a_1 \sum_{i=1}^n x_i^2 = \sum_{i=1}^n x_i y_i$$

z którego natychmiast dostaje się ogólne wzory na współczynniki definiujące linię prostą:

$$a_1 = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - \left( \sum_{i=1}^n x_i \right)^2} \quad a_0 = \frac{\sum_{i=1}^n y_i \sum_{i=1}^n x_i^2 - \sum_{i=1}^n x_i \sum_{i=1}^n x_i y_i}{n \sum_{i=1}^n x_i^2 - \left( \sum_{i=1}^n x_i \right)^2} \quad (11.6)$$

Spróbujmy określić niepewności pomiarowe w problemie regresji liniowej. Zakładamy, że niepewność w pomiarach  $x$ -ów jest zanedbywalna. Założenie to jest w pełni uzasadnione ponieważ nawet jeśli jakaś niepewność istnieje, to jest ona niewielka w porównaniu z niepewnościami  $y$ -ów. Dalej zakładamy, że niepewności wszystkich wartości  $y$  mają taką samą wielkość (jeśli tak nie jest możemy ratować się stosowaniem różnych wag statystycznych). Obliczając odpowiednie średnie odchylenie standardowe (błąd standardowy) powinniśmy uwzględnić fakt, że jest ono podwyższone z powodu błędów z jakimi wyliczyliśmy stałe  $a_0$  i  $a_1$ . Inaczej można powiedzieć, że wyznaczenie parametrów prostej obniża liczbę stopni swobody naszego układu. Znajduje to odzwierciedlenie w fakcie, że licząc średnią dzielimy przez  $(n-2)$  zamiast przez  $n$ , co powoduje odpowiednie podwyższenie wartości błędu.

$$s = \sqrt{\frac{1}{n-2} \left( \sum_{i=1}^n \varepsilon_i^2 \right)} \quad \text{gdzie } \varepsilon_i = y_i - a_0 - a_1 x_i \quad (11.7)$$

Po podstawieniu do wzorów (11.5) danych z Tabeli 11.1 otrzymujemy:

$$a_0 = -14,6 \quad a_1 = 7,25 \quad s = 4.151305$$

Takie same wartości współczynników prostej obliczyliśmy poprzednio, rozwiązując w prosty sposób układ równań (11.2).

Niepewności parametrów prostej znajdujemy metodami przenoszenia błędów, traktując te parametry jako funkcje zmierzonych wartości  $y$  (patrz rozdz. 1). Otrzymamy w ten sposób następujące wzory na błędy standardowe wyznaczonych parametrów  $a_0$  oraz  $a_1$ :

$$s_{a_0} = s \sqrt{\frac{\sum_{i=1}^n x_i^2}{n \sum_{i=1}^n x_i^2 - \left( \sum_{i=1}^n x_i \right)^2}} \quad s_{a_1} = s \sqrt{\frac{n}{n \sum_{i=1}^n x_i^2 - \left( \sum_{i=1}^n x_i \right)^2}} \quad (11.8)$$

gdzie wielkość  $s$  jest średnim standardowym odchyleniem od prostej zdefiniowanym wzorem (11.7).

Zastosowanie tych wzorów do danych z Tabeli 11.1 daje:

$$s_{a_0} = 4,353925; \quad s_{a_1} = 0,656379$$

Po wykreśleniu punktów danych w Tabeli 11.1 (Rysunek 11.1) oceniliśmy „na oko”, że są one powiązane zależnością liniową. W praktyce tego typu ocena raczej nie wystarcza i dlatego oblicza się tzw. **współczynnik korelacji**, którego wartość jest miarą *korelacji* między zmiennymi  $x$  i  $y$ . Jeśli współczynnik ten jest bliski  $\pm 1$  oznacza to, że wielkości są dobrze skorelowane (lub całkowicie, gdy współczynnik jest równy  $\pm 1$ ), co oznacza, że między tymi zmiennymi prawdopodobnie istnieje jakaś zależność funkcyjną. W przypadku regresji liniowej współczynnik ten nosi oczywiście nazwę współczynnika korelacji liniowej i oblicza się go zgodnie ze wzorem:

$$r = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{\sqrt{n \sum_{i=1}^n x_i^2 - \left( \sum_{i=1}^n x_i \right)^2} \sqrt{n \sum_{i=1}^n y_i^2 - \left( \sum_{i=1}^n y_i \right)^2}} \quad (11.9)$$

Wzór ten podajemy w postaci najlepiej nadającej się do obliczeń numerycznych. Warto jednak przyrzeć mu się bliżej, by zrozumieć jego sens.

Pokaż, że wyrażenie:

$$\sum_{i=1}^n (x_i - \bar{x})^2$$

w którym średnia wartość  $x$  jest:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

można przekształcić do postaci:

$$\sum_{i=1}^n (x_i)^2 - n \bar{x}^2$$

Po podstawieniu do wzoru (11.9) danych z Tabeli 11.1 otrzymujemy  $r = 0,9879$ . Zmienne  $x$  i  $y$  są zatem dobrze skorelowane i założenie, że są powiązane zależnością liniową, jest uzasadnione.

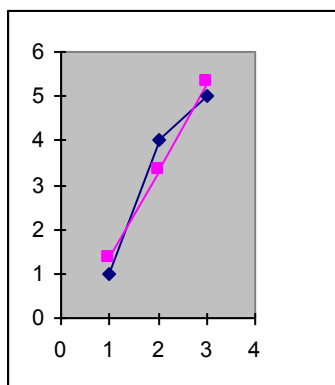
Omówione wyżej parametry regresji liniowej (oprócz samych współczynników  $a_0$  oraz  $a_1$ ) nazywa się **statystykami**

## Regresja liniowa w Excelu

Excel wyposażony jest w narzędzia i funkcje statystyczne służące do obliczeń związanych z regresją liniową i wykładniczą. Jeden ze sposobów szybkiego uzyskania wartości  $y$  wynikających z regresji liniowej lub wykładniczej polega na wybraniu obszaru komórek zawierających dane  $y$  ewentualnie plus dodatkowe komórki i polecenia *Edycja / Wypełnij / Serie danych...*, *Trend*, *Wiersze* lub *Kolumny*, *Arytmetyczny* lub *Geometryczny*. Program zamienia dane na wartości  $y$  leżące na prostej a puste komórki wypełnia następnymi wartościami  $y$  leżącymi na tej prostej. Oto tabelka i wykres zrobiony w oparciu o tę tabelkę (I6:K8) dla regresji liniowej: dane są wpisane w komórki I6:K7. Obok (M6:Q8) pokazano te same obliczenia wykonane przy pomocy funkcji statystycznych:

	I	J	K	L	M	N	O	P	Q
6	1	2	3		$y = a x + b$				
7	1	4	5		a	b	1,333333	3,333333	5,333333
8	1,333333	3,333333	5,333333		2	-0,666667	1,333333	3,333333	5,333333

	M	N	O	P	Q
6	$y = a x + b$				
7	a	b	$=\$M\$8*\$I6+\$N\$8$	$=\$M\$8*\$J6+\$N\$8$	$=\$M\$8*\$K6+\$N\$8$
8	$=\text{REGLINP}(I7:K7;I6:K6)$	$=\text{REGLINP}(I7:K7;I6:K6)$	$=\text{REGLINW}(I7:K7;I6:K6)$	$=\text{REGLINW}(I7:K7;I6:K6)$	$=\text{REGLINW}(I7:K7;I6:K6)$



W komórkach M8:N8 zastosowano formułę tablicową REGLINP, a w komórkach O8:Q8 formułę tablicową REGLINW. W Excelu mamy do regresji liniowej jeszcze jedną formułę tablicową REGLINX, oraz następujące formuły działające na jednej komórce: REGBŁSTD (liczy  $s$ ), NACHYLENIE (liczy  $a_0$ ), ODCIĘTA (liczy  $a_1$ ), WSP.KORELACJI (liczy  $r$ ), PEARSON (liczy  $r$ ), R.KWADRAT (liczy  $r^2$ ), oraz KOWARIANCJA (liczy  $s_{xy}$ ). Następnym przykładem pokazujemy zastosowanie niektórych z nich. Funkcja REGLINP może liczyć aż 10 parametrów, w przykładzie wybrano tylko 6.

	A	B	C	D	E	F	G	H
1	$i$	$x_i$	$y_i$	$x_i^2$	$y_i^2$	$x_i y_i$	$p_1(x_i)$	$\varepsilon_i^2$
2	1	2	2,5	4	6,25	5	-0,1	6,76
3	2	4	10	16	100	40	14,4	19,36
4	3	6	32	36	1024	192	28,9	9,61
5	4	8	40	64	1600	320	43,4	11,56
6	5	10	60	100	3600	600	57,9	4,41
7	Sumy:	30	144,5	220	6330,25	1157		51,7
8								
9	$a_0$ =	-14,6	NACHYLENIE=		7,25			
10	$a_1$ =	7,25	ODCIĘTA=		-14,6			
11	$s$ =	4,151305	REGBŁSTD=		4,151305			
12	$s_{a0}$ =	4,353925	PEARSON=		0,987927			
13	$s_{a1}$ =	0,656379						
14	$r$ =	0,987927		7,25	-14,6			
15				0,656379	4,353925			
16				0,976000	4,151305			
17								

Arkusz 11.1

W kolumnach B i C Arkusza 11.1 znajdują się dane z Tabeli 11.1, w kolumnach D, E, F obliczono odpowiednio kwadraty i iloczyny  $x_i$  oraz  $y_i$ . W kolumnie G mamy wartości  $y$  obliczone z równania prostej dla kolejnych  $x_i$ , a w kolumnie H kwadraty  $\varepsilon_i$  (patrz wzór (11.7)). W kolumnie B od wiersza 8 do 13 są współczynniki prostej oraz część statystyk regresji obliczone na podstawie wzorów (11.6 – 11.9). W komórkach od ósmej do jedenastej kolumny E obliczono ponownie część tych wielkości z użyciem odpowiednich funkcji Excela. Wyróżniona w komórkach D7:E15 tabelka jest macierzą zawierającą wyniki działania funkcji REGLINP. Jak zawsze, gdy mamy do czynienia z funkcją dającą wynik w postaci macierzy, najpierw należy zaznaczyć komórki, w których mają się znaleźć wyniki, wywołać funkcję, zadać jej parametry, a następnie nacisnąć jednocześnie klawisze Shift, Ctrl oraz Enter. Funkcja REGLINP oblicza więcej statystyk niż pokazano w Arkuszu 11.1, ale na tym etapie ograniczyliśmy się tylko do tych, które omówiono wyżej. Należy również dodać, że w



wyniku działania tej funkcji otrzymujemy kwadrat współczynnika korelacji liniowej, a nie sam współczynnik ( $0,987927^2 = 0,9760$ ).

Regresję liniową stosuje się również do zależności wyrażonych funkcją wykładniczą, ponieważ logarytmując równanie wykładnicze otrzymujemy równanie prostej. Jest to ważne zastosowanie regresji liniowej, bowiem zależności wykładnicze są powszechne w fizyce, chemii i biologii. Pokażemy jak to się robi na przykładzie rozpadu jąder promieniotwórczych (Rozdział 9). Rozwiązanie równania opisującego rozpad można zapisać jako:

$$N(t) = N_0 e^{-t/\tau}$$

gdzie  $\tau$  jest średnim czasem życia związanym z okresem połowicznego zaniku relacją:

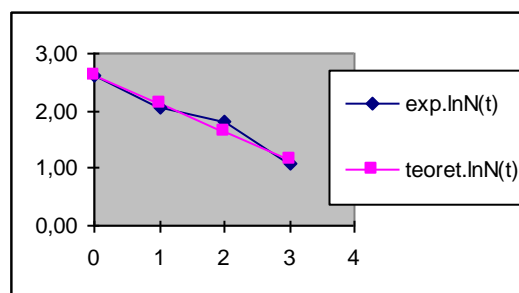
$$T_{1/2} = 0,693 \tau$$

Po zlogarytmowaniu mamy:

$$\ln N(t) = \ln N_0 - \frac{t}{\tau}$$

gdzie zmienną jest oczywiście  $t$ , czyli czas. Wobec tego parametrami prostej będą  $\ln N_0$  oraz  $-1/\tau$ . W tabelce zamieszczamy odpowiednie obliczenia dla danych umieszczonych w dwóch pierwszych kolumnach. Są to: czas np. w godzinach oraz szybkość zliczania rozpadów w jednostkach umownych. Wykres zrobiony jest w oparciu o zaznaczone kolumny. W zaznaczonej na szaro komórce znajduje się wyliczone najlepsze przybliżenie średniego czasu życia wynikające z metody regresji liniowej.

