Metoda mnożników Lagrange'a

Przemysław Ryś

1. Motywacja i założenia

W analizie mikroekonomicznej spotykamy się często z problemem znalezienia miejsca, gdzie zadana funkcja przyjmuje największą lub najmniejszą wartość przy pewnym ograniczeniu narzuconym na zbiór. Przykładem może być szukanie koszyka o największej użyteczności, leżącego w zbiorze budżetowym albo szukanie najmniejszych kosztów przy ustalonej ilości produkcji (czyli najmniejszej wartości funkcji kosztów na ustalonej izokwancie). W oczywisty sposób takie miejsce, zwane ***ekstremum warunkowym*** nie musi być tożsame z ekstremum lokalnym. Funkcja użyteczności doskonałych substytutów $U(x,y)=x+y$ ma pochodne kierunkowe równe 1 w każdym punkcie. Pochodna nigdzie się nie zeruje, nie ma więc ekstremum lokalnego, a jednak znajdujemy najlepsze możliwe wybory przy ograniczonym budżecie. Metodą przydatną przy rozwiązywaniu sporej części takich problemów okazuje się ***metoda mnożników Lagrange'a***.

Kiedy możemy z niej skorzystać?

- Funkcja g, której ekstremów szukamy musi być określona na pewnym **zbiorze otwartym** Ω w R^n i przyjmować wartości rzeczywiste. Taka funkcja przypisuje ciągowi n-liczb jedną wartość liczbową. g: $(x_1,x_2,...,x_n) \rightarrow y$. Przykładowo funkcja produkcji przyporządkowuje maksymalną wielkość produkcji do ciągu zadanych ilości różnych czynników produkcji.

- Funkcja g musi mieć **ciągłe pochodne cząstkowe**, co implikuje różniczkowalność dla funkcji wielu zmiennych.

- Zbiór M na którym optymalizujemy funkcje składa się z punktów zbioru Ω, spełniających określone równania $\mathbf{F_1(x_1,x_2,...,x_n)=0}$, $\mathbf{F_1(x_1,x_2,...,x_n)=0}$, **...**, $\mathbf{F_m(x_1,x_2,...,x_n)=0}$.
  W praktyce może być tylko jedno równanie.

  Przykładem może być równanie ograniczenia budżetowego $p_1x_1+p_2x_2+...+p_nx_n-m=0$, gdzie m- budżet, p_i- cena i-tego dobra, x_i- ilość i-tego dobra. W tej sytuacji nasza funkcja wiążąca zmienne to $\mathbf{F(x_1,x_2,...,x_n)=p_1x_1+p_2x_2+...+p_nx_n-m}$, a zbiór na którym szukamy ekstremum opisany jest równaniem $F(x_1,x_2,...,x_n)=0$ czyli $p_1x_1+p_2x_2+...+p_nx_n-m=0$.

- Funkcje wiążące zmienne, podobnie jak funkcja g muszą być określone w każdym punkcie zbioru Ω i mieć ciągłe pochodne cząstkowe, czyli być klasy C^1.

- Funkcja $F=(F_1,F_2,...,F_m)$, czyli funkcja która każdemu ciągowi n liczb przyporządkowuje ciąg m liczb- wartości kolejnych funkcji $F_1,F_2,...,F_m$, ma w każdym punkcie zbioru Ω różniczkę o **maksymalnym rzędzie macierzy** (czyli różniczkę będącą epimorfizmem liniowym- przekształceniem surjektywnym).

O co tu chodzi?

Różniczka funkcji wielu zmiennych F w punkcie x_0 to przekształcenie liniowe o macierzy (zwanej jakobianem) postaci:

$$\begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix}$$

czyli macierzy takiej, że na miejscu o **i-tym wierszu** i **j-tej kolumnie** stoi pochodna cząstkowa funkcji $\mathbf{F_i}$ po zmiennej $\mathbf{x_j}$. Chcemy, żeby ilość liniowo niezależnych wierszy (równoważnie kolumn) była równa **min(m,n)** - gdzie m,n to wymiary macierzy (odpowiednio ilość funkcji F_i i zmiennych x_i). Oczywiście ilość liniowo niezależnych wierszy nie może być większa od ilości wszystkich wierszy- więc ten warunek oznacza **maksymalną ilość liniowo niezależnych wierszy (kolumn)**

Aby sprawdzić to założenie można skorzystać z metody eliminacji Gaussa (tzw. schodkowanie). Dla macierzy kwadratowych spełnienie założeń gwarantuje niezerowy wyznacznik, a dla zbioru opisanego przez jedną funkcję F_1 wystarczy pokazać, że w rozpatrywanym zbiorze Ω nie ma punktu, w którym wszystkie pochodne cząstkowe F_1 są równe zero.

