Maciej Grzesiak

Metoda mnożników Lagrange'a i jej zastosowania w ekonomii

1. Metoda mnożników Lagrange'a znajdowania ekstremum warunkowego

Pochodna kierunkowa i gradient.
Dla prostoty ograniczymy się do funkcji dwóch zmiennych.
Pochodną kierunkową funkcji f w punkcie (x_0, y_0) w kierunku $\vec{v}$ określamy wzorem:

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

gdzie $\vec{v} = (v_x, v_y)$.
Jeżeli α jest kątem jaki wektor $\vec{v}$ tworzy z osią Ox, to $\cos\alpha = \frac{v_x}{|v|}$, $\sin\alpha = \frac{v_y}{|v|}$. Stąd wynika wzór:

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

stosowany do obliczania pochodnej kierunkowej.
Gradient (oznaczenia $\operatorname{grad} f$ lub ∇f) definiujemy wzorem:

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

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

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

a więc maksymalna wartość wynosi $|\nabla f|$ i jest osiągana wtedy, gdy $\cos\angle(\nabla f, \vec{v}) = 1$, czyli gdy wektor $\vec{v}$ ma kierunek gradientu.
Gradient jest tu kluczowym pojęciem. Dla funkcji $f : \mathbb{R}^n \to \mathbb{R}$ definiujemy go wzorem:

$$\operatorname{grad} f(x_1, x_2, \ldots, x_n) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n}\right).$$

Problem optymalizacji: znaleźć minimalną wartość funkcji $f(\mathbf{x})$ dla $\mathbf{x} \in \mathbb{R}^n$ takich, że

$$g_1(\mathbf{x}) = 0,\ g_2(\mathbf{x}) = 0,\ \ldots,\ g_k(\mathbf{x}) = 0.$$

Należy więc znaleźć minimum funkcji na zbiorze określonym ograniczeniami równościowymi (więzami).
Jak wiadomo, gradient $\nabla f(\mathbf{x})$ wskazuje kierunek, w którym funkcja najszybciej rośnie (analogicznie, $-\nabla f(\mathbf{x})$ wskazuje kierunek, w którym funkcja najszybciej maleje).
Intuicyjnie więc czujemy, że w ekstremum, gdzie nie ma czego wskazywać, gradient powinien być równy 0.
Jeśli jednak musimy się ograniczyć do pewnej powierzchni, to gradient mógłby wskazywać kierunek na zewnątrz. Ale jego rzut na przestrzeń styczną S do powierzchni będzie wskazywał kierunek najszybszego wzrostu spośród kierunków dopuszczalnych. Punkty, w których funkcja będzie mogła mieć ekstremum to takie punkty, w których rzut gradientu na przestrzeń styczną jest równy 0.
Przykład. Dla sfery $x^2 + y^2 + z^2 = r^2$ w $\mathbb{R}^3$ mamy $g(x, y, z) = x^2 + y^2 + z^2 - r^2$, $\nabla g(x_0, y_0, z_0) = (2x_0, 2y_0, 2z_0)$.
W punkcie (x_0, y_0, z_0) płaszczyzna styczna ma więc równanie

$$x_0(x - x_0) + y_0(y - y_0) + z_0(z - z_0) = 0,$$

czyli

$$xx_0 + yy_0 + zz_0 = r^2.$$

Przykład. Dla "powierzchni" określonej funkcjami: $g_1(x, y, z) = x^2 + y^2 + z^2 - 4$, $g_2(x, y, z) = x + y + z - 2$ (jest to oczywiście okrąg) obliczamy:

$$\nabla g_1 = (2x, 2y, 2z), \quad \nabla g_2 = (1, 1, 1).$$

Gradienty te wyznaczają płaszczyznę ("przestrzeń" wektorów prostopadłych do okręgu). Np. w punkcie $(2, 0, 0)$ iloczyn wektorowy gradientów wynosi $\nabla g_1 \times \nabla g_2 = (0, -4, 4)$, więc ta płaszczyzna ma równanie $y - z = 0$. Natomiast "przestrzeń styczna" jest w tym przypadku prostą prostopadłą do tej płaszczyzny; jej równania parametryczne to $x = 2, y = -t, z = t$.
Ogólniej, jeśli powierzchnia jest określona warunkami

$$g_1(\mathbf{x}) = 0,\ g_2(\mathbf{x}) = 0,\ \ldots,\ g_k(\mathbf{x}) = 0.$$

to gradienty ∇g_i, $i = 1, 2, \ldots, k$ wyznaczają przestrzeń wektorów prostopadłych do tej powierzchni. Aby gradient ∇f należał do tej przestrzeni (a więc by jego rzut prostopadły na przestrzeń styczną był równy 0), musi on być kombinacją liniową wektorów ∇g_i.
Powyższe rozumowanie prowadzi do następującego twierdzenia.