	A	B	C	D
26	t	N(t)	exp.lnN(t)	teoret.lnN(t)
27	0	13,8	2,62	2,63
28	1	7,9	2,07	2,14
29	2	6,1	1,81	1,64
30	3	2,9	1,06	1,15
31				
32			-0,49	2,63
33			2,02	13,90



	A	B	C	D
26	t	N(t)	exp.lnN(t)	teoret.lnN(t)
27	0	13,8	=LN(B27)	=REGLINW(C27:C30;A27:A30;A27:A30)
28	1	7,9	=LN(B28)	=REGLINW(C27:C30;A27:A30;A27:A30)
29	2	6,1	=LN(B29)	=REGLINW(C27:C30;A27:A30;A27:A30)
30	3	2,9	=LN(B30)	=REGLINW(C27:C30;A27:A30;A27:A30)
31				
32			=REGLINP(C27:C30;A27:A30)	=REGLINP(C27:C30;A27:A30)
33			=-1/C32	=EXP(D32)

## Aproksymacja wielomianami wyższych stopni.

Oczywiście wiele zależności spotykanych w przyrodzie nie ma charakteru liniowego. Staramy się wówczas znaleźć wielomian odpowiedniego stopnia, który przebiegałby możliwie blisko wszystkich punktów, które chcemy przybliżyć (doświadczalnych lub otrzymanych z jakiejś bardziej skomplikowanej funkcji).

Dla uproszczenia założmy, że mamy wielomian drugiego stopnia i wykonajmy formalnie wszystkie etapy metody najmniejszych kwadratów.

$$p_2(x) = a_0 + a_1x + a_2x^2 = y(x)$$

Funkcja  $S(a_0, a_1, a_2)$  dla  $\{(x_i, y_i) [i = 1, 2, \dots, n]\}$  danych ma teraz postać:

$$S(a_0, a_1, a_2) = \sum_{i=1}^n (y_i - a_0 - a_1x_i - a_2x_i^2)^2$$

a układ równań normalnych wygląda następująco:

$$\begin{aligned}\frac{\partial S(a_0, a_1, a_2)}{\partial a_0} &= 2 \sum_{i=1}^n (y_i - a_0 - a_1x_i - a_2x_i^2)(-1) = 0 \\ \frac{\partial S(a_0, a_1, a_2)}{\partial a_1} &= 2 \sum_{i=1}^n (y_i - a_0 - a_1x_i - a_2x_i^2)(-x_i) = 0 \\ \frac{\partial S(a_0, a_1, a_2)}{\partial a_2} &= 2 \sum_{i=1}^n (y_i - a_0 - a_1x_i - a_2x_i^2)(-x_i^2) = 0\end{aligned}\tag{11.10}$$

Po uporządkowaniu, otrzymamy:

$$\begin{aligned}a_0n + a_1 \sum_{i=1}^n x_i + a_2 \sum_{i=1}^n x_i^2 &= \sum_{i=1}^n y_i \\ a_0 \sum_{i=1}^n x_i + a_1 \sum_{i=1}^n x_i^2 + a_2 \sum_{i=1}^n x_i^3 &= \sum_{i=1}^n x_i y_i \\ a_0 \sum_{i=1}^n x_i^2 + a_1 \sum_{i=1}^n x_i^3 + a_2 \sum_{i=1}^n x_i^4 &= \sum_{i=1}^n x_i^2 y_i\end{aligned}\tag{11.11}$$

Tak wygląda układ równań normalnych w wypadku, gdy wielomian aproksymacyjny jest stopnia drugiego. Dla stopnia trzeciego otrzymamy układ:

$$\begin{aligned}a_0n + a_1 \sum_{i=1}^n x_i + a_2 \sum_{i=1}^n x_i^2 + a_3 \sum_{i=1}^n x_i^3 &= \sum_{i=1}^n y_i \\ a_0 \sum_{i=1}^n x_i + a_1 \sum_{i=1}^n x_i^2 + a_2 \sum_{i=1}^n x_i^3 + a_3 \sum_{i=1}^n x_i^4 &= \sum_{i=1}^n x_i y_i \\ a_0 \sum_{i=1}^n x_i^2 + a_1 \sum_{i=1}^n x_i^3 + a_2 \sum_{i=1}^n x_i^4 + a_3 \sum_{i=1}^n x_i^5 &= \sum_{i=1}^n x_i^2 y_i\end{aligned}\tag{11.12}$$

$$a_0 \sum_{i=1}^n x_i^3 + a_1 \sum_{i=1}^n x_i^4 + a_2 \sum_{i=1}^n x_i^5 + a_3 \sum_{i=1}^n x_i^6 = \sum_{i=1}^n x_i^3 y_i$$

Układy równań normalnych odpowiadające wielomianom aproksymacyjnym wyższych stopni rozbudowuje się w sposób analogiczny.

Układy równań normalnych są układami równań liniowych, w których niewiadomymi są  $a_0, a_1, a_2, \dots$ , a sumy różnych potęg  $x_i^k$  oraz iloczynów  $x_i^k y_i$  grają rolę wolnych wyrazów oraz współczynników przy niewiadomych. Można to zapisać w postaci macierzowej jako:

$$\mathbf{X} \cdot \mathbf{A} = \mathbf{Y} \quad (11.13)$$

Postać macierzy  $\mathbf{X}$ ,  $\mathbf{A}$  oraz  $\mathbf{Y}$  jest określona przez układ równań normalnych dla danego problemu i na przykład w przypadku wielomianu drugiego stopnia wyglądają one następująco:

$$\mathbf{X} = \begin{bmatrix} n & \sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 \\ \sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i^3 \\ \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i^3 & \sum_{i=1}^n x_i^4 \end{bmatrix} \quad \mathbf{A} = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} \quad \mathbf{Y} = \begin{bmatrix} \sum_{i=1}^n y_i \\ \sum_{i=1}^n x_i y_i \\ \sum_{i=1}^n x_i^2 y_i \end{bmatrix} \quad (11.14)$$