Zwróćmy uwagę na szczególny rodzaj ograniczeń- **ograniczenia liniowe.** Są to ograniczenia, gdzie zmienne występują w pierwszej potędze, tj. $p_1x_1+p_2x_2+...+p_nx_n$«m. (m>0). Przykładem może być ograniczenie budżetowe, Dla zbioru o jednym ograniczeniu liniowym macierz różniczki to wektor $[p_1,p_2,...,p_n]$, który może być zerowy tylko, jeśli wszystkie p_i byłyby zerowe, ale wtedy nie byłoby ograniczenia budżetowego- nierówność 0«m spełniają wszystkie koszyki.

Dla kilku ograniczeń liniowych macierz różniczki również ma maksymalny rząd, o ile równania są liniowo niezależne. Jeśli nie to możemy nasz układ równań zapisać za pomocą mniejszej ilości równań liniowo niezależnych. Dla jednego ograniczenia liniowego lub kilku liniowo niezależnych ograniczeń warunek ten jest zawsze spełniony, stąd w wielu przypadkach analizy mikroekonomicznej ten punkt pomijamy.

2. Opis metody

Przy spełnieniu powyższych założeń możemy mówić o **warunku koniecznym ekstremum warunkowego:** Jeśli funkcja g osiąga w punkcie p ∈ M swój kres górny lub dolny na zbiorze M to istnieją liczby rzeczywiste $\boldsymbol{\lambda_1, \ldots, \lambda_m}$ nazywane **mnożnikami Lagrange'a**, takie że macierz różniczki funkcji g jest kombinacją liniową macierzy różniczek funkcji F_i, czyli zachodzi:

$$g'(x_1,...,x_n)_{x1} = \lambda_1 F_1'(x_1,...,x_n)x_1 + \lambda_2 F_2'(x_1,...,x_n)x_1 + ... + \lambda_n F_n'(x_1,...,x_n)x_1$$

$$g'(x_1,...,x_n)_{x2} = \lambda_1 F_1'(x_1,...,x_n)x_2 + \lambda_2 F_2'(x_1,...,x_n)x_2 + ... + \lambda_n F_n'(x_1,...,x_n)x_2$$

...

$$g'(x_1,...,x_n)_{x1} = \lambda_1 F_1'(x_1,...,x_n)x_n + \lambda_2 F_2'(x_1,...,x_n)x_n + ... + \lambda_n F_n'(x_1,...,x_n)x_n$$

gdzie $g'(x_1,\ldots,x_n)_{\mathbf{xi}}$ oznacza pochodną kierunkową funkcji g po zmiennej $\mathbf{x_i}$.

Równoważne jest to oczywiście układowi postaci:

$$g'(x_1,\ldots,x_n)_{x1} - \lambda_1 F_1'(x_1,\ldots,x_n)x_1 - \lambda_2 F_2'(x_1,\ldots,x_n)x_1 - \ldots - \lambda_n F_n'(x_1,\ldots,x_n)x_1 = 0$$

$$g'(x_1,\ldots,x_n)_{x2} - \lambda_1 F_1'(x_1,\ldots,x_n)x_2 - \lambda_2 F_2'(x_1,\ldots,x_n)x_2 - \ldots - \lambda_n F_n'(x_1,\ldots,x_n)x_2 = 0$$

...

$$g'(x_1,\ldots,x_n)_x n - \lambda_1 F_1'(x_1,\ldots,x_n)x_n - \lambda_2 F_2'(x_1,\ldots,x_n)x_n - \ldots - \lambda_n F_n'(x_1,\ldots,x_n)x_n = 0$$

Metoda polega na znalezieniu mnożników $\lambda_1, \ldots, \lambda_m$ i punktów $(x_1,\ldots,x_n)$ spełniających ten układ równań.