Twierdzenie 1. *Niech A będzie podzbiorem $\mathbb{R}^n$ zadanym warunkami*

$$g_1(\mathbf{x}) = 0,\ g_2(\mathbf{x}) = 0,\ \ldots,\ g_k(\mathbf{x}) = 0,$$

Załóżmy, że funkcja f przyjmuje w $\mathbf{x}_0 \in A$ ekstremum. Jeżeli f oraz wszystkie funkcje g_i są różniczkowalne w $\mathbf{x}_0$ oraz wektory $\nabla g_1(\mathbf{x}_0), \nabla g_2(\mathbf{x}_0), \ldots, \nabla g_k(\mathbf{x}_0)$ są liniowo niezależne, to istnieją takie liczby rzeczywiste $\lambda_1, \lambda_2, \ldots, \lambda_k$, że funkcja

$$L(\mathbf{x}) = f(\mathbf{x}) - \lambda_1 g_1(\mathbf{x}) - \lambda_2 g_2(\mathbf{x}) - \cdots - \lambda_k g_k(\mathbf{x})$$

ma w punkcie $\mathbf{x}_0$ punkt krytyczny.

Funkcję L nazywamy funkcją Lagrange'a, a liczby λ_i — mnożnikami Lagrange'a. Rodzaj punktu krytycznego (maksimum, minimum, punkt siodłowy) można w przypadku funkcji różniczkowalnej ustalić badając macierz Hessego w tym punkcie. Macierz Hessego zbudowana jest z drugich pochodnych cząstkowych funkcji.

Twierdzenie 2. *Funkcja f przyjmuje w punkcie krytycznym $\mathbf{x}_0 \in A$:*
– maksimum wtedy i tylko wtedy, gdy macierz Hessego jest ujemnie półokreślona;
– mininimum wtedy i tylko wtedy, gdy macierz Hessego jest dodatnio półokreślona.
Ponadto jeśli macierz jest określona, to odpowiednie ekstremum jest ścisłe.

Uwagi.
(1) Warunek liniowej niezależności gradientów warunków (czyli wektorów $\nabla g_1(\mathbf{x}_0)$, $\nabla g_2(\mathbf{x}_0), \ldots, \nabla g_k(\mathbf{x}_0)$) nazywamy warunkiem jakości więzów w punkcie $\mathbf{x}_0$. W szczególności, gdy mamy do czynienia tylko z jedną funkcją g, oznacza on że $\nabla g(\mathbf{x}_0) \neq 0$. Warto też zauważyć, że warunek liniowej niezależności nie może być spełniony, gdy liczba warunków jest większa, niż wymiar przestrzeni.
(2) Punkty, w których funkcja f osiąga ekstrema są (przy spełnieniu podanych w twierdzeniu powyżej warunków) punktami krytycznymi funkcji Lagrange'a. Nie oznacza to jednak, że zachowany jest charakter tych punktów krytycznych. Następujący przykład pokazuje, że w punkcie, w którym f ma maksimum, funkcja L wcale maksimum mieć nie musi.

Przykład. Rozważmy funkcję $f(x,y) = xy$ i poszukajmy jej maksimum na zbiorze $A = \{(x,y) \in \mathbb{R}^2 : \ x + y = 2\}$. Zatem $g(x,y) = x + y - 2$. Dla punktów zbioru A jest $f(x,y) = f(x, 2-x) = x(2-x)$. Ta funkcja przyjmuje maksimum równe 1 w punkcie $x_0 = 1$. Wtedy $y_0 = 1$, zatem szukanym maksimum jest 1 osiągane w $(x_0, y_0) = (1,1)$. Z drugiej strony $L(x,y) = xy - \lambda(x + y - 2)$, $\nabla L(x,y) = (y - \lambda, x - \lambda)$, więc punkty krytyczne L leżące w A spełniają układ równań

$$\begin{array}{ccccccc} & & y & - & \lambda & = & 0 \\ x & & & & -\lambda & = & 0 \\ x & + & y & & & = & 2 \end{array}$$