Rozwiązanie tego układu można uzyskać, mnożąc z lewej strony obie strony równania (11.13) razy macierz odwrotną  $\mathbf{X}^{-1}$ :

$$\mathbf{A} = \mathbf{X}^{-1} \cdot \mathbf{Y} \quad (11.15)$$

### Przykład 11.1

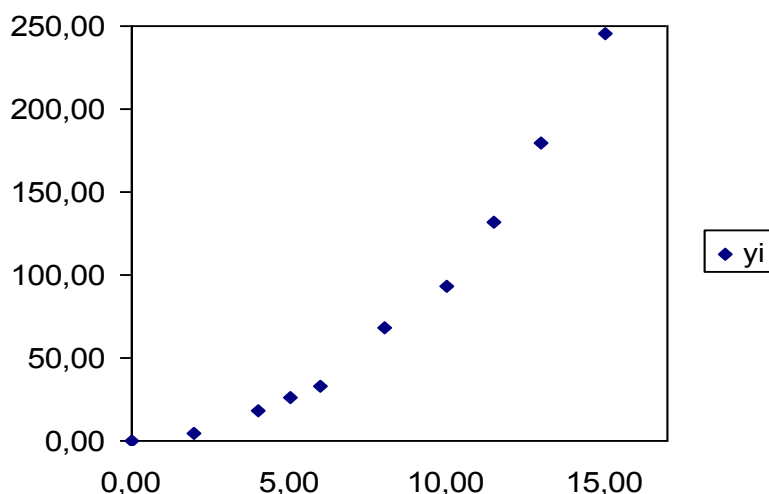
Znaleźć wielomian aproksymujący punkty zapisane w Tabeli 11.2.

Tabela 11.2

$i$	$x_i$	$y_i$
1	0,00	0,30
2	2,00	4,50
3	4,00	18,00
4	5,00	26,00
5	6,00	33,00
6	8,00	68,00
7	10,00	93,00
8	11,50	132,00

9	13,00	179,00
10	15,00	245,00

Z Rysunku 11.2, przedstawiającego wykres tych punktów, wynika, że mamy tu do czynienia z zależnością typu kwadratowego, a zatem funkcją, która najlepiej nadaje się do aproksymacji, będzie wielomian stopnia przynajmniej drugiego.



Rysunek 11.2

Stwórzmy teraz arkusz, który będzie rozwiązywał problem znalezienia współczynników wielomianu aproksymującego drugiego stopnia.

	A	B	C	D	E	F	G	H	I	J
1	Aproksymacja funkcji danej w Tabeli 10.2 wielomianem stopnia drugiego									
2	n= 10		m= 2							
3	$i$	$x_i$	$y_i$	$x_i^2$	$x_i^3$	$x_i^4$	$x_i y_i$	$x_i^2 y_i$	$p_2(x_i)$	$\varepsilon_i^2$
4	1	0,00	0,30	0,00	0,00	0,00	0,00	0,00	3,8923	12,9046
5	2	2,00	4,50	4,00	8,00	16,00	9,00	18,00	4,1465	0,1250
6	3	4,00	18,00	16,00	64,00	256,00	72,00	288,00	14,0483	15,6156
7	4	5,00	26,00	25,00	125,00	625,00	130,00	650,00	22,6172	11,4436
8	5	6,00	33,00	36,00	216,00	1296,00	198,00	1188,00	33,5979	0,3575
9	6	8,00	68,00	64,00	512,00	4096,00	544,00	4352,00	62,7951	27,0908
10	7	10,00	93,00	100,00	1000,00	10000,00	930,00	9300,00	101,6400	74,6501
11	8	11,50	132,00	132,25	1520,88	17490,06	1518,00	17457,00	137,1050	26,0610
12	9	13,00	179,00	169,00	2197,00	28561,00	2327,00	30251,00	177,9968	1,0064
13	10	15,00	245,00	225,00	3375,00	50625,00	3675,00	55125,00	240,9609	16,3143
14	Sumy=	74,50	798,80	771,25	9017,88	112965,06	9403,00	118629,00		185,5688
15										
16				10	74,50	771,25			798,80	
17			$X=$	74,50	771,25	9017,88		$Y=$	9403,00	
18				771,25	9017,88	112965,06			118629,00	
19										
20				0,676377	-0,170295	0,008977			3,892295	
21			$X^{-1}=$	-0,170295	0,062345	-0,003814		$A=X^{-1}Y=$	-2,284829	
22				0,008977	-0,003814	0,000252			1,205960	
23										
24						$s=$	5,1487702			

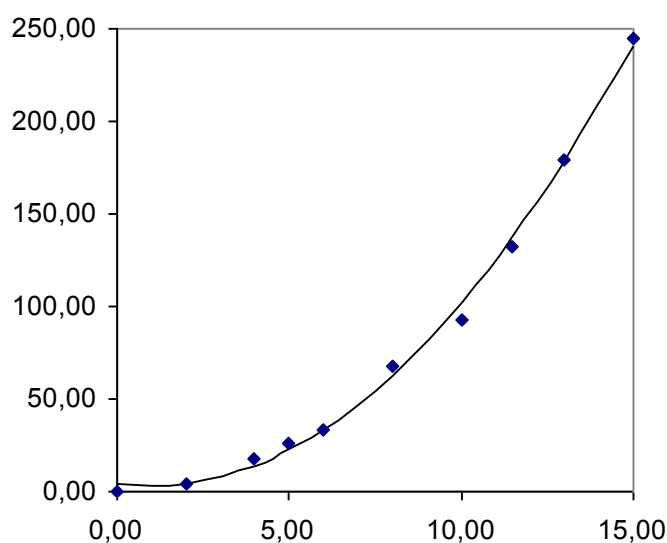
Arkusz 11.2

W wierszach od czwartego do trzynastego Arkusza 11.2 i kolumnach od B do H są współrzędne punktów i ich różne potęgi oraz iloczyny, które zgodnie ze wzorem (11.14) służą do zbudowania macierzy  $X$  oraz  $Y$ . W Excelu jest funkcja odwracająca macierze o nazwie MACIERZ.ODW, która została tu wykorzystana do otrzymania  $X^{-1}$ . Jak zwykle, kiedy używa się funkcji dającej w wyniku macierz, należy najpierw zaznaczyć komórki, w których ma zostać zapisany rezultat, wywołać odpowiednią funkcję ( w tym wypadku MACIERZ.ODW), zadać jej parametry i nacisnąć jednocześnie klawisze Ctrl, Shift oraz Enter. W następnym kroku oblicza się (zgodnie ze wzorem (11.15)) iloczyn macierzy  $X^{-1}$  oraz  $Y$ . Tym razem także można się posłużyć funkcją Excela o nazwie MACIERZ.ILOCZYN. W rezultacie tego mnożenia otrzymuje się macierz A, której elementami są współczynniki wielomianu aproksymującego  $p_2(x)$ , a mianowicie  $a_0, a_1, a_2$ . W kolumnie I umieszczono wartości tego wielomianu obliczone dla wszystkich wartości  $x_i$ . W ostatniej kolumnie J mamy kwadraty wielkości  $\varepsilon_i$ , czyli różnic między prawdziwymi wartościami  $y_i$  a wartościami wielomianu aproksymującego. Służą one do obliczenia średniego odchylenia od krzywej, które w ogólnym wypadku (porównaj ze wzorem (11.7)) definiuje się następująco:

$$s = \sqrt{\frac{\sum_{i=1}^n \left( y_i - \sum_{l=0}^m a_l x_i^l \right)^2}{n - m - 1}} = \sqrt{\frac{\sum_{i=1}^n \varepsilon_i^2}{n - m - 1}} \quad (11.16)$$

gdzie  $n$  to liczba punktów aproksymowanych, a  $m$  to stopień wielomianu aproksymującego.

Na Rysunku 11.3 przedstawiono wykresy paraboli określonej przez tak obliczone w Arkuszu 11.2 współczynniki i punktów z Rysunku 11.2.



Rysunek 11.3

W sytuacji, gdy chcemy zastosować w aproksymacji wielomiany wyższych od drugiego stopni, postępujemy analogicznie. Najpierw zgodnie ze schematem omówionym podczas dyskusji wzorów (11.11) i (11.12) rozbudowujemy macierze  $X$  oraz  $Y$ , a potem powtarzamy obliczenia przedstawione w Arkuszu 11.2.

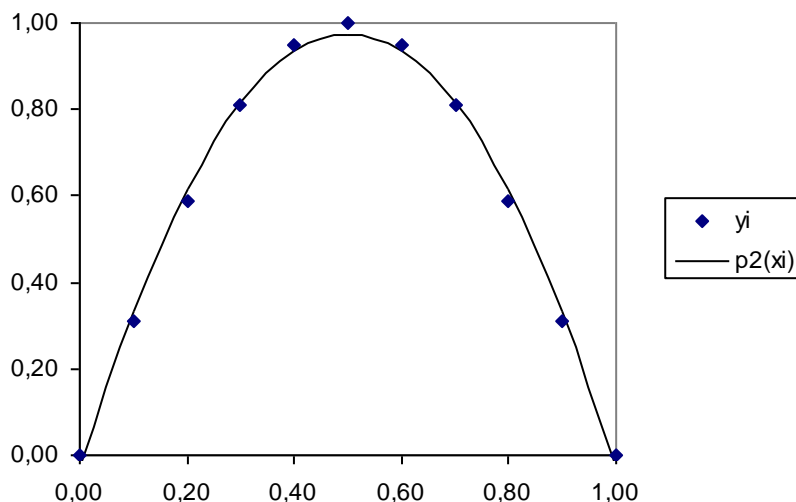
### Przykład 11.2

Wykorzystując Arkusz 11.2 obliczyć współczynniki wielomianu drugiego stopnia aproksymującego funkcję  $f(x) = \sin(\pi x)$  w przedziale  $0 \in [0,1]$ . Tabelaizując funkcję  $f(x)$ , przyjąć krok  $h = 0,1$ .

	A	B	C	D	E	F	G	H	I	J
1	Aproksymacja wielomianem stopnia $k$ funkcji $f(x)=\sin(\pi x)$									
2	n=	11	k=		2					
3	$i$	$x_i$	$y_i$	$x_i^2$	$x_i^3$	$x_i^4$	$x_i y_i$	$x_i^2 y_i$	$p_2(x_i)$	$\varepsilon_i^2$
4	1	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	-0,02558	0,00065
5	2	0,10000	0,30902	0,01000	0,00100	0,00010	0,03090	0,00309	0,33416	0,00063
6	3	0,20000	0,58779	0,04000	0,00800	0,00160	0,11756	0,02351	0,61395	0,00068
7	4	0,30000	0,80902	0,09000	0,02700	0,00810	0,24271	0,07281	0,81380	0,00002
8	5	0,40000	0,95106	0,16000	0,06400	0,02560	0,38042	0,15217	0,93371	0,00030
9	6	0,50000	1,00000	0,25000	0,12500	0,06250	0,50000	0,25000	0,97368	0,00069
10	7	0,60000	0,95106	0,36000	0,21600	0,12960	0,57063	0,34238	0,93371	0,00030
11	8	0,70000	0,80902	0,49000	0,34300	0,24010	0,56631	0,39642	0,81380	0,00002
12	9	0,80000	0,58779	0,64000	0,51200	0,40960	0,47023	0,37618	0,61395	0,00068
13	10	0,90000	0,30902	0,81000	0,72900	0,65610	0,27812	0,25030	0,33416	0,00063
14	11	1,00000	0,00000	1,00000	1,00000	1,00000	0,00000	0,00000	-0,02558	0,00065
15	Sumy=	5,50000	6,31375	3,85000	3,02500	2,53330	3,15688	1,86687		0,00528
16										
17			11	5,50	3,85			6,31		
18		$X=$	5,50	3,85	3,03		$Y=$	3,16		
19			3,85	3,03	2,53			1,87		
20										
21			0,5804	-2,2028	1,7483			-0,0256		
22		$X^{-1}=$	-2,2028	12,5641	-11,6550		$A=X^{-1}Y=$	3,9970		
23			1,7483	-11,6550	11,6550			-3,9970		
24										
25				s=	0,025694					

Arkusz 11.3

Na Rysunku 11.4 przedstawiono wykres punktów otrzymanych w wyniku tabelaizacji funkcji  $f(x) = \sin(\pi x)$  oraz wykres wielomianu, którego współczynniki obliczono w Arkuszu 11.3.



