

Maciej Gonet

# ZROZUMIEĆ EXCELA

OBLICZENIA I WYKRESY



Helion 

Wszelkie prawa zastrzeżone. Nieautoryzowane rozpowszechnianie całości lub fragmentu niniejszej publikacji w jakiegokolwiek postaci jest zabronione. Wykonywanie kopii metodą kserograficzną, fotograficzną, a także kopiowanie książki na nośniku filmowym, magnetycznym lub innym powoduje naruszenie praw autorskich niniejszej publikacji.

Wszystkie znaki występujące w tekście są zastrzeżonymi znakami firmowymi bądź towarowymi ich właścicieli.

Autor oraz Wydawnictwo HELION dołożyli wszelkich starań, by zawarte w tej książce informacje były kompletne i rzetelne. Nie biorą jednak żadnej odpowiedzialności ani za ich wykorzystanie, ani za związane z tym ewentualne naruszenie praw patentowych lub autorskich. Autor oraz Wydawnictwo HELION nie ponoszą również żadnej odpowiedzialności za ewentualne szkody wynikłe z wykorzystania informacji zawartych w książce.

Redaktor prowadzący: Małgorzata Kulik

Projekt okładki: Studio Gravite / Olsztyn  
Obarek, Pokoński, Pazdrijowski, Zaprucki

Grafika na okładce została wykorzystana za zgodą Shutterstock.com

Wydawnictwo HELION  
ul. Kościuszki 1c, 44-100 GLIWICE  
tel. 32 231 22 19, 32 230 98 63  
e-mail: [helion@helion.pl](mailto:helion@helion.pl)  
WWW: <http://helion.pl> (księgarnia internetowa, katalog książek)

Drogi Czytelniku!

Jeżeli chcesz ocenić tę książkę, zajrzyj pod adres

<http://helion.pl/user/opinie/zrexow>

Możesz tam wpisać swoje uwagi, spostrzeżenia, recenzję.

ISBN: 978-83-283-5693-1

Copyright © Helion 2020

Printed in Poland.

- [Kup książkę](#)
- [Poleć książkę](#)
- [Oceń książkę](#)

- [Księgarnia internetowa](#)
- [Lubię to! » Nasza społeczność](#)

# Spis treści

<b>Wstęp .....</b>	<b>15</b>
<b>Rozdział 1. Graficzna prezentacja danych .....</b>	<b>23</b>
Rodzaje wykresów dostępnych w Excelu .....	23
<i>Wykres punktowy a wykres liniowy — podobieństwa i różnice .....</i>	<i>24</i>
Przygotowanie danych do wykresu punktowego .....	26
Wykonanie wykresu punktowego .....	26
<i>Zastosowanie skali logarytmicznej .....</i>	<i>28</i>
<i>Serie danych różniące się o rzędy wielkości na jednym wykresie .....</i>	<i>29</i>
<i>Modyfikacja zakresu danych wykresu .....</i>	<i>29</i>
<i>Bezpośrednie definiowanie zakresów danych wykresu .....</i>	<i>30</i>
<i>Wykorzystanie funkcji SERIE do modyfikacji danych wykresu .....</i>	<i>30</i>
<i>Punkty wykresu połączone linią .....</i>	<i>31</i>
<i>Wykorzystanie zakresów złożonych przy definiowaniu danych do wykresu .....</i>	<i>32</i>
<i>Niestandardowe użycie funkcji UDF przy definiowaniu danych do wykresu .....</i>	<i>36</i>
<i>Dynamiczne definiowanie zakresu danych do wykresu .....</i>	<i>36</i>
<i>Uwaga na błędy w przypadku, gdy dane do wykresu określają formuły .....</i>	<i>37</i>
<i>Wykresy a zmiana sposobu wyświetlania liczb .....</i>	<i>38</i>
Kopiowanie wykresów z arkusza do arkusza .....	38
Automatyczne skalowanie osi wykresu .....	40
Dodawanie słupków błędów do wykresu .....	41
<i>Sposób obliczania błędu i odchylenia standardowego .....</i>	<i>42</i>
<i>Wskazówki praktyczne .....</i>	<i>44</i>
<i>Wykresy niestandardowe wykonywane z wykorzystaniem słupków błędów .....</i>	<i>45</i>
Dodawanie linii trendu (regresji) do wykresu .....	47
Wykresy radarowe .....	50
Wykresy powierzchniowe .....	52
Małe wykresy przebiegu w czasie .....	54

<b>Rozdział 2. Wykresy punktowe w przykładach .....</b>	<b>57</b>
Typowy wykres funkcji opisanej wzorem .....	57
Użycie formuł nazwanych przy definiowaniu danych wykresu .....	58
Wykres dynamiczny prezentujący najnowsze dane .....	58
Dane wykresu określone formułami nazwanymi .....	59
<i>Funkcja ciągła</i> .....	59
<i>Funkcja nieciągła</i> .....	61
Wykres prezentujący wybrane serie danych .....	62
Stopniowe ujawnianie danych na wykresie .....	63
<i>Wyświetlanie kolejnych punktów w serii danych</i> .....	63
<i>Wyświetlanie kolejnych serii danych</i> .....	64
Wykres z zaznaczenia .....	64
Wykres prezentujący proste figury geometryczne .....	65
Wykres przedstawiający wielokąty foremne .....	65
Wykresy dynamiczne .....	68
Umieszczanie etykiet tekstowych przy punktach wykresu .....	70
Usuwanie szumów z danych przed wykonaniem wykresu .....	72
Cieniowanie obszaru między osią a wykresem lub między dwiema liniami na wykresie .....	74
Nierównomierna skala na osiach wykresu punktowego .....	76
<b>Rozdział 3. Zaokrąglanie liczb i obliczenia przybliżone .....</b>	<b>81</b>
Pojęcie błędu pomiaru .....	81
<i>Błąd bezwzględny i błąd względny</i> .....	81
<i>Błąd maksymalny wielkości zależnej</i> .....	82
<i>Błąd średni wielkości zależnej</i> .....	83
Przybliżenia dziesiętne .....	84
<i>Cyfry znaczące</i> .....	84
<i>Cyfry dokładne</i> .....	85
<i>Zasady zaokrąglania liczb</i> .....	86
Zasady obliczeń przybliżonych .....	86
<i>Dodawanie i odejmowanie</i> .....	87
<i>Mnożenie i dzielenie</i> .....	88
<i>Potęgowanie i pierwiastkowanie</i> .....	88
<i>Obliczanie logarytmu i antylogarytmu liczby (funkcji wykładniczej)</i> .....	88
<i>Rozwiązanie równania kwadratowego</i> .....	88
Rachunki przybliżone bez szacowania błędów .....	90
Dokładność obliczeń w Excelu i jej konsekwencje .....	90
Formatowanie komórek zawierających liczby a zaokrąglanie .....	92
Przegląd funkcji służących do zaokrąglania .....	94
<i>Funkcje zaokrągleń, które wymagają podania liczby cyfr po przecinku</i> .....	95
<i>Funkcje zaokrągleń, w których dokładność jest określona w sposób bezwzględny</i> .....	96

<i>Funkcja, w której dokładność zaokrąglenia jest określona przez kod formatu</i> .....	98
<i>Zaokrąglenie liczb z uwzględnieniem cyfr znaczących</i> .....	100
<i>Nietypowy sposób zaokrąglenia</i> .....	101
<i>Zaokrąglenie wartości czasu</i> .....	101
Wyodrębnianie części całkowitej i części ułamkowej liczby .....	102
<b>Rozdział 4. Różniczkowanie numeryczne</b> .....	<b>105</b>
Pojęcie pochodnej .....	105
Numeryczne obliczanie pochodnej funkcji danej wzorem .....	108
Numeryczne różniczkowanie danych pomiarowych .....	110
<i>Bezpośrednie oszacowanie wartości pochodnych na podstawie danych pomiarowych</i> .....	110
<i>Obliczenie pochodnej za pośrednictwem funkcji aproksymującej</i> .....	111
<b>Rozdział 5. Rozwiązywanie równań nieliniowych</b> .....	<b>115</b>
Równanie kwadratowe .....	115
Graficzne oszacowanie pierwiastków .....	118
Metoda iteracji prostej .....	119
Metoda siecznych (regula falsi) .....	123
Przyspieszenie zbieżności metody siecznych .....	125
Metoda stycznych (Newtona) .....	125
Zmodyfikowana metoda stycznych .....	126
Inne warianty metody siecznych .....	128
Zamierzone użycie odwołań cyklicznych przy rozwiązywaniu równań .....	129
Wykorzystanie narzędzi Szukaj wyniku i Solver .....	131
<i>Procedura Szukaj wyniku</i> .....	131
<i>Dodatek Solver</i> .....	133
Specyfika rozwiązywania równań wielomianowych .....	134
<i>Liczba pierwiastków</i> .....	135
<i>Rozkład na czynniki</i> .....	135
<i>Obliczanie wartości wielomianu</i> .....	135
<i>Wzory Viéte'a</i> .....	137
<i>Inne użyteczne obserwacje</i> .....	138
<i>Pochodna i całka wielomianu</i> .....	138
<i>Równania trzeciego stopnia</i> .....	140
<i>Równania czwartego stopnia</i> .....	142
<b>Rozdział 6. Układy równań liniowych i nieliniowych</b> .....	<b>145</b>
Zapis układu równań liniowych i jego rozwiązania w formie macierzowej .....	145
Przebieg rozwiązania w Excelu .....	146
Wykorzystanie Solvera do rozwiązywania układu równań liniowych .....	149
<i>Minimalizacja sumy kwadratów odchyleń</i> .....	149
<i>Zerowanie odchyleń w warunkach ograniczających</i> .....	150

Układy złożone z dwóch równań nieliniowych .....	151
<i>Wstępna analiza układu równań</i> .....	151
<i>Próby rozdzielenia zmiennych w równaniu z dwiema niewiadomymi</i> .....	152
<i>Procedura rozwiązywania układu równań</i> .....	153
Układy złożone z co najmniej trzech równań nieliniowych .....	160
<i>Rozwiązanie z wykorzystaniem Solvera</i> .....	160
<i>Metoda Newtona-Raphsona</i> .....	161
<b>Rozdział 7. Regresja liniowa jednowymiarowa .....</b>	<b>165</b>
Zakres zastosowania regresji liniowej .....	165
Ocena statystyczna danych pomiarowych i jakości dopasowania równania regresji .....	168
<i>Charakterystyka zbiorów danych pomiarowych</i> .....	168
<i>Ocena zależności między zbiorami danych pomiarowych</i> .....	169
<i>Zasady doboru współczynników w równaniu regresji</i> .....	171
Regresja liniowa w Excelu .....	172
<i>Znaczenie współczynników w równaniu regresji i sposoby ich obliczania</i> .....	172
<i>Równanie ze znanym wyrazem wolnym lub bez niego</i> .....	177
<i>Równanie ze znanym współczynnikiem kierunkowym</i> .....	180
<i>Specyfika działania funkcji REGLINP przy braku wyrazu wolnego</i> .....	181
<i>Nietypowe zastosowania funkcji REGLINP</i> .....	182
<i>Przykłady wyznaczania równań regresji</i> .....	185
<b>Rozdział 8. Regresja liniowa wielowymiarowa .....</b>	<b>193</b>
Regresja wielomianowa dla jednej zmiennej .....	193
<i>Funkcja REGLINW</i> .....	194
<i>Funkcja REGLINX</i> .....	195
Regresja liniowa w przypadku równania w postaci zlogarytmowanej .....	196
<i>Funkcja REGEXPP</i> .....	197
<i>Funkcja REGEXPW</i> .....	198
<i>Zależności typu potęgowego</i> .....	200
Regresja liniowa w przypadku wielomianu dwu zmiennych .....	202
<i>Przygotowanie danych do użycia funkcji REGLINP lub REGLINW</i> .....	202
<i>Regresja wielomianem</i> .....	203
Współczynniki korelacji i determinacji w modelach wielowymiarowych .....	205
Skorygowany współczynnik determinacji .....	207
Regresja danych o niejednakowej dokładności .....	209
<b>Rozdział 9. Regresja nieliniowa .....</b>	<b>213</b>
Zakres zastosowania regresji nieliniowej .....	213
<i>Linearyzacja równania przez rozwinięcie w szereg Taylora</i> .....	213
<i>Regresja quasi-liniowa</i> .....	215

Ocena statystyczna jakości dopasowania równania regresji .....	217
Regresja nieliniowa z wykorzystaniem Solvera .....	217
<i>Regresja logistyczna</i> .....	217
<i>Porównanie regresji quasi-liniowej i nieliniowej</i> .....	222
Regresja nieliniowa metodą iteracyjną .....	229
Korelacja liniowa i korelacja nieliniowa w Excelu .....	231
<i>Korelacja liniowa</i> .....	233
<i>Korelacja nieliniowa z wykorzystaniem Solvera</i> .....	235
Prognozowanie i funkcje służące do tego celu .....	236
<b>Rozdział 10. Interpolacja w tablicach danych .....</b>	<b>241</b>
Interpolacja liniowa .....	241
Trend liniowy .....	244
Interpolacja kubiczna (wielomianem trzeciego stopnia) .....	245
Interpolacja za pomocą funkcji sklepanych .....	246
<i>Implementacja w Excelu</i> .....	249
Interpolacja liniowa w przypadku funkcji dwu zmiennych .....	251
<i>Interpolacja etapowa bez wykorzystania funkcji</i> .....	251
<i>Interpolacja z użyciem funkcji REGLINW</i> .....	252
Interpolacja kubiczna w przypadku funkcji dwu zmiennych .....	254
<i>Interpolacja bez wykorzystania funkcji</i> .....	254
<i>Interpolacja z użyciem funkcji REGLINP lub REGLINW</i> .....	256
<b>Rozdział 11. Obliczanie całek oznaczonych metodami numerycznymi .....</b>	<b>257</b>
Interpretacja geometryczna całki oznaczonej .....	257
Sposoby obliczania całek metodą kwadratur .....	258
Obliczenia całek w Excelu .....	260
Wykorzystanie operacji tablicowych do skrócenia zapisu obliczeń .....	262
<i>Rozwiązanie ogólne</i> .....	262
<i>Rozwiązanie szczególne z zastosowaniem wybranej metody całkowania</i> .....	266
<i>Zastosowanie funkcji makr SZACUJ do obliczania wartości funkcji podcałkowej</i> .....	268
<i>Uzmiennienie granic całkowania</i> .....	270
Wykorzystanie iteracji i odwołań cyklicznych do obliczania całek .....	274
Obliczanie całek niewłaściwych .....	277
Przykład obliczenia całki metodą ekstrapolacji .....	277
<i>Obliczanie całki Arrheniusa</i> .....	278
Obliczanie górnej granicy całkowania .....	281
<i>Funkcja podcałkowa opisana wzorem</i> .....	281
Funkcja podcałkowa w postaci zbioru punktów pomiarowych .....	283

<b>Rozdział 12. Obliczenia z wykorzystaniem szeregów potęgowych .....</b>	<b>289</b>
Podstawowe pojęcia dotyczące ciągów i szeregów .....	289
Obliczenia sum szeregów w Excelu .....	291
<i>Całkowanie funkcji <math>EXP(x)/x</math> z ręcznym określaniem liczby</i>	
<i>sumowanych wyrazów szeregu .....</i>	<i>291</i>
<i>Zastosowanie formuł tablicowych .....</i>	<i>292</i>
<i>Automatyzacja procesu określania liczby sumowanych wyrazów szeregu .....</i>	<i>292</i>
<i>Alternatywny wzór na sumę szeregu .....</i>	<i>296</i>
<i>Sumowanie z wykorzystaniem obliczeń iteracyjnych .....</i>	<i>297</i>
<i>Suma szeregu zdefiniowanego rekurencyjnie .....</i>	<i>299</i>
<i>Problemy występujące przy obliczaniu szeregów .....</i>	<i>300</i>
<i>Obliczanie całki temperaturowej metodą sumowania szeregów .....</i>	<i>302</i>
<b>Rozdział 13. Równania różniczkowe zwyczajne .....</b>	<b>307</b>
Pojęcie równania różniczkowego .....	307
Rozwiązywanie zagadnienia początkowego metodą Eulera .....	308
<i>Typowe równania kinetyki chemicznej .....</i>	<i>308</i>
Modyfikacje metody Eulera .....	316
<i>Idea rozwiązania .....</i>	<i>316</i>
<i>Implementacje rozwiązania w Excelu .....</i>	<i>317</i>
Metoda Rungego-Kutty .....	320
<i>Opis metody .....</i>	<i>320</i>
<i>Implementacja rozwiązania w Excelu .....</i>	<i>321</i>
Porównanie dokładności metod całkowania równań różniczkowych .....	323
Specjalne metody całkowania równań kinetycznych .....	324
<i>Opis metody .....</i>	<i>324</i>
<i>Implementacja rozwiązania w Excelu .....</i>	<i>326</i>
<i>Inne metody całkowania równań kinetycznych .....</i>	<i>328</i>
Zagadnienie brzegowe równania różniczkowego zwyczajnego .....	329
Metoda strzałów .....	330
Metoda różnic skończonych .....	333
<b>Rozdział 14. Równania różniczkowe cząstkowe .....</b>	<b>337</b>
Przykłady równań różniczkowych cząstkowych .....	337
Równania eliptyczne .....	338
Równania paraboliczne .....	344
<i>Rozwiązanie w postaci szeregu potęgowego .....</i>	<i>346</i>
<i>Schematy różnicowe .....</i>	<i>347</i>



