Урок 5 – Градиент и частни производни

1 Цели на урока

В този урок ще:

- изградим интуиция за промяна на величина, която зависи от няколко фактора;
- започнем от производната на една променлива като мост към многомерния случай;
- дефинираме частни производни, насочена производна и градиент;
- тълкуваме геометрично градиента и линейната апроксимация;
- разгледаме верижното правило за функции на няколко променливи;
- разгледаме втори производни, Хесиева матрица и тест за екстремум;
- видим връзката с градиентния спуск (gradient descent) в AI;
- решим 25 задачи с различна трудност и подробни решения.

2 Въведение

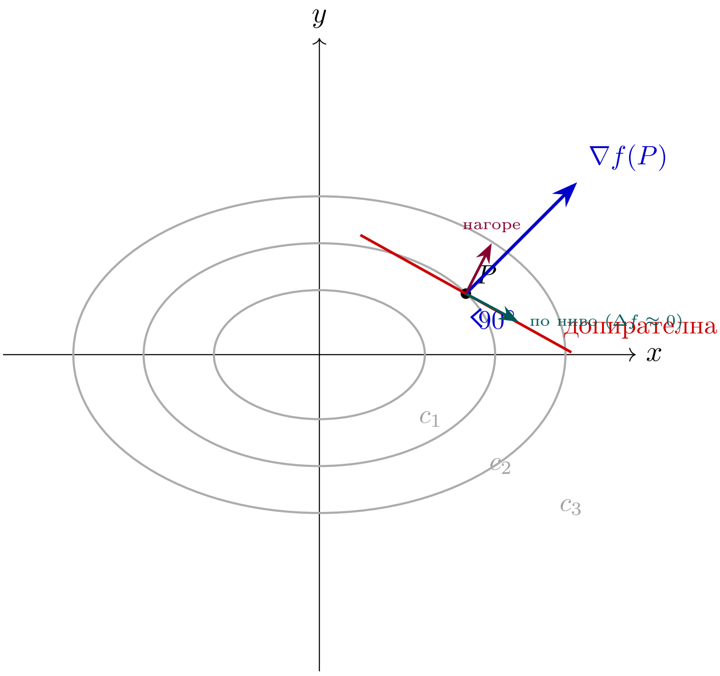
В реални проблеми една величина често зависи от няколко независими параметъра: позиция в равнината, температура и налягане, цена и реклама и т.н. Вместо да разглеждаме само една посока на промяна, трябва да умеем да измерваме ефекта от отделните фактори и да определяме посоката на най-бързо нарастване. Именно тук се появяват **частните производни** и **градиентът**.

Започваме от производната на една променлива, защото тя е най-простата и ясна интерпретация на “локален наклон”, а многомерните понятия я обобщават по естествен начин.

3 Интуитивни примери (без формули)

- **Карта на височини.** Представяме си хълм. Ако се придвижим на север или изток, височината се променя с различна скорост. Интересува ни посоката, в която височината расте *най-бързо.*
- **Температура в стая.** Температурата зависи от мястото. Малка крачка наляво може да даде различна промяна от същата крачка напред.
- **Цена и продажби.** Печалбата зависи от цена и реклама. Ако променим само цената, ефектът е различен от този при промяна само на рекламата.

- **Невронна мрежа.** Функцията на загуба зависи от хиляди тегла. Градиентът показва как да променим всяко тегло, за да намалим грешката.



Фигура 1: Ниви линии и градиент: ∇f сочи в посоката на най-бързо нарастване и е *перпендикулярен* на допирателната към нивото в точка P.

4 От едномерни към многомерни производни

4.1 Припомняне: производна на функция от една променлива

Нека $y = f(x)$. Производната в точка x_0 се дефинира чрез граница:

$$f'(x_0) = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h},$$

ако границата съществува. Геометрично: наклонът на допирателната към графиката в точка $(x_0, f(x_0))$.

Мини-задача. Намерете $f'(2)$ за $f(x) = 3x^2 - 1$.

Отговор: $f'(x) = 6x$, следователно $f'(2) = \boxed{12}$.

Мини-задача. Намерете $f'(x)$ за $f(x) = x^3 - 5x + 2$ и изчислете $f'(1)$.

Отговор: $f'(x) = 3x^2 - 5$, следователно $f'(1) = 3 - 5 = \boxed{-2}$.

