Maciej Grzesiak
Instytut Matematyki Politechniki Poznańskiej

Funkcje dwóch zmiennych

1. Funkcje dwóch zmiennych: pojęcia podstawowe

Definicja 1. Funkcją dwóch zmiennych *określoną na zbiorze* $A \subset \mathbb{R}^2$ *o wartościach w zbiorze* $\mathbb{R}$ *nazywamy przyporządkowanie każdemu punktowi ze zbioru* A *dokładnie jednej liczby rzeczywistej. Piszemy* $f : A \longrightarrow \mathbb{R}$ *lub* $z = f(x, y)$. *Wartość funkcji* f *w punkcie* (x, y) *oznaczamy* $f(x, y)$.

Przykład. $f(x, y) = xy^3 + \sin(x + y)$.
Z definicji wynika, że zbiór A na którym funkcja f jest określona (*dziedzina funkcji*, oznaczenie D_f) powinien być podany. Jeżeli podany jest tylko wzór określający funkcję, to zbiór punktów płaszczyzny dla których ten wzór ma sens nazywamy *dziedziną naturalną funkcji.*
Przykłady. Znaleźć i narysować dziedziny naturalne funkcji:
1. $z = \sqrt{xy}$;
2. $z = \arccos \frac{x}{x+y}$;
3. $z = \ln(1 - x^2 - y^2)$.

Definicja 2. *Wykresem funkcji* $z = f(x, y)$ *nazywamy zbiór:*

$$\{(x, y, z) \in \mathbb{R}^3 :\ (x, y) \in D_f,\ z = f(x, y)\}.$$

Poziomicą wykresu funkcji f *dla poziomu* $h \in \mathbb{R}$ *nazywamy zbiór:*

$$\{(x, y) \in D_f :\ f(x, y) = h\}.$$

Poziomica jest krzywą zawartą w dziedzinie funkcji.
Przykłady. Wyznaczyć poziomice danych funkcji i na ich podstawie naszkicować odpowiednią powierzchnię w przestrzeni będącą wykresem funkcji.
1. $z = \frac{x}{y}$;
2. $z = |x| + |y|$;
3. $z = x^2 + y^2$;
4. $z = \sqrt{x^2 + y^2}$.

Wykresy ważniejszych funkcji dwóch zmiennych

1. Wykresem funkcji

$$z = Ax + By + C$$

jest płaszczyzna. Jej wektorem normalnym jest $[-A, -B, 1]$ i przechodzi ona przez punkt $(0, 0, C)$.
2. Wykresem funkcji

$$z = a(x^2 + y^2)$$

jest *paraboloida obrotowa*, tj. powierzchnia powstała przez obrót paraboli $z = ax^2$ wokół osi Oz.
3. Wykresem funkcji

$$z = \pm\sqrt{R^2 - x^2 - y^2}$$

jest górna $(+)$ lub dolna $(-)$ półsfera o środku w początku układu współrzędnych i promieniu R.

4. Wykresem funkcji

$$z = k\sqrt{x^2 + y^2}$$

jest stożek.
5. Wykresem funkcji $z = g(x)$ lub $z = h(y)$ jest powierzchnia walcowa utworzona z wszystkich prostych przechodzących przez punkty krzywej $z = g(x)$ (odpowiednio $z = h(y)$) na płaszczyźnie Oxz (odpowiednio Oyz) i równoległych do osi Oy (odp. Ox).

2. Pochodne cząstkowe funkcji dwóch zmiennych

Definicja 3. *Niech funkcja f będzie określona przynajmniej na otoczeniu punktu (x_0, y_0).* Pochodną cząstkową *rzędu pierwszego funkcji f względem x w punkcie (x_0, y_0) określamy wzorem:*

$$\frac{\partial f}{\partial x}(x_0, y_0) \stackrel{\text{def}}{=} \lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x, y_0) - f(x_0, y_0)}{\Delta x}.$$

Inne oznaczenia pochodnej: $f_x(x_0, y_0)$, $D_1 f(x_0, y_0)$.
Analogicznie określamy *pochodną cząstkową* funkcji f względem y w punkcie (x_0, y_0):

$$\frac{\partial f}{\partial y}(x_0, y_0) \stackrel{\text{def}}{=} \lim_{\Delta y \to 0} \frac{f(x_0, y_0 + \Delta y) - f(x_0, y_0)}{\Delta y}.$$