<b>Rozdział 15. Zagadnienia kombinatoryki .....</b>	<b>355</b>
Ogólne prawa kombinatoryki .....	355
Liczba wariacji, permutacji i kombinacji .....	356
<i>Wariacje z powtórzeniami</i> .....	356
<i>Wariacje bez powtórzeń</i> .....	357
<i>Permutacje</i> .....	357
<i>Nieporządki (przestawienia)</i> .....	358
<i>Permutacje z powtórzeniami</i> .....	359
<i>Kombinacje</i> .....	360
<i>Kombinacje z powtórzeniami</i> .....	361
<i>Problemy terminologiczne</i> .....	362
Generowanie wyników kombinatorycznych .....	363
<i>Lista wariacji z powtórzeniami</i> .....	363
<i>Lista wariacji bez powtórzeń</i> .....	366
<i>Lista permutacji</i> .....	368
<i>Ciąg permutacji z powtórzeniami</i> .....	369
<i>Generowanie kombinacji przez eliminację z listy wariacji</i> .....	369
<i>Bezpośrednie generowanie kombinacji</i> .....	372
<i>Generowanie kombinacji z powtórzeniami</i> .....	377
<b>Rozdział 16. Konwersja liczb i jednostek .....</b>	<b>379</b>
Dostosowanie stylu wyświetlania liczb do lokalnych ustawień systemowych .....	379
<i>Przyczyna problemu</i> .....	379
<i>Odróżnienie liczby od tekstu wyglądającego jak liczba</i> .....	380
<i>Konwersja stylu liczby za pomocą narzędzi dostępnych w arkuszu</i> .....	381
Konwersja liczb w różnych systemach pozycyjnych .....	384
Konwersja jednostek miar .....	387
<i>Konwersja jednostek złożonych</i> .....	390
<i>Zmiany i uzupełnienia wprowadzone w edycji 2013</i> .....	390
<b>Rozdział 17. Liczby i funkcje zespolone .....</b>	<b>395</b>
Podstawowe wzory i definicje .....	396
Interpretacja geometryczna liczby zespolonej i jej postać trygonometryczna .....	397
Wzór Eulera i wykładnicza postać liczby zespolonej .....	398
Równość liczb zespolonych .....	398
Działania arytmetyczne na liczbach zespolonych .....	398
Potęgowanie i pierwiastkowanie .....	399
Funkcje elementarne z argumentem zespolonym .....	401
<i>Funkcja wykładnicza i logarytm</i> .....	401
<i>Funkcja potęgowa</i> .....	402
<i>Funkcje trygonometryczne i hiperboliczne</i> .....	403

Format wyświetlania liczb zespolonych .....	404
Zastosowanie liczb zespolonych w elektrotechnice .....	406
<b>Rozdział 18. Operacje na danych oznaczających datę i czas .....</b>	<b>411</b>
Wprowadzanie czasu .....	412
Wprowadzanie dat .....	414
Interpretacja tekstu jako daty .....	417
<i>Użycie funkcji do zmiany postaci daty</i> .....	417
Przegląd funkcji Excela do operacji na datach .....	418
<i>Funkcja DATA</i> .....	419
<i>Funkcja DATA.WARTOŚĆ</i> .....	420
<i>Funkcja DZIEŃ</i> .....	421
<i>Funkcja DNI</i> .....	421
<i>Funkcja DNI.360</i> .....	421
<i>Funkcja NR.SER.DATY</i> .....	422
<i>Funkcja NR.SER.OST.DN.MIES</i> .....	422
<i>Funkcja GODZINA</i> .....	423
<i>Funkcja MINUTA</i> .....	423
<i>Funkcja MIESIĄC</i> .....	423
<i>Funkcja DNI.ROBOCZE</i> .....	423
<i>Funkcja DNI.ROBOCZE.NIESTAND</i> .....	424
<i>Funkcja TERAZ</i> .....	425
<i>Funkcja SEKUNDA</i> .....	426
<i>Funkcja CZAS</i> .....	426
<i>Funkcja CZAS.WARTOŚĆ</i> .....	427
<i>Funkcja DZIŚ</i> .....	427
<i>Funkcja DZIEŃ.TYG</i> .....	427
<i>Funkcja NUM.TYG</i> .....	428
<i>Funkcja ISO.NUM.TYG</i> .....	429
<i>Funkcja DZIEŃ.ROBOCZY</i> .....	429
<i>Funkcja DZIEŃ.ROBOCZY.NIESTAND</i> .....	430
<i>Funkcja ROK</i> .....	431
<i>Funkcja CZĘŚĆ.ROKU</i> .....	431
Typowe operacje na danych oznaczających datę i czas .....	431
<i>Dodawanie określonych jednostek czasu do daty lub czasu albo ich odejmowanie</i> .....	431
<i>Konwersja pomiędzy długościami okresów czasu wyrażonymi w różnych jednostkach</i> .....	433
<i>Obliczanie różnicy między datami</i> .....	433
<i>Obliczanie daty początku następnego tygodnia</i> .....	434
<i>Najbliższy dzień roboczy</i> .....	435
<i>Obliczanie daty i godziny zakończenia czynności z uwzględnieniem dni roboczych i godzin pracy</i> .....	435

Wyświetlanie nazw miesięcy i dni tygodnia .....	436
Wyświetlanie numeru miesiąca i dnia tygodnia .....	436
Niestandardowy sposób zapisu czasu .....	437
Funkcja CENA.UŁAM .....	437
Funkcja CENA.DZIES .....	438
Konwersja czasu zapisanego jako godziny, minuty na godziny i ułamki dziesiętne godzin .....	439
Niestandardowe minuty w indeksie górnym .....	440
<b>Rozdział 19. Operacje z udziałem tekstów .....</b>	<b>443</b>
Kodowanie znaków w systemie Windows .....	443
Przegląd funkcji standardowych do operacji na tekstach .....	443
Funkcja KOD .....	444
Funkcja ZNAK .....	444
Funkcja OCZYŚĆ .....	444
Wprowadzanie i usuwanie znaku LF .....	445
Funkcja USUŃ.ZBĘDNE.ODSTĘPY .....	446
Funkcja DŁ .....	447
Funkcja PORÓWNAJ .....	447
Funkcja LEWY .....	448
Funkcja PRAWY .....	448
Funkcja FRAGMENT.TEKSTU .....	448
Funkcja ZNAJDŹ .....	450
Funkcja PODSTAW .....	451
Funkcja ZASTĄP .....	452
Funkcja SZUKAJ.TEKST .....	454
Funkcja POWT .....	456
Funkcja LITERY.MAŁE .....	456
Funkcja LITERY.WIELKIE .....	457
Funkcja Z.WIELKIEJ.LITERY .....	457
Funkcja WARTOŚĆ .....	457
Funkcja ZŁĄCZ.TEKSTY .....	458
Funkcja ZŁĄCZ.TEKST .....	459
Funkcja POŁĄCZ.TEKSTY .....	459
Funkcje KWOTA i ZAOKR.DO.TEKST .....	460
Funkcja TEKST .....	460
Funkcja T .....	460
Używanie znaków zastępczych .....	461
Wybrane operacje na danych tekstowych .....	462
Sprawdzanie, czy tekst zawiera określoną frazę .....	462
Zliczanie określonych znaków lub fraz w tekście .....	462

<i>Zliczanie słów w tekście</i> .....	463
<i>Podział tekstu na dwie części na granicy słów</i> .....	464
<i>Wybór pierwszego lub ostatniego słowa w tekście</i> .....	465
<i>Podział tekstu na słowa lub inne fragmenty na podstawie położenia separatorów</i> .....	465
<i>Wydzielenie liczby z tekstu</i> .....	466
<i>Oddzielenie cyfr od innych znaków w tekście</i> .....	467
<i>Wykonanie funkcji tekstowych w komórkach sformatowanych tekstowo</i> .....	468
<i>Ciągi cyfr z wiodącymi zerami</i> .....	468
Łączenie tekstów .....	470
<i>Wykorzystywanie obliczeń iteracyjnych</i> .....	470
<i>Przekształcanie zakresu w stałą tablicową, a następnie w tekst</i> .....	472
<b>Rozdział 20. Użycie funkcji INDEKS z argumentami indeksowymi w postaci tablic</b> .....	<b>477</b>
Ograniczenia w użyciu tablic, które są wynikami formuł .....	477
Specyfika użycia funkcji INDEKS z argumentami indeksowymi w postaci tablic .....	479
<i>Sposób 1.</i> .....	479
<i>Sposób 2.</i> .....	480
<i>Sposób 3.</i> .....	481
<i>Sposób 4.</i> .....	482
Przegląd problemów wynikających z używania wyrażeń indeksowych o różnej postaci .....	482
<i>Jedno z wyrażeń indeksowych jest tablicą, a drugie liczbą różną od zera</i> .....	483
<i>Oba wyrażenia indeksowe są tablicami bez elementów zerowych</i> .....	483
<i>Jedno z wyrażeń indeksowych jest zwykłą tablicą, a drugie SWI</i> .....	484
<i>Oba wyrażenia indeksowe są typu SWI bez zer</i> .....	485
<i>Jedno z wyrażeń indeksowych jest tablicą, a drugie zerem</i> .....	485
<i>Użycie zera jako wartości elementu wyrażenia indeksowego</i> .....	486
<i>Jedno z wyrażeń indeksowych jest tablicą zawierającą zero, a drugie zerem</i> .....	487
<i>Niedopasowanie wielkości wyrażeń indeksowych</i> .....	487
<i>Argument główny jest tablicą jednowymiarową,</i> <i>a wyrażenie indeksowe tablicą dwuwymiarową</i> .....	488
<i>Struktura tablicowa wyniku funkcji INDEKS</i> .....	488
<i>Funkcja INDEKS z argumentami tablicowymi skojarzona z funkcją</i> <i>agregującą lub tablicową</i> .....	488
<i>Pominięcie indeksu lub użycie zera jako jego wartości przy sumowaniu</i> .....	491
<i>Funkcja INDEKS z zakresami nieregularnymi</i> .....	493
Wybrane niestandardowe zastosowania funkcji INDEKS .....	498
<i>Łączenie fragmentów tablicy dwuwymiarowej</i> .....	498
<i>Zastosowania odwołaniowej formy funkcji INDEKS</i> .....	499
<i>Odwracanie porządku elementów w zakresie</i> .....	499

Wykorzystanie tablic bazowych do indeksowania tablic .....	500
<i>Odwracanie kolejności elementów w tablicach</i> .....	500
<i>Rozmieszczenie danych w tablicy o założonych rozmiarach</i> .....	501
<i>Sumowanie wyrazów w wyodrębnionym fragmencie tablicy</i> .....	501
<i>Selektywne sumowanie elementów tablic dwuwymiarowych</i> .....	502
<i>Usuwanie wybranego wiersza lub wybranej kolumny z tablicy</i> .....	503
<i>Łączenie fragmentów zakresów i tablic w tablicę ciągłą</i> .....	503
<b>Rozdział 21. Pseudozakresy .....</b>	<b>507</b>
Pojęcie pseudozakresu .....	507
<i>Pseudozakresy jako wyniki funkcji ADR.POŚR</i> .....	509
<i>Pseudozakresy jako wyniki funkcji PRZESUNIĘCIE</i> .....	510
Pseudozakresy jako argumenty funkcji L(N) lub T .....	511
Pseudozakres jako drugi (lub kolejny) argument funkcji JEŻELI lub WYBIERZ .....	514
Pseudozakres jako pierwszy argument funkcji INDEKS .....	515
Pseudozakresy jako argumenty funkcji SUMY.CZĘŚCIOWE i AGREGUJ .....	518
Pseudozakresy jako argumenty funkcji SUMA.JEŻELI .....	520
Pseudozakresy jako argumenty funkcji agregujących .....	523
<i>Pseudozakresy I rodzaju</i> .....	523
<i>Pseudozakresy II rodzaju</i> .....	527
Pseudozakresy jako argumenty funkcji tablicowych .....	528
Pseudozakresy jako argumenty funkcji baz danych .....	530
Pseudozakresy jako argumenty funkcji KOMÓRKA .....	533
Użycie pseudozakresów z funkcją LOS .....	535
Pseudozakresy jako argumenty funkcji makr XLM .....	536
Szczególne przypadki użycia pseudozakresów .....	536
<i>Pseudozakres jako drugi argument funkcji INDEKS</i> .....	536
<i>Pseudozakres jako pierwszy argument funkcji PRZESUNIĘCIE</i> .....	537
<i>Indeksowanie funkcji PRZESUNIĘCIE z pseudozakresem</i> .....	538
<i>Użycie funkcji PRZESUNIĘCIE do danych przefiltrowanych funkcją INDEKS</i> .....	539
<i>Formuły tablicowe z pseudozakresami uwzględniające faktyczne rozmiary obszaru formuły</i> .....	542
<i>Operacje agregujące na zakresach dynamicznych</i> .....	546
<i>Równoległe tablice bazowe</i> .....	547
<b>Rozdział 22. Ciekawostki i pomysły niewymagające użycia VBA .....</b>	<b>549</b>
Wyznaczanie długości części wspólnej dwu odcinków .....	549
Suma cyfr liczby całkowitej .....	550
Budowanie liczby z cyfr .....	551

Obliczenia uwzględniające ukryte kolumny .....	552
<i>Z wierszem pomocniczym</i> .....	552
<i>Z formułą nazwaną (rozwiązanie egzotyczne)</i> .....	553
<i>Z formułą nazwaną (z użyciem pseudozakresu)</i> .....	554
Funkcja PROCENT.POZYCJA — specyfika i możliwości wykorzystania .....	554
<i>Zastosowania funkcji PROCENT.POZYCJA</i> .....	558
Funkcja PRAWDPD — specyfika i możliwości wykorzystania .....	560
Użycie funkcji generujących liczby pseudolosowe .....	562
Formuły „zatrzaskowe” .....	565
<i>Wykonanie jednorazowe</i> .....	565
<i>Wykonanie wielokrotne</i> .....	567
Cykliczna formuła warunkowa .....	568
Suma iloczynów według formuły narastającej .....	569
Wstawianie i usuwanie wierszy w zakresie sumowania .....	570
Obliczanie wartości wyrażeń danych w postaci tekstu .....	571
Operacje bitowe na danych liczbowych .....	573
<b>Literatura cytowana i uzupełniająca .....</b>	<b>577</b>
<b>Skorowidz .....</b>	<b>579</b>

## Rozdział 7.

# Regresja liniowa jednowymiarowa

## Zakres zastosowania regresji liniowej

Eksperymenty, których rezultatem są duże zbiory danych liczbowych reprezentujących zależności pomiędzy zmiennymi wejściowymi a wynikami, staramy się zwykle opracować w postaci wykresów lub równań. Postać równań może wynikać z przesłanek teoretycznych, lecz najczęściej stosuje się do opisu typowe, stosunkowo proste równania funkcji liniowej, wielomianowej, potęgowej, wykładniczej lub logarytmicznej. Najczęściej stosowanym kryterium dopasowania równania do danych doświadczalnych jest **minimalna suma kwadratów** różnic między wartościami wynikowymi zmierzonymi a wyznaczonymi z równania. Kryterium to stanowi podstawę metody regresji, której szczególnym przypadkiem jest regresja liniowa.

**Regresja liniowa** nie ogranicza się do zależności liniowej między wynikami a danymi wejściowymi, lecz obejmuje wszystkie równania, które są liniowe ze względu na występujące w nich współczynniki. Jeżeli równanie opisujące wyniki eksperymentu przedstawimy w postaci  $Y = f(X, A)$ , gdzie  $Y$  reprezentuje wielkości wynikowe,  $X$  — zmienne wejściowe, a  $A$  — współczynniki, to równanie jest liniowe ze względu na współczynniki, jeśli pochodna  $\frac{\partial Y}{\partial A}$  nie zależy od  $A$ . Do tej kategorii zaliczymy więc zarówno równania typu:

$$y = a_0 + a_1x_1 + a_2x_2 + a_3x_3$$

w których wielkość wynikowa zależy od trzech niezależnych zmiennych wejściowych, jak i równania typu:

$$y = a_0 + a_1x + a_2x^2 + a_3x^{-1} + a_4 \ln x$$

gdzie zależność od jednej zmiennej wejściowej przedstawiona jest za pomocą równania nieliniowego ze względu na  $x$ , ale liniowego ze względu na współczynniki. Możliwe są też równania typu:

$$y = a_0 + a_1x_1 + a_2x_2 + a_{12}x_1x_2 + a'_{12}x_1/x_2$$

reprezentujące zależność od dwóch zmiennych wejściowych.

Istnieje także kategoria równań, które wprawdzie nie są liniowe, ale dają się przekształcić do postaci liniowej przez odpowiednią zamianę zmiennych. Do tej grupy zaliczamy na przykład równanie wykładnicze:

$$y = ae^{bx}$$

Po obustronnym zlogarytmowaniu przyjmie ono postać:

$$\ln y = \ln a + bx$$

która jest już liniowa ze względu na współczynniki. Pamiętać jednak należy, że zastosowanie metody regresji do tak przekształconego równania prowadzi do innych wartości współczynników, niż gdyby brać pod uwagę równanie w pierwotnej postaci. Czym innym jest suma kwadratów odchyleń liczona dla  $y$ , a czym innym ta suma liczona dla  $\ln y$ . W tym drugim przypadku relatywnie zmniejsza się waga punktów pomiarowych, w których wartości  $y$  były większe, a zwiększa się waga punktów z małymi wartościami  $y$ . Niemniej metodę linearyzacji równań stosuje się dość powszechnie, gdy tylko jest to możliwe.

Podobnie daje się linearyzować zależność potęgowa:

$$y = ax^b$$

Po obustronnym zlogarytmowaniu otrzymujemy:

$$\ln y = \ln a + b \ln x$$

czyli równanie liniowe, ale — podobnie jak poprzednio — nie dla  $y$ , lecz dla  $\ln y$ .

Kolejny typ równania podlegający linearyzacji to dwuparametrowa funkcja homograficzna:

$$y = \frac{ax}{x+b}$$

Po inwersji równanie przekształca się do postaci:

$$\frac{1}{y} = \frac{1}{a} + \frac{b}{a} \cdot \frac{1}{x}$$

która reprezentuje liniową formę zależności  $1/y = f(1/x)$ .

Inne możliwe sposoby linearyzacji tego równania to:

$$\frac{x}{y} = \frac{b}{a} + \frac{1}{a}x, \quad y = a - b \frac{y}{x} \quad \text{lub} \quad xy = ax - by$$

Pierwszy reprezentuje liniową formę zależności  $x/y = f(x)$ , drugi  $y = f(y/x)$ , a trzeci  $xy = f(x, y)$  bez wyrazu wolnego. Do każdej z tych zależności można dołączyć zależność odwrotną, która też jest liniowa.



Zastosowanie każdej z form linearyzacji prowadzi do innych wartości współczynników  $a$  i  $b$ .

Linearyzacji podlega również funkcja Gaussa, ważna w rachunku prawdopodobieństwa i statystyce matematycznej, ale także w obróbce danych doświadczalnych, gdyż wykresy otrzymywane w różnych technikach badawczych przyjmują często kształty zbliżone do krzywej Gaussa. Krzywą Gaussa można opisać równaniem z trzema parametrami,  $A$ ,  $B$  i  $m$ :

$$y = Ae^{-B(x-m)^2}$$

Po zlogarytmowaniu otrzymamy zależność:

$$\ln y = \ln A - B(x - m)^2 = (\ln A - Bm^2) + 2Bmx - Bx^2$$

Jest to równanie kwadratowe ze względu na  $x$ . Można do niego zastosować regresję liniową wielowymiarową.