Funkcję $\mathbf{L(x_1,\ldots,x_n,\lambda_1,\ldots,\lambda_m) = g(x_1,\ldots,x_n) - \lambda_1 F_1(x_1,\ldots,x_n) - \ldots - \lambda_n F_n(x_1,\ldots,x_n)}$ nazywamy **Lagrangianem**, a rozwiązanie powyższego układu równań jest równoważne znajdywaniu miejsc, gdzie zerują się jej wszystkie pochodne cząstkowe.

Dla zbioru M opisanego jednym równaniem F_1 dostajemy prostszy układ równań:

$$g'(x_1,\ldots,x_n)_{x1} = \lambda_1 F_1'(x_1,\ldots,x_n)x_1$$

$$g'(x_1,\ldots,x_n)_{x2} = \lambda_1 F_1'(x_1,\ldots,x_n)x_2$$

...

$$g'(x_1,\ldots,x_n)_{x1} = \lambda_1 F_1'(x_1,\ldots,x_n)x_n$$

W tej sytuacji Lagrangian to $L(x_1,\ldots,x_n,\lambda_1) = g(x_1,\ldots,x_n) - \lambda_1 F_1(x_1,\ldots,x_n)$

Spełnienie tego układu (zerowanie się pochodnych cząstkowych Lagrangianu) jest warunkiem koniecznym ekstremum związanego, więc po znalezieniu punktów **nie mamy pewności,** czy są ekstremami. W szczególności możemy znaleźć więcej niż 2 takie punkty o różnych wartościach, mamy przecież co najwyżej jedną największą i najmniejszą wartość.

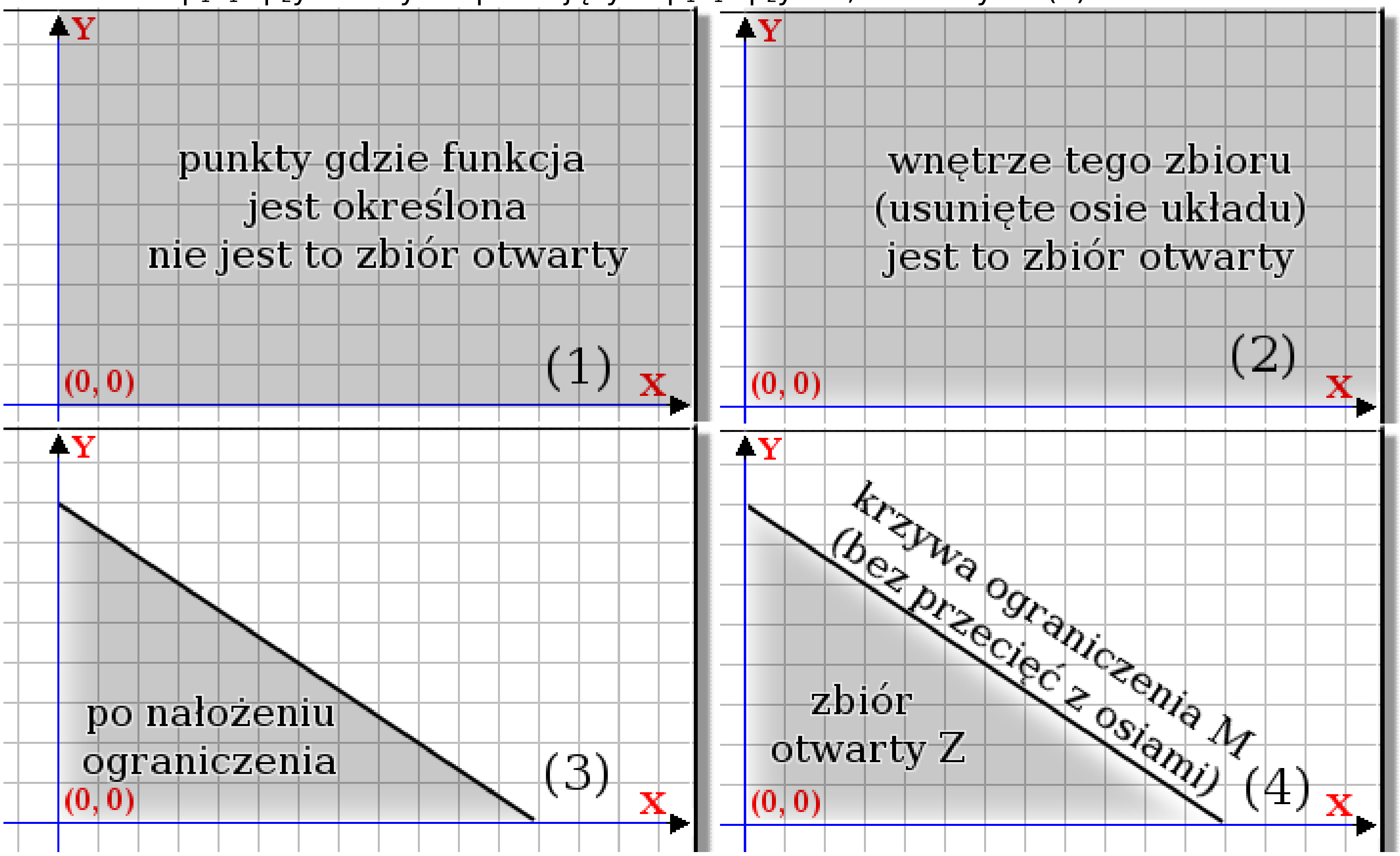
Jest to sytuacja analogiczna do szukaniu ekstremum funkcji jednej zmiennej - zerowa pochodna jest konieczna, aby w miejscu było ekstremum lokalne, ale **nie jest wystarczającym warunkiem**!