Inne oznaczenia tej pochodnej: $f_y(x_0, y_0)$, $D_2 f(x_0, y_0)$.
Z definicji wynika, że pochodne cząstkowe obliczamy według tych samych zasad, co zwykłe (można stosować wzory na pochodną sumy, iloczynu, itd.). Przy obliczaniu pochodnej cząstkowej funkcji f względem x należy y traktować jak stałą, a przy obliczaniu pochodnej cząstkowej funkcji f względem y należy x traktować jak stałą.
Przykłady. Obliczyć pochodne cząstkowe rzędu pierwszego:
1. $z = x^3 y - xy^3$;
2. $z = \frac{x}{y} + \frac{y^2}{x}$;
3. $z = \ln(x + \sqrt{x^2 + y^2})$;
4. $z = x^y$;
5. $z = \operatorname{arc\,tg} \frac{x}{y}$.
Jeżeli pochodne cząstkowe rzędu pierwszego istnieją dla wszystkich punktów pewnego zbioru A, to w tym zbiorze można rozpatrywać funkcje pochodne. Obliczając pochodne cząstkowe tych funkcji otrzymamy *pochodne cząstkowe drugiego rzędu.* Dokładniej, określamy:

$$\frac{\partial^2 f}{\partial x^2}(x_0, y_0) \stackrel{\text{def}}{=} \left(\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right)\right)(x_0, y_0),$$

$$\frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) \stackrel{\text{def}}{=} \left(\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)\right)(x_0, y_0),$$

$$\frac{\partial^2 f}{\partial y \partial x}(x_0, y_0) \stackrel{\text{def}}{=} \left(\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)\right)(x_0, y_0),$$

$$\frac{\partial^2 f}{\partial y^2}(x_0, y_0) \stackrel{\text{def}}{=} \left(\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right)\right)(x_0, y_0).$$

Inne oznaczenia tych pochodnych: $f_{xx}(x_0, y_0)$, $f_{xy}(x_0, y_0)$, $f_{yx}(x_0, y_0)$, $f_{yy}(x_0, y_0)$ lub $D_{11} f(x_0, y_0)$, $D_{12} f(x_0, y_0)$, $D_{21} f(x_0, y_0)$, $D_{22} f(x_0, y_0)$.

Twierdzenie 1. (Schwarza o pochodnych mieszanych) *Jeżeli pochodne cząstkowe $\partial^2 f/\partial x \partial y$, $\partial^2 f/\partial y \partial x$ są ciągłe w punkcie (x_0, y_0), to są równe, tj.*

$$\frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) = \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0).$$

Przykłady. Obliczyć pochodne rzędu drugiego:
1. $z = \cos^2(2x + 3y)$;
2. $z = \frac{x+y}{x-y}$;
3. $z = \operatorname{arc\,tg} \frac{x+y}{1-xy}$.

3. Różniczka funkcji dwóch zmiennych

Niech dana będzie funkcja $z = f(x, y)$ i punkty $P_0(x_0, y_0)$, $P_1(x_1, y_1)$ należące do D_f. Przyrosty argumentów funkcji to $\Delta x = x_1 - x_0$ i $\Delta y = y_1 - y_0$. Nazywa się je również różniczkami i oznacza $\mathrm{d}x$, $\mathrm{d}y$. Natomiast dla wartości funkcji "przyrost" i "różniczka" oznaczają coś innego:

$$\Delta f = f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0),$$

$$\mathrm{d}f = \frac{\partial f}{\partial x}(x_0, y_0)\,\mathrm{d}x + \frac{\partial f}{\partial y}(x_0, y_0)\,\mathrm{d}y.$$

Różniczka $\mathrm{d}f$ (oznaczana także $\mathrm{d}z$) jest funkcją czterech zmiennych: $x_0, y_0, \mathrm{d}x, \mathrm{d}y$.
Przykład. Obliczyć przyrost Δf i różniczkę $\mathrm{d}f$, gdy $f(x, y) = x^2y - x + 1$, w punkcie $P_0(2, 2)$ dla przyrostów: a) $\mathrm{d}x = \mathrm{d}y = 1$; b) $\mathrm{d}x = 1, \mathrm{d}y = -1$; c) $\mathrm{d}x = \mathrm{d}y = 0,1$.
Przykłady. Napisać wzór na różniczkę dla funkcji
1. $z = \operatorname{arc\,tg} \frac{y}{x}$;
2. $z = e^{3x} \sin y$.