To najbardziej typowe przykłady linearyzacji. W podobny sposób linearyzuje się bardziej złożone funkcje, na przykład zależność potęgowo-wykładniczą:

$$y = ax^b e^{cx}$$

$$\text{lub } y = \frac{ax^n}{x + b}$$

pod warunkiem że  $n$  jest znaną stałą, a nie współczynnikiem do wyznaczenia.

Istnieje jednak wiele zależności ważnych z punktu widzenia obróbki danych doświadczalnych, które nie dają się linearyzować i wymagają nieco innego podejścia, zwanego regresją nieliniową. Przykłady takich zależności to:

$$y = a(1 - e^{-bx})$$

$$y = \frac{ax^{1/2}}{1 + bx^{1/2}} + cx$$

W tym rozdziale ograniczę się do omówienia przypadków regresji liniowej jednej zmiennej, które da się sprowadzić do najprostszej zależności:

$$y = a + bx$$

W następnym rozdziale omówione będą pozostałe przypadki regresji liniowej, a w kolejnym — regresja nieliniowa.

# Ocena statystyczna danych pomiarowych i jakości dopasowania równania regresji

## Charakterystyka zbiorów danych pomiarowych

Poszukując zależności między wielkościami  $Y$  i  $X$ , zgromadziliśmy dane pomiarowe w postaci zbioru  $N$  par wartości  $(x_i, y_i)$ . Zbiór ten można scharakteryzować za pomocą różnych wielkości zagregowanych, to znaczy obliczonych dla wszystkich punktów zbioru łącznie. Wielkości te mogą być obliczane osobno dla każdej zmiennej lub mogą obejmować obie zmienne równocześnie.

Dla każdej zmiennej można obliczyć wartość średnią (średnia arytmetyczna) według wzorów:

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i \qquad \bar{y} = \frac{1}{N} \sum_{i=1}^N y_i$$

W tym tekście kreska nad symbolem oznacza wartość średnią. W Excelu służy do tego funkcja ŚREDNIA, która może przyjmować dowolną liczbę argumentów w postaci stałych, adresów komórek i zakresów oraz nazw odnoszących się do tych obiektów.

Do określenia liczby  $N$  można użyć funkcji ILE.LICZB, przydatnej przy dużej liczbie pomiarów i wtedy, gdy ich wyniki są rozmieszczone w kilku miejscach w arkuszu.

Jedną z miar rozproszenia danych jest ich odchylenie kwadratowe, czyli suma kwadratów odchyleń zmiennej od wartości średniej:

$$SS_x = \sum_{i=1}^N (x_i - \bar{x})^2 \qquad SS_y = \sum_{i=1}^N (y_i - \bar{y})^2$$

W Excelu wielkość tę oblicza funkcja ODCH.KWADRATOWE.

Ponieważ odchylenie kwadratowe rośnie ze wzrostem liczby punktów, to aby można było porównywać różne zbiory danych, często się je standaryzuje, dzieląc przez liczbę danych. Otrzymuje się wielkość zwaną **wariancją**, która funkcjonuje w dwóch postaciach — **wariancji próbki**, gdy w mianowniku jest  $N-1$ , i **wariancji populacji**, gdy w mianowniku jest  $N$ :

$$S_{xpr} = \frac{\sum_{i=1}^N (x_i - \bar{x})^2}{N-1} \qquad S_{ypr} = \frac{\sum_{i=1}^N (y_i - \bar{y})^2}{N-1}$$

$$S_{xpop} = \frac{\sum_{i=1}^N (x_i - \bar{x})^2}{N} \qquad S_{ypop} = \frac{\sum_{i=1}^N (y_i - \bar{y})^2}{N}$$

W Excelu obliczenia te wykonują funkcje WARIANCJA i WARIANCJA.POPUL. Poczawszy od Excela 2010, zaleca się używanie nowych nazw: WARIANCJA.PRÓBK I WARIANCJA.POP, jakkolwiek stare nazwy nadal mogą być stosowane.

Pierwiastek z wariancji ma ten sam wymiar co wielkość mierzona i może być miarą odchylenia wartości zmierzonej w punkcie od średniej. Analogicznie do wariancji stosowane są dwie wielkości — **odchylenie standardowe próbki**, gdy w mianowniku jest  $N-1$ , i **odchylenie standardowe populacji**, gdy w mianowniku jest  $N$ :

$$s_{xpr} = \sqrt{\frac{\sum_{i=1}^N (x_i - \bar{x})^2}{N-1}} \quad s_{ypr} = \sqrt{\frac{\sum_{i=1}^N (y_i - \bar{y})^2}{N-1}}$$

$$s_{xpop} = \sqrt{\frac{\sum_{i=1}^N (x_i - \bar{x})^2}{N}} \quad s_{ypop} = \sqrt{\frac{\sum_{i=1}^N (y_i - \bar{y})^2}{N}}$$

W Excelu używa się funkcji ODCH.STANDARDOWE i ODCH.STANDARD.POPUL. Poczawszy od Excela 2010, zaleca się stosowanie nowych nazw: ODCH.STANDARD.PRÓBK I ODCH.STAND.POPUL, jakkolwiek stare nazwy nadal mogą być używane.

Rzadziej stosowaną miarą odchylenia zbioru danych od wartości średniej jest **odchylenie średnie**, liczone na podstawie sumy wartości bezwzględnych odchyleń:

$$SS'_x = \frac{\sum_{i=1}^N |x_i - \bar{x}|}{N}$$

W Excelu obliczenia te wykonuje funkcja ODCH.ŚREDNIE.

## Ocena zależności między zbiorami danych pomiarowych

### Kowariancja

Omówione powyżej wielkości charakteryzują oddzielnie zbiory wartości  $x$  i  $y$ . Aby ocenić zależność między tymi wielkościami, należy powiązać ich wartości we wspólnym agregacie. Pierwszą z takich wielkości jest **kowariancja**, występująca — podobnie jak wariancja — w dwóch odmianach: **kowariancji z próbki** i **kowariancji populacji**. Znając definicję wariancji, łatwo się domyślić, jak jest zbudowana ta funkcja:

$$s_{xypr} = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{N-1}$$

$$S_{xypop} = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{N}$$

Kowariancja służy do oceny, czy pomiędzy zmiennymi  $x$  i  $y$  istnieje zależność liniowa. Jeżeli zmienne są niezależne, kowariancja jest bliska zeru. Wariancje i kowariancję można wyrazić równoważnie także w inny sposób:

$$S_{xpop} = \frac{\sum_{i=1}^N x_i^2}{N} - \bar{x}^2 \quad S_{ypop} = \frac{\sum_{i=1}^N y_i^2}{N} - \bar{y}^2 \quad S_{xypop} = \frac{\sum_{i=1}^N x_i y_i}{N} - \bar{x} \cdot \bar{y}$$

Stosując konsekwentnie oznaczenie średniej arytmetycznej w postaci kreski nad symbolem, można te wzory zapisać jeszcze prościej:

$$S_{xpop} = \overline{x^2} - \bar{x}^2 \quad S_{ypop} = \overline{y^2} - \bar{y}^2 \quad S_{xypop} = \overline{xy} - \bar{x} \cdot \bar{y}$$

Należy pamiętać, że średnia wartość iloczynów to nie to samo co iloczyn średnich. W starszych wersjach Excela kowariancję obliczymy za pomocą funkcji KOWARIANCJA. Oblicza ona kowariancję populacji. Począwszy od Excela 2010, dostępne są dwie funkcje, KOWARIANCJA.PRÓBK I KOWARIANCJA.POPUL, ale można też używać starych nazw funkcji w dotychczasowym znaczeniu.

## Współczynnik korelacji

Do oceny zależności liniowej między zmiennymi  $x$  i  $y$  wykorzystuje się często kowariancję znormalizowaną poprzez podzielenie jej przez iloczyn odchyłeń standardowych. Współczynnik ten nosi nazwę **współczynnika korelacji Pearsona**, zwany jest często krótko współczynnikiem korelacji  $R$ .

$$R = \frac{S_{xypop}}{S_{xpop} S_{ypop}} = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\sqrt{(\overline{x^2} - \bar{x}^2)(\overline{y^2} - \bar{y}^2)}} = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2 \sum_{i=1}^N (y_i - \bar{y})^2}}$$

Współczynnik ten może przyjmować wartości z przedziału  $[-1, +1]$ . Wartości skrajne odpowiadają doskonałej korelacji liniowej, odpowiednio malejącej lub rosnącej.

W Excelu współczynnik korelacji oblicza się za pomocą funkcji WSP.KORELACJI lub PEARSON.

Równie często, a może nawet częściej, zamiast współczynnika korelacji oblicza się jego kwadrat. W Excelu służy do tego funkcja R.KWADRAT.

## Zasady doboru współczynników w równaniu regresji

Jeżeli chcemy opisać zależność pomiędzy zmiennymi  $y$  i  $x$  w sposób sformalizowany, czyli w postaci równania ze współczynnikami liczbowymi (**równania regresji**), musimy określić kryterium doboru wartości tych współczynników. Jak już wspominałem, kryterium jakości dopasowania równania regresji do danych pomiarowych stanowi najczęściej minimalna wartość sumy kwadratów różnic między wartościami  $y$  zmierzonymi a wartościami  $\hat{y}$  obliczonymi z równania regresji. Tę sumę określa się najczęściej nazwą **sumy resztkowej** i oznacza symbolem  $SS_{resz}$ :

$$SS_{resz} = \sum_{i=1}^N (y_i - \hat{y}_i)^2$$

$N$  oznacza tu liczbę danych odnoszących się do zmiennej  $y$ .

Do oceny jakości dopasowania niezbędna jest też tzw. **liczba stopni swobody**,  $f = N - l$ , czyli różnica między liczbą pomiarów  $N$  a liczbą współczynników  $l$ , które trzeba wyznaczyć. Jeżeli dysponujemy tylko tyloma wynikami pomiarów, ile współczynników jest do wyznaczenia, liczba stopni swobody  $f = 0$ , wtedy również  $SS_{resz} = 0$ , ale nie ma możliwości statystycznej oceny jakości otrzymanego równania. Na przykład przez dwa punkty zawsze można przeprowadzić prostą, ale gdyby dodać trzeci punkt pomiarowy, to nie wiadomo, czy też leżałby na tej prostej. Miarą jakości dopasowania równania jest stosunek sumy resztkowej  $SS_{resz}$  do liczby stopni swobody  $f$ , zwany **wariancją resztkową**.

Pierwiastek z wariancji resztkowej zwany jest **średnim błędem oszacowania wartości  $y$** ,  $s_y$ . To kolejna charakterystyka statystyczna równania regresji. Wartość tę umieszcza się czasem na wykresach w postaci słupków błędu wartości  $y$ . Jest ona obliczana ze wzoru:

$$s_y = \sqrt{\frac{SS_{resz}}{N - l}} = \sqrt{\frac{SS_{resz}}{f}}$$

Każdy z wyznaczonych współczynników w równaniu regresji obarczony jest pewnym błędem, który również można oszacować. Im większa jest wartość współczynnika (liczona jako wartość bezwzględna) w stosunku do błędu jego oszacowania, tym bardziej istotny jest dany współczynnik. Współczynniki mało istotne, tzn. takie, których błąd jest relatywnie duży, powinny być usuwane z równania regresji. Pomocne jest tu **kryterium statystyczne  $t$  Studenta**, które określa graniczną wartość tego stosunku, zależną od liczby stopni swobody powtarzalności pomiarów oraz od przyjętego tzw. poziomu istotności.

Suma resztkowa i wariancja resztkowa nie wystarczają jednak do pełnej charakterystyki danych pomiarowych, bo jeżeli wartości  $y$  zmieniają się wyraźnie od punktu do punktu (zależność jest stroma), to i suma resztkowa będzie stosunkowo duża. Jeżeli natomiast zmienność  $y$  jest niewielka (zależność jest płaska), to i suma resztkowa będzie mała, chociaż równanie niekoniecznie musi być dobrze dobrane. Dlatego sensownie jest porównać sumę resztkową z sumą kwadratów odchyień wyników w punktach pomiarowych odniesionych do ogólnej średniej z wszystkich

wyników pomiarów. Tę ogólną sumę kwadratów wokół średniej pomniejszoną o sumę resztkową nazwano **sumą regresyjną**. Oblicza się ją ze wzoru:

$$SS_{regr} = \sum_{i=1}^N (y_i - \bar{y})^2 - SS_{resz} = SS_y - SS_{resz}$$

W Excelu (od edycji 2003) zależność tę stosuje się tylko do pełnego równania regresji (gdy wyraz wolny jest obliczany). Kiedy wyraz wolny jest równy zero, stosowany jest zmodyfikowany wzór (kwestię tę wyjaśniam dokładnie dalej). Sumy kwadratów resztkowa i regresyjna są porównywane w dwojaki sposób. Na ich podstawie oblicza się tzw. **współczynnik determinacji D**:

$$D = \frac{SS_{regr}}{SS_{regr} + SS_{resz}} = 1 - \frac{SS_{resz}}{SS_y}$$

Wielkość ta jest bardzo często wykorzystywana do oceny jakości równania regresji. Jeżeli dane nie są obciążone dużymi błędami pomiarów, a równanie regresji zostało dobrze dobrane, jej wartość jest bardzo bliska 1. Wartości poniżej 0,9 świadczą już o stosunkowo mało dokładnym równaniu. Nie jest to jednak całkiem jednoznaczne — w pomiarach fizycznych (szerzej w naukach ścisłych) współczynniki determinacji przyjmują zwykle wartości dużo bliższe 1 niż na przykład w naukach medycznych, rolniczych czy ekonomicznych.

Inną charakterystyką statystyczną wyliczaną na podstawie powyższych sum kwadratów jest stosunek  $F$  wariancji regresji do wariancji resztkowej. Jest to wielkość, którą porównuje się z **kryterium statystycznym Fishera-Snedecora**:

$$F = \frac{SS_{regr} / (l - 1)}{SS_{resz} / f} \quad \text{lub} \quad F = \frac{SS'_{regr} / l}{SS_{resz} / f}$$

Pierwszy wariant wzoru stosuje się, gdy stała w równaniu regresji jest obliczana, a drugi — gdy stała jest równa zero. Jednak w tym drugim przypadku suma regresyjna obliczana jest ze zmodyfikowanego wzoru (zobacz dalej). Wartość  $F$  powinna być możliwie jak największa, a w każdym razie większa od wartości granicznej, określonej w zależności od liczby stopni swobody licznika i mianownika oraz przyjętego poziomu istotności.

## Regresja liniowa w Excelu

### Znaczenie współczynników w równaniu regresji i sposoby ich obliczania

W arkuszu kalkulacyjnym dostępnych jest kilka funkcji, które służą do wyznaczania równań regresji. Ponadto dostępne są narzędzia do analizy regresji powiązane z kreatorem wykresów (linia trendu) oraz niezależne, wchodzące w skład dodatku *Analysis Toolpak*, których nie będę tu omawiał.

Najprostszy przypadek dotyczy dopasowania funkcji liniowej jednej zmiennej. Jeśli poszukiwana zależność da się opisać równaniem:

$$y = ax + b$$

to zastosowanie metody najmniejszej sumy kwadratów (określonej dalej skrótowo jako NK) prowadzi do następujących wzorów na współczynniki  $a$  i  $b$ :

$$a = \frac{S_{xy\text{pop}}}{S_{x\text{pop}}} = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\overline{x^2} - \bar{x}^2} = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^N (x_i - \bar{x})^2}$$

$$b = \bar{y} - a\bar{x} = \frac{\bar{y} \cdot \overline{x^2} - \bar{x} \cdot \overline{xy}}{\overline{x^2} - \bar{x}^2}$$

$$R^2 = \frac{S_{xy\text{pop}}^2}{S_{x\text{pop}} S_{y\text{pop}}} = \frac{(\overline{xy} - \bar{x} \cdot \bar{y})^2}{(\overline{x^2} - \bar{x}^2)(\overline{y^2} - \bar{y}^2)} = a^2 \frac{\overline{x^2} - \bar{x}^2}{\overline{y^2} - \bar{y}^2} = a^2 \frac{S_{x\text{pop}}}{S_{y\text{pop}}}$$

$$SS_{\text{reszt}} = \sum_{i=1}^N (y_i - \bar{y})^2 - \frac{\left( \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y}) \right)^2}{\sum_{i=1}^N (x_i - \bar{x})^2} = (1 - R^2) \sum_{i=1}^N (y_i - \bar{y})^2 = \sum_{i=1}^N (y_i - \bar{y})^2 - a^2 \sum_{i=1}^N (x_i - \bar{x})^2$$

$$s_y = \sqrt{\frac{SS_{\text{reszt}}}{N-2}} = \sqrt{\frac{SS_{\text{reszt}}}{f}}$$

## Funkcje NACHYLENIE, ODCIĘTA, R.KWADRAT i REGBŁSTD

W Excelu można wykorzystać cztery funkcje o podobnej składni, które umożliwiają obliczenie współczynników  $a$  i  $b$ , współczynnika determinacji  $D$ , który w tym przypadku jest równy kwadratowi współczynnika korelacji  $R^2$ , oraz standardowego błędu oceny wartości  $y$ ,  $s_y$ . Są to funkcje:

NACHYLENIE(**znane\_y**; **znane\_x**) — oblicza współczynnik  $a$

ODCIĘTA(**znane\_y**; **znane\_x**) — oblicza współczynnik  $b$

R.KWADRAT(**znane\_y**; **znane\_x**) — oblicza współczynnik  $R^2$

REGBŁSTD(**znane\_y**; **znane\_x**) — oblicza oszacowanie błędu  $s_y$

**Znane\_y** i **znane\_x** to tablice lub zakresy komórek o identycznych rozmiarach zawierające dane liczbowe. Jeżeli zakresy lub tablice zawierają dane nieliczbowe lub puste komórki, te wartości są ignorowane (jeśli jedna z wartości  $x$  lub  $y$  jest nieprawidłowa, ignorowana jest cała para). Jeżeli rozmiary tablic lub zakresów są różne, otrzymujemy komunikat błędu #N/D!. Orientacja **znane\_y**

i *znane\_x* jest obojętna, ale zakresy muszą być spójne. Kolejność punktów *znane\_y* i *znane\_x* ustala się, odczytując te tablice lub zakresy wierszami. *Znane\_x* nie musi być uporządkowane. Zakresy lub tablice *znane\_y* i *znane\_x* muszą zawierać przynajmniej dwa różne punkty.

