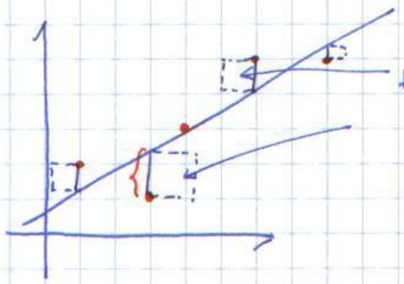
Problem, którym będziemy się dzisiaj zajmować jest następujący: Mając n punktów pomiarowych $(x_i, y_i)_{i=1}^{n}$ znaleźć prostą o równaniu $y=ax+b$ tak aby jak najlepiej przybliżała ona punkty pomiarowe. Dokładniej; aby ~~różnica~~ funkcja

$$F(a,b)=\sum_{i=1}^{n}\left(y_i-(ax_i+b)\right)^2 \qquad (1)$$

była możliwie mała. Jest to jedna z wersji tak zwanej METODY NAJMNIEJSZYCH KWADRATÓW. Na obrazku wyglądałoby to tak

I

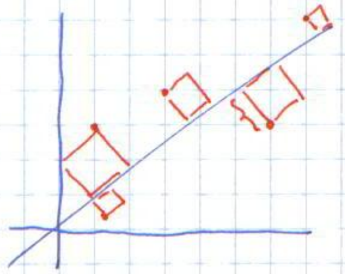


kwadraty, które mają być najmniejsze

Dokładniej - najmniejsza ma być suma pól tych kwadratów.

Inna wersja metody najmniejszych kwadratów jest taka:

II



Różnica polega na sposobie mierzenia odległości od prostej – w I mierzymy pionowo zaś w II prostopadle

My zajmiemy się poszukiwaniem wzorów w przypadku I, bo jest łatwiejszy rachunkowo.

Matematycznie problem możemy postawić następująco: Znaleźć liczby a i b takie, aby funkcja

$$F(a,b)=\sum_{i=1}^{n}\left(y_i-(ax_i+b)\right)^2$$

była możliwie mała. Szukamy więc minimum tej funkcji

(2)

$$F(a,b) = \sum_{i=1}^{n} \left(y_i - (ax_i + b) \right)^2 =$$

$$= \sum_{i=1}^{n} \left(y_i^2 - 2y_i(ax_i+b) + (ax_i+b)^2 \right) =$$

$$= \left(y_i^2 - 2y_i x_i a - 2y_i b + a^2 x_i^2 + 2abx_i + b^2 \right) =$$

$$= \sum y_i^2 - 2a \sum y_i x_i - \left(\sum 2y_i\right) b + a^2 \sum x_i^2 +$$
$$+ 2ab\left(\sum_i x_i\right) + \sum_i b^2 =$$

$$= n\,E(Y^2) - 2a\,n\,E(XY) - 2bn\,E(Y) + a^2 n E(X^2) +$$
$$+ 2abn\left(E(X)\right) + n b^2 =$$

$$= n \left(E(Y^2) + 2a\,E(XY) - 2b\,E(Y) + a^2 E(X^2) + \right.$$
$$\left. + 2ab\,E(X) + b^2 \right)$$

liczby wyznaczane z danych doświadczalnych

średnia X : $E(X) = \frac{1}{n} \sum x_i$ średnia Y $E(Y) = \frac{1}{n} \sum y_i$

średnia $X \cdot Y$: $E(XY) = \frac{1}{n} \sum x_i y_i$ ... itd.

Funkcja F jest wielomianem drugiego stopnia od dwóch zmiennych a i b.

Jak znaleźć minimum funkcji dwóch zmiennych?

Dla funkcji jednej zmiennej wiemy jak to robić: ekstremum (minimum lub maksimum) jest zawsze w punkcie w którym pochodna funkcji jest równa 0 (zakładając oczywiście, że funkcja jest różniczkowalna) jeśli funkcja nie jest różniczkowalna w w jakimś punkcie trzeba oddzielnie sprawdzić, czy nie ma w nim ekstremum. Jeśli ~~w ty~~ tego rodzaju punktów jest dużo sprawa robi się kłopotliwa. My zajmować się będziemy teraz jedynie bardzo przyzwoitymi funkcjami, tzn. funkcjami różniczkowalnymi, których pochodna też jest ciągła.

(3)

Dla funkcji jednej zmiennej mamy więc jasność — ekstrema mogą być tam, gdzie pochodna się zeruje. Pamiętamy oczywiście, że to jedynie warunek konieczny bo np. pochodna jest zero w następujących ~~tu~~ trzech przypadkach (w $x=0$)

f	$x \longmapsto x^2$	$x \longmapsto -x^2$	$x \longmapsto x^3$
f'	$x \longmapsto 2x$	$x \longmapsto -2x$	$x \longmapsto 3x^2$

minimum — maksimum — punkt krytyczny ale nie ekstremum.

