11 Ekstrema: funkcji uwikłanych, warunkowe

Zacznijmy od funkcji uwikłanych. Są to np. funkcje typu $y = y(x)$ stanowiące rozwiązanie równania $g(x, y) = 0$. Przykładem może tu być równanie okręgu: $x^2 + y^2 - 1 = 0$, gdzie mamy (z wyjątkiem 2 punktów) -rozwiązania: $y = \pm\sqrt{1 - x^2}$.

Twierdzenie o funkcjach uwikłanych. Gdy funkcja g ma ciągłe pochodne cząstkowe, a w punkcie $P_0 = (x_0, y_0)$, gdzie $g(P_0) = 0$ spełnia ona **założenie** $g'_y(P_0) \neq 0$, to w pewnym otoczeniu x_0 (np. dla $|x - x_0| < \delta$) istnieje dokładnie jedna funkcja ciągła $y(x)$ taka, że $y(x_0) = y_0$ oraz $g(x, y(x)) = 0$. Ponadto wówczas y jest funkcją klasy C^1 zmiennej x w tym otoczeniu punktu x_0 oraz

$$g'(x) = -\frac{g'_x(x, y(x))}{g'_y(x, y(x))} \tag{1}$$

Analogicznie wygląda wzór na pochodne cząstkowe $z'_x(x, y)$ w przypadku funkcji dwu zmiennych $z(x, y)$ uwikłanej w równaniu $G(x, y, z(x, y)) = 0$. Wtedy zakładamy, że G'_z jest w danym punkcie (x_0, y_0, z_0) różna od zera i ta niezerowa pochodna będzie w mianowniku wzoru na z'_x, licznikiem ułamka (mnożonego przez (-1)) będzie $g'_x(x, y, z(x, y))$. Przy szukaniu ekstremów mamy na pierwszym etapie sprawdzać warunek konieczny- czyli zerowanie się y' (odpowiednio gradientu funkcji $z(x, y)$). Dzięki wzorowi na pochodne funkcji uwikłanej, jest to równoważne zerowaniu się liczników po prawej stronie (1) (albo liczników analogicznych wzorów na z'_x, z'_y) . Jeśli już wiemy, że funkcja uwikłana jest różniczkowalna (to można uzyskać np. z twierdzenia o lokalnej odwracalności, lub bezpośrednio), to wzór (1) łatwo wynika z reguły łańcucha stosowanej do pochodnej (stale równej zero) funkcji złożonej: $g(x, y(x))' = g'_x + g'_y(x, y(x)) \cdot y'(x)$ Rożniczkując ponownie tę równość otrzwmamy nieco dłuższy wzór, który w punkcie, gdzie y' jest równa zero -przyjmie znacznie prostszą postać, dającą analogiczny do (1) wzór na druga pochodna w punktach "podejrzanych o ekstrema, czyli tam, gdzie $y' = 0$. Mamy mianowicie wzory (również w przypadku funkcji z dwu zmiennych):

$$y''(x) = -\frac{g''_{xx}(x, y(x))}{g'_y(x, y(x))}, \quad \texttt{odpowiednio } z''_{xx}(x, y) = -\frac{G''_{xx}(x, y, z(xy))}{G'_z(x, y, z(x, y))}. \tag{2}$$

Podobnie jest w przypadku pozostałych pochodnych cząstkowych 2. rzędu dla funkcji $z(x, y)$:

$$z''_{yy}(x, y) = -\frac{G''_{yy}(x, y, z(xy))}{G'_z(x, y, z(x, y))} \quad \texttt{oraz} \quad z''_{xy}(x, y) = -\frac{G''_{xy}(x, y, z(xy))}{G'_z(x, y, z(x, y))}. \tag{3}$$

Te wzory pozwalają na sprawdzanie warunku wystarczającego dla istnienia ekstremów w punktach stacjonarnych, uwzględniamy zarówno znak minus przed ułamkiem, jak i znak (wspólnego) mianownika oraz dodatnią lub ujemną określoność macierzy drugich pochodnych cząstkowych funkcji G -ale tylko tych dotyczących zmiennych x oraz y.