Zastosowanie różniczki do obliczeń przybliżonych

Twierdzenie 2. *Jeżeli funkcja f ma ciągłe pochodne cząstkowe pierwszego rzędu w (x_0, y_0), to*

$$f(x_0 + \Delta x, y_0 + \Delta y) \approx f(x_0, y_0) + \mathrm{d}f(x_0, y_0)(\Delta x, \Delta y),$$

przy czym błąd $\delta(\Delta x, \Delta y)$ tego przybliżenia dąży do 0 szybciej niż $\sqrt{\Delta x^2 + \Delta y^2}$.

Jeżeli maksymalne błędy pomiarów wielkości x, y oznaczymy Δ_x, Δ_y, a Δ_z będzie maksymalnym błędem obliczeń, to

$$\Delta_z = \left|\frac{\partial f}{\partial x}\right| \Delta_x + \left|\frac{\partial f}{\partial y}\right| \Delta_y.$$

Przykład. Kąt φ pod jakim widzimy drzewo zmierzony z dokładnością 0,01 radiana jest równy $\pi/4$, a odległość d miejsca pomiaru od pnia drzewa zmierzona z dokładnością $\Delta_d = 0,1$ m jest równa 30 m. Z jaką dokładnością możemy obliczyć wysokość tego drzewa?
Rozwiązanie. $h = d \operatorname{tg} \varphi = 30$ m. Dokładność:

$$\Delta_z = \left|\frac{\partial h}{\partial d}\right| \Delta_d + \left|\frac{\partial h}{\partial \varphi}\right| \Delta_\varphi = \operatorname{tg} \varphi \cdot \Delta_d + \frac{d}{\cos^2 \varphi}\Delta_\varphi = 1 \cdot 0,1 + \frac{30}{0,5} \cdot 0,01 = 0,7.$$

4. Pochodna kierunkowa i gradient funkcji

Definicja 4. *Niech funkcja f będzie określona przynajmniej na otoczeniu punktu (x_0, y_0) i niech $\vec{v} = (v_x, v_y)$ będzie wektorem jednostkowym.* Pochodną kierunkową *funkcji f w punkcie (x_0, y_0) w kierunku $\vec{v}$ określamy wzorem:*

$$\frac{\partial f}{\partial \vec{v}}(x_0, y_0) \stackrel{\text{def}}{=} \lim_{t \to 0} \frac{f(x_0 + tv_x, y_0 + tv_y) - f(x_0, y_0)}{t}.$$

Pochodną kierunkową można traktować jak uogólnienie pochodnej cząstkowej, bo jeśli $\vec{v} = (1, 0)$, to

$$\frac{\partial f}{\partial \vec{v}} = \frac{\partial f}{\partial x},$$

a jeśli $\vec{v} = (0, 1)$, to

$$\frac{\partial f}{\partial \vec{v}} = \frac{\partial f}{\partial y}.$$

Do obliczania pochodnej kierunkowej stosujemy wzór:

$$\frac{\partial f}{\partial \vec{v}} = \frac{\partial f}{\partial x} v_x + \frac{\partial f}{\partial y} v_y.$$

Jeśli α oznacza kąt jaki wektor $\vec{v}$ tworzy z osią Ox, to $v_x = \cos\alpha$, $v_y = \sin\alpha$. zatem

$$\frac{\partial f}{\partial \vec{v}} = \frac{\partial f}{\partial x} \cos\alpha + \frac{\partial f}{\partial y} \sin\alpha$$

Definicja 5. *Gradientem funkcji f w punkcie (x_0, y_0) nazywamy wektor określony wzorem:*

$$\operatorname{grad} f(x_0, y_0) \stackrel{\text{def}}{=} \left(\frac{\partial f}{\partial x}(x_0, y_0), \frac{\partial f}{\partial y}(x_0, y_0) \right).$$

Gradient oznacza się też używając symbolu ∇, nazywanego *nabla.*[1]

$$\operatorname{grad} f = \nabla f$$

Przy użyciu pojęcia gradientu wzór na pochodną kierunkową przyjmie postać:

$$\frac{\partial f}{\partial \vec{v}} = \operatorname{grad} f \circ \frac{\vec{v}}{|\vec{v}|}.$$