Rozwiązaniem jest $x = y = \lambda = 1$. Dla $\lambda = 1$ macierz Hessego wynosi

$$D^2L(1,1) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Jest to macierz nieokreślona, zatem $L(x,y)$ ma w $(1,1)$ nie maksimum, lecz punkt siodłowy.
Przykład. Znaleźć minimum funkcji $x^2 + 2y^2 + z^2 + w^2$ z więzami

$$x + y + z + 3w = 1$$
$$x + y + 2z + w = 2$$

Funkcję Lagrange'a definiujemy jako:

$$L(x,y,z,w) = \frac{x^2 + 2y^2 + z^2 + w^2}{2} - \lambda(x+y+z+3w-1) - \mu(x+y+2z+w-2)$$

(czynnik $\frac{1}{2}$ upraszcza rachunki).

Ma ona pochodne cząstkowe

$$
\begin{aligned}
L_x &= x - \lambda - \mu \\
L_y &= 2y - \lambda - \mu \\
L_z &= z - \lambda - 2\mu \\
L_w &= w - 3\lambda - \mu
\end{aligned}
$$

a więc punktem krytycznym jest

$$x_0 = \lambda + \mu, \quad y_0 = \frac{1}{2}\lambda + \frac{1}{2}\mu, \quad z_0 = \lambda + 2\mu, \quad w_0 = 3\lambda + \mu.$$

Podstawiając do równań więzów otrzymamy równania

$$
\begin{aligned}
23\lambda + 13\mu &= 2 \\
13\lambda + 13\mu &= 4.
\end{aligned}
$$

Obliczamy $\lambda = -\frac{13}{65} = -\frac{1}{5}$, $\mu = -\frac{33}{65}$, a następnie punkt krytyczny

$$x_0 = \frac{4}{13}, \quad y_0 = \frac{2}{13}, \quad z_0 = \frac{53}{65}, \quad w_0 = -\frac{6}{65}.$$

W tym punkcie jest minimum równe $\frac{53}{65}$ (macierz Hessego jest diagonalna, określona dodatnio).
Przykład. Znaleźć największą i najmniejszą wartość formy kwadratowej

$$f(\mathbf{x}) = \mathbf{x}\mathbf{A}\mathbf{x}^T,$$

gdzie $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ oraz $\mathbf{A}$ jest macierzą symetryczną,
przy warunku $\|\mathbf{x}\| = 1$ (równoważnie: $\|\mathbf{x}\|^2 - 1 = 0$).
Tworzymy funkcję Lagrange'a

$$L(\mathbf{x}, \lambda) = \mathbf{x}\mathbf{A}\mathbf{x}^T - \lambda(\|\mathbf{x}\|^2 - 1),$$

czyli

$$L(x_1, \ldots, x_n) = \sum_{i,j=1}^{n} a_{ik}x_ix_k - \lambda(\sum_{i=1}^{n} x_i^2 - 1).$$

Dla $j = 1, 2, \ldots, n$ obliczamy

$$\frac{\partial L}{\partial x_j} = 2\sum_{k=1}^{n} a_{jk}x_k - 2\lambda x_j$$

(uwzględniamy fakt, że $a_{jk} = a_{kj}$).
Przyrównując pochodne do 0 otrzymujemy układ równań:

$$
\begin{array}{ccccccccc}
(a_{11} - \lambda)x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & 0 \\
a_{21}x_1 & + & (a_{22} - \lambda)x_2 & + & \cdots & + & a_{2n}x_n & = & 0 \\
\multicolumn{9}{c}{\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots} \\
a_{n1}x_1 & + & a_{n2}x_2 & + & \cdots & + & (a_{nn} - \lambda)x_n & = & 0.
\end{array}
$$