Pomiędzy funkcjami wymienionymi wcześniej istnieją zależności:

$$\begin{aligned} \text{NACHYLENIE}(\text{znane\_y}; \text{znane\_x}) &= \text{KOWARIANCJA}(\text{znane\_y}; \text{znane\_x}) / \text{WARIANCJA.POPUL}(\text{znane\_x}) \\ \text{ODCIĘTA}(\text{znane\_y}; \text{znane\_x}) &= \text{ŚREDNIA}(\text{znane\_y}) - \text{ŚREDNIA}(\text{znane\_x}) * \text{NACHYLENIE}(\text{znane\_y}; \text{znane\_x}) \\ \text{R.KWADRAT}(\text{znane\_y}; \text{znane\_x}) &= \text{NACHYLENIE}(\text{znane\_y}; \text{znane\_x}) * \text{NACHYLENIE}(\text{znane\_x}; \text{znane\_y}) \end{aligned}$$

Zależności te są spełnione pod warunkiem, że wszystkie dane są poprawne, to znaczy zakresy *znane\_y* i *znane\_x* mają takie same rozmiary i zawierają wyłącznie liczby.

## Funkcja REGLINP

Jeśli w równaniu regresji występuje więcej wyrazów, a zarazem więcej współczynników, albo jeśli wyraz wolny (odcięta) jest równy zeru, albo jeśli potrzebny jest pełniejszy statystyczny opis zbioru danych — należy skorzystać z funkcji tablicowej:

$$\text{REGLINP}(\text{znane\_y}; \text{znane\_x}; \text{czy\_stała}; \text{czy\_stat})$$

*Znane\_y* i *znane\_x*, jak poprzednio, reprezentują zbiory danych pomiarowych zmiennych *y* i *x*, ale teraz obszary te nie muszą mieć identycznych rozmiarów, jednak muszą być dopasowane. Jeżeli tablica *znane\_y* jest pojedynczą kolumną, każda kolumna zbioru *znane\_x* jest interpretowana jako oddzielna zmienna. Analogicznie jeżeli tablica *znane\_y* jest pojedynczym wierszem, każdy wiersz zbioru *znane\_x* jest interpretowany jako oddzielna zmienna.

Tablica *znane\_x* może zawierać jeden lub więcej zbiorów zmiennych. Jeżeli używana jest tylko jedna zmienna, *znane\_y* i *znane\_x* mogą być obszarami lub tablicami o dowolnym kształcie, ale o jednakowych wymiarach.

Jeżeli stosuje się więcej niż jedną zmienną *x*, *znane\_y* musi być wektorem (tzn. zakresem lub tablicą o wysokości jednego wiersza bądź o szerokości jednej kolumny). Tablice *znane\_y* i *znane\_x* muszą być w całości wypełnione danymi liczbowymi (**nie mogą zawierać pustych komórek ani danych nieliczbowych** — inaczej niż w przypadku funkcji NACHYLENIE, ODCIĘTA, R.KWADRAT i REGBŁSTD), ponadto jeśli *znane\_x* reprezentuje kilka zmiennych, musi stanowić **zakres spójny**, czyli musi być zawarte w jednym obszarze lub jednej tablicy.

Jeżeli pominięto *znane\_x*, argument ten jest przyjmowany jako tablica {1;2;3;...} lub {1\2\3\...}, która ma taką samą wielkość jak *znane\_y*.

Funkcja toleruje sytuację, w której liczba punktów jest zbyt mała do jednoznacznego wyznaczenia współczynników regresji. Przyjmuje wówczas zerowe współczynniki dla zmiennych, dla których nie jest możliwe wyznaczenie jednoznacznego wyniku.

*Czy\_stala* jest to wartość logiczna określająca, czy stała *b* ma być równa 0, czy obliczana normalnie. Jeżeli *czy\_stala* ma wartość PRAWDA lub jest pominięta, *b* jest obliczane normalnie. Jeżeli *czy\_stala* ma wartość FAŁSZ, *b* jest równe 0, a wartości *a* są dopasowywane do równania  $y = ax$  lub  $y = \sum a_i x_i$ .





Przy dodawaniu linii trendu do wykresu punktowego można zadeklarować dowolną wartość stałej  $b$  w równaniu regresji. W przypadku funkcji REGLINP można jedynie zadeklarować wartość 0. Jeżeli znana wartość stałej  $b$  jest różna od zera, należy od zmierzonych wartości  $y$  odjąć  $b$  i w charakterze zmiennej **znane\_y** użyć tej różnicy, nadając równocześnie argumentowi **czy\_stat** wartość FAŁSZ.

Jeżeli linia regresji ma przechodzić przez punkt stały o współrzędnych  $(x0, y0)$ , należy postąpić analogicznie, to znaczy wywołanie funkcji powinno mieć postać:

```
REGLINP(znane_y - y0; znane_x - x0; FAŁSZ; czy_stat)
```

a równanie linii regresji:

$$y = y0 + a*(x - x0).$$

**Czy\_stat** jest to wartość logiczna określająca, czy ma być dodatkowo obliczana statystyka regresji. Jeżeli **czy\_stat** ma wartość PRAWDA, funkcja REGLINP daje w wyniku dodatkowe statystyki regresji, tak że uzyskiwana tablica ma postać  $\{a_n; a_{n-1}; \dots; a_1; a_0 \setminus s_{an}; s_{an-1}; \dots; s_{a1}; s_{a0} \setminus D; s_y \setminus F; f \setminus SS_{regr}; SS_{resz}\}$  (separatory w konwencji Excela 2007–). Wynik składa się z pięciu wierszy i tylu kolumn, ile współczynników jest obliczanych. W pierwszym wierszu jest lista współczynników, ale w **kolejności odwrotnej** w stosunku do kolejności kolumn z danymi, zakończona wyrazem wolnym  $a_0$ . W drugim wierszu są średnie błędy współczynników z pierwszego wiersza. Od trzeciego do piątego wiersza tylko po dwie wartości — kolejno: współczynnik determinacji  $D$ , średni błąd oszacowania wartości  $y$ , statystyka Fishera-Snedecora  $F$ , obok liczba stopni swobody wariancji resztkowej  $f$  i wreszcie w piątym wierszu suma regresyjna  $SS_{regr}$  i suma resztkowa  $SS_{resz}$ . Znaczenie tych wielkości i sposób ich obliczania podałem wcześniej, w punkcie „Zasady doboru współczynników w równaniu regresji”. Jeżeli **czy\_stat** ma wartość FAŁSZ lub jest pominięty, funkcja REGLINP daje w wyniku tylko współczynniki  $\{a_n; a_{n-1}; \dots; a_1; a_0\}$ . Wyniki mieszczą się w jednym wierszu.

Funkcja REGLINP jest funkcją tablicową, więc wymaga wcześniejszego zaznaczenia miejsca na wynik i zatwierdzenia klawiszami **Ctrl+Shift+Enter**. Chociaż funkcja w swojej pełnej wersji może wyświetlać wyniki w pięciu wierszach, nie jest konieczne wyświetlanie ich wszystkich. Można na przykład zaznaczyć miejsce na wyświetlenie tylko trzech pierwszych wierszy, a zrezygnować z dwu ostatnich.

W przypadku równania z jedną zmienną współczynnik nachylenia i odciętą można odczytać z formuł:

Nachylenie  $a$ : = INDEKS(REGLINP(**znane\_y**; **znane\_x**); 1)

Punkt przecięcia z osią  $b$ : = INDEKS(REGLINP(**znane\_y**; **znane\_x**); 2)

## Relacja między współczynnikiem determinacji a kwadratem współczynnika korelacji

Należy zwrócić uwagę, że relacja:

$$\sum_{i=1}^N (y_i - \bar{y})^2 = \sum_{i=1}^N (\hat{y}_i - \bar{y})^2 + \sum_{i=1}^N (y_i - \hat{y}_i)^2$$

jest spełniona tylko wtedy, gdy wartości funkcji  $\hat{y}$  są wyznaczone z pełnego równania prostej z wyrazem wolnym, a współczynniki są obliczane zgodnie z zasadą NK.

Człony tego wzoru są interpretowane kolejno jako: całkowite zróżnicowanie zmiennej  $y$ , zróżnicowanie zmiennej  $y$  objaśnione modelem i zróżnicowanie zmiennej  $y$  nieobjaśnione modelem.

W oparciu o ten wzór określa się współczynnik determinacji  $D$  jako stosunek sum:

$$D = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2} = 1 - \frac{SS_{resz}}{SS_y}$$

Po spełnieniu warunków, o których była mowa wyżej, prawdziwa jest także równość:

$$D = \frac{\sum_{i=1}^N (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^N (y_i - \bar{y})^2}$$

Uwzględniając wartości współczynników  $a$  i  $b$  obliczone metodą NK, otrzymujemy:

$$D = \frac{\sum_{i=1}^N (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^N (y_i - \bar{y})^2} = \frac{\sum_{i=1}^N (ax_i + b - \bar{y})^2}{\sum_{i=1}^N (y_i - \bar{y})^2} = \frac{\sum_{i=1}^N (ax_i - a\bar{x})^2}{\sum_{i=1}^N (y_i - \bar{y})^2} = a^2 \frac{\sum_{i=1}^N (x_i - \bar{x})^2}{\sum_{i=1}^N (y_i - \bar{y})^2} = R^2$$

Zatem w tym przypadku współczynnik determinacji jest równy kwadratowi współczynnika korelacji.

Należy zauważyć, że według podanej definicji wartość **współczynnika determinacji zależy** od przyjętego **równania modelowego** opisującego dane, natomiast wartość **współczynnika korelacji zależy** tylko od samych **danych**.

## Współczynnik korelacji między zmienną $y$ a funkcją liniową $x$

Załóżmy, że mamy zbiór par wartości  $(x, y)$ . Dla tego zbioru możemy określić współczynnik korelacji pomiędzy wartościami  $x$  i  $y$ . Możemy obliczyć go za pomocą funkcji WSP.KORELACJI, a jego kwadrat za pomocą funkcji R.KWADRAT. Przekształcimy teraz wartości  $x$  w wartości  $\hat{y}$ , korzystając z dowolnej funkcji liniowej:

$$\hat{y} = ax + b$$

Obliczymy współczynnik korelacji między wartościami  $y$  i  $\hat{y}$ :

$$R(y, \hat{y}) = \frac{\sum_{i=1}^N (\hat{y}_i - \bar{\hat{y}})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (\hat{y}_i - \bar{\hat{y}})^2 \sum_{i=1}^N (y_i - \bar{y})^2}}$$

Mamy:

$$\hat{y}_i = ax_i + b \quad \text{i} \quad \bar{\hat{y}} = a\bar{x} + b$$

Zatem:

$$\hat{y}_i - \bar{\hat{y}} = (ax_i + b) - (a\bar{x} + b) = a(x_i - \bar{x})$$

Po podstawieniu otrzymamy:

$$R(y, \hat{y}) = \frac{a \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{a^2 \sum_{i=1}^N (x_i - \bar{x})^2 \sum_{i=1}^N (y_i - \bar{y})^2}} = \text{sgn}(a)R(x, y)$$

Widzimy zatem, że współczynnik korelacji dowolnej funkcji liniowej zmiennej  $x$  ze zmienną  $y$  jest taki sam jak współczynnik korelacji samych zmiennych  $x$  i  $y$  z dokładnością do znaku  $a$ . Kwadraty współczynników korelacji są równe. Zależność ta obowiązuje dla dowolnych  $a$  i  $b$ , a nie tylko tych wyznaczonych metodą NK.

## Równanie ze znanym wyrazem wolnym lub bez niego

Przypadkiem szczególnym jest model liniowy ze stałym współczynnikiem  $b$  (prosta przechodzi przez punkt stały) — wtedy zastosowanie metody NK prowadzi do poniższego wzoru na współczynnik kierunkowy  $a_N$ :

$$a_N = \frac{\sum_{i=1}^N x_i y_i - b \sum_{i=1}^N x_i}{\sum_{i=1}^N x_i^2} = \frac{\sum_{i=1}^N x_i (y_i - b)}{\sum_{i=1}^N x_i^2} = \frac{\overline{xy} - b\bar{x}}{\overline{x^2}}$$

Model procesu może nie zawierać w ogóle wyrazu wolnego, czyli  $b = 0$ , a prosta wówczas przechodzi przez początek układu współrzędnych. W tym przypadku wzór upraszcza się do:

$$a_N = \frac{\overline{xy}}{\overline{x^2}}$$

Jeżeli rozpatrujemy pełne równanie prostej  $y = ax + b$ , to zgodnie z logiką narzucenie warunku  $b = 0$  powoduje na ogół pogorszenie jakości dopasowania równania regresji do danych i w związku z tym należy oczekiwać niższych (w każdym razie nie wyższych) wartości współczynnika determinacji  $D$ . W pewnych przypadkach można otrzymać nawet wartości ujemne. Dotyczy to zwykle sytuacji, gdy dane są nieskorelowane lub słabo skorelowane, tzn. współczynnik korelacji  $R$  jest równy 0 lub bliski zera. Równocześnie współczynnik korelacji i jego kwadrat zachowują tę samą wartość (zależą tylko od danych, a nie od postaci równania). W każdym razie jeśli narzucimy wartość któregośkolwiek współczynnika lub dobierzemy współczynniki według algorytmu innego niż metoda NK, zamiast relacji  $D = R^2$  mamy  $D < R^2$ .

## Niescentrowany współczynnik determinacji

Aby uniknąć ujemnych wartości współczynnika  $D$ , w przypadku równania **bez wyrazu wolnego** zaproponowano użycie zmodyfikowanego współczynnika, zwanego w literaturze **niescentrowanym współczynnikiem determinacji**  $D_N$ :

$$D_N = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N y_i^2}$$

Wzór ten wynika z relacji:

$$\sum_{i=1}^N y_i^2 = \sum_{i=1}^N \hat{y}_i^2 + \sum_{i=1}^N (y_i - \hat{y}_i)^2$$

Jest ona spełniona tylko wtedy, gdy wartości funkcji  $\hat{y}$  są wyznaczone z równania prostej bez wyrazu wolnego, a współczynnik  $a_N$  jest obliczany zgodnie z zasadą NK.

Wzór opisujący  $D_N$  można uogólnić na przypadek dowolnej (lecz ustalonej) wartości wyrazu wolnego  $b$ :

$$D_N = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - b)^2}$$

Jeżeli wzór ten zastosujemy do modelu liniowego z ustaloną wartością wyrazu wolnego  $b$  i ze współczynnikiem  $a_N$  określonym według metody NK, otrzymamy:

$$D_N = \frac{\left( \sum_{i=1}^N x_i (y_i - b) \right)^2}{\sum_{i=1}^N x_i^2 \sum_{i=1}^N (y_i - b)^2} = \frac{(\overline{xy} - b\bar{x})^2}{x^2 \cdot (y^2 - 2b\bar{y} + b^2)}$$

W szczególnym przypadku, gdy  $b = 0$ , mamy:

$$D_N = \frac{\left( \sum_{i=1}^N x_i y_i \right)^2}{\sum_{i=1}^N x_i^2 \sum_{i=1}^N y_i^2} = \frac{(\overline{xy})^2}{\overline{x^2} \cdot \overline{y^2}}$$

Z powyższych zależności widać, że tak zdefiniowany współczynnik będzie przyjmował wartości z przedziału  $[0, 1]$ . Należy jednak zauważyć, że w przypadku arbitralnego równania modelowego współczynnik  $D_N$  również może przyjmować wartości ujemne.

## Związek między niescentrowanym a tradycyjnym współczynnikiem determinacji

Związek między niescentrowanym (przy  $b = 0$ ) a tradycyjnym współczynnikiem determinacji można określić, porównując wzory:

$$D = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2} = 1 - \frac{SS_{resz}}{SS_y} \quad \text{i} \quad D_N = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N y_i^2} = 1 - \frac{SS'_{resz}}{SS'_y}$$

Sumę z mianownika pierwszego wzoru można zapisać jako:

$$\sum_{i=1}^N (y_i - \bar{y})^2 = \sum_{i=1}^N y_i^2 - N\bar{y}^2 \quad \text{czyli} \quad SS_y = SS'_y - N\bar{y}^2$$

Z porównania wzorów wynika zależność:

$$\frac{1}{1-D} = \frac{1}{1-D_N} - \frac{N\bar{y}^2}{SS_{resz}}$$

## Niescentrowany współczynnik korelacji

Można zdefiniować wielkość zwaną **niescentrowanym współczynnikiem korelacji**  $R_N$  według wzoru:

$$R_N = \frac{\overline{xy}}{\sqrt{\overline{x^2} \cdot \overline{y^2}}} = \frac{\sum_{i=1}^N x_i y_i}{\sqrt{\sum_{i=1}^N (x_i)^2 \sum_{i=1}^N (y_i)^2}}$$

Postać wzoru jest analogiczna do wzoru definiującego współczynnik korelacji  $R$ , lecz jego sens jest inny. Podobnie jak w przypadku  $R$  wartość tego współczynnika mieści się w przedziale  $[-1, +1]$ . Wartości skrajne odpowiadają doskonałej zależności proporcjonalnej, odpowiednio ze współczynnikiem ujemnym lub dodatnim. Jeżeli wyraz wolny nie jest równy zeru, należałoby zastosować wzór:

$$R_N = \frac{\sum_{i=1}^N x_i (y_i - b)}{\sqrt{\sum_{i=1}^N x_i^2 \sum_{i=1}^N (y_i - b)^2}} = \frac{\overline{xy} - b\bar{x}}{\sqrt{\overline{x^2} \cdot \sqrt{y^2 - 2b\bar{y} + b^2}}}$$

Definicja współczynnika  $R_N$  zależy od wartości  $b$ , czyli od postaci równania regresji. Jeżeli współczynnik proporcjonalności  $a_N$  jest wyznaczony według metody NK, spełniona jest zależność  $D_N = R_N^2$ .