Przykład. Zbadamy ekstrema $z(x, y)$ jako funkcji uwikłanej w równaniu

$$F(x, y, z) = x^4 + y^4 + z^4 - 2a^2(x^2 + y^2 + z^2) = 0, a > 0.$$

Od razu widać, że $F(x, y, z) = F(x, y, -z)$, więc mamy lokalnie parę funkcji rozwikłujących: $\pm z(x, y)$, gdzie wybór znaku dyktuje nam znak zadanej wartości z_0. Ponieważ $F'_x = 4x^3 - 4a^2x = 0$, mamy albo $x = 0$ albo $x = \pm a$, takie same możliwości-dla y. Ponieważ ma być $F'_z = 4z^3 - 4a^2z \neq 0$, dla $(x, y) = (0, 0)$ wartość $z = 0$ musimy odrzucić i pozostają 2 możliwości: $F(0, 0, z) = 0$ dla

$z = \pm a\sqrt{2}$. Ponumerujmy te rozwiązania P_1, P_2. Dalej, gdy $|x| = a = |y|$, mamy rozwiązania $\pm z$ spełniające $|z| = a\sqrt{1+\sqrt{3}}$, w kórych F_z' jest niezerowe. Dla niezerowych x, y mamy więc jeszcze 8 rozwiązań poprzez dowolność doboru znaków dla każdej z 3 zmiennych - oznaczmy je przez $P_j, 3 \leqslant j \leqslant 10$. Pochodne mieszane są zerowe, więc macierz Hessego dla z będzie diagonalna. Mamy $F_{xx}^{''} = 12x^2 - 4a^2$, $F_{yy}^{''} = 12y^2 - 4a^2$ i te wartości będą jednakowe, bo $|x| = |y|$ w naszych punktach stacjonarnych -więc stałego znaku. Określoność macierzy drugiej różniczki dla z bedzie więc dodatnia dla punktu $P_1 = (0, 0, a\sqrt{2})$ (minimum lokalne funkcji uwikłanej z) i maksimum lokalne w $(0, 0)$ dla $z = -a\sqrt{2}$ (punkt P_2). Podobnie w punktach $P_3, \ldots, P_{10}$ liczniki wzorów (3), czyli pochodne 2. rzędu "niemieszane" z F są takie same, zmienia się znak mianowników - w zależności od doboru znaku rozwiązania z. I tak, dla znaku "plus" wartość $z(x, y)$ równa $a\sqrt{1+\sqrt{3}}$ jest maksimum lokalnym finkcji uwikłanej z, zaś wartość $-a\sqrt{1+\sqrt{3}}$ jest jej minimum lokalnym. Jak widać po punktach P_1, P_2 samo pozównanie wartości w punktach ekstremów nie wystarcza do określenia, kóra z nich jest minimum lokalnym.

11.1 Ekstrema warunkowe.

Jak już wiemy, w pewnych punktach zbiór M rozwiązań równania $G(x) = 0$, gdzie $G : D \subset \mathbb{R}^d \to \mathbb{R}$ lokalnie wygląda jak wykres funkcji. Ale na ogół jest to jednak pewien zwarty podzbiór $\mathbb{R}^n$ -np. sfera, elipsoida. Tak jest zwłaszcza, gdy przy $\|x\| \to +\infty$ jest $|G(x)| \to \infty$ - bo wówczas przeciwobrazy zbiorów zwartych są zwarte. Dla regularnych w odpowiednim sensie funkcji G takie zbiory stanowią pewne "gładkie powierzchnie zakrzywione" zwane rozmaitościami (ang. *manifolds*). Można formalnie wprowadzić ich definicję i np. w przypadku $d = 3$ wektorem normalnym (prostopadłym do płaszczyzny stycznej) dla takiej powierzchni M będzie gradient G. Nas interesują ekstrema funkcji określonych na M, gdzie wartości w punktach spoza M nie uwzględniamy. Takie funkcje f mogą być określone na znacznie większej,niż M dziedzinie- np. na jakimś zbiorze otwartym $U \subset \mathbb{R}^d$, jednak wartości w punktach $x \in U \setminus M$ nas nie interesują. Tak jest np. gdy liczymy kres górny wartości f po sferze jednostkowej -a więc jest to dość często występujące zagadnienie. W takim punkcie $P_0 \in M$ gradient f **nie musi być równy zero**, bo w otoczeniu tego punktu znajdą się punkty spoza M, gdzie f może mieć wartości istotnie większe, niż $\sup\{f(x) : x \in M\}$. Dobrym przykładem jest $f(x, y, z) = x + y + z$- funkcja ciągła osiągająca na sferze swoją wartość największą w pewnym punkcie P_*, ale jej gradient w każdym punkcie $\mathbb{R}^3$ jest taki sam, różny od zera. Pokażemy, jak taki punkt znaleźć, nie odwołując się do teorii rozmaitości. Zacznijmy jednak od definicji.