Jeśli pochodna jest ciągła możemy sprawdzić, czy w podejrzanym punkcie zmienia ~~z~~ znak:

z − na +	z + na −	cały czas +
min	max	
↘ ↗	↗ ↘	↗

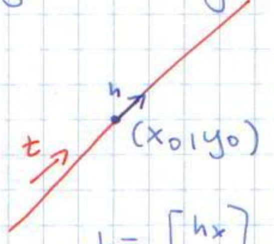
Możemy też użyć drugiej pochodnej jeśli istnieje

	dodatnie	ujemna	zero → znaczy niewiadomo!
f''	2	-2	$6x\|_{x=0} = 0$

Super! Ale co robić z funkcją dwóch lub więcej zmiennych? np. Jeśli

$$f(x,y) = x^2 + xy + y^2 \quad ?$$

Załóżmy że f ma ekstremum (np. minimum) w jakimś punkcie (x_0, y_0). Wtedy ekstremum tego samego rodzaju będzie miała funkcja



$$h = \begin{bmatrix} h_x \\ h_y \end{bmatrix}$$

obcięta do prostej przechodzącej przez punkt (x_0, y_0). Każda prosta jest dobra do sprawdzania. Tak naprawdę trzebaby sprawdzić na każdej prostej!

Tworzymy więc funkcję

4

$$t \longmapsto f\left(x_0 + t h_x,\ y_0 + t h_y\right)$$

to już jest funkcja jednej zmiennej.

Zobaczmy jak wyglądają tego rodzaju funkcje zmiennej t wyprodukowane dla naszej funkcji f i różnych h. w dwóch różnych punktach $(0,0)$ i $(0,1)$

W punkcie $(0,1)$ weźmiemy $h_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $h_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ i $h_3 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$
(podobnie w $(0,0)$)

h_1 : $f(0+t, 1+0) = t^2 + t + 1$

h_2 : $f(0+0, 1+t) = 0 + 0(1+t) + (1+t)^2 =$
$= (1+t)$

$h_3 \quad f(0+t, 1+t) = t^2 + t(1+t) + (1+t)^2 =$
$= t^2 + t + t^2 + 1 + 2t + t^2 =$
$= 3t^2 + 3t + 1$

5

$t = 0.$

$(0,0) \quad h_1 \quad f(0+t, 0) = t^2$

$h_2 \quad f(0, 0+t) = t^2$

$h_3 \quad f(0+t, 0+t) = t^2 + \cancel{2t} t^2 + t^2 =$
$= 3t^2$

Wszystkie trzy funkcje zmiennej t mają w $t=0$ minimum. Może więc nasza funkcja też ma w $(0,0)$ minimum?

Badanie funkcji jednej zmiennej postaci

$\phi_h : t \longmapsto f(x_0 + th_x, y_0 + th_y)$

można przeprowadzić tradycyjnie licząc pochodną po t. Nazwijmy powyższą funkcję ϕ_h

$$\phi_h'(t) = \frac{\partial f}{\partial x} h_x + \frac{\partial f}{\partial y} h_x = \underbrace{\begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{bmatrix}}_{f'(x_0, y_0)} \begin{bmatrix} h_x \\ h_y \end{bmatrix} = 0$$

Bo ϕ ma mieć ekstremum!

Punkt podejrzany o bycie ekstremum ma więc tę własność, że pochodna w tym punkcie jest równa zero — oznacza to znikanie wszystkich pochodnych cząstkowych!

Okazuje się więc, że kryterium konieczne istnienia w danym punkcie ekstremum funkcji wielu zmiennych jest takie samo jak dla funkcji jednej zmiennej — (6)

– POCHODNA W TYM PUNKCIE MA BYĆ RÓWNA ZERO!

Podobnie jak dla funkcji jednej zmiennej jest to jednak warunek niewystarczający. Wykresy poniższych funkcji pokazują, jakie mogą się pojawić kłopoty.

siodło

$f_1(x,y) = x^2 - y^2$ $\quad f'(x,y) = [\,2x\,,\ -2y\,]$ $\quad (x_0,y_0) = (0,0)$

$f_2(x,y) = x(x^2-3y^2) =$

$= x^3 - 3y^2x$ $\quad f_2'(x,y) = [\,3x^2-3y^2\,,\ 6yx\,]$

małpie siodło $\quad (x_0,y_0) = (0,0)$

$h_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \Rightarrow f(0,t) = 0$

$h_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \Rightarrow f(t,0) = t^3$

$h_3 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \Rightarrow f(0+t,\ 0+t) = t(t^2-3t^2) = -2t^3$

$h_4 = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \Rightarrow f(-t,t) = -t(t^2-3t^2) = 2t^3$

$h_5 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \Rightarrow f(t,2t) = t(t^2-12t^2 = -11t^3$

Przyjrzyjmy się jeszcze jednej okiconej dości funkcji

$$f_3(x,y) = (x-y^2)(x-3y^2)$$

$$f_3'(x,y) = \left[2x - 4y^2 \qquad -4y(2x+3y^2)\right]$$

$$(x_0, y_0) = (0,0)$$

W punkcie $(0,0)$ jest punkt krytyczny.