## Równanie ze znanym współczynnikiem kierunkowym

Analogicznie gdyby w modelu liniowym wartość współczynnika  $a$  była ustalona, wartość współczynnika  $b_M$  wyznaczona metodą NK wyniosłaby:

$$b_M = \frac{\sum_{i=1}^N y_i - a \sum_{i=1}^N x_i}{N} = \frac{\sum_{i=1}^N (y_i - ax_i)}{N} = \bar{y} - a\bar{x}$$

W takim przypadku obowiązuje relacja:

$$\sum_{i=1}^N (y_i - ax_i)^2 = \sum_{i=1}^N (\hat{y}_i - ax_i)^2 + \sum_{i=1}^N (y_i - \hat{y}_i)^2$$

Na jej podstawie można zdefiniować jeszcze inny zmodyfikowany współczynnik determinacji  $D_M$ , specyficzny dla tej postaci równania:

$$D_M = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - ax_i)^2}$$

Po podstawieniu wartości współczynnika  $b_M$  obliczonej metodą NK otrzymujemy:

$$D_M = \frac{\sum_{i=1}^N (\hat{y}_i - ax_i)^2}{\sum_{i=1}^N (y_i - ax_i)^2} = \frac{\sum_{i=1}^N (ax_i + b_N - ax_i)^2}{\sum_{i=1}^N (y_i - ax_i)^2} = \frac{\sum_{i=1}^N b_N^2}{\sum_{i=1}^N (y_i - ax_i)^2} = \frac{(\bar{y} - a\bar{x})^2}{(y - ax)^2} = R_M^2$$

Zmodyfikowany współczynnik korelacji  $R_M$ , który mógłby mieć zastosowanie w tym przypadku, wyraża się wzorem:

$$R_M = \frac{\bar{y} - a\bar{x}}{\sqrt{(y - ax)^2}}$$

Ten przypadek nie został uwzględniony w Excelu; przytaczam go, aby zwrócić uwagę, że do różnych postaci równania regresji można dopasować odpowiednio zmodyfikowane definicje współczynników korelacji i determinacji, jednakże nie ma pomiędzy nimi prostych zależności, które umożliwiałyby porównania.

Pomiędzy współczynnikami można zapisać relacje:

$$\frac{D}{a^2} = \frac{\overline{x^2} - \bar{x}^2}{\overline{y^2} - \bar{y}^2} = \frac{\overline{(x - \bar{x})^2}}{\overline{(y - \bar{y})^2}}, \quad \frac{D_N}{a_N^2} = \frac{\overline{x^2}}{\overline{(y - b)^2}} \quad \text{oraz} \quad \frac{D_M}{a^2} = \frac{(\bar{y}/a - \bar{x})^2}{\overline{(y - ax)^2}}$$

## Specyfika działania funkcji REGLINP przy braku wyrazu wolnego

W opisie działania funkcji REGLINP podano, że gdy wyraz wolny ma być równy **zeru**, zamiast współczynnika determinacji  $D = R^2$  oblicza się **niescentrowany** współczynnik determinacji  $D_N = R_N^2$ . Do obliczania wartości  $F$  i  $R^2$  zamiast sumy regresyjnej liczonej według wzoru:

$$SS_{regr} = \sum_{i=1}^N (y_i - \bar{y})^2 - SS_{resz} = SS_y - SS_{resz}$$

wykorzystuje się wielkość:

$$SS'_{regr} = \sum_{i=1}^N y_i^2 - SS_{resz} = SS'_y - SS_{resz}$$

Dotyczy to wersji 2003+. W wersji 2000 zawsze liczony jest tradycyjny współczynnik determinacji.

Przyjęcie takiego rozwiązania jest co najmniej dyskusyjne, gdyż prowadzi do różnych niekonsekwencji, o których mowa poniżej. Ponadto rozwiązanie to jest niezgodne z przyjętym w kreatorze wykresów. Gdy na wykresie, do którego dodano linię trendu, zażądamy wyświetlenia  $R^2$ , otrzymujemy w istocie wartości współczynnika determinacji  $D$  obliczone ze wzoru:

$$D = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2}$$

niezależnie od tego, czy wartość stałej  $b$  jest obliczana czy ustalona. Wartości  $D$  czasem bywają ujemne, ale nie powinno to stwarzać problemów interpretacyjnych, poza tym, że niefortunnie określono ten współczynnik mianem  $R^2$ . Należy pamiętać, że na wykresie można zażądać dowolnej wartości współczynnika  $b$ , a nie tylko wartości zerowej, jak w przypadku funkcji REGLINP. Co wtedy należałoby wyświetlić jako  $D$ ? Gdyby chcieć utrzymać konwencję przyjętą w funkcji REGLINP, należałoby użyć podanej wcześniej uogólnionej definicji niescentrowanego współczynnika determinacji:

$$D_N = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - b)^2}$$

W wersji 2000 funkcja REGLINP daje taką samą wartość współczynnika determinacji jak kreator wykresu, do którego dodano linię trendu.

Na rysunku 7.1 pokazano, na czym polega (od wersji 2003) problem z funkcją REGLINP. Obliczenia przy użyciu tej funkcji wykonano trzykrotnie. Raz jako **znane\_x** użyto jednej kolumny danych, z argumentem *czy\_stala* o wartości FAŁSZ — wówczas funkcja zwraca (moim zdaniem niesłusznie) niescentrowany współczynnik determinacji, natomiast na wykresie współczynnik determinacji obliczony jest poprawnie. W drugim przypadku dodano drugą kolumnę  $x$  złożoną z samych jedynek i obliczono dwa współczynniki nadal z argumentem *czy\_stala* o wartości FAŁSZ — tu również uzyskano niescentrowany współczynnik determinacji. W trzecim przypadku użyto jednej kolumny danych jako **znane\_x**, z argumentem *czy\_stala* o wartości PRAWDA — i uzyskano klasyczny współczynnik determinacji (poprawnie). Obliczone współczynniki równania prostej w drugim i trzecim przypadku są liczbowo równe, gdyż dotyczą faktycznie tych samych danych, różnica polega tylko na innym sposobie obliczania wskaźników statystycznych.

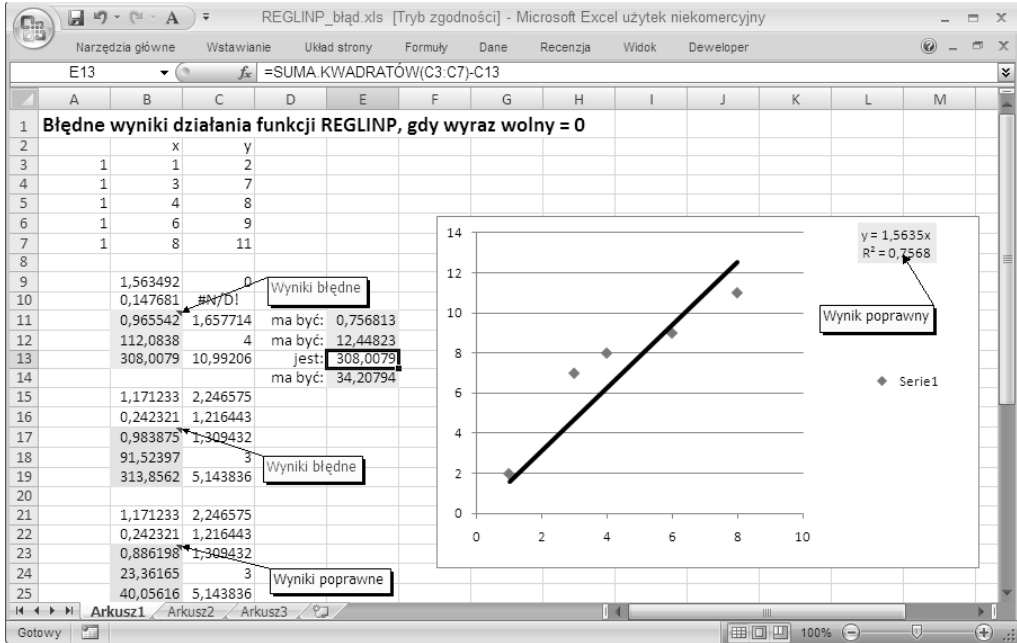
Reasumując, uważam, że sposób wyliczenia współczynnika determinacji zastosowany w kreatorze wykresów jest bardziej poprawny i spójny logicznie, a podejście zastosowane w funkcji REGLINP od wersji 2003 jest niekonsekwentne, prowadzi bowiem do niejednoznaczności i niezbyt logicznych wniosków dotyczących jakości dopasowania wzoru.

## Nietypowe zastosowania funkcji REGLINP

### Rozwiązanie układu równań liniowych

Kwestii rozwiązywania układów równań liniowych poświęciłem początek poprzedniego rozdziału. Przedstawiłem tam typowe sposoby rozwiązania tego zagadnienia: wykorzystanie rachunku macierzowego lub dodatku *Solver*. W matematyce znane są również inne sposoby osiągnięcia tego celu, takie jak na przykład zastosowanie wzorów Cramera czy metody eliminacji Gaussa, których użycie w Excelu wymaga bardziej złożonych obliczeń, dlatego ich omówienie pominąłem. Skoro jednak znamy już funkcję REGLINP, warto powrócić do tego zagadnienia. Istotą działania funkcji REGLINP jest przecież rozwiązanie układu równań liniowych, tylko tych równań jest





**RYSUNEK 7.1.** Przykład ilustrujący nieprawidłowe wartości parametrów statystycznych podawane przez funkcję REGLINP w przypadku, gdy wyraz wolny jest równy zero (od wersji 2003)

z reguły więcej niż niewiadomych, więc funkcja REGLINP nie może znaleźć dokładnego rozwiązania układu wszystkich równań, lecz szuka rozwiązania przybliżonego, najlepiej dopasowanego do danych. Jeśli jednak liczba równań jest równa liczbie niewiadomych, to może istnieć rozwiązanie dokładne (o ile układ jest oznaczony). Dla funkcji REGLINP jest to przypadek nietypowy, bo liczba stopni swobody jest równa zero, więc funkcja nie może obliczyć żadnych statystyk regresji, ale same współczynniki, które odpowiadają w tym przypadku naszym niewiadomym — obliczyć może. Jedyne, o czym musimy pamiętać, to odwrócona kolejność niewiadomych w wynikach. Jak już pisałem, jest to specyfika funkcji REGLINP, którą zresztą też można wykorzystać (zobacz następny podpunkt).

Jeśli mamy do rozwiązania układ równań (w zapisie macierzowym):

$$A \times X = B$$

i chcemy go rozwiązać za pomocą funkcji REGLINP, musimy potraktować macierz współczynników  $A$  jako **znane\_x**, natomiast wektor wyrazów wolnych  $B$  jako **znane\_y**. Należy przyjąć, że funkcja nie oblicza wyrazu wolnego, dlatego trzeci argument funkcji REGLINP powinien być równy FAŁSZ (lub 0), a czwarty argument można pominąć (statystyka jest w tym przypadku nieistotna). Wyniki otrzymamy w wierszu (przy tradycyjnym sposobie rozwiązania była to kolumna) i w kolejności odwróconej w stosunku do kolejności kolumn w macierzy  $A$ .

Na przykład jeśli mamy układ czterech równań z czterema niewiadomymi  $x$ ,  $y$ ,  $z$ ,  $w$  i kolejność kolumn w macierzy  $A$  odpowiada tej kolejności niewiadomych, to w wyniku funkcja REGLINP

zwróci nam wiersz z wynikami w kolejności:  $w$ ,  $z$ ,  $y$ ,  $x$ . Ilustrację tego sposobu rozwiązania stanowią rysunek 7.2. W obliczeniach wykorzystano ten sam układ równań, który był pokazany na rysunku 6.1.

**RYСУNEK 7.2.**  
Rozwiązanie układu  
równań liniowych  
za pomocą funkcji  
REGLINP

	A	B	C	D	E	F	G	H	I
1		<b>Układ równań liniowych</b>							
2									
3									
4									
5									
6									
7									
8									
9									
10									
11									
12									
13									
14									
15									
16									
17									
18									

## Odwracanie kolejności elementów w zakresie lub tablicy

To zagadnienie również było już wcześniej poruszane. W rozdziale 6. pokazałem, jak można je rozwiązać za pomocą funkcji INDEKS i PRZESUNIĘCIE. Temat będzie jeszcze kontynuowany w rozdziale 20. Z przedstawionych opisów wynika, że stosunkowo łatwo można zapisać formułę generującą rozwiązanie sekwencyjnie (formuła wymaga kopiowania do kolejnych komórek). Natomiast uzyskanie rozwiązania w formie wielokomórkowej formuły tablicowej zwracającej od razu cały wynik przydatny do dalszych obliczeń jest trudniejsze i wymaga użycia sposobów niestandardowych.

Tymczasem, jak wiadomo, funkcja REGLINP została tak napisana, że zwraca wyniki w odwrotnej kolejności. Zapiszmy przykładowy układ równań w postaci:

$$1 \cdot x + 0 \cdot y + 0 \cdot z = 2$$

$$0 \cdot x + 1 \cdot y + 0 \cdot z = 5$$

$$0 \cdot x + 0 \cdot y + 1 \cdot z = 1$$

W formalnym zapisie macierzowym można to zapisać jako:

$$I \cdot X = B$$

$I$  oznacza macierz jednostkową rzędu równego liczbie niewiadomych. Oczywistym rozwiązaniem tego równania jest:

$$X = B$$

Możemy zastosować do tego układu równań metodę opisaną w poprzednim podpunkcie. Jeśli jako argumentu *znane\_x* użyjemy macierzy jednostkowej, otrzymamy wynik w postaci wektora **B**, czyli *znane\_y*, w odwróconej kolejności. Ponadto, ponieważ macierz jednostkowa jest symetryczna, wektor przetwarzany (*znane\_y*) może być zarówno w układzie pionowym, jak i poziomym. Natomiast wynik otrzymamy zawsze w układzie poziomym. W razie potrzeby można zastosować do niego transpozycję. Wykorzystanie tej metody ogranicza się wyłącznie do danych **liczbowych**.

W Excelu 2013+ sprawa jest prosta, bo istnieje funkcja generująca macierz jednostkową i wystarczy zapisać:

```
= REGLINP(znane_y; MACIERZ.JEDNOSTKOWA(ILE.NIEPUSTYCH(znane_y))); FAŁSZ)
```

We wcześniejszych wersjach programu trzeba wygenerować macierz jednostkową odpowiednią formułą [A4].

Jeśli *znane\_y* składa się z wartości unikatowych, wystarczy taki zapis:

```
= REGLINP(znane_y; (znane_y = TRANSPONUJ(znane_y))*1; FAŁSZ)
```

Gdy przewidujemy możliwość wystąpienia duplikatów, przydatne może być zdefiniowanie i nazwanie tablicy bazowej:

```
wd : = WIERSZ(ADR.POŚR("1:"&ILE.NIEPUSTYCH(znane_y)))
```

i wykorzystanie jej w formule analogicznej do poprzedniej:

```
= REGLINP(znane_y; (wd = TRANSPONUJ(wd))*1; FAŁSZ)
```

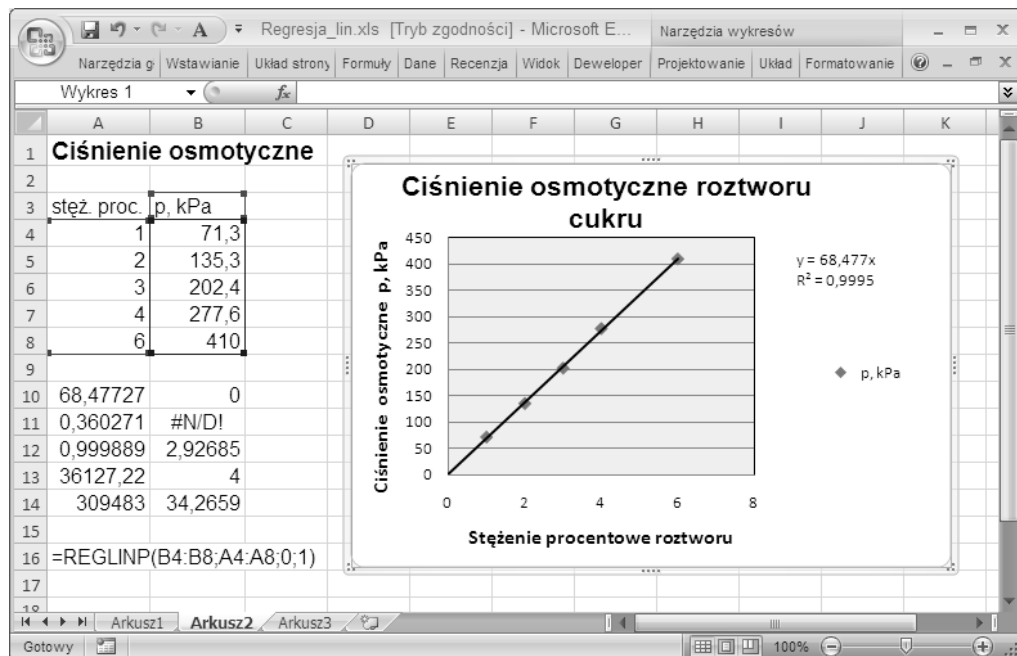
Wyniki uzyskane przy zastosowaniu funkcji MACIERZ.JEDNOSTKOWA są tablicami **TPW**, natomiast wyrażenia zawierające funkcję TRANSPONUJ wymagają w przypadku wykorzystania do dalszych obliczeń zatwierdzenia tablicowego.

## Przykłady wyznaczania równań regresji

### Przykład 7.1.

Dana jest zależność ciśnienia osmotycznego roztworu cukru od jego stężenia procentowego. Wykonamy wykres tej zależności i dopasujemy linię trendu — na wykresie wyświetlimy jej równanie i kwadrat współczynnika korelacji. Wykres powinien przechodzić przez początek układu współrzędnych.

Współczynnik proporcjonalności i współczynnik determinacji można wyznaczyć również za pomocą funkcji REGLINP. Niestety, jak już pisałem, w tym przypadku od wersji 2003 funkcja zwraca kwadrat niescentrowanego współczynnika korelacji zamiast tradycyjnego współczynnika. Z kolei w wersji 2007 Excela w przypadku korelacji wielomianem niekiedy nie jest wyświetlany współczynnik przy najwyższej potędze *x*. (We wcześniejszych wersjach ten błąd nie występował). Dane i wyniki pokazano na rysunku 7.3. Dane do przykładu pochodzą z książki [10].



**RYSUNEK 7.3.** Zależność ciśnienia osmotycznego roztworu cukru od jego stężenia procentowego jako przykład regresji liniowej

Przy prowadzeniu obliczeń dotyczących regresji, począwszy od wersji 2007, można dodatkowo wykorzystać tabele Excela wraz z odwołaniami strukturalnymi. Nie uzyskamy przez to jakichś szczególnych ułatwień, lecz unikniemy kopiowania w kolumnach, które będzie się odbywać automatycznie, i nieco szybciej wprowadzimy nagłówki kolumn niż nazwy odpowiednich zakresów. Ponadto jeśli w jednym arkuszu chcemy umieścić kilka podobnych obliczeń, łatwiej zachować podobne schematy oznaczeń ze względu na to, że nagłówki kolumn w różnych tabelach mogą się powtarzać, a nazwy zakresów muszą być unikatowe w arkuszu.