Definicja. Jeżeli $G_1, \ldots, G_k : D \subset \mathbb{R}^n \to \mathbb{R}$ są odwzorowaniami "wyznaczającymi warunek $G_j = 0$", to mówimy, że przy tym warunku funkcja $f : U \to \mathbb{R}$ ma maksimum lokalne warunkowe w punkcie P_0, jeśli po pierwsze $P_0 \in M := \{P \in D : \forall_{j \leqslant k} G_j(P) = 0\}$ a po drugie istnieje otoczenie W punktu P_0 takie, że gdy $P \in W \cap M$, to $f(P_0) \geqslant f(P)$. Analogicznie definiujemy minimum warunkowe, ścisłe (= silne) maksimum warunkowe.
Funkcją Lagrange'a dla powyższego zagadnienia nazwiemy funkcję

$$\Phi(x) := f(x) + \lambda_1 G_1(x) + \lambda_2 G_2(x) + \cdots + \lambda_k G_k(x).$$

Tu liczby $\lambda_j \in \mathbb{R}$ są pewnymi stałymi, zwanymi mnożnikami Lagrange'a (i na początku traktujemy je jako niewiadome).

Uwaga: Można wykazać (używając tw. o lokalnej odwracalności) twierdzenie o funkcyjnej zależności, które mówi, że gdy gradienty funkcji $G_1, \ldots, G_k$ są liniowo zależne, to w otoczeniu danego punktu którąś z tych funkcji da się wyrazić jako funkcję od pozostałych funkcji. W pewnym sensie, można ją wyeliminować z zestawu definiującego zbiór M. Będziemy więc zakładać liniową niezależność gradientów ∇G_j w interesującym nas punkcie.

Twierdzenie (Warunek konieczny na ekstremum warunkowe). Jeśli w punkcie $P_0 \in M$ funkcja f osiąga lokalne ekstremum warunkowe, to w tym punkcie gradienty funkcji f oraz G_j są liniowo zależne. (Ze względu na niezależność samych $\nabla_{P_0}G_j$ oznacza to, że $\nabla_{P_0}F$ jest liniową kombinacją typu $\sum_{j=1}^{k}(-\lambda_j)\nabla_{P_0}G_j$, czyli P_0 **jest punktem stacjonarnym dla funkcji Lagrange'a Φ.**

Zauważmy, że jeśli M jest zbiorem zwartym, to często już znajomość punktów stacjonarnych wystarcza, by znaleźć punkt, w którym f osiąga któryś ze swoich kresów (górny lub dolny) na zbiorze M. Jeśli jest tylko skończenie wiele punktów stacjonarnych dla Φ, możemy porównać wartości f w tych punktach i wybrać tę największą (odp. najmniejszą).