Punkty **,,podejrzane o bycie ekstremum"** sprawdzamy, badając określoność drugiej różniczki Lagrangianu- dodatnia oznacza minimum, ujemna maksimum. Jest to analogiczne do badania znaku drugiej pochodnej funkcji jednej zmiennej. Dla zbioru, który jest zwarty (domknięty i ograniczony) zachodzi twierdzenie Weierstrassa - funkcja osiąga swoje kresy. Istnieją wtedy punkty o największej i najmniejszej wartości. Wtedy sprawdzamy wartości podejrzanych punktów i patrzymy na to, które z nich są największe i najmniejsze.

3. Przykład - wyznaczenie największej wartości funkcji Cobba-Douglassa $Q(x,y)=x^{\alpha}y^{\beta}$ na zbiorze produkcyjnym ($\alpha,\beta>0$).

1. Zadbanie o odpowiedni zbiór

Funkcja produkcji określona jest dla każdej **nieujemnej kombinacji produktów (1).** Zbiór punktów na płaszczyźnie im odpowiadający nie jest niestety zbiorem otwartym. Nie możemy więc oczekiwać, że metoda Lagrange'a da oczekiwany rezultat. Ograniczmy się więc jedynie do dodatnich liczb (2), a brzeg, czyli te koszyki gdzie rezygnujemy z zakupu jednego dobra pomijamy. Skoro produkcja jest tam zerowa, nie ma tam więc na pewno największej wartości. Teraz wprowadzamy ograniczenie budżetowe, pozwalając sobie na wykorzystanie jedynie koszyków z prawej górnej ćwiartki układu spełniających nierówność $p_1x_1+p_2y$«m, gdzie wszystkie stałe są dodatnie (3). Rozpatrzmy je jako dwa zbiory, zbiór Z: punktów spełniających nierówność $p_1x_1+p_2y<m$ i tych spełniających $p_1x_1+p_2y=m$, nazwany M. (4)



Dla pierwszego z nich - Z, który jest otwarty, ekstremum globalne musi być w szczególności ekstremum lokalnym - patrzymy więc czy gdziekolwiek zerują się pochodne cząstkowe.

$$Q'(x,y)_x=\alpha x^{\alpha-1}y^{ß}\,, \qquad Q'(x,y)_y=\alpha x^{\alpha-1}y$$

Jak widać, pochodne cząstkowe nie zerują się na naszym zbiorze Z, ponieważ zarówno x jak i y są dodatnie- czyli w szczególności niezerowe. Nie ma tam więc ekstremum lokalnego. Możemy przejść do ostatniego punktu, czyli szukania największej wartości na zbiorze M.