Zatem pochodna kierunkowa jest iloczynem skalarnym gradientu i wersora o kierunku wektora $\vec{v}$.
Przykład. Znaleźć pochodną kierunkową funkcji $z = 2x^2 - 3y^2$ w punkcie $P(1, 0)$ w kierunku, który z osią Ox tworzy kąt $\frac{2}{3}\pi$.
Rozwiązanie. Mamy

$$\frac{\partial z}{\partial x} = 4x, \quad \left(\frac{\partial z}{\partial x}\right)_P = 4,$$

$$\frac{\partial z}{\partial y} = -6y, \quad \left(\frac{\partial z}{\partial y}\right)_P = 0,$$

$$\cos\alpha = \cos\frac{2}{3}\pi = -\frac{1}{2}, \quad \sin\alpha = \sin\frac{2}{3}\pi = \frac{\sqrt{3}}{2},$$

więc

$$\frac{\partial z}{\partial \vec{v}} = 4 \cdot (-\frac{1}{2}) + 0 \cdot \frac{\sqrt{3}}{2} = -2.$$

Przykład. Znaleźć gradient funkcji $z = x^3 + y^3 - 3xy$ w punkcie $P(2, 1)$.
Rozwiązanie. Mamy

$$\frac{\partial z}{\partial x} = 3x^2 - 3y, \quad \left(\frac{\partial z}{\partial x}\right)_{(2,1)} = 9,$$

$$\frac{\partial z}{\partial y} = 3y^2 - 3x, \quad \left(\frac{\partial z}{\partial y}\right)_{(2,1)} = -3,$$

więc

$$\operatorname{grad} f = (9, -3).$$

[1] Nazwa pochodzi od greckiego wyrazu określającego hebrajską harfę o podobnym kształcie.

Znajdziemy teraz wektor $\vec{v}$ wyznaczający kierunek, w którym pochodna ma największą wartość. Mamy:

$$\frac{\partial f}{\partial \vec{v}} = \nabla f \circ \frac{\vec{v}}{|\vec{v}|} = |\nabla f| \cdot \cos(\nabla f, \vec{v}),$$

a więc maksymalna wartość wynosi $|\nabla f|$ i jest osiągana wtedy, gdy $\cos(\nabla f, \vec{v}) = 1$, czyli gdy wektor $\vec{v}$ ma kierunek gradientu. Mamy więc wniosek.

Wniosek 1. *Gradient wskazuje kierunek najszybszego wzrostu wartości funkcji. Szybkość wzrostu funkcji f w punkcie (x, y) jest długością wektora* $(\operatorname{grad} f)(x, y)$.

5. Ekstremum funkcji dwóch zmiennych

Tak samo jak w przypadku funkcji jednej zmiennej, maksimum i minimum funkcji są pojęciami lokalnymi, tzn. odnoszącymi się do zachowania funkcji w pewnym małym otoczeniu punktu.

Definicja 6. (minimum lokalnego) *Funkcja f ma w punkcie (x_0, y_0)* minimum, *jeżeli istnieje otoczenie tego punktu takie, że dla dowolnego (x, y) z tego otoczenia zachodzi nierówność*

$$f(x, y) \geqslant f(x_0, y_0).$$

Jeżeli powyższa nierówność jest ostra, to minimum nazywamy właściwym.

Definicja 7. (maksimum lokalnego) *Funkcja f ma w punkcie (x_0, y_0)* maksimum, *jeżeli istnieje otoczenie tego punktu takie, że dla dowolnego (x, y) z tego otoczenia zachodzi nierówność*

$$f(x, y) \leqslant f(x_0, y_0).$$

Jeżeli powyższa nierówność jest ostra, to maksimum nazywamy właściwym.

Przykłady. 1. Funkcja $f(x, y) = 5 - x^4 - y^2$ ma maksimum równe 5 w punkcie $(0, 0)$.
2. Funkcja $f(x, y) = 2x^2 - y^2$ nie ma ekstremum w punkcie $(0, 0)$, bo $f(0, 0) = 0$, ale w dowolnym otoczeniu punktu $(0, 0)$ są punkty, w których wartość funkcji jest dodatnia (są to punkty osi Ox), i są punkty, w których wartość funkcji jest ujemna (punkty osi Oy).