Rysunek 11.4

## Jakobiany

Problem poszukiwania współczynników wielomianu aproksymacyjnego można uprościć, wprowadzając tzw. jacobiany. Zanim to zrobimy, napiszmy najpierw wielomian aproksymacyjny stopnia  $m$ :

$$y(x) = p_m(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + \dots + a_mx^m \quad (11.17)$$

Jeżeli teraz policzymy pochodne tak zapisanego wielomianu po wszystkich współczynnikach (traktowanych jako zmienne), otrzymamy:

$$\frac{\partial y(x)}{\partial a_0} = 1; \quad \frac{\partial y(x)}{\partial a_1} = x; \quad \frac{\partial y(x)}{\partial a_2} = x^2; \quad \dots \quad \frac{\partial y(x)}{\partial a_m} = x^m$$

Te pochodne (funkcje) przyjmują dla konkretnych  $x_i$  (współrzędnych  $x$  punktów, które chcemy aproksymować) wartości:

$$\frac{\partial y(x_i)}{\partial a_0} = 1; \quad \frac{\partial y(x_i)}{\partial a_1} = x_i; \quad \frac{\partial y(x_i)}{\partial a_2} = x_i^2; \quad \dots \quad \frac{\partial y(x_i)}{\partial a_m} = x_i^m$$

Zbudujmy teraz macierz, której elementy można zapisać ogólnie w postaci:  $\frac{\partial y(x_i)}{\partial a_j}$ , gdzie

wskaźnik  $i$ , numerujący wiersze, zmienia się od 1 do  $n$  (liczba punktów aproksymowanych), a  $j$ , numerujący kolumny, zmienia się od 0 do  $m$ .

	$\frac{\partial y(x_i)}{\partial a_0}$	$\frac{\partial y(x_i)}{\partial a_1}$	$\frac{\partial y(x_i)}{\partial a_2}$	...	$\frac{\partial y(x_i)}{\partial a_m}$
$x_1$	1	$x_1$	$x_1^2$	...	$x_1^m$
$x_2$	1	$x_2$	$x_2^2$	...	$x_2^m$
$x_3$	1	$x_3$	$x_3^2$	...	$x_3^m$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$x_n$	1	$x_n$	$x_n^2$	...	$x_n^m$

(11.18)

Elementy tabeli (11.14) tworzą macierz zwaną jacobianem i oznaczaną literą **J**. Dla przykładu jacobian odpowiadający problemowi rozwiązanemu w Arkuszu 11.2 (wielomian drugiego stopnia aproksymujący 10 punktów ma postać macierzy o wymiarach 10 na 3:

$$\mathbf{J} = \begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \\ 1 & x_4 & x_4^2 \\ 1 & x_5 & x_5^2 \\ 1 & x_6 & x_6^2 \\ 1 & x_7 & x_7^2 \\ 1 & x_8 & x_8^2 \\ 1 & x_9 & x_9^2 \\ 1 & x_{10} & x_{10}^2 \end{bmatrix}$$

Łatwo teraz zauważyć, że po pomnożeniu jacobianu transponowanego razy jacobian otrzymamy macierz **X** ze wzoru (11.14)

$$\mathbf{J}^T \mathbf{J} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & x_7 & x_8 & x_9 & x_{10} \\ x_1^2 & x_2^2 & x_3^2 & x_4^2 & x_5^2 & x_6^2 & x_7^2 & x_8^2 & x_9^2 & x_{10}^2 \end{bmatrix} * \begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \\ 1 & x_4 & x_4^2 \\ 1 & x_5 & x_5^2 \\ 1 & x_6 & x_6^2 \\ 1 & x_7 & x_7^2 \\ 1 & x_8 & x_8^2 \\ 1 & x_9 & x_9^2 \\ 1 & x_{10} & x_{10}^2 \end{bmatrix} = \begin{bmatrix} 10 & \sum_{i=1}^{10} x_i & \sum_{i=1}^{10} x_i^2 \\ \sum_{i=1}^{10} x_i & \sum_{i=1}^{10} x_i^2 & \sum_{i=1}^{10} x_i^3 \\ \sum_{i=1}^{10} x_i^2 & \sum_{i=1}^{10} x_i^3 & \sum_{i=1}^{10} x_i^4 \end{bmatrix}$$

Dzięki wprowadzeniu jacobianów możemy zatem znacznie uprościć tworzenie macierzy **X**, co było najbardziej pracochłonną częścią przedstawionego w Arkuszu 11.2 sposobu



rozwiązywania problemu aproksymacji wielomianem drugiego stopnia. Zmodyfikowany w ten sposób Arkusz 11.2 wygląda następująco:

	A	B	C	D	E	F	G	H	I	J
1	Aproksymacja funkcji danej w Tabeli 10.2 wielomianem stopnia drugiego									
2	n= 10		m= 2							
3			$x_i$	$x_i^2$		$y_i$	$x_i y_i$	$x_i^2 y_i$	$p_2(x_i)$	$\varepsilon_i^2$
4		1	0,00	0,00		0,30	0,00	0,00	3,8923	12,9046
5		1	2,00	4,00		4,50	9,00	18,00	4,1465	0,1250
6		1	4,00	16,00		18,00	72,00	288,00	14,0483	15,6156
7		1	5,00	25,00		26,00	130,00	650,00	22,6172	11,4436
8	<b>J=</b>	1	6,00	36,00		33,00	198,00	1188,00	33,5979	0,3575
9		1	8,00	64,00		68,00	544,00	4352,00	62,7951	27,0908
10		1	10,00	100,00		93,00	930,00	9300,00	101,6400	74,6501
11		1	11,50	132,25		132,00	1518,00	17457,00	137,1050	26,0610
12		1	13,00	169,00		179,00	2327,00	30251,00	177,9968	1,0064
13		1	15,00	225,00		245,00	3675,00	55125,00	240,9609	16,3143
14					Sumy=	798,80	9403,00	118629,00		185,5688
15										
16										
17			10	74,50	771,25			798,80		
18		<b>X=J<sup>T</sup>J=</b>	74,50	771,25	9017,88		<b>Y=</b>	9403,00		
19			771,25	9017,88	112965,06			118629,00		
20										
21			0,676377	-0,170295	0,008977			3,892295		
22		<b>X<sup>-1</sup>=</b>	-0,170295	0,062345	-0,003814		<b>A=X<sup>-1</sup>Y=</b>	-2,284829		
23			0,008977	-0,003814	0,000252			1,205960		
24										
25					s=	5,1487702				

Arkusz 11.4

Widać, że w Arkuszu 11.2 i Arkuszu 11.4 otrzymaliśmy identyczne wyniki, tyle że w Arkuszu 11.3 udało nam się tego dokonać mniejszym kosztem. Po pierwsze nie było potrzeby tworzenia kolumn z  $x_i^3$  i  $x_i^4$ , ale jeszcze ważniejsze jest to, że macierzy  $X$  nie tworzyliśmy tu „ręcznie”, wpisując do niej sumy odpowiednich potęg  $x_i$ , lecz obliczając ją jako  $J^T J$  za pomocą jednej formuły excelowskiej:

=MACIERZ.ILOCZYN (TRANSPONUJ(B4:D13);B4:D13)

Korzyści ze stosowania jacobianów są jeszcze bardziej widoczne, kiedy wielomian aproksymacyjny jest wyższego stopnia niż drugi.