2. **Sprawdzenie założeń**

Nasza funkcja **$Q(x,y)=x^{\alpha}y^{ß}$** jest określona na zbiorze otwartym par liczb dodatnich - R_+^2, podobnie jak funkcja **$F(x,y)=p_1x+p_2y-m$,** której miejsca zerowe są krzywą ograniczenia budżetowego.

Obie mają też ciągłe pochodne cząstkowe, dla funkcji Q(x,y) wyznaczyliśmy je wcześniej. Dla funkcji F(x,y): $F'(x,y)_x=p_1$, $F'(x,y)_y=p_2$ - pochodne cząstkowe są funkcjami stałymi- które są oczywiście ciągłe.

Ostatnie założenie mówi o maksymalnym rządzie macierzy różniczki funkcji F(x,y). Macierz ta składa się z pochodnych cząstkowych $DF(x,y)=[p_1,p_2]$ i ma jeden liniowo niezależny wiersz niezależnie od wartości x i y- założenie jest spełnione. Przypomnę, że dla jednego ograniczenia maksymalny rząd różniczki tożsamy jest z tym, że pochodne cząstkowe się nie zerują w jednym miejscu.

3**. Wyznaczenie układu równań i obliczenie wyniku**

Zapisujemy lagrangian **$L(x,y,\lambda) = Q(x,y)-F(x,y)=x^{\alpha}y^{ß}-\lambda(p_1x_1+p_2y-m)$** , wyznaczamy pochodne cząstkowe i przyrównujemy do zera:

$$L'(x,y,\lambda)_x=\alpha x^{\alpha-1}y^{ß} - \lambda p_1 = 0 \quad \rightarrow \quad \lambda = \frac{\alpha x^{\alpha-1} yß}{p_1}$$

$$L'(x,y,\lambda)_y=ßx^{\alpha}y^{ß-1} - \lambda p_2 = 0 \quad \rightarrow \quad \lambda = \frac{ß x^{\alpha} y^{ß-1}}{p_2}$$

przyrównujemy do siebie te równania stronami

$$\frac{\alpha x^{\alpha-1} y^{ß}}{p_1} = \frac{ß x^{\alpha} y^{ß-1}}{p_2} \quad \rightarrow \quad x= \frac{\alpha y p_2}{p_1 ß} \quad (*)$$

ale przecież x i y związane są też równaniem ograniczenia budżetowego:

$$p_1x+p_2y-m=0 \quad \rightarrow \quad x=\frac{m-yp_2}{p_1} \quad (**)$$

przyrównajmy do siebie dwa równania (*) i (**):

$$\frac{\alpha y p_2}{p_1 ß}=\frac{m-yp_2}{p_1} \quad \rightarrow \quad y=\frac{mß}{(\alpha+ß)p_2}$$

skoro znamy ilość y, z równania ograniczenia możemy dowiedzieć się ile wykorzystamy x:

$$x=m-\frac{yp_2}{p_1}=m-\frac{mß}{(\alpha+ß)p_2}\frac{p_2}{p1}=\frac{m\alpha}{(\alpha+ß)p_1}$$

4. Weryfikacja wyniku

Dostaliśmy rozwiązanie układu - pozostaje pytanie, czy jest ono rzeczywiście ekstremum naszej funkcji na zbiorze M. Możemy podstawić wyznaczoną wartość λ do lagrangianu i zbadać drugą różniczkę. Możemy też zauważyć, że zbiór *M u {(* $\frac{m}{p_1}$ *,0),(0,* $\frac{m}{p_2}$ *)}*, czyli **cała krzywa ograniczenia budżetowego** (włącznie z sytuacją gdzie jeden czynnik nie jest używany) jest domknięta i ograniczona - funkcja musi na niej osiągać swoje kresy. Kres dolny równy 0 osiąga przy rezygnacji z użycia jednego z czynników. Kres górny osiągnąć więc musi w wyznaczonym punkcie.