Twierdzenie 3. (warunek konieczny istnienia ekstremum) *Jeżeli funkcja $f(x, y)$ ma pochodne cząstkowe w punkcie (x_0, y_0) i ma w tym punkcie ekstremum, to*

$$\frac{\partial f}{\partial x}(x_0, y_0) = 0, \quad \frac{\partial f}{\partial y}(x_0, y_0) = 0. \tag{1}$$

Zatem punkty, w których może być ekstremum znajdziemy rozwiązując układ równań

$$\frac{\partial f}{\partial x}(x, y) = 0, \quad \frac{\partial f}{\partial y}(x, y) = 0. \tag{2}$$

Rozwiązania układu (2) nazywamy *punktami stacjonarnymi* lub *krytycznymi* funkcji.

Twierdzenie 4. (warunek dostateczny istnienia ekstremum) *Jeżeli funkcja $f(x, y)$ ma w pewnym otoczeniu punktu $P(x_0, y_0)$ ciągłe pochodne cząstkowe drugiego rzędu, przy czym*

$$\frac{\partial f}{\partial x}(x_0, y_0) = 0, \quad \frac{\partial f}{\partial y}(x_0, y_0) = 0,$$

oraz

$$H(x_0, y_0) = \begin{vmatrix} \frac{\partial^2 f}{\partial x^2}(x_0, y_0) & \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) \\ \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) & \frac{\partial^2 f}{\partial y^2}(x_0, y_0) \end{vmatrix} > 0,$$

to funkcja $f(x, y)$ ma w punkcie $P_0(x_0, y_0)$ maksimum właściwe, gdy $\frac{\partial^2 f}{\partial x^2}(x_0, y_0) < 0$, a minimum właściwe, gdy $\frac{\partial^2 f}{\partial x^2}(x_0, y_0) > 0$

Wyznacznik $H(x, y)$ nazywamy *hesjanem.*
Przykłady. Znaleźć ekstrema funkcji:
1. $z = e^{x-y}(x^2 - 2y^2)$.
Wyniki: Punkty stacjonarne: $P_1(0,0), P_2(-4,-2)$. W P_1 nie ma ekstremum, w P_2 jest maksimum.
2. $z = x^3 + 3xy^2 - 30x - 18y$.
Wyniki: Punkty stacjonarne: $P_1(3,1), P_2(1,3), P_3(-1,-3), P_4(-3,-1)$, hesjan $H(x, y) = 36(x^2 - y^2)$. W P_1 jest minimum (równe -72), w P_2 i P_3 nie ma ekstremum, w P_4 jest maksimum (równe 72).
3. $z = x^2 + xy + y^2 - 2x - y$ (minimum równe -1 w $(1, 0)$).
4. Na płaszczyźnie Oxy znaleźć punkt dla którego suma kwadratów odległości od trzech prostych $x = 0$, $y = 0$ i $x - y + 1 = 0$ jest najmniejsza. (Odp.: $P(-\frac{1}{4}, \frac{1}{4})$)
5. Znaleźć wymiary a, b, c prostopadłościanu o ustalonej objętości V, który ma najmniejsze pole powierzchni całkowitej.

6. Metoda najmniejszych kwadratów

W naukach doświadczalnych istotne jest wyznaczanie zależności między różnymi wielkościami na podstawie wyników pomiarów. Można to robić rozmaicie. Naturalnym odruchem jest poszukiwanie w miarę prostych zależności między wielkościami. Na przykład w pewnym eksperymencie mierzymy jednocześnie dwie wielkości fizyczne x i y. Wynik n pomiarów, to zbiór par

$$\{(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)\}.$$

Zmienną x traktujemy jako niezależną i przypuszczamy, że wartość drugiej zmiennej y jest funkcją liniową x, tj. $y = ax + b$.
Gdyby punkty były tylko dwa, to prosta byłaby określona jednoznacznie, bo współczynniki a, b obliczylibyśmy z układu

$$\begin{array}{ccccc} ax_1 & + & b & = & y_1 \\ ax_2 & + & b & = & y_2. \end{array}$$

Przy większej liczbie, np. n punktów, układ

$$\begin{array}{ccccc} ax_1 & + & b & = & y_1 \\ ax_2 & + & b & = & y_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ ax_n & + & b & = & y_n \end{array}$$

będzie na ogół sprzeczny.
Nie będziemy więc szukać rozwiązania, ale liczb a, b dla których powyższy układ jest "jak najmniej" sprzeczny.
Szukamy a i b takich, że wyrażenie

$$S(a, b) = \sum_{k=1}^{n} (y_k - ax_k - b)^2$$

ma najmniejszą wartość. Aby znaleźć minimum funkcji obliczamy pochodne cząstkowe

$$\begin{aligned} \frac{\partial S}{\partial a} &= -2 \sum_{k=1}^{n} (y_k - ax_k - b) \cdot x_k \\ \frac{\partial S}{\partial b} &= -2 \sum_{k=1}^{n} (y_k - ax_k - b). \end{aligned}$$