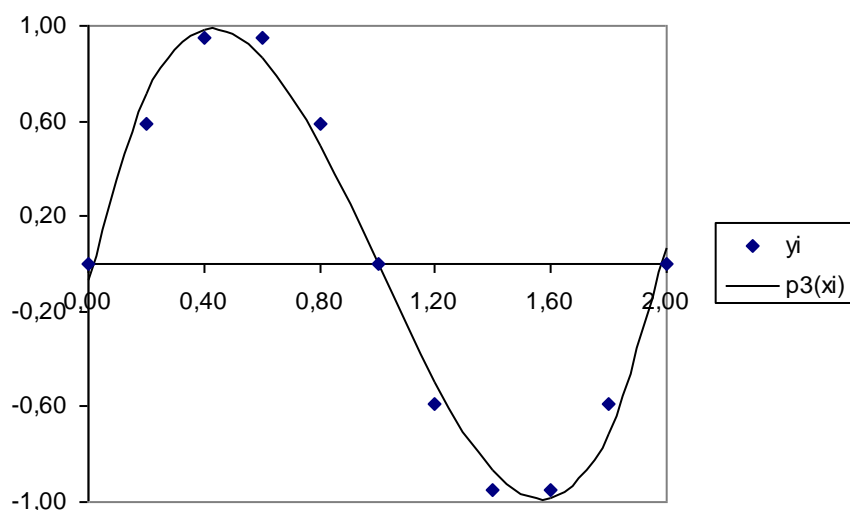
### Przykład 11.3

Stworzyć arkusz, który będzie obliczał współczynniki wielomianu trzeciego stopnia aproksymującego funkcję  $f(x) = \sin(\pi x)$  w przedziale  $0 \in [0,2]$ . Tabelaryzując funkcję  $f(x)$ , przyjęć krok  $h = 0,2$ .

	A	B	C	D	E	F	G	H	I	J	K	L	
1	Aproksymacja wielomianem stopnia trzeciego funkcji $f(x)=\sin(\pi x)$												
2	n=	11	m=			3							
3			$x_i$	$x_i^2$	$x_i^3$		$y_i$	$x_i y_i$	$x_i^2 y_i$	$x_i^3 y_i$	$p_3(x_i)$	$\epsilon_i^2$	
4		1	0,00000	0,00000	0,00000		0,00000	0,00000	0,00000	0	-0,06351	0,00403	
5		1	0,20000	0,04000	0,00800		0,58779	0,11756	0,02351	0,0047023	0,71218	0,01547	
6		1	0,40000	0,16000	0,06400		0,95106	0,38042	0,15217	0,0608676	0,97920	0,00079	
7		1	0,60000	0,36000	0,21600		0,95106	0,57063	0,34238	0,2054282	0,86474	0,00745	
8	J=	1	0,80000	0,64000	0,51200		0,58779	0,47023	0,37618	0,300946	0,49595	0,00843	
9		1	1,00000	1,00000	1,00000		0,00000	0,00000	0,00000	1,225E-16	0,00000	0,00000	
10		1	1,20000	1,44000	1,72800		-0,58779	-0,70534	-0,84641	-1,015693	-0,49595	0,00843	
11		1	1,40000	1,96000	2,74400		-0,95106	-1,33148	-1,86407	-2,609699	-0,86474	0,00745	
12		1	1,60000	2,56000	4,09600		-0,95106	-1,52169	-2,43470	-3,895527	-0,97920	0,00079	
13		1	1,80000	3,24000	5,83200		-0,58779	-1,05801	-1,90442	-3,427964	-0,71218	0,01547	
14		1	2,00000	4,00000	8,00000		0,00000	0,00000	0,00000	-9,07E-15	0,06351	0,00403	
15						Sumy=	0,00000	-3,07768	-6,15537	-10,37694		0,07236	
16													
17		11,00000	11,00000	15,40000	24,20000			0,00000					
18	X=	11,00000	15,40000	24,20000	40,53280		Y=	-3,07768					
19		15,40000	24,20000	40,53280	70,66400			-6,15537					
20		24,20000	40,53280	70,66400	126,61792			-10,37694					
21										s=	0,10167		
22		0,79021	-2,76807	2,62238	-0,72844			-0,06351					
24	X <sup>-1</sup> =	-2,76807	16,38177	-18,81799	5,78704		A=X <sup>-1</sup> Y=	5,36200					
25		2,62238	-18,81799	23,49213	-7,58790			-7,94773					
26		-0,72844	5,78704	-7,58790	2,52930			2,64924					

#### Arkusz 11.4

Na Rysunku 11.5 przedstawiono wykres punktów otrzymanych w wyniku tabelaryzacji funkcji  $f(x) = \sin(\pi x)$  w przedziale  $[0,1]$  oraz wykres wielomianu, którego współczynniki obliczono w Arkuszu 11.4.



Rysunek 11.5

## Interpolacja a aproksymacja

Na interpolację można spojrzeć jak na wyznaczanie współczynników wielomianu na podstawie reprezentacji tego wielomianu przez wartości w punktach, które nazywaliśmy węzłami interpolacji. Dla  $n$  węzłów można wyznaczyć wielomian stopnia  $n-1$  (ściślej: wielomian o ograniczeniu stopnia  $n$ ). Jednoznaczność tak określonego wielomianu interpolacyjnego można matematycznie udowodnić [2]. Pamiętajmy, że wielomian interpolacyjny musi spełniać warunki interpolacji, czyli przyjmować w węzłach dokładnie takie wartości, jakie wynikają z danych punktów. Można to zapisać w eleganckiej postaci macierzowej w następujący sposób:

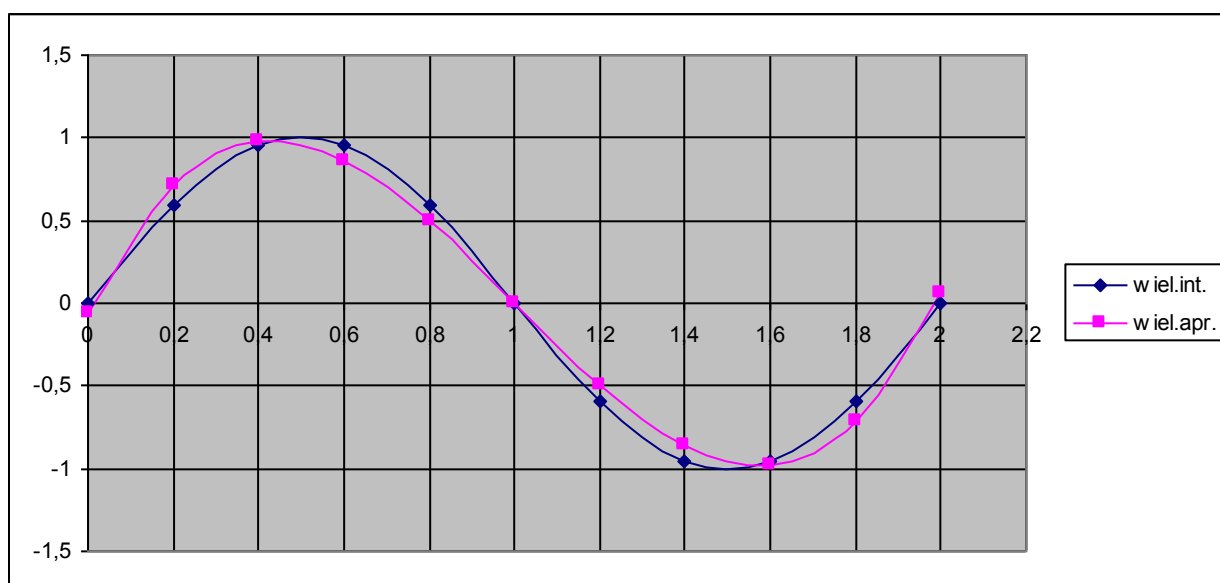
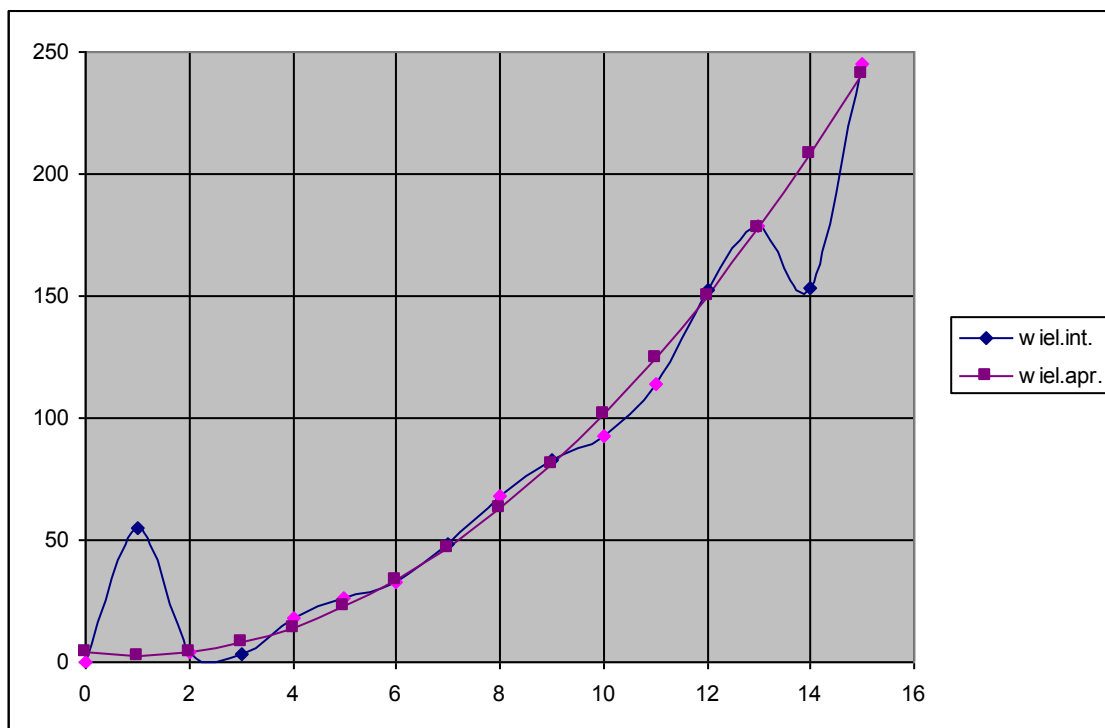
$$\begin{bmatrix} 1 & x_0 & x_0^2 & \cdots & x_0^{n-1} \\ 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n-1} & x_{n-1}^2 & \cdots & x_{n-1}^{n-1} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_{n-1} \end{bmatrix} = \begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_{n-1} \end{bmatrix}$$

$$\mathbf{V}\mathbf{A} = \mathbf{Y}$$

Pojawiła się w tym zapisie macierz, którą nazwaliśmy jacobianem. Jest to szczególnie przypadek jacobianu, macierz o wymiarach wynikających z liczby węzłów: dla  $n$  węzłów numerowanych od 0 do  $n-1$  (zachowaliśmy numerację od 0 ponieważ tak zwykliśmy numerować współczynniki wielomianu) wymiar tej macierzy jest  $n$  razy  $n+1$ . W literaturze nazywa się ją macierzą Vandermonde'a. Można przekonać się, że dla różnych węzłów jest ona macierzą nieosobliwą, a więc odwracalną. Wobec tego macierz współczynników wielomianu interpolacyjnego można wyznaczyć w następujący sposób:

$$\mathbf{A} = \mathbf{V}^{-1} \mathbf{Y}$$

Nie jest to najszybszy algorytm interpolacji, ale czasami jest wygodny i warto o nim pamiętać. Jeśli porównamy go z algorytmem aproksymacji wielomianowej z poprzedniego rozdziału, to zauważymy, że wielomian aproksymacyjny może być niższego stopnia niż interpolacyjny i przebiega pomiędzy węzłami w sposób wyznaczony przez metodę najmniejszych kwadratów, podczas gdy wielomian interpolacyjny przechodzi dokładnie przez wszystkie węzły interpolacji, ale za to ma często dość skomplikowaną postać. Spójrzmy jak wypadnie graficzne porównanie obu wielomianów dla obu badanych poprzednio przykładów.



## Literatura

1. John R. Taylor, *Wstęp do analizy błędów pomiarowych*, PWN Warszawa 1999.
2. Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, *Wprowadzenie do algorytmów*, WNT, 2000.