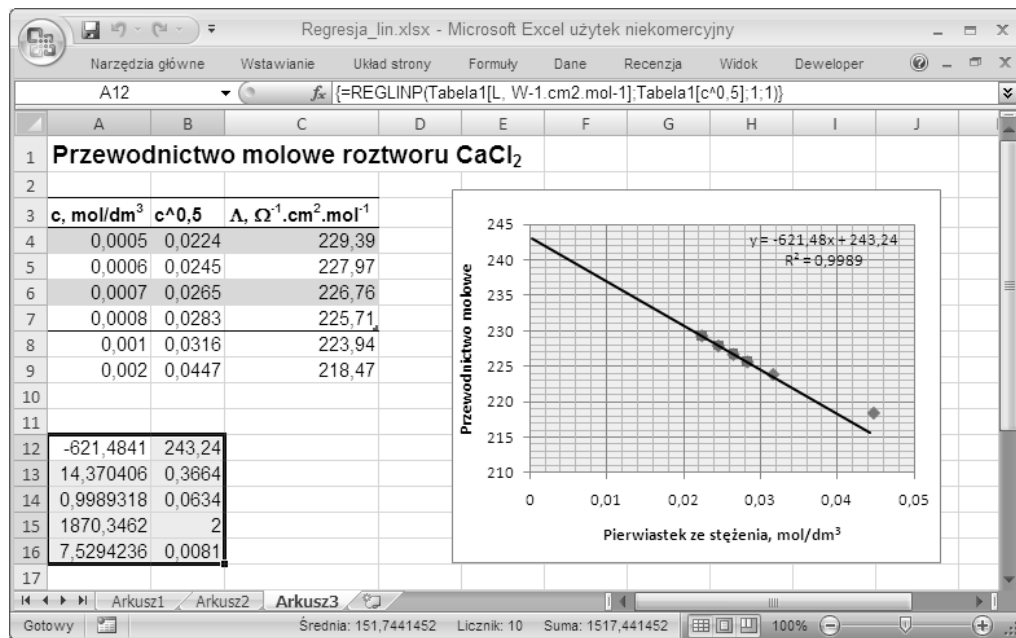
### Przykład 7.2.

Dana jest zależność przewodnictwa molowego rozcieńczonych roztworów chlorku wapnia od stężenia molowego w temperaturze 25°C. (Dane pochodzą z książki [10]). Zgodnie z teorią rozcieńczonych roztworów elektrolitów przewodnictwo molowe powinno być liniową funkcją pierwiastka ze stężenia:

$$\Lambda = \Lambda_0 - k\sqrt{C}$$

Należy zweryfikować to równanie, obliczając jednocześnie jego współczynniki. Jeżeli chcemy wyznaczyć współczynniki na podstawie równania linii trendu z wykresu, to trzeba przygotować w arkuszu kolumnę z wartościami pierwiastka ze stężenia i wykorzystać tę kolumnę jako zmienną niezależną. W standardowym zestawie funkcji opisujących linię trendu nie przewidziano bowiem zależności tego typu. Korzystając z funkcji REGLINP, nie musimy przygotowywać kolumn pomocniczych. Można zastosować notację tablicową i jako *znane\_x* podać PIERWIASTEK(stężenie) lub  $\text{stężenie}^0,5$ .

Dane i wyniki pokazano na rysunku 7.4. Zależność liniowa jest dobrze spełniana przez cztery pierwsze punkty, dlatego tylko te punkty wykorzystano do obliczeń.



**RYСУNEK 7.4.** Zależność przewodnictwa molowego rozcieńczonych roztworów chlorku wapnia od stężenia molowego w temperaturze 25 °C jako przykład regresji liniowej ze zmianą zmiennej niezależnej

### Przykład 7.3.

Adsorpcję par na stałych adsorbentach opisuje się często równaniem izotermy Langmuira:

$$x = \frac{abp}{1 + bp}$$

gdzie:

$x$  — ułamek wagowy określający ilość substancji zaadsorbowanej na jednostkę masy sorbenta;

$p$  — cząstkowa prężność pary substancji adsorbowanej w równowadze z adsorbentem.

Równanie izotermy Langmuira jest nieliniowe i przed jego opracowaniem metodą regresji liniowej trzeba je przekształcić do postaci liniowej względem nowych zmiennych.

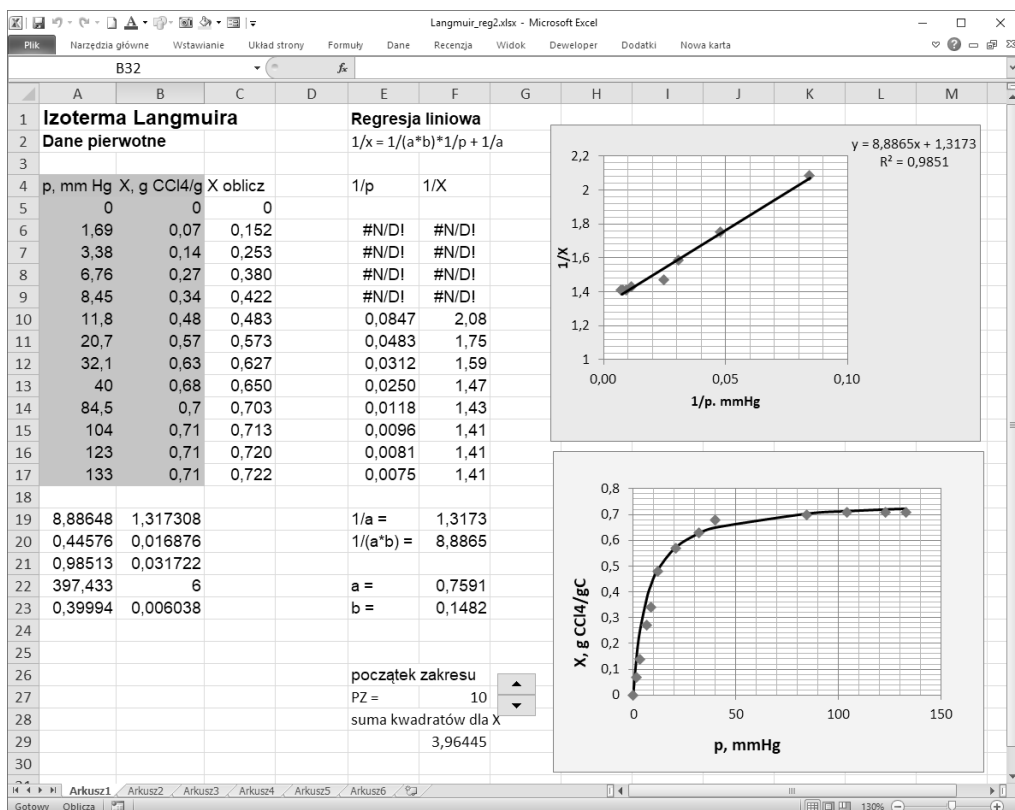
Sposoby przekształcenia podałem na początku rozdziału. Podobnie jak w poprzednim przykładzie, można albo przekształcić zmienne w arkuszu w sposób jawny, wykonać wykres linearyzowany i odczytać współczynniki z równania linii trendu, albo zapisać przekształcenie zmiennych w argumentach funkcji REGLINP i odczytać współczynniki z wyniku funkcji.

Należy pamiętać, że przekształcenie zmiennych zmienia wagi poszczególnych punktów pomiarowych, na przykład po obliczeniu  $1/x$  małe wartości  $x$  zyskują na znaczeniu, a duże tracą.

Po wstawieniu obliczonych współczynników do pierwotnego równania jakość dopasowania równania ze względu na początkowe zmienne może być niezadowalająca. W takim przypadku pomaga czasem rezygnacja z kilku skrajnych punktów.

Przykładowe wyniki obliczeń przedstawiono na rysunku 7.5. Dane do tego przykładu pochodzą z książki [7]. W zaprezentowanym przykładzie przy obliczaniu współczynników równania linearyzowanego pominięto cztery pierwsze punkty, aby uzyskać zadowalające wizualnie dopasowanie linii trendu do punktów z końca zakresu — po powrocie do pierwotnych zmiennych. Aby można było łatwo modyfikować zakresy punktów branych pod uwagę w obliczeniach, zdefiniowano komórkę przechowującą numer wiersza z początkowymi danymi. Komórce tej nadano nazwę **PZ** (komórka F27). Wszystkie adresy potrzebne do zdefiniowania zakresów roboczych są konstruowane z użyciem funkcji ADR.POŚR, której argumentem jest adres w postaci tekstowej. Tekst ten może być złożony z elementów oznaczających kolumnę i wiersz. Alternatywnie można użyć funkcji ADRES, która również pozwala na konstrukcję adresu w formie tekstowej z numerów wiersza i kolumny. Jeżeli na przykład chcemy odwołać się do komórki z kolumny B (jest to druga kolumna) i z wiersza o numerze pobranym z komórki **PZ**, to możemy użyć jednego z wyrażeń:

$$= \text{ADR.POŚR}(\text{"B"}\&\text{PZ}) \text{ lub } = \text{ADR.POŚR}(\text{ADRES}(\text{PZ};2))$$



**RYСУNEK 7.5.** Stopień adsorpcji par czterochloru węgla na węglu aktywnym jako przykład regresji według równania linearyzowanego

Funkcja `ADR.POŚR` wymaga argumentu tekstowego, dlatego użyto operatora połączenia tekstów `&`. Zawartość komórki `PZ` jest liczbą, ale zostanie automatycznie skonwertowana na tekst. Funkcja `ADRES [A3]` wymaga co najmniej dwu argumentów liczbowych: pierwszym jest numer wiersza, drugim numer kolumny. Można podać jeszcze argumenty opcjonalne, ale nie będą nam tu potrzebne.

Do wyznaczenia wartości zmiennych pomocniczych  $1/p$  i  $1/X$  użyto wyrażeń typu:

```
= JEŻELI(WIERSZ())>=PZ;1/A6;BRAK()
```

Funkcja `WIERSZ()` bez argumentów zwraca numer wiersza z miejsca wywołania. Wyrażenie oblicza wartości  $1/p$ , w tym przypadku  $1/A6$ , jeżeli wywołanie jest w wierszu o numerze `PZ` lub większym, w przeciwnym razie zwraca kod `#N/D!`.

Współczynniki w linearyzowanym równaniu regresji ( $1/ab$  i  $1/a$ ) obliczono ze wzoru:

```
= REGLINP(1/ADR.POŚR("B"&PZ):B17;1/ADR.POŚR("A"&PZ):A17;PRAWDA;PRAWDA)
```

Jest to oczywiście formuła tablicowa.

W analogiczny sposób można przeprowadzić obliczenia dla pozostałych wzorów linearyzujących izotermę Langmuira.

### Przykład 7.4.

Często dane doświadczalne zawierają punkty, co do których można mieć wątpliwości, czy wartości zostały poprawnie zmierzone. Chcielibyśmy móc sprawdzić, jak zmieniłyby się wyniki regresji po wyeliminowaniu wątpliwych punktów. Można te punkty trwale wyeliminować, ale korzystniej byłoby mieć możliwość ich tymczasowego (odwracalnego) zamaskowania. Gdy korzystamy z regresji (linii trendu) na wykresie, sprawa jest dość prosta: wystarczy przenieść wartość  $y$  do sąsiedniej kolumny, pozostawiając pustą komórkę — równanie linii trendu zostanie wyznaczone z pominięciem usuniętych punktów. Jednak kiedy obliczamy współczynniki równania regresji lub wartości prognozowane za pomocą funkcji, sprawa nie jest już taka prosta. Funkcje `REGLINP` i `REGLINW` (oraz podobne) nie tolerują pustych komórek w zakresach danych, natomiast funkcje `NACHYLENIE`, `ODCIĘTA`, `R.KWADRAT` i `REGLINX` radzą sobie z tym problemem doskonale, ale za to zakres ich zastosowania nie jest tak szeroki jak wymienionych wcześniej funkcji. (Opisy funkcji `REGLINW` i `REGLINX` podałem w następnym rozdziale).

Jeśli chcemy zamaskować niektóre punkty w zakresie danych funkcji `REGLINP` lub `REGLINW`, musimy zastosować procedurę bardziej złożoną i utworzyć nowe ciągłe zakresy danych, w których zamaskowane punkty będą pominięte.

Zakresom pierwotnych danych nadajemy nazwy  $x$  i  $y$ . Zakładamy, że dane są zapisane w kolumnach. Wartości  $y$ , które mają być zamaskowane, przenosimy do sąsiedniej kolumny. Sprawdzamy liczebność danych (`ILE.LICZB(y)`) i tworzymy listę z numerami wierszy w zakresie  $y$ . Wykorzystujemy do tego znaną konstrukcję z funkcją `WIERSZ` (tablicę bazową):

```
z : = WIERSZ(ADR.POŚR("1:"&ILE.LICZB(y)))
```

Ustalamy położenie zamaskowanych punktów w zakresie  $y$ . W tym celu użyjemy funkcji PODAJ.  
 ↪ POZYCJĘ. Pierwszej pustej komórki szukamy za pomocą formuły tablicowej:

$$= \text{PODAJ.POZYCJĘ}(\text{""; } y \& \text{""; } 0)$$

Alternatywnie można użyć formuły zwykłej z funkcją INDEKS:

$$= \text{PODAJ.POZYCJĘ}(\text{""; INDEKS}(y \& \text{""; }); 0)$$

Pusta komórka musi być poszukiwana jako pusty tekst, po przekształceniu zakresu  $y$  w teksty. Służy do tego operacja łączenia zakresu  $y$  z pustym tekstem.

Aby znaleźć położenie kolejnych pustych komórek, trzeba zmodyfikować zakres do przeszukania tak, by zaczynał się za znalezioną pustą komórką. Odpowiednia formuła (tablicowa) może mieć postać:

$$= \text{PODAJ.POZYCJĘ}(\text{""; PRZESUNIĘCIE}(y;p;0) \& \text{""; } 0) + p$$

W tej formule  $p$  oznacza odwołanie (względne) do pozycji poprzedniej pustej komórki. Wzór można kopiować w dół lub w prawo; każde jego użycie wskazywać będzie kolejne puste komórki. Gdy uzyskamy wynik większy od długości zakresu  $y$ , będzie to oznaczać koniec zakresu i zarazem brak dalszych pustych komórek.

Wzór można też zmodyfikować tak, aby po osiągnięciu końca zakresu sygnalizował błąd:

$$= \text{JEŻELI.BŁĄD}(\text{PODAJ.POZYCJĘ}(\text{""; PRZESUNIĘCIE}(y;p;0); \text{ILE.WIERSZY}(y) - p) \& \text{""; } 0) + p; \text{"koniec"}$$

Dalej można wykorzystać funkcję MACIERZ.ILOCZYN. Jeśli dane  $x$  i  $y$  są w kolumnach, to numery punktów, które chcemy zamaskować, umieścimy w wierszu i nazwiemy ten wiersz  $wp$ . Zdefiniujemy nową tablicę bazową (pionową) opartą na numerach punktów zamaskowanych i nadamy jej nazwę  $wz$ :

$$wz : = \text{WIERSZ}(\text{ADR.POŚR}(\text{"1:"} \& \text{ILE.LICZB}(wp)))$$

Po zdefiniowaniu formuły indeksowej:

$$fi : = z + \text{MACIERZ.ILOCZYN}(\text{—}(z > wp - \text{TRANSPONUJ}(wz)); wz^0)$$

wartości współrzędnych punktów niezamaskowanych otrzymamy ze wzorów:

$$\begin{aligned} &= \text{INDEKS}(x; fi) \\ &= \text{INDEKS}(y; fi) \end{aligned}$$

Te wyniki mają postać tablic TDW. Piszę o tym dokładniej w rozdziale 20. Aby użyć ich jako argumentów funkcji REGLINP lub REGLINW, należy wykorzystać jeden ze sposobów podanych we wspomnianym rozdziale.

Według sposobu 1. można użyć zapisu:

$$= \text{REGLINP}(+\text{INDEKS}(y; fi); \text{INDEKS}(x; fi)^{\{1;2\}}; ; \text{PRAWDA})$$

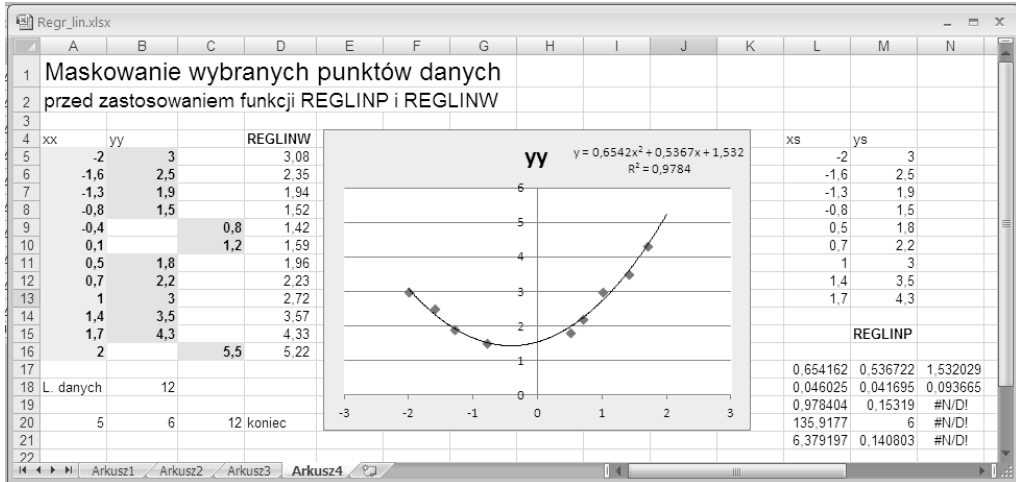
Według sposobu 2. należy obudować wyrażenie indeksowe  $fi$  funkcją L (lub N w wersji 2010+):

$$= \text{REGLINP}(\text{INDEKS}(y; L(fi)); \text{INDEKS}(x; L(fi))^{\{1;2\}}; ; \text{PRAWDA})$$



Funkcja INDEKS w takiej formie zwraca wynik pełnowartościowy, który może zostać użyty jako argument funkcji REGLINP, ale lepiej nadać formule nazwę, by potem móc ją wykorzystać również do zdefiniowania wykresu.

Przykład obliczeń wykonanych w ten sposób pokazano na rysunku 7.6.