Ostatecznie funkcja osiąga największą wartość w punkcie:

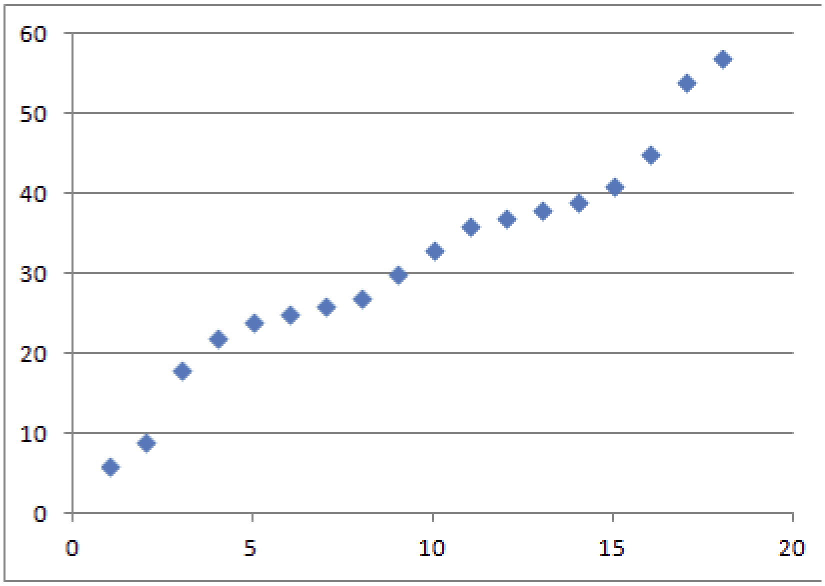
$$x=\frac{m\alpha}{(\alpha+ß)p_1} \qquad y=\frac{mß}{(\alpha+ß)p_2}$$

Rozwiązanie tego problemu dobrze obrazuje zastosowanie mnożników Lagrange'a. Ta metoda okazuje się szczególnie użyteczna dla większej ilości zmiennych- w sytuacji z dwoma dobrami istnieją szybsze rozwiązania, które nie odwołują się do analizy wielowymiarowej. Można zastosować następujące rozumowanie: Funkcja jest monotoniczna (rosnąca), więc ekstremum przyjmie na brzegu zbioru. Szukanie ekstremum na odcinku można za pomocą parametryzacji (uzależnienia jednej zmiennej od drugiej, albo obu od zmiennej niezależnej) sprowadzić do szukania ekstremum funkcji jednej zmiennej na pewnym przedziale.

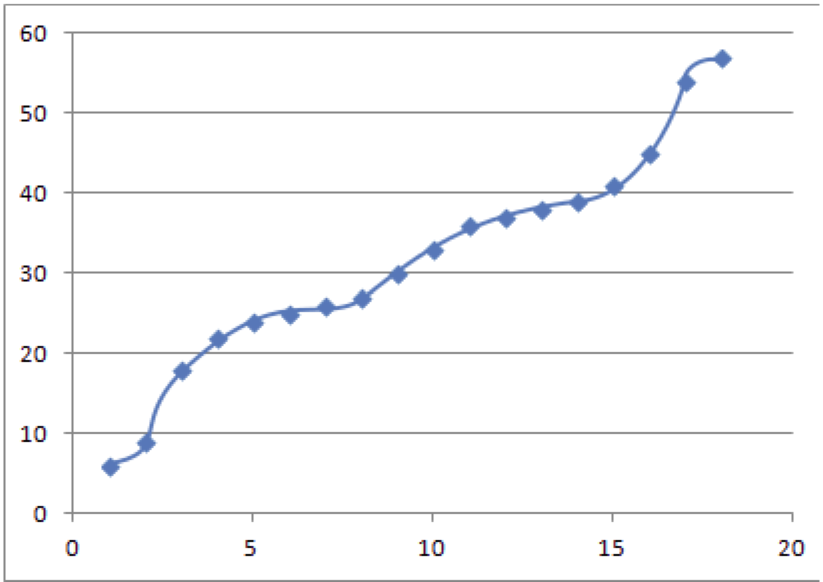
4. Dlaczego te dziwne założenia są konieczne?

Praktycznie każda funkcja z jaką się spotkamy będzie miała **ciągłe pochodne cząstkowe** - bez istnienia pochodnych metoda traci zupełnie sens. Ich ciągłość gwarantuje nam **różniczkowalność funkcji wielu zmiennych**.