Przyrównując je do zera otrzymujemy układ

$$\begin{aligned} a\sum_{k=1}^{n} x_k^2 + b\sum_{k=1}^{n} x_k &= \sum_{k=1}^{n} x_k y_k \\ a\sum_{k=1}^{n} x_k + bn &= \sum_{k=1}^{n} y_k \end{aligned} \tag{3}$$

z którego możemy obliczyć a i b, a więc wyznaczymy punkt krytyczny funkcji $S(a, b)$.
Bardziej zwięzły zapis uzyskamy wprowadzając oznaczenia

$$\bar{x} = \frac{1}{n}\sum_{k=1}^{n} x_k$$

$$\bar{y} = \frac{1}{n}\sum_{k=1}^{n} y_k,$$

a więc $\bar{x}$, $\bar{y}$ są średnimi arytmetycznymi wyników pomiarów.
Układ (3) zapisujemy teraz jako

$$\begin{aligned} a\sum_{k=1}^{n} x_k^2 + bn\bar{x} &= \sum_{k=1}^{n} x_k y_k \\ an\bar{x} + bn &= n\bar{y} \end{aligned} \tag{4}$$

Rozwiązania tego układu (otrzymane np. przy pomocy wzorów Cramera) zapisuje się zwykle w postaci

$$\begin{aligned} a &= \frac{\sum_{k=1}^{n} x_k y_k - n\bar{x}\bar{y}}{\sum_{k=1}^{n} x_k^2 - n\bar{x}^2} \\ b &= \bar{y} - a\bar{x}. \end{aligned}$$

Para (a, b) określa punkt stacjonarny funkcji S.
Można standardowo (obliczając hesjan) sprawdzić, że w tym punkcie funkcja ma rzeczywiście minimum. Zatem $y = ax + b$ jest szukaną zależnością.
Ze względu na wzór na S taki sposób wyznaczania zależności nazywamy *metodą najmniejszych kwadratów.*
Przykład. Metodą najmniejszych kwadratów znaleźć równanie prostej, która najlepiej przybliża dane:

x_i	1	2	3	4
y_i	2	4	5	7

Rozwiązanie. Obliczamy $\bar{x} = \frac{5}{2}$, $\bar{y} = \frac{9}{2}$, $\sum_{i=1}^{4} x_i^2 = 30$, $\sum_{i=1}^{4} x_i y_i = 53$, więc

$$a = \frac{53 - 4 \cdot \frac{5}{2} \cdot \frac{9}{2}}{30 - 4 \cdot \frac{25}{4}} = \frac{8}{5},$$

$$b = \frac{9}{2} - \frac{8}{5} \cdot \frac{5}{2} = \frac{1}{2}$$

Równanie prostej to $y = \frac{8}{5}x + \frac{1}{2}$.
Metoda najmniejszych kwadratów została wprowadzona na początku XIX wieku. Opiera się na *postulacie Legendre'a.* Można go sformułować następująco.
Wartością najbardziej prawdopodobną, otrzymaną z szeregu wyników tak samo dokładnych pomiarów, jest taka, od której obliczone odchylenia tych wyników, po podniesieniu do drugiej potęgi i zsumowaniu dają wielkość najmniejszą z możliwych.

Z postulatu Legendre'a wynika, że najbardziej prawdopodobną wielkością z szeregu jednakowo dokładnych pomiarów jednej wielkości jest ich średnia zwykła. W przypadku pomiarów niejednakowo dokładnych postulat ten brzmi podobnie, stosuje się jednak do odchyleń równoważonych „wagami", tj wartość ma tym większą wagę im bardziej dokładny jest pomiar. W tym przypadku najbardziej prawdopodobną okazuje się wielkość zwana średnią ważoną.
Omówiliśmy tu najprostszy przypadek: wyznaczanie "najbardziej pasującej do danych" funkcji liniowej. Można w analogiczny sposób wyznaczać najlepiej pasującą funkcję kwadratową, lub wielomian wyższego stopnia. Historycznie, jednym z pierwszych zastosowań tej metody było wyznaczenie przez Gaussa w 1801 roku orbity planetoidy Ceres.