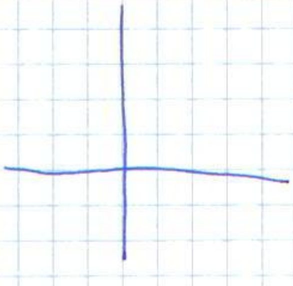
Sprawdźmy jak funkcja zachowuje się wzdłuż prostej

$$t \longmapsto \begin{bmatrix} h_x \\ h_y \end{bmatrix} t$$

$$f_3(th_x, th_y) = (th_x - t^2 h_y^2)(th_x - 3t^2 h_y^2) =$$

$$= t^2 (h_x - t h_y^2)(\cancel{t} h_x - 3t h_y^2) =$$

$$= 3h_y^4 t^2 \left(\frac{h_x}{h_y^2} - t\right)\left(\frac{h_x}{3h_y^2} - t\right)$$



istotny jest stosunek $\frac{h_x}{h_y^2}$

$1/2$

1

2

3

-1

-2

Widać, że dla każdego kierunku w $t=0$ mamy minimum

$\to$ mimo to funkcja nie ma minimum!

Ze wzoru definiującego funkcję wynika, że funkcja ta przyjmuje wartości równe zero na krzywych

$x = y^2$ i $x = 3x^2$ a pomiędzy nimi jest ujemna! (8)

Widać więc, że potrzebujemy więcej informacji o punkcie krytycznym niż tylko fakt, że jest on krytyczny tzn. znika pochodna.

Będzie to tematem następnego wykładu.
Teraz wrócimy na chwilę do metody najmniejszych kwadratów i poszukajmy punktu krytycznego funkcji $F(a,b)$:

$$F(a,b) = \sum_{i=1}^{n} \left(y_i - (ax_i + b) \right)^2$$

$$\frac{\partial F}{\partial a} = \sum_{i=1}^{n} 2\left(y_i - (ax_i+b) \right)(-x_i) =$$

$$= 2\sum_{i=1}^{n} (-x_i y_i) + a x_i^2 + b x_i =$$

$$= 2\left(-\sum_{i=1}^{n} x_i y_i + a \sum_i x_i^2 + b \sum_i x_i \right) =$$

$$= 2n\left(-E(XY) + aE(X^2) + bE(X) \right)$$

$$\frac{\partial F}{\partial b} = \sum_{i=1}^{n} 2\left(y_i - (ax_i+b) \right)(-1) =$$

$$= -2\left(\sum_{i=1}^{n} (y_i - ax_i - b) \right) =$$

$$= -2\left(\sum y_i - a\sum x_i - b\sum 1 \right) =$$

$$= -2n\left(E(Y) - aE(X) - b \right)$$

$$0 = -E(XY) + aE(X^2) + bE(X)$$
$$0 = E(Y) - aE(X) - b$$ 9

$$b = E(Y) - aE(X)$$

$$E(XY) = aE(X^2) + \cancel{b}\left(E(Y) - aE(X)\right)E(X) =$$
$$= aE(X^2) + E(Y)E(X) - aE^2(X) =$$
$$= a\left(E(X^2) - E^2(X)\right) + E(Y)E(X)$$
$$\underbrace{E(XY) - E(Y)E(X)}_{\text{kowariancja } C(X,Y)} = a\left(\underbrace{E(X^2) - E^2(X)}_{\text{wariancja } D^2(X)}\right)$$

współczynniki regresji liniowej

$$a = \frac{C(X,Y)}{D^2(X)}$$
$$b = E(Y) - aE(X) \quad : \quad E(Y) = aE(X) + b$$
$$y = ax + b$$

pojęcie statystyczne

punkt $(E(X), E(Y))$ spełnia równanie prostej!

kowariancja

$$C(X,Y) = \frac{1}{n}\sum_{i=1}^{n}(x_i - E(X))(y_i - E(Y)) =$$
$$= \frac{1}{n}\sum_{i=1}^{n} x_i y_i - \left(\frac{1}{n}\sum x_i\right)\left(\frac{1}{n}\sum y_i\right)$$

wariancja

$$D^2(X) = \frac{1}{n}\sum_{i=1}^{n}(x_i - E(X))^2 =$$
$$= \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \left(\frac{1}{n}\sum x_i\right)^2$$