Oczywiście bez otwartości zbioru **nie możemy mówić o różniczkowalności** - jeśli w żadnym otoczeniu punktu funkcja nie jest określona, to nie możemy badać przyrostów na nim. Innymi słowy jeśli zbiór nie jest otwarty to mamy w nim punkt, który nie ma żadnego otoczenia w tym zbiorze. W takiej sytuacji nie możemy zapisać ilorazów różnicowych, więc tym bardziej wyznaczyć ich granicy, którą jest pochodna.



Tutaj mamy funkcję dyskretną, określoną na skończonej ilości punktów - nie jest to zbiór otwarty. Jak wiemy wyznaczenie pochodnej nie jest możliwe w żadnym punkcie. Często w ekonomii rozpatrujemy model ciągły dopasowany do pomiarów- ale nawet wtedy mówienie o pochodnej w pewnych punktach może nie mieć sensu.



Problemem w tej sytuacji jest wskazanie pochodnej w punkcie 1 - czym miałaby być styczna do wykresu?

Dla każdego punktu poza 1 i 18, czyli z wnętrza naszego przedziału - (1,18) możemy rozpatrywać przyrosty funkcji prawo i lewostronne- bez tego nie ma pochodnych.

Najmniej naturalnym założeniem wydaje się być warunek o **maksymalnym rzędzie macierzy różniczki** funkcji F, której miejsca zerowe opisują zbiór.

Przypomnijmy, że różniczka funkcji ***F=(F₁,F₂,...,Fₘ)*** w punkcie ***a*** (*gdzie* $F_1,F_2,...,F_m$ przyjmują wartości rzeczywiste) to przekształcenie liniowe, które przyporządkowuje wektorowi ciąg **pochodnych kierunkowych** funkcji $F_1,F_2,...,F_n$. Dla wektora bazowego (np. [1,0,0]) różniczka zwraca nam ciąg pochodnych cząstkowych funkcji $F_1,F_2,...,F_n$. Dla dowolnego wektora **u** zwraca nam pochodne kierunkowe, czyli granice przyrostu funkcji po wektorze u.

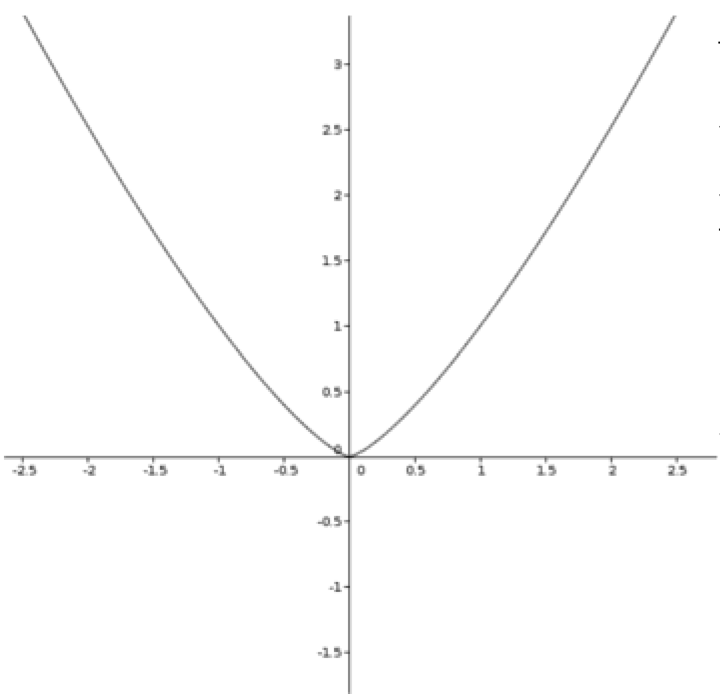
$$\frac{\partial f(\mathbf{x})}{\partial \mathbf{u}} = \lim_{t\to 0} \frac{f(\mathbf{x}+t\mathbf{u}) - f(\mathbf{x})}{t},$$

Pochodna kierunkowa funkcji f(x) po wektorze u.