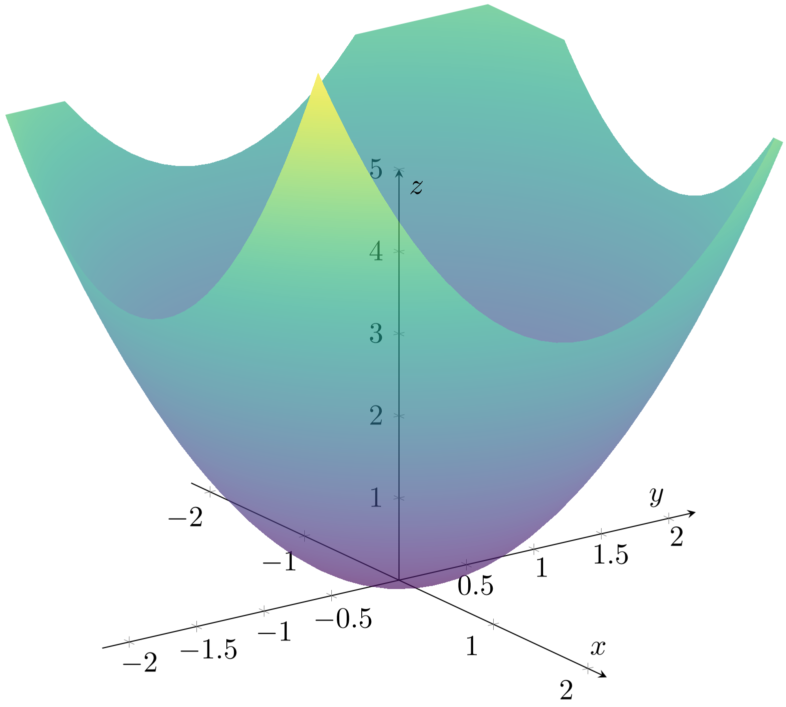
4.2 Функции от две променливи

Функция от две променливи $f : \mathbb{R}^2 \to \mathbb{R}$ съпоставя на двойка (x, y) число $f(x, y)$.

Графиката е **повърхност** в $\mathbb{R}^3$: множеството от точки $(x, y, f(x, y))$.

Линия на ниво: Кривата $f(x, y) = c$ за фиксирано c. Тя е “хоризонтален разрез” на повърхността на височина c (като контурните линии на топографска карта).

Връзка с предишната секция: В 1D имаме една посока (наляво/надясно), а в 2D имаме безкрайно много посоки. Частните производни ще измерят промяната по координатните оси, а после градиентът ще "сглоби" цялата картина.



Фигура 2: Повърхността $z = x^2 + y^2$ (параболоид). Нивата $z = c$ са концентрични окръжности.

4.3 Частни производни

За да видим как се променя f само при изменение на x (с фиксирано y), дефинираме **частната производна по** x:

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

Аналогично за y:

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

Практически: За да намерим $\frac{\partial f}{\partial x}$, диференцираме по x и третираме y като константа (и обратно).

Означения: f_x, $\frac{\partial f}{\partial x}$, $\partial_x f$, $D_x f$ — всичко означава едно и също.

Мини-задача. За $f(x, y) = x^2 + 3y$ намерете f_x и f_y в точка $(1, 2)$.

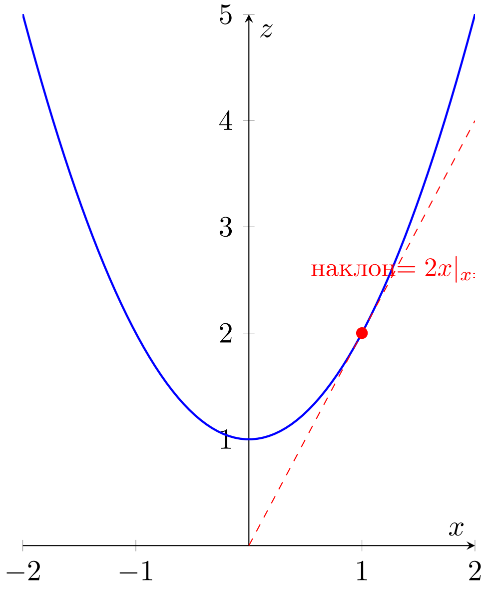
Отговор: $f_x = 2x$, $f_y = 3$, следователно $f_x(1, 2) = \boxed{2}$, $f_y(1, 2) = \boxed{3}$.

Мини-задача. За $f(x, y) = x^3y^2 - 4xy + 7$ намерете f_x и f_y.