**RYСУNEK 7.6.** Maskowanie wybranych punktów pomiarowych w celu umożliwienia obróbki danych za pomocą funkcji REGLINP i REGLINW

Można ten problem rozwiązać również inaczej. Po usunięciu wybranych punktów z zakresu  $y$  tworzymy listę indeksów pozostałych (nieusuniętych) punktów. Użyjemy do tego formuły, której można nadać nazwę  $zs$ :

$$zs := \text{MIN.K}(\text{JEŻELI}(y <> " "; \text{WIERSZ}(y) - \text{KOMÓRKA}(\text{"wiersz"; } y) + 1); \\ \text{WIERSZ}(\text{ADR.PÓŚR}(\text{"1:" & \text{ILE.LICZB}(y))))$$

Formuła ta zastosowana jako argument funkcji INDEKS utworzy listę współrzędnych nieusuniętych punktów. Listy tej można użyć w funkcji REGLINP. Przykładowo wywołanie funkcji może mieć postać:

$$= \text{REGLINP}(\text{INDEKS}(y; zs) * 1; \text{INDEKS}(x; zs) \wedge \{1; 2\}; ; \text{PRAWDA})$$

Wynikiem funkcji mają być współczynniki równania kwadratowego. Pierwszy argument funkcji REGLINP jest pomnożony przez 1, aby wyeliminować błąd wynikający z zastosowania funkcji INDEKS z argumentem tablicowym. W przypadku drugiego argumentu rolę tę spełnia podniesienie do potęgi.

### Przykład 7.5.

Do problemu przedstawionego w poprzednim przykładzie można podejść również z drugiej strony. Zamiast usuwać niepotrzebne punkty można stworzyć listę numerów (indeksów) punktów, które chcemy pozostawić do analizy lub do prezentacji na wykresie. Sposób ten będzie wygodniejszy, gdy mamy stosunkowo dużo punktów do dyspozycji i chcemy wybrać tylko niektóre. Wyboru

można dokonać tradycyjnie, za pomocą funkcji INDEKS, ale takich przykładów było już dużo, więc tym razem wykorzystamy pseudozakresy. Wyrażenia te omawiam dość szczegółowo w rozdziale 21. Celem jest uzyskanie formuły pełnowartościowej, którą można wykorzystać zarówno do zdefiniowania wykresu, jak i do obliczeń za pomocą funkcji z grupy REGLINP.

Załóżmy, że pierwotne dane znajdują się w kolumnach, których pierwsze komórki oznaczono nazwami **px** i **py**. Utworzymy dodatkową kolumnę, **wyb** (można ją zdefiniować z zapasem), przeznaczoną do wpisywania indeksów punktów, które mają być uwzględnione w obliczeniach. Pierwszą komórkę tego zakresu oznaczymy jako **pwymb**. Należy teraz zdefiniować w *Menedżerze nazw* kilka formuł pomocniczych. Pierwsza definiuje listę wybranych punktów:

**lista** := PRZESUNIĘCIE(**pwymb**; ; ; ILE.NIEPUSTYCH(**wyb**)) - 1

Zakładam, że indeksy punktów w kolumnie **wyb** będą liczone od 1, jak w funkcji INDEKS, ponieważ jednak zamierzam wykorzystać funkcję PRZESUNIĘCIE, która liczy od 0, stąd potrzeba odjęcia jedynki w powyższym wzorze.

Jak już wspomniałem, do utworzenia roboczych (przefiltrowanych) tablic wartości **x** i **y** wykorzystamy pseudozakresy bazujące na funkcji PRZESUNIĘCIE. Mamy do wyboru trzy równoważne formuły, analogiczne do obu zmiennych. Oto te trzy warianty dla zmiennej **x** (w przypadku zmiennej **y** wystarczy zamienić **px** na **py**):

**x.1** := L(PRZESUNIĘCIE(**px**; **lista**; 0))

**x.2** := SUMA.JEŻELI(PRZESUNIĘCIE(**px**; **lista**; 0); "<1E100")

**x.3** := SUMY.CZĘŚCIOWE(9; PRZESUNIĘCIE(**px**; **lista**; 0))

W wariantcie drugim należy podać warunek, który będzie spełniony przez wszystkie wybrane punkty.

Tak zdefiniowane tablice **x** i **y** mogą zostać wykorzystane zarówno w formułach, jak i przy określaniu danych do wykresów.

Jeszcze jeden sposób rozwiązania tego problemu przedstawiam w następnym rozdziale, w podrozdziale „Regresja danych o niejednakowej dokładności”.

# Skorowidz

## A

- adres
  - mieszany, 312
  - w postaci tekstowej, 188
  - względny, 58, 313, 314
- aproksymacja
  - dokładność, 260, 261
  - metoda
    - $\frac{3}{8}$ , 259, 260, 263
    - kwadratur, 258, 260
    - parabol, 259, 260, 263, 271
    - prostokątów, 258, 260, 261, 262, 263
    - Simpsona, 258, 259, 261, 262, 263, 267, 271
    - trapezów, 258, 260, 261, 262, 263, 267, 274

## B

- błąd, 244
  - #ADR!, 482
  - #ARG!, 483, 507
  - #LICZBA!, 97, 561
  - #N/D!, 173, 482, 525, 561
  - bezwzględny, 81
  - losowy, 563
  - maksymalny, 81, 82
  - słupek, *Patrz:* słupek błędu
  - standardowy, 42
  - średni
    - kwadratowy, 84
    - oszacowania wartości, 171, 217
  - względny, 81, 82, 88
    - maksymalny, 82
    - różnicy, 87
    - sumy, 87
    - zaokrąglenia, 92
- B-splajn, *Patrz:* splajn

## C

- całka
  - Arrheniusa, 278, 279, 280, 302
  - niewłaściwa, 277
    - ekstrapolacja, 277
    - przekształcenie w całkę właściwą, 277, 278
  - obliczanie granicy całkowania, 281, 282, 283, 286
  - oznaczona, 257, 258, 260, 262, 266, 291
    - iteracja, 274, 275
    - odwołanie cykliczne, 274, 275
    - zmiana granic, 270, 271
  - temperaturowa, *Patrz:* całka Arrheniusa
- ciąg liczbowy, 289
- cyfra
  - dokładna, 85
  - znacząca, 85, 90, 100
- czas, 101, 102, 411, 423, 426
- dokładność, 413
- konwersja
  - jednostek, 433
  - ułamki dziesiętne, 439
- operacje, 431
- wprowadzanie, 412
- zapis, 412, 413
  - minuty w indeksie górnym, 440
  - niestandardowy, 437

## D

- dane
  - do wykresu punktowego, *Patrz:* wykres punktowy dane
  - duplikat, 559
  - o niejednakowej dokładności, 210

## dane

- pomiarowe, 72, 73, 110
  - charakterystyka, 168, 171
  - obciążone błędami, 244
- standaryzowanie, 168

## data, 102, 411

- długa, 414, 415
- krótka, 414, 415
- obliczanie
  - początku następnego tygodnia, 434
  - różnic, 433
- operacje, 431, 433, 434, 435
- wprowadzanie, 414, 416
- wyświetlanie
  - dnia tygodnia, 436
  - nazwy miesiąca, 436
  - numeru dnia tygodnia, 436
  - numeru miesiąca, 436

## dokładność

- obliczeń, 90, 91, 92
- pomiaru, 81

**F**

## formuła

- arkuszowa, 35
- nazwana, 35, 58, 59, 269, 317, 319, 553, 554
- tablicowa, 292, 501
  - z funkcją ADR.POŚR, 542, 544
  - z funkcją PRZESUNIĘCIE, 545
- w postaci tekstowej, *Patrz:* wyrażenie w postaci tekstowej
- warunkowa cykliczna, 568
- zatraskowa, 563, 565, 566
  - faza czuwania, 566
  - faza inicjacji, 566, 568
  - wykonanie wielokrotne, 567

## funkcja

- ADR.POŚR, 188, 189, 507, 509, 510, 514, 527, 542, 544
- ADRES, 188
- AGREGUJ, 507, 508, 518, 552
- agregująca, 380, 483, 502, 503, 508, 523, 526, 546
- aproksymacja wielomianem, 110
- aproksymująca, 111, 113
- BD.MAX, 532
- BIT.PRZESUNIĘCIE.W.LEWO, 573
- BIT.PRZESUNIĘCIE.W.PRAWO, 573

## bitowa, 573

- celu, 154
- CENA.DZIES, 438, 439
- CENA.UŁAM, 437
- ciągła, 59, 105
  - wykres, *Patrz:* wykres funkcji ciągłej
- COMPLEX, 116, 117
- CONVERT, 387, 389, 390, 433
- CZAS, 426, 431
- CZAS.WARTOŚĆ, 427
- CZĘSTOŚĆ, 360, 463, 508, 529
- CZĘŚĆ.ROKU, 431
- CZY.ADR, 479, 507, 509
- CZY.LICZBA, 380
- CZY.TEKST, 380
- DATA, 102, 417, 418, 419, 420, 431, 432
- DATA.WARTOŚĆ, 417, 420
- DEC2BIN, 385
- DEC2HEX, 385
- DŁ, 443, 447
- DNI, 421
- DNI.360, 421
- DNI.ROBOCZE, 423
- DNI.ROBOCZE.NIESTAND, 424
- DOLLARDE, 438
- DOLLARFR, 437
- dwu zmiennych, 68
  - interpolacja, 251, 252, 254
- DZIEŃ, 421, 431
- DZIEŃ.ROBOCZY, 429, 435
- DZIEŃ.ROBOCZY.NIESTAND, 430
- DZIEŃ.TYG, 427
- DZIES.NA.DWÓJK, 385
- DZIES.NA.SZESN, 385
- DZIŚ, 427
- ekstremum, 133
- EOMONTH, 422
- Evaluate, 60, 270
- EXP, 91, 291
- FACTDOUBLE, 300
- FRAGMENT.TEKSTU, 443, 444, 448, 449, 450, 456
- GODZINA, 423, 431
- homograficzna, 166
- ILE.LICZB, 168, 479, 483, 507, 508
- ILE.NIEPUSTYCH, 479, 483, 507, 508
- ILE.WIERSZY, 479, 483, 508, 529
- ILOCZYN, 508

- INDEKS, 62, 311, 507, 515, 516, 529, 536, 538,  
539, 541, 546  
błąd, 482  
odwracanie porządku elementów  
w tablicy, 500  
odwracanie porządku elementów  
w zakresie, 499  
tworzenie ciągłego zakresu, 499  
z argumentami indeksowymi w postaci  
tablic, 477, 479, 480, 481, 482, 500
- ISO.NUM.TYG, 429  
ISO.ZAOKR.W.GÓRĘ, 98  
JEŻELI, 116, 117, 478, 507, 508, 514, 550  
KOD, 444  
KOMBINACJE, 362  
KOMBINACJE.A, 362  
KOMÓRKA, 38, 58, 507, 508, 514, 533, 534  
KONWERTUJ, 387, 389, 390, 392, 433  
KOWARIANCJA, 170, 174  
KOWARIANCJA.POPUL, 170  
KOWARIANCJA.PRÓBK, 170  
KWARTYL, 508, 529  
KWOTA, 96, 460  
L, 507, 508, 511, 513  
LEWY, 443, 444, 448, 465  
LICZ.JEŻELI, 507, 508  
LICZ.PUSTE, 507, 508  
LICZ.WARUNKI, 508  
LICZBA.CAŁK, 93, 95, 439  
LICZBA.KOLUMN, 479, 508, 529  
LICZBA.ZESP, 116, 117  
LITERY.MAŁE, 444, 456  
LITERY.WIELKIE, 444, 457  
logarytmiczna, 157  
LOS, 508, 535, 562, 563, 564  
LOS.ZAKR, 562, 563  
MACIERZ.ILOCZYN, 136, 147, 149, 546, 552  
MACIERZ.JEDNOSTKOWA, 185  
MACIERZ.ODW, 147  
makr, 64, 269, 276, 507, 508  
MAKS.WARUNKÓW, 508  
MAX, 508, 550  
MAX.K, 508, 529  
MEDIANA, 549, 550  
MIESIĄC, 102, 423, 431  
MIN, 508, 550  
MIN.K, 508, 529  
MIN.WARUNKÓW, 508  
minimum, 154  
MINUTA, 423, 431  
MOD, 102, 439, 551  
MODUŁ.LICZBY, 116, 550  
modułowa, 17  
MULTINOMIAL, 360, 362  
N, 507, 508, 511, 513  
NACHYLENIE, 173, 174, 189  
NETWORKDAYS, 423  
nieciągła, 58, 61, 556  
NR.KOLUMNY, 503, 508, 510, 545  
NR.SER.DATY, 422, 432  
NR.SER.OST.DN.MIES, 422  
NUM.TYG, 428  
O.FORMULE, 507, 508, 514, 536  
O.KOMÓRCE, 507, 508, 514, 536  
OBLICZ, 270, 275  
OBSZARY, 507, 508  
OCT2DEC, 385  
OCZYŚĆ, 444, 446  
ODCH.KWADRATOWE, 168  
ODCH.STAND.POPUL, 169  
ODCH.STANDARD.POPUL, 84, 169  
ODCH.STANDARD.PRÓBK, 43, 84, 169  
ODCH.STANDARDOWE, 43, 84, 169  
ODCIĘTA, 173, 174, 189  
odwołaniowa, 483, 530, 531, 532, 533, 534, 536  
ÓSM.NA.DZIES, 385  
PEARSON, 170  
PERCENTYL, 508, 529  
PERMUTACJE, 356, 357  
PIERWIASTEK, 116  
pochodna, 105  
drugiego rzędu, 108  
obliczanie numeryczne, 108, 110  
obliczanie za pośrednictwem funkcji  
aproksymującej, 111, 113  
punkt krytyczny, 108  
punkt przegięcia, 108  
szacowanie, 106  
PODAJ.POZYCJĘ, 58, 62, 461, 508, 529  
podcałkowa, 258, 281, 283, 286  
aproksymacja, 258, *Patrz też:* aproksymacja  
obliczanie, 268  
PODSTAW, 444, 446, 451  
POŁĄCZ.TEKSTY, 444, 459, 470  
PORÓWNAJ, 447  
POTĘGA, 356

## funkcja

- POWT, 444, 456  
 POZYCJA, 508, 559, 564  
 PRAWDPD, 560, 561  
 PRAWY, 443, 444, 448  
 PROC.POZ.PRZEDZ.OTW, 557  
 PROC.POZ.PRZEDZ.ZAMK, 557  
 PROCENT.POZYCJA, 508, 529, 554, 555, 556, 557, 558  
   nieciągłość, 556  
   zastosowania, 558  
 prognozująca, 111  
 PRZESUNIĘCIE, 58, 62, 310, 507, 510, 514, 527, 537, 538, 539, 541, 545  
 punkt nieciągłości, 58  
 R.KWADRAT, 170, 173, 174, 176, 189  
 RANDBETWEEN, 562  
 REGBŁSTD, 173, 174  
 REGEXPP, 196, 197, 198, 229  
 REGEXPW, 196, 198, 199, 200, 201  
 REGLINP, 111, 174, 175, 189, 193, 197, 200, 202, 245, 246, 256, 508, 529  
   nietypowe zastosowania, 182, 183, 184, 185, 186  
   przy braku wyrazu wolnego, 181, 182  
   z wagami, 211  
 REGLINW, 111, 189, 194, 195, 201, 202, 245, 246, 252, 253, 256  
 REGLINX, 189, 195, 242, 243  
   rozwiniecie, 236  
 REGLINX.ETS, 237  
 REGLINX.ETS.CONFINT, 238  
 REGLINX.ETS.SEZONOWOŚĆ, 237  
 REGLINX.ETS.STATYSTYKA, 238  
 REGLINX.LINIOWA, 236  
 ROK, 102, 431  
 Round, 86  
 ROZKŁAD.LIN.GAMMA, 278, 300  
 SEKUNDA, 426, 431  
 SERIE, 30  
 SERIESSUM, 136, 508, 529  
 SILNIA, 91, 300  
 SILNIA.DWUKR, 300  
 sklejana, *Patrz:* splajn  
 SUMA, 300, 380, 508, 551, 552  
 SUMA.ILOCZYNÓW, 265, 300, 439, 551  
   formuła narastająca, 569  
 SUMA.JEŻELI, 507, 508  
 SUMA.KWADRATÓW, 149, 508  
 SUMA.SZER.POT, 508, 529  
 SUMY.CZEŚCIOWE, 507, 508, 518, 519, 520, 546, 552, 571  
 SUMY.JEŻELI, 520, 522  
 SZACUJ, 60, 269, 276  
 SZUKAJ.TEKST, 443, 454, 455, 461, 462, 463  
 ŚREDNIA, 168, 174, 508, 552  
 ŚREDNIA.JEŻELI, 508  
 ŚREDNIA.WARUNKÓW, 508  
 ŚREDNIA.WEWN, 508, 529  
 T, 460, 507, 508, 511  
 tablicowa, 175, 198, 262, 481, 483, 502, 503, 508, 528, 529, 530  
 TEKST, 98, 99, 100, 103, 382, 436, 460, 474  
 TERAZ, 425  
 TEST.Z, 508, 529  
 TRANSPONUJ, 185, 473, 481, 503  
 TYP, 479, 483  
 UDF, 36, 40, 270, 481, 566  
 ulotna, 508, 562  
   zamrażanie, 563, 565  
 UMA.WARUNKÓW, 508  
 USUŃ.ZBĘDNE.ODSTĘPY, 444, 446, 464  
 użytkownika, 481  
 WARIANCJA, 169  
 WARIANCJA.POP, 169  
 WARIANCJA.POPUL, 169, 174  
 WARIANCJA.PRÓBK, 169  
 WARTOŚĆ, 96, 457  
 WARTOŚĆ.LICZBOWA, 382  
 WIELOMIAN, 360  
 WIERSZ, 503, 508, 510, 545  
 WORKDAY, 429  
 WSP.KORELACJI, 170, 176  
 WYBIERZ, 116, 117, 507, 508, 514  
 WYST.NAJCZĘŚCIEJ, 366, 508, 527  
 WYSZUKAJ, 508, 529, 530  
 WYSZUKAJ.PIONOWO, 461, 508, 529  
 WYSZUKAJ.POZIOMO, 461, 508, 529  
 Z.WIELKIEJ.LITERY, 444, 457  
 ZAOKR, 93, 95, 101  
 ZAOKR.DO.CAŁK, 96, 102  
 ZAOKR.DO.NPARZ, 97  
 ZAOKR.DO.PARZ, 96  
 ZAOKR.DO.TEKST, 95, 96, 100, 382, 444, 460  
 ZAOKR.DO.WIELOKR, 97, 102  
 ZAOKR.DÓŁ, 95