Widzimy tutaj analogie do przypadku jednowymiarowego z tą różnicą, że liczymy tu krańcowy przyrost przesuwając się nie po osi OX, ale po prostej **x+tu**, czyli prostej zawierającej punkt x o kierunku wyznaczonym przez wektor u. Dla wektora u=(1,0,...,0) dostajemy po prostu pochodną cząstkową po pierwszej zmiennej.

Rozpatrujemy jakobian, czyli macierz różniczki funkcji ***F=(F₁,F₂,...,Fₘ)*** w punkcie $(x_1,...,x_n)$.
Co oznacza dla nas jej maksymalny rząd? Dokładnie to, że nie istnieje wektor, po którym pochodne kierunkowe każdej z funkcji ***F₁,F₂,...,Fₘ*** w tym punkcie są zerowe.

Dla funkcji F przyjmującej wartości rzeczywiste macierz **ma tylko jeden wiersz**. Nie osiąga więc maksymalnej ilości liniowo niezależnych wierszy (równej 1) tylko w sytuacji, kiedy wiersz jest zerowy, czyli wszystkie pochodne cząstkowe są zerowe. Maksymalny rząd różniczki takiej funkcji F jest równoznaczny z tym, że nie ma punktów, gdzie wszystkie pochodne się zerują.



Przykład:

Weźmy podzbiór płaszczyzny taki, że $y^3=x^4$. czyli $F(x,y)=y^3-x^4=0$. Chcemy znaleźć punkt z tego zbioru o najmniejszej współrzędnej y. Szukamy więc **minimum warunkowego** funkcji g(x,y)=y na zadanym zbiorze.

Skoro x^4 jest nieujemne to y również. g(x,y)=y przyjmuje najmniejszą wartość równą 0 w punkcie (0,0). Sprawdźmy, czy metoda mnożników Lagrange'a da nam spodziewany wynik.

Funkcje **F i g** mają ciągłe pochodne cząstkowe (wyznaczone poniżej), określone są na całej płaszczyźnie, która jest zbiorem otwartym.

Wyznaczmy różniczkę funkcji F:
$DF(x,y)=[F'(x,y)_x,F'(x,y)_y,]=[4x^3,3y^2]$

Wyznaczmy różniczkę funkcji g:
$Dg(x,y)[g'(x,y)_x,g'(x,y)_y,]=[0,1]$

Zapisujemy równanie warunku koniecznego ekstremum warunkowego: $Dg(x,y)=\lambda DF(x,y)$
$[0,1]=\lambda[4x^3,3y^2]$, czyli w punkcie ekstremum musi zachodzić:

$0=\lambda 4x^3 \rightarrow x=0 \rightarrow y=0$
$1=\lambda 3y^2 \rightarrow y=\frac{1}{\sqrt{\lambda 3}}$ lub $y=\frac{-1}{\sqrt{\lambda 3}}$

Ale ten układ jest sprzeczny niezależnie od doboru λ! Pierwsze równanie chce zerowego y, a drugie y o konkretnej- niezerowej wartości. Wynika z tego, że nasza funkcja g nie ma ekstremum warunkowego na zbiorze $y^3=x^4$.

Jednocześnie wiemy, że $0 \ll x^4=y^3 \rightarrow 0 \ll y$, ale mamy punkt (0,0) spełniający $y^3=x^4$. y jest nieujemny i osiąga 0, więc zero to jego najmniejsza możliwa wartość. Funkcja g(x,y)=y ma swoje minimum warunkowe, co widać na rysunku.

Co poszło nie tak? Zapomnieliśmy o kluczowym założeniu maksymalnego rzędu macierzy DF(x,y).

$$DF(x,y)=[4x^3,3y^2]$$

Jak widzimy macierz jest niezerowa wszędzie poza punktem (0,0). Metoda mnożników Lagrange'a działa więc w każdym punkcie płaszczyzny, poza punktem (0,0), gdzie było nasze ekstremum.

Dziękuje za poświęcenie czasu na przeczytanie tekstu. Mam nadzieję, że okazał się przydatny i dzięki niemu udało się nieco lepiej zrozumieć metodę mnożników. Uwagi i sugestie można zgłaszać drogą mailową: *przrys@gmail.com.*