Układ ma rozwiązanie niezerowe, gdy wyznacznik jest równy 0, czyli gdy równanie charakterystyczne

$$\begin{vmatrix} a_{11}-\lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22}-\lambda & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn}-\lambda \end{vmatrix} = 0$$

ma rozwiązanie niezerowe, czyli gdy λ jest wartością własną macierzy $\mathbf{A}$. Zatem wektory $\mathbf{x} = (x_1, x_2, \dots, x_n)$ będące punktami krytycznymi są wektorami własnymi macierzy $\mathbf{A}$, i można je wyznaczyć w zwykły sposób. Jednak jeśli chcemy tylko znać wartość największą i najmniejszą formy (bez informacji gdzie jest osiągana), to można to uzyskać dość prosto. Mianowicie mnożąc równania układu odpowiednio przez $x_1, x_2, \dots, x_n$ i dodając stronami otrzymamy równanie

$$f(x_1, x_2, \dots, x_n) - \lambda(x_1^2 + x_2^2 + \dots + x_n^2) = 0.$$

Ale $x_1^2 + x_2^2 + \dots + x_n^2 = 1$, więc

$$f(x_1, x_2, \dots, x_n) = \lambda.$$

a więc wartość funkcji w punkcie krytycznym jest równa wartości własnej λ. Zatem wartość największa i najmniejsza formy pokrywa się z największą i najmniejszą wartością własną macierzy $\mathbf{A}$.
Np. rozważmy formę $x^2 - 2y^2 + z^2 + 6xy - 2yz$. Forma ta określona jest macierzą

$$\begin{bmatrix} 1 & 3 & 0 \\ 3 & -2 & -1 \\ 0 & -1 & 1 \end{bmatrix},$$

która ma wartości własne -4, 1, 3. Największą wartością formy na sferze jednostkowej $x^2 + y^2 + z^2 = 1$ jest więc 3, a najmniejszą wartością -4.

2. Ekonomiczna interpretacja mnożnika Lagrange'a

Rozważmy problem znalezienia maksimum funkcji $f(\mathbf{x})$ przy warunku $g(\mathbf{x}) = w$. Liczbę w będziemy traktować jak parametr problemu. Funkcja $f(\mathbf{x})$ może być funkcją zysku określonego wartością wejściową w. Dla ustalonej wielkości w niech $\mathbf{x}^*(w)$ oznacza wartość $\mathbf{x}$ maksymalizującą f, czyli $f(\mathbf{x}^*(w))$ jest maksymalnym zyskiem dla ustalonej wartości wejściowej w. Pochodna

$$\frac{\mathrm{d}}{\mathrm{d}\,w} f(\mathbf{x}^*(w))$$

określa zmienność optymalnego zysku w zależności od zmiany parametru w.
Wartość $\mathbf{x}^*(w)$ możemy znaleźć przy pomocy funkcji Lagrange'a:

$$L(\mathbf{x}) = f(\mathbf{x}) - \lambda g(\mathbf{x}).$$

W punkcie krytycznym gradient funkcji Lagrange'a wynosi 0, więc

$$\nabla f(\mathbf{x}^*(w)) = \lambda^*(w) \nabla g(\mathbf{x}^*(w)),$$

gdzie λ^* jest odpowiednią wartością mnożnika Lagrange'a.

Twierdzenie 3. *Mamy*

$$\lambda^*(w) = \frac{\mathrm{d}}{\mathrm{d}\,w} f(\mathbf{x}^*(w)),$$

czyli mnożnik Lagrange'a jest miarą zmiany zysku maksymalnego spowodowanej zmianą parametru w.

D o w ó d. Aby uprościć zapis wykonamy obliczenia na przykładzie funkcji dwóch zmiennych, tzn. gdy $\mathbf{x} = (x_1, x_2)$. Dla n zmiennych dowód przebiega tak samo. Obliczamy pochodną

$$\frac{\mathrm{d}}{\mathrm{d}\,w} f(\mathbf{x}^*(w)) = \frac{\partial f}{\partial x_1}(\mathbf{x}^*(w)) \cdot \frac{\mathrm{d}\,x_1^*}{\mathrm{d}\,w}(w) + \frac{\partial f}{\partial x_2}(\mathbf{x}^*(w)) \cdot \frac{\mathrm{d}\,x_2^*}{\mathrm{d}\,w}(w). \tag{1}$$

Ale $\mathbf{x}^*(w)$ jest rozwiązaniem problemu Lagrange'a, więc dla $i = 1, 2$:

$$\frac{\partial f}{\partial x_i}(\mathbf{x}^*(w)) = \lambda^*(w) \frac{\partial g}{\partial x_i}(\mathbf{x}^*(w)).$$

Podstawiając do równości (1) otrzymujemy

$$\frac{\mathrm{d}}{\mathrm{d}\,w} f(\mathbf{x}^*(w)) = \lambda^*(w) \Big(\frac{\partial g}{\partial x_1}(\mathbf{x}^*(w)) \cdot \frac{\mathrm{d}\,x_1^*}{\mathrm{d}\,w}(w) + \frac{\partial g}{\partial x_2}(\mathbf{x}^*(w)) \cdot \frac{\mathrm{d}\,x_2^*}{\mathrm{d}\,w}(w) \Big). \tag{2}$$