Przykład. Niech S będzie sferą jednostkową w $\mathbb{R}^d$ i rozważmy jako f formę kwadratową wyznaczoną przez macierz symetryczną $A = (a_{mj})_{m,j\leqslant d}$. Szukamy jej kresów na sferze S, w tym celu badamy ekstrema warunkowe

$$f(x) = \sum_{m=1}^{d}\sum_{j=1}^{d} a_{mj}x_m x_j, \quad \texttt{warunek to} \quad G(x) := \sum_j x_j^2 - 1 = 0.$$

Ponieważ w punkcie ekstremum warunkowego będzie

$$\forall_{p\in\{1,\dots,d\}}\ 0 = \Phi_p'(x) = \sum_{j=1, j\neq p}^{d} (a_{jp}x_j + a_{pj}x_j) + 2a_{pp}x_p - 2\lambda x_p,$$

oznacza to, że λ jest taką liczbą, że dla pewnego wektora $x \in S$ (a więc niezerowego) jest $(A - \lambda I)x = 0$. To nic innego, jak fakt, że λ jest wartością własną macierzy A, zaś x -związanym z nią wektorem własnym. Czyli $Ax = \lambda x$. Natomiast (w zapisie macierzowym) $f(x) = x^T Ax = \lambda x^T x = \lambda$, gdy $\|x\| = 1$. Stąd maksimum formy f na sferze jednostkowej, to jej największa wartość własna. Te wartości możemy znaleźć rozwiązując równanie charakterystyczne: $\det(A - \lambda I) = 0$. Następnie układ równań liniowych pozwolli na wyznaczenie podprzestrzeni własnej. Wektory własne są bowiem wyznaczone z dokładnością do stałego współczynnika. A ten współczynnik tak dobieramy, by umieścić wektor na sferze jednostkowej. Mamy przynajmniej dwa takie wektory na sferze S -bo $f(-x) = (-1)^2 f(x) = f(x)$.

Na koniec tego wykładu przyjrzyjmy się warunkom opisanym nie przez równania (bądź ich układy), lecz przez słabe nierówności. Na przykład mamy jakiś zbiór zwarty K opisany przez układ nierówności: $G_1(x) \leqslant 0, \dots, G_k(x) \leqslant 0$ oraz określoną na nim funkcję klasy C^1, powiedzmy funkcję f. Wiemy, że już z samej jej ciągłości wynika, że w pewnych punktach P_1, P_2 funkcja f osiąga swoje minimum i maksimum na zbiorze K. Dla każdej z wartości $j = 1, 2$ mamy dwie możliwości: albo P_j leży na brzegu zbioru K, albo w jego wnętrzu. W tym drugim przypadku mamy zwykłe ekstremum lokalne. Na brzegu (np. dla $k = 1$ oznacza to , że $G_1(P_j) = 0$) mamy do czynienia z ekstremum warunkowym. Może się zdarzyć w wyższych wymiarach, że w tym punkcie częś funkcji G_m przyjmuje wartości silnie mniejsze od zera, a tylko część -równe zero. Wtedy ostre nierówności wyznaczają zbiory otwarte (ich skończone przecięcie też jest otwarte), a te pozostałe (np. $G_{m+1}, \dots, G_k$) przyjmują w badanym punkcie wartość zero. Tylko te funkcje bierzemy jako układ wyznaczający warunek w sensie ekstremum warunkowego. Pozostałe -mogą jedynie zawężać otwarty zbiór, na którym zdefiniowane jest f.

Przykład: Wyznaczmy maksimum i minimum z funkcji $f(x, y) = x^2 - y^2 + 2R^2$ w kole $K = \{(x, y) : x^2 + y^2 \leqslant R^2\}$. Punkty stacjonarne dla f, to tylko początek układu, czyli punkt $P_0 = (0, 0)$, co sprawdzamy przyrównując gradient f do zera. Mamy $f(0, 0) = 2R^2$. Następnie badamy kresy f na okręgu $\partial K = \{(x, y) : x^2 + y^2 = R^2\}$. Można to sprawdzić rozwikłując równanie okręgu, czyli wstawiając $y(x) = \pm\sqrt{R^2 - x^2}$. Wtedy mamy $f(x, y(x)) = 2x^2 + R^2$ dla $|x| \leqslant R$. Pochodna z tej funkcji w przedziale otwartym ma jedyne miejsce zetowe dla $x = 0$, wtedy $f(0, R) = R^2 = f(0, -R)$. Nie możemy też zapomnieć o końcach przedziału, czyli dla $x = \pm R$. Wtedy $y(x) = 0$ oraz $f(\pm R, 0) = 3R^2$. Obydwa ekstrema osiągnięte są więc na brzegu.