ZAOKR.GÓRA, 95  
 ZAOKR.W.DÓŁ, 97, 101  
 ZAOKR.W.DÓŁ.DOKŁ, 97  
 ZAOKR.W.DÓŁ.MATEMATYCZNE, 97  
 ZAOKR.W.GÓRĘ, 97  
 ZAOKR.W.GÓRĘ.DOKŁ, 98  
 ZAOKR.W.GÓRĘ.MATEMATYCZNE, 98  
 ZASTĄP, 443, 444, 452, 453  
 ZAZNACZENIE, 64  
 ZŁĄCZ.TEKST, 444, 459, 470  
 ZŁĄCZ.TEKSTY, 444, 458, 470, 473  
 ZNAJDŹ, 443, 450, 451, 462, 463  
 ZNAK, 444

**I**

iloraz różnicowy, 105, 108  
 interpolacja  
   funkcja  
     REGLINP, 256  
     REGLINW, 252, 253, 256  
 kubiczna, 245  
   funkcji dwu zmiennych, 254  
 liniowa, 241  
   funkcji dwu zmiennych, 251, 252  
   odwrócona, 558  
   za pomocą funkcji sklejaných, 246, 249  
 izoterma Langmuira, 187

**K**

karta  
   Formatowanie, 27  
   Projektowanie, 27  
   Słupki błędów, 42  
   Układ, 27, 42  
 kod  
   ANSI, *Patrz:* kod ASCII  
   ASCII, 443, 445, 447  
   Unicode, 443, 444, 445, 446, 447  
 kombinacja, 360, 362  
   generowanie, 370, 371, 372, 373, 374, 375, 377  
   numer, 371  
   z powtórzeniami, 361, 362  
     generowanie, 377  
 kombinatoryka, 355  
   terminologia, 362  
 komórka, 380

korelacja, 175, 176  
   liniowa, 231, 233  
   nieliniowa, 231, 235  
   Pearsona, 170  
 kowariancja, 169  
 kreator, 26  
 kryterium statystyczne Fishera-Snedecora, 172  
 krzywa  
   Lissajous, 68, 69  
   miareczkowania kwasowo-zasadowego, 110

**L**

liczba  
   budowanie z cyfr, 551  
   całkowita suma cyfr, 550  
   część całkowita, 102  
   długość, 90  
   dwójkowa, 384, 385  
   formatowanie, 92, 93  
   kolejna, 411  
   konwersja, 381  
   pseudolosowa, 562  
   rząd, 85  
   seryjna, 411  
   sposób wyświetlania, 38  
   szesnastkowa, 384, 385  
   wygląd, 380, 450, 456  
   zaokrąglanie, 92, 93, 94, 95, 96, 98, 101  
     wartości czasu, 101  
     z uwzględnieniem cyfr znaczących, 100  
   zasady, 86  
   zapis naukowy, 85  
   zespolona, 116, 135  
 liczbotekst, 380, 444, 457  
   konwersja na liczbę, 383, 384  
 linia  
   najlepszego dopasowania, 47  
   regresji, *Patrz:* linia trendu  
   trendu, 24, 25, 28, 175, *Patrz też:* trend  
     nachylenie, 231  
     tworzenie, 48

**M**

macierz, 145  
   iloczyn, 146, 147  
   odwrotna, 146, 147, 184  
 mapa w formie drzewa, 23

## metoda

- Cranka-Nicolsona, *Patrz:* schemat różnicowy niejawny
  - eliminacji Gaussa, 182
  - Eulera, 308, 315, 323
    - modyfikacje, 316, 317, 318, 319
  - gaussowska, *Patrz:* zasada cyfry parzystej
  - iteracji prostej, 119, 120, 121, 123
  - kupiecka, *Patrz:* zasada cyfry parzystej
  - nadrelaksacji, 340
  - najmniejszej sumy kwadratów, 173, 213, 218, 235
  - Newtona, *Patrz:* metoda stycznych
  - Newtona-Raphsona, 161
  - regresji, *Patrz:* regresja
  - różnic skończonych, 333, 338, 342
  - różniczki zupełnej, 82, 83, 86
  - Rungego-Kutty, 320, 321, 322, 323
  - siatek, *Patrz:* metoda różnic skończonych
  - siecznych, 123, 128
    - zbieżność, 125
  - strzałów, 331
  - stycznych, 125, 126, 282
- Microsoft Query, 17

**N**

- nadrelaksacja, 341
- narzędzie
  - Analysis ToolPak, 385
  - Dane/Tekst jako kolumny, 381
  - Solver, 133, 140, 149, 154, 157, 215, 219, 343
  - Sprawdzanie błędów, 383
  - Szukaj wyniku, 131, 157
  - Wklej specjalnie, 383
  - Znajdź i zaznacz/Zamień, 381
- nazwa przypisana do adresu
  - mieszanego, 312
  - względego, 313, 314

**O**

- odchylenie
  - kwadratowe, 168
  - standardowe, 42, 43, 169
  - średnie, 169
- odwołanie
  - absolutne, 30
  - cykliczne, 129, 131, 158, 274, 340, 568
  - strukturalne, 314, 318

## oś

- dodatkowa, *Patrz:* oś pomocnicza
- głębokości, 52
- główna, 29
- kategori, *Patrz:* oś pozioma
- odciętych, *Patrz:* oś pozioma
- pionowa, 24
- pomocnicza, 29
- pozioma, 24, 26
- rzędnych, *Patrz:* oś pionowa
- skalowanie
  - automatyczne, 40
  - nierównomierne, 76, 77
  - ręczne, 40, 77
- w skali logarytmicznej, 28
- wartości, *Patrz:* oś pionowa

**P**

- permutacja, 357, 362
  - generowanie, 368, 369
  - nieporządk, *Patrz:* przestawienie
  - z powtórzeniami, 359, 362
    - generowanie, 369
- podsilnia, 358, 359
- pomiar
  - bezpośredni, 81
  - błąd, *Patrz:* błąd
  - dokładność, 81
- Power Query, 17
- prawdopodobieństwo, 560
- prawo
  - iloczynu, 356
  - rozcieńczeń Ostwalda, 228
  - sumy, 355
- przestawienie, 358
- przybliżenie startowe, 154, 157, 158, 160, 161, 214
- pseudofunkcja, 425
- pseudozakres, *Patrz:* PZ
- PZ, 32, 507, 510
  - argument, 508
  - I rodzaju, 510, 511, 518, 523, 533
  - II rodzaju, 510, 513, 518, 527, 533
  - jako argument funkcji
    - AGREGUJ, 518
    - agregującej, 523, 526
    - baz danych, 530, 531, 532
    - INDEKS, 515, 516, 517, 536
    - JEŻELI, 514



KOMÓRKA, 533, 534  
 L, 511  
 makr XLM, 536  
 PRZESUNIĘCIE, 537, 538, 539, 541  
 SUMY.CZĘŚCIOWE, 518, 519, 520  
 SUMY.JEŻELI, 520, 522  
 tablicowej, 528, 529, 530  
 WYBIERZ, 514

nazwany, 508

z funkcją

ADR.POŚR, *Patrz:* PZ I rodzaju

PRZESUNIĘCIE, *Patrz:* PZ II rodzaju

## R

regresja

liniowa, 111, 165, 167, 236

wielomianu dwu zmiennych, 202, 203,  
 205, 207

wielowymiarowa, *Patrz:* regresja  
 wielokrotna

logistyczna, 217, 218, 220

nieliniowa, 167, 213, 217, 219, 224  
 metodą iteracyjną, 229, 230

postać wykładnicza, 196

postać zlogarytmowana, 196

quasi-liniowa, 215, 222

wielokrotna, 194

wielomianowa, *Patrz:* regresja wielokrotna

reguła fałsi, *Patrz:* metoda siecznych

rozkład dyskretny, 563

równanie

izotermy Langmuira, 187

kinetyczne, 324, 326, 328

kwadratowe, 115, 116, 118

Laplace'a, 337

liniowe, 165, 166

układ, *Patrz:* układ równań liniowych

ze znanym współczynnikiem kierunkowym,  
 180, 181

nieliniowe

linearyzacja, 166, 167, 214

układ, *Patrz:* układ równań nieliniowych

regresji, 171, 174

ocena jakości, 172

postać wykładnicza, 196

postać zlogarytmowana, 196

wyznaczanie, 172

różniczkowe, 307, 320, 321, 322, 323, 324, 326, 328

cząstkowe, 337, 338, 344, 346, 347, 352

eliptyczne, 337, 338

hiperboliczne, 337

hiperboliczno-eliptyczne, 337

paraboliczne, 337, 344, 346, 347, 352

rozwiązanie, 307

rząd, 307

warunki początkowe, 307

zagadnienie początkowe, *Patrz:* zagadnienie  
 początkowe

zwyczajne, 307, 329, 330, 333

wielomianowe, 115, 119, 120, 121, 123, 125,

126, 131, 133, 134, 140, 141

metoda iteracji prostej, *Patrz:* metoda

iteracji prostej

metoda siecznych, *Patrz:* metoda siecznych

oblicznie, 135

pierwiastek, 135, 137, 138, 140

wykładnicze, 166

różnica

centralna, 105, 106, 319

przednia, 105

wsteczna, 105

różniczkowanie, 105

numeryczne, 105, 108, 110

## S

schemat różnicowy

jawny, 348, 349, 352

niejawny, 348, 352

separator, 379

seria, 25, 27

dodawanie, 27

modyfikowanie, 30

nazwa, 27, 30

wybór, 62

zróżnicowanie skali, 29

skala

logarytmiczna, 28, 77

odwrotności, 76, 77

pierwiastkowa, 76

słupek błędu, 41, 43, 45, 76

opcje, 41

zastosowanie niestandardowe, 45, 46

splajn, 246, 249

stała  
 czasowa, 412  
 dysocjacji, 228  
 tablicowa, 65, 472

statystyka matematyczna, 17

stopień swobody, 171

subsilnia, *Patrz:* podsilnia

suma  
 kwadratów, 166, 165  
 regresyjna, 172  
 resztkowa, 171, 172, 210  
 z wagami, 210, 211

SWI, 480, 481, 483, 514, 534

system  
 daty 1900, 411, 412  
 daty 1904, 411, 413  
 jednostek  
 CGS, 387  
 niemetrycznych, 387  
 przeliczanie, 387, 389, 390  
 SI, 387  
 temperatura, 390  
 pozycyjny, 384  
 operacje arytmetyczne, 386

szereg, 289, 300  
 asymptotyczny, 303  
 geometryczny, 289  
 hipergeometryczny, 301  
 Maclaurina, 290, 302  
 naprzemienny, 289  
 potęgowy, 290, 296, 346  
 potęgowy Taylora, 214  
 suma, 291, 292, 296, 299  
 liczba wyrazów, 292, 293, 295  
 obliczenia iteracyjne, 297  
 Taylora, 162, 290  
 zbieżny, 289  
 zdefiniowany rekurencyjnie, 299

szumu usuwanie, 72, 73

## Ś

średnia arytmetyczna, 168

## T

tabela przestawna, 17

tablica  
 bazowa, 500  
 równoległa, 547

do obliczeń, *Patrz:* TDO  
 do wyświetlania, *Patrz:* TDW  
 dwuwymiarowa, 501  
 element sumowanie, 501, 502

kolumna  
 łączenie, 503  
 ukryta, 552, 553, 554  
 usuwanie, 503

nazwana, 479, *Patrz:* TN  
 pełnowartościowa, *Patrz:* TPW  
 quasi-pełnowartościowa, *Patrz:* TQP

wiersz  
 łączenie, 503  
 usuwanie, 503  
 wstawianie, 570, 571

TDO, 508, 509

TDW, 478, 479, 484, 503, 509, 512, 514, 516, 518,  
 533, 543

tekst, 443

fragment, 448, 453

konwersja

na datę, 417

na liczbę, 457

liczba znaków, 447

łączenie, 458, 459, 470, 472

obliczenia iteracyjne, 470, 471, 472

podział, 464, 465

wygląd, 380

wyodrębnianie

cyfr, 467

liczby, 466, 467

zliczanie słów, 463

TN, 479, 509

TPA, 478

TPW, 478, 480, 509, 514, 530

TQP, 479, 481, 509

trend

linia, *Patrz:* linia trendu

liniowy, 244

logarytmiczny, 28

potęgowy, 28

wykładniczy, 28

## U

układ równań

liniowych, 145, 182

rozwiązanie, 146, 149, 150

nieliniowych, 151, 152, 153, 154, 156, 157, 158, 160

nieokreślony, 146

ułamek

- łańcuchowy, 304
- zwykły, 437

ustawienia lokalne, 379, 383

- data, 414

## V

VBA, 16, 270, 481, 482

Visual Basic for Applications, *Patrz:* VBA

## W

wariacja, 362

- bez powtórzeń, 357
  - generowanie, 366, 367
- numer, 365, 368
- z powtórzeniami, 356, 362, 363
  - generowanie, 363, 364, 365, 366

wariancja

- populacji, 168
- próbki, 168
- regresji, 172
- resztkowa, 171, 172
  - pierwiastek, *Patrz:* błąd średni oszacowania
  - wartości

wielomian, 135, 137, 138

- całka, 138
- Lagrange'a, 245, 256
- pochodna, 138
- rozkład na czynniki liniowy i kwadratowy, 142
- stopnia

- czwartego, 142

- drugiego, *Patrz:* równanie kwadratowe

- trzeciego, 119, 140, 141, 142, 245, 246

wielozbiór, 362, 369

współczynnik

- determinacji, 172, 175, 176, 179, 181, 185, 206, 207
- niescentrowany, 178, 179, 181, 182
- skorygowany, 207, 208
- korelacji, 175, 176, 181, 206
- Pearsona, 170

- regresji, 193

- nieliniowej, 217

współrzędna biegunowa, 50

wykres, 23

- animowany, 68
- Arrheniusa, 76, 77, 79
- bąbelkowy, 23

cieniowany, 74, 75

cyldryczny, 23, 26

etykieta przy punktach wykresu, 70, 71

figura geometryczna, 65

funkcji

- ciągłej, 26

- dwu zmiennych, 68

giełdowy, 23, 26

hierarchiczny, 23

histogram, 23

jako szablon, 38

kaskadowy, 23

kolumnowy, 23, 26, 46

kołowy, 23

kopiowanie z arkusza do arkusza, 38, 39

kreator, *Patrz:* kreator

krzywych Lissajous, 68

liniowy, 23, 24, 25

nieciągły, 32, 33, 34, 35, 36, 58, 61

ostrosłupowy, 23, 26

Pareto, 23

pierścieniowy, 23

powierzchniowy, 23, 24, 52

przebiegu w czasie, 54, 55

przestawny, 17

pudełkowy, 23

punktowy, 18, 23, 24, 25, 57, 74

- dane, 26

- nierównomierna skala, 76, 77, 79

- wykonanie, 26, 27

punkty połączone linią, 31

radarowy, 23, 24, 50

schodkowy, 45

słupkowy, 23, 26

statystyczny, 23

stopniowe ujawnianie danych, 63

stożkowy, 23, 26

w komórce, 54, 55

warstwowy, 23, 26, 74

- formatowanie, 75

widmowy, 45, 46

wielokąt foremny, 65

wodospadowy, *Patrz:* wykres kaskadowy

XY, *Patrz:* wykres punktowy

z zaznaczenia, 64

zakres danych, 30, 32

- dynamiczny, 36, 37

- określany przez formułę, 37

- złożony, 32, 33

## wyrażenie

indeksowe, 482, 483, 484, 485

specjalne, *Patrz:* SWI

w postaci tekstowej, 571, 572, 573

## wzór

Cramera, 182

Maclaurina, 290

Newtona, 282

Taylora, 290

Viéte'a, 137, 140

**Z**

zagadnienie początkowe, 308, 324, 329, 330, 333

## zakres

do obliczeń, 508, *Patrz:* ZDO

odwracanie porządku elementów, 499

przekształcanie na stałą tablicową, 472

## zakres danych

dynamiczny, 59

z zaznaczenia, 64

zakres do wyświetlania, *Patrz:* ZDW

zasada cyfry parzystej, 86

ZDO, 518, 530

ZDW, 478, 479, 514, 515, 516, 518, 530, 533

struktura pasiasta, 507

## znak

gwiazdki, 461

LF, 445

niedrukowalny, 446

## spacja

nierozdzielająca, 446

wiodąca, 446

wieloznaczny, 455, *Patrz też:* znak zastępczy

zapytania, 461

zastępczy, 461

zero wiodące, 468, 469

# PROGRAM PARTNERSKI

— GRUPY HELION —



1. ZAREJESTRUJ SIĘ
2. PREZENTUJ KSIĄŻKI
3. ZBIERAJ PROWIZJĘ

Zmień swoją stronę WWW w działający bankomat!

**Dowiedz się więcej i dołącz już dzisiaj!**

<http://program-partnerski.helion.pl>

GRUPA  
**Helion** 

## Obliczenia i wykresy w Excelu? Nic trudnego!

- Wykonuj obliczenia
- Twórz wykresy
- Zrozum Excela

Na rynku nie brakuje książek opisujących obsługę i zastosowania arkusza kalkulacyjnego MS Excel, żadna jednak nie wprowadzi Cię w tę tematykę tak skutecznie jak ta! Omiń rafa i białe plamy dokumentacji, skorzystaj z doświadczenia autora i śmiało wkrocz w świat niesamowitych możliwości Excela.

Książka przedstawia sposoby tworzenia i zastosowania wykresów w Excelu. Przybliży zagadnienie zaokrąglania liczb, a także prezentuje metody przeprowadzania w arkuszu kalkulacyjnym różnego rodzaju obliczeń matematycznych, w tym różniczkowania, całkowania, rozwiązywania równań i układów równań oraz przetwarzania liczb zespolonych.

- Graficzna reprezentacja danych
- Zaokrąglanie liczb i obliczenia przybliżone
- Równania i układy równań nieliniowych
- Regresja i interpolacja
- Obliczanie całek oznaczonych
- Równania różniczkowe zwyczajne i cząstkowe
- Szeregi potęgowe i kombinatoryka
- Konwersja liczb i jednostek
- Liczby i funkcje zespolone
- Przetwarzanie tekstów, dat i czasów

**Dowiedz się, jak wykorzystać Excela do obliczeń i wizualizowania wyników!**

**Helion** 

 **helion.pl**

 **HELION SA**  
ul. Kościuszki 1c  
44-100 Gliwice  
tel.: 32 230 98 63  
helion@helion.pl

*Sprawdź nasze szkolenia!*



**AKADEMIA IT & BUSINESS**

**WWW.SZKOLENIA.HELION.PL**

**KOD KORZYŚCI**  
Sięgnij po więcej! ▶



ISBN 978-83-283-5693-1



9 788328 356931

**INFORMATYKA W NAJLEPSZYM WYDANIU**

Cena: 99,00 zł