Ale $g(\mathbf{x}^*(w)) = w$ dla każdego w. Różniczkując tę równość otrzymujemy

$$\frac{\mathrm{d}}{\mathrm{d}\,w} g(\mathbf{x}^*(w)) = 1,$$

czyli

$$\frac{\partial g}{\partial x_1}(\mathbf{x}^*(w)) \cdot \frac{\mathrm{d}\,x_1^*}{\mathrm{d}\,w}(w) + \frac{\partial g}{\partial x_2}(\mathbf{x}^*(w)) \cdot \frac{\mathrm{d}\,x_2^*}{\mathrm{d}\,w}(w) = 1,$$

i podstawiając do (2) uzyskujemy tezę.
Gdy f oznacza zysk, a parametr w jest wartością nakładów, to mnożnik Lagrange'a jest zyskiem krańcowym nakładów (*marginal profit of money*). Inaczej, jest to dodatkowy zysk gdy nakłady wzrosną o jednostkę. Gdyby w było ilością komponentu produkcji, to mnożnik Lagrange'a jest dodatkowym zyskiem gdy ilość komponentu wzrośnie o 1; wtedy można to interpretować jako maksymalną cenę, którą producent byłby skłonny zapłacić za jednostkę komponentu.
Gdy f oznacza wielkość produkcji, a parametr w nakłady, to mnożnik Lagrange'a jest produkcją krańcową (*marginal product of money*).

3. Badanie popytu przy pomocy funkcji użyteczności

Przestrzeń towarów i przestrzeń cen

W problemie badania popytu konsumenta rozważa się *przestrzeń towarów* i *przestrzeń cen*.
Załóżmy, że mamy n towarów. Przestrzenią towarów nazywamy zbiór

$$X = \mathbb{R}^n_+ = \{(x_1, x_2, \ldots, x_n) : \; x_i \geqslant 0, \; i = 1, 2, \ldots, n\}$$

gdzie współrzędna x_i określa ilość i-tego towaru.
Przyjmujemy, że wektory przestrzeni X reprezentują koszyki towarów.

Definicja 1. Funkcja użyteczności $u : X \to \mathbb{R}$ jest funkcją, która koszykowi $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ przyporządkowuje pewną liczbę, umożliwiając w ten sposób porównanie użyteczności różnych koszyków dóbr. Zakładamy, że

1. $u(\mathbf{x})$ ma ciągłe pochodne cząstkowe pierwszego i drugiego rzędu;
2. $\frac{\partial u}{\partial x_i}(\mathbf{x}) > 0$ dla $i = 1, 2, \ldots, n$;
3. $\frac{\partial^2 u}{\partial x_i^2}(\mathbf{x}) < 0$ dla $i = 1, 2, \ldots, n$;

Warunek 2 nazywamy *postulatem niedosytu.* W terminologii ekonomicznej mówi się, że krańcowa użyteczność i-tego towaru w koszyku jest dodatnia.
Warunek 3 nazywamy *prawem Gossena.* W terminologii ekonomicznej: krańcowa użyteczność każdego towaru maleje w miarę jak wzrasta jego spożycie.

Maksymalizacja funkcji użyteczności

Załóżmy, że użyteczność dwóch dóbr X, Y w zależności od ich ilości x oraz y wyraża się funkcją różniczkowalną $u(x, y)$.
Jeśli u_0 jest ustaloną wartością funkcji użyteczności $u(x, y)$, to *krzywa obojętności* zdefiniowana jest jako zbiór tych punktów (x, y), które spełniają warunek $u(x, y) = u_0$. Różniczkując tę równość otrzymujemy:

$$\frac{\partial u}{\partial x}\,\mathrm{d}x + \frac{\partial u}{\partial y}\,\mathrm{d}y = 0,$$

a zatem

$$\left.\frac{\mathrm{d}y}{\mathrm{d}x}\right|_{u=u_0} = -\frac{\frac{\partial u}{\partial x}}{\frac{\partial u}{\partial y}}$$