7. Funkcje uwikłane

Jeżeli równość $f(x, y) = 0$, gdzie $f(x, y)$ jest różniczkowalną funkcją zmiennych x, y określa y jako funkcję x, to pochodną tej *funkcji uwikłanej* można obliczyć ze wzoru

$$\frac{\mathrm{d}y}{\mathrm{d}x} = -\frac{f'_x(x, y)}{f'_y(x, y)}, \tag{5}$$

pod warunkiem, że $f'_y(x, y) \neq 0$.
Pochodne wyższych rzędów znajdujemy różniczkując kolejny raz wzór (5)
Przykład. Obliczyć $\frac{\mathrm{d}y}{\mathrm{d}x}$ i $\frac{d^2y}{\mathrm{d}x^2}$ gdy

$$(x^2 + y^2)^3 - 3(x^2 + y^2) + 1 = 0.$$

Rozwiązanie. Oznaczając lewą stronę przez $f(x, y)$ obliczamy pochodne cząstkowe:

$$f'_x(x, y) = 6x\big((x^2 + y^2)^2 - 1\big),$$

$$f'_y(x, y) = 6y\big((x^2 + y^2)^2 - 1\big).$$

Stosując wzór (5) otrzymujemy

$$\frac{\mathrm{d}y}{\mathrm{d}x} = -\frac{f'_x(x, y)}{f'_y(x, y)} = -\frac{x}{y}.$$

Aby obliczyć drugą pochodną różniczkujemy tę pochodną względem x, **uwzględniając, że y jest funkcją x:**

$$\frac{d}{\mathrm{d}x}\left(-\frac{x}{y}\right) = -\frac{1 \cdot y - x\frac{\mathrm{d}y}{\mathrm{d}x}}{y^2} = -\frac{y - x(-\frac{x}{y})}{y^2} = -\frac{y^2 + x^2}{y^3}.$$

Ogólny wzór na drugą pochodną ma postać:

$$\frac{d^2y}{\mathrm{d}x^2} = -\frac{f''_{xx}(f'_y)^2 - 2f''_{xy}f'_xf'_y + f''_{yy}(f'_x)^2}{(f'_y)^3}. \tag{6}$$

Ekstremum funkcji uwikłanej

Ekstremum może wystąpić w punktach dla których $\frac{\mathrm{d}y}{\mathrm{d}x} = 0$. Uwzględniając wzór (5) otrzymujemy warunki:

$$f(x, y) = 0, \quad f'_x(x, y) = 0, \quad f'_y(x, y) \neq 0. \tag{7}$$

Z powyższych równości obliczamy x, y. Maksimum wystąpi, gdy $\frac{d^2y}{\mathrm{d}x^2} < 0$, a minimum gdy $\frac{d^2y}{\mathrm{d}x^2} > 0$. Posłużymy się wzorem (6), który przy warunkach (7) przyjmie postać uproszczoną

$$\frac{d^2y}{\mathrm{d}x^2} = -\frac{f''_{xx}}{f'_y}.$$

Zatem gdy $\frac{f''_{xx}}{f'_y} < 0$, to mamy minimum, a gdy $\frac{f''_{xx}}{f'_y} > 0$, to mamy maksimum.
Przykład. Znaleźć ekstrema funkcji uwikłanej $xy^2 - x^2y = 2$.
Rozwiązanie. Obliczamy:

$$f'_x = y^2 - 2xy, \quad f'_y = 2xy - x^2, \quad f''_{xx} = -2y.$$

Najpierw rozwiązujemy układ $f(x,y) = 0$, $f'_x(x,y) = 0$, tj.

$$\begin{cases} xy^2 - x^2y &= 2 \\ y^2 - 2xy &= 0 \end{cases}$$

Otrzymujemy $x = 1$, $y = 2$. Sprawdzamy teraz, czy f'_y jest różna od 0:

$$f'_y(1,2) = 2 \cdot 1 \cdot 2 - 1^2 = 3 \neq 0.$$

Następnie badamy znak ułamka $\frac{f''_{xx}}{f'_y}$:

$$\frac{f''_{xx}}{f'_y}(1,2) = \frac{-4}{3} < 0.$$

Zatem dla $x = 1$ funkcja y ma minimum równe 2.