Iloraz $-\frac{\frac{\partial u}{\partial x}}{\frac{\partial u}{\partial y}}$ nazywamy *krańcową stopę substytucji* dobra Y dobrem X (marginal rate of substitution, $MRS_{Y/X}$).
Geometrycznie jest to współczynnik kierunkowy stycznej do krzywej obojętności.
Ekonomicznie jest to taki stosunek przyrostu konsumpcji jednego dobra do ubytku konsumpcji innego dobra, że konsument nie zmienia osiąganej użyteczności (przy założeniu, że jego krzywa obojętności pozostaje niezmieniona).
Niech p_x, p_y będą cenami dóbr X, Y, a m niech będzie budżetem konsumenta.
Znajdziemy maksimum funkcji użyteczności przy ograniczeniu

$$xp_x + yp_y = m.$$

Tworzymy funkcję Lagrange'a

$$L(x, y, \lambda) = u(x, y) + \lambda(m - xp_x - yp_y)$$

Punkt krytyczny tej funkcji spełnia układ równań:

$$\begin{aligned} \frac{\partial u}{\partial x} - \lambda p_x &= 0 \\ \frac{\partial u}{\partial y} - \lambda p_y &= 0 \\ m - xp_x - yp_y &= 0. \end{aligned}$$

Z dwóch pierwszych równań otrzymujemy

$$\frac{\frac{\partial u}{\partial x}}{\frac{\partial u}{\partial y}} = \frac{p_x}{p_y},$$

a więc punkt krytyczny określony jest warunkiem

$$MRS_{Y/X} = -\frac{p_x}{p_y},$$

Macierz Hessego wynosi

$$\begin{bmatrix} u''_{xx} & u''_{xy} \\ u''_{yx} & u''_{yy} \end{bmatrix}.$$

Jeśli jest ona ujemnie półokreślona, to w punkcie krytycznym mamy maksimum. Ale, jak już wiemy, w punkcie, w którym u ma maksimum, funkcja L wcale maksimum mieć nie musi (forma może być nieokreślona)

Definicja 2. Funkcję n zmiennych postaci

$$u(x_1, x_2, \dots, x_n) = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}, \quad \alpha_i > 0,\ x_i > 0$$

nazywamy funkcją Cobba-Douglasa.

Funkcja ta jest wykorzystywana jako funkcja użyteczności, ale także w problemach analizy produkcji (wtedy zmiennymi mogą być np. kapitał, praca i surowce).
Przykład. Załóżmy, że funkcja użyteczności z poprzedniego przykładu jest funkcją Cobba-Douglasa postaci $u = x^\alpha y^{1-\alpha}$, gdzie $0 < \alpha < 1$.
Wtedy funkcja Lagrange'a to

$$L(x, y, \lambda) = x^\alpha y^{1-\alpha} + \lambda(m - xp_x - yp_y)$$

Punkt krytyczny tej funkcji spełnia układ równań:

$$\begin{aligned} \alpha x^{\alpha-1} y^{1-\alpha} - \lambda p_x &= 0 \\ (1-\alpha) x^\alpha y^{-\alpha} - \lambda p_y &= 0 \\ m - xp_x - yp_y &= 0. \end{aligned}$$

Rugując λ z dwóch pierwszych równań otrzymujemy

$$\frac{\alpha}{1-\alpha} \cdot \frac{y}{x} = \frac{p_x}{p_y}.$$

Podstawiając do trzeciego równania otrzymujemy punkt krytyczny:

$$x = \frac{\alpha m}{p_x}, \quad y = \frac{(1-\alpha)m}{p_y}.$$

Widzimy, że wydatki na dobro x wynoszą $p_x x = \alpha m$, a więc wykładnik α określający preferencje konsumenta determinuje jaką część budżetu należy przeznaczyć na dobro x. Analogicznie, $p_y y = (1-\alpha)m$.
Ciekawe jest to, że cena dóbr wydaje się nie mieć znaczenia; jeżeli $\alpha = 0,3$, $m = 1000$, to na dobro x wydamy 300 a na y kwotę 700 niezależnie od ich cen. Wytłumaczenie jest takie: jeśli cena dobra x maleje, to następuje substytucja dobra y dobrem x. Inny aspekt: jeśli ceny rosną, to konsument jest realnie biedniejszy, więc ogranicza wydatki.
Przykład. Konsument może wydać 300 zł na dwa dobra: dobro X kosztuje 5 zł za jednostkę, a dobro Y kosztuje 10 zł za jednostkę. Jakich zakupów powinien dokonać, jeśli jego funkcja użyteczności wynosi $u(x, y) = x^2 y$?

Rozwiązanie. Należy znaleźć maksimum funkcji $u(x, y) = x^2 y$ przy warunku $5x + 10y = 300$, tj. $x + 2y = 60$. Tworzymy funkcję Lagrange'a

$$L(x, y, \lambda) = x^2 y + \lambda(x + 2y - 60)$$

Obliczamy pochodne cząstkowe i punkt krytyczny znajdujemy z układu równań:

$$\begin{aligned} 2xy + \lambda &= 0 \\ x^2 + 2\lambda &= 0 \\ x + 2y - 60 &= 0. \end{aligned}$$

Stąd $\lambda = -800$, $x = 40$, $y = 10$.
Macierz Hessego funkcji Lagrange'a w punkcie (x, y) to:

$$\begin{bmatrix} 2y & 2x \\ 2x & 0 \end{bmatrix}.$$

W punkcie krytycznym $(40, 10)$ jest to macierz nieokreślona:

$$\begin{bmatrix} 20 & 80 \\ 80 & 0 \end{bmatrix}.$$

Mimo to funkcja $u = x^2 y$ ma maksimum.
Ogólniej, niech m będzie budżetem konsumenta, $X = \mathbb{R}^n_+$ przestrzenią towarów oraz $\mathbf{p} = (p_1, p_2, \ldots, p_n)$ wektorem cen tych towarów. Dla koszyka $\mathbf{x} \in X$ iloczyn skalarny

$$< \mathbf{p}, \mathbf{x} >= \sum_{i=1}^{n} p_i x_i$$

jest wartością koszyka. Zbiór

$$B(\mathbf{p}, m) = \{\mathbf{x} \in X : \sum_{i=1}^{n} p_i x_i \leqslant m\}$$

nazywamy *zbiorem dopuszczalnych planów konsumpcji* (krócej: zbiorem budżetowym). Jest to oczywiście zbiór wypukły.
Jeśli $u(\mathbf{x})$ jest użytecznością koszyka $\mathbf{x}$, to *zagadnieniem maksymalizacji użyteczności konsumenta* nazywamy problem wyznaczenia punktu $\mathbf{x}^*$ dla którego

$$u(\mathbf{x}^*) = \max u(\mathbf{x}) \quad \text{przy warunku} \ \sum_{i=1}^{n} p_i x_i \leqslant m.$$

Zbiór

$$\{\mathbf{x} \in X : \ u(\mathbf{x}) = u_0\}$$

nazywamy *powierzchnią obojętności.*
Różniczkując równość $u(\mathbf{x}) = u_0$ otrzymujemy:

$$\frac{\partial u}{\partial x_1}\,\mathrm{d}x_1 + \frac{\partial u}{\partial x_2}\,\mathrm{d}x_2 + \cdots + \frac{\partial u}{\partial x_n}\,\mathrm{d}x_n = 0,$$

więc dla $\mathbf{x}$ należących do powierzchni obojętności gradient ∇u jest prostopadły do wektora $(\mathrm{d}x_1, \mathrm{d}x_2, \ldots, \mathrm{d}x_n)$ stycznego do tej powierzchni. Jak wiadomo, jeśli $u(\mathbf{x})$ osiąga w $\mathbf{x}^*$ maksimum, to $\mathbf{x}^*$ jest punktem krytycznym funkcji Lagrange'a:

$$L(\mathbf{x}) = u(\mathbf{x}) + \lambda(m - \sum_{i=1}^{n} p_i x_i),$$

którą można zapisać jako

$$L(\mathbf{x}) = u(\mathbf{x}) + \lambda m - \lambda < \mathbf{p}, \mathbf{x} > .$$

Przy obliczaniu gradientu zauważmy, że $\nabla(\langle \mathbf{p}, \mathbf{x} \rangle) = \mathbf{p}$. Zatem

$$\nabla L(\mathbf{x}) = \nabla u(\mathbf{x}) - \lambda \mathbf{p}.$$

W punkcie krytycznym $\nabla u(\mathbf{x}^*) - \lambda \mathbf{p} = 0$ lub

$$\nabla u(\mathbf{x}^*) = \lambda \mathbf{p}.$$

A więc jeśli gradient $\nabla u(\mathbf{x}^*)$ ma kierunek wektora cen, to $\mathbf{x}^*$ jest punktem krytycznym i użyteczność osiąga maksimum.