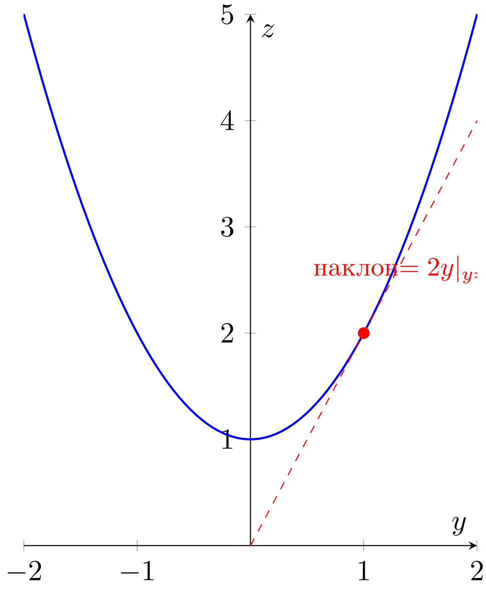
Отговор: $f_x = 3x^2y^2 - 4y$, $f_y = 2x^3y - 4x$.

Мини-задача. За $f(x, y) = \sin(xy)$ намерете $f_x(1, \pi)$.

Отговор: $f_x = y\cos(xy)$, $f_x(1, \pi) = \pi\cos(\pi) = \boxed{-\pi}$.



Фигура 3: *
Сечение при $y = 1$: $z = x^2 + 1$.



Фигура 4: *
Сечение при $x = 1$: $z = 1 + y^2$.

Фигура 5: Сечения на повърхността $z = x^2 + y^2$: фиксирано $y = 1$ (ляво) и фиксирано $x = 1$ (дясно). Наклоните на допирателните са частните производни.

Подробен пример (с връзка към 1D): Нека

$$f(x, y) = x^2 y + y^3.$$

Ако фиксираме $y = 2$, получаваме едномерна функция $g(x) = f(x, 2) = 2x^2 + 8$, и тогава $g'(x) = 4x$. Това е точно $f_x(x, 2)$.

Ако фиксираме $x = 1$, получаваме $h(y) = f(1, y) = y + y^3$, и тогава $h'(y) = 1 + 3y^2$. Това е точно $f_y(1, y)$.

Тоест: частните производни не са нова магия — те са обикновени производни, но по “срезове” на повърхността.

4.4 Таблица: правила за частно диференциране

Правилата са **същите** като за обикновено диференциране — просто третираме другата променлива като константа.

Правило	**Формула (по x, y е конст.)**
Константен множител	$\dfrac{\partial}{\partial x}[c\,f] = c\,f_x$
Сума/разлика	$\dfrac{\partial}{\partial x}[f \pm g] = f_x \pm g_x$
Произведение	$\dfrac{\partial}{\partial x}[fg] = f_x g + f g_x$
Частно	$\dfrac{\partial}{\partial x}\left[\dfrac{f}{g}\right] = \dfrac{f_x g - f g_x}{g^2}$
Верижно правило	$\dfrac{\partial}{\partial x}[h(f(x,y))] = h'(f) \cdot f_x$
Степен	$\dfrac{\partial}{\partial x}[x^n] = nx^{n-1}$
e-степен	$\dfrac{\partial}{\partial x}[e^{f(x,y)}] = e^f \cdot f_x$
Логаритъм	$\dfrac{\partial}{\partial x}[\ln f(x,y)] = \dfrac{f_x}{f}$

4.5 Градиент

Градиентът събира частните производни в един **вектор**:

$$\nabla f(x,y) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right).$$

За n променливи:

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

Ключови свойства:

- ∇f сочи в посоката на **най-бързо нарастване** на f.
- $-\nabla f$ сочи в посоката на **най-бързо намаляване** на f.
- $\|\nabla f\|$ е **скоростта** на нарастване в тази посока.
- ∇f е **перпендикулярен** на нивата $f = \mathrm{const}$.
- Ако $\nabla f = \mathbf{0}$, точката е **критична** (кандидат за екстремум или седлова точка).

Мини-задача. Намерете $\nabla f(2,3)$ за $f(x,y) = xy$.

Отговор: $\nabla f = (y, x)$, следователно $\nabla f(2,3) = \boxed{(3,2)}$.

Мини-задача. Намерете $\nabla f(1,0)$ за $f(x,y) = e^x \cos y$.

Отговор: $f_x = e^x \cos y$, $f_y = -e^x \sin y$. $\nabla f(1,0) = (e \cdot 1, -e \cdot 0) = \boxed{(e,0)}$.

Мини-задача. Намерете $\nabla f(1,-1)$ за $f(x,y) = x^2 - 2xy + y^3$.

Отговор: $f_x = 2x - 2y = 2 + 2 = 4$, $f_y = -2x + 3y^2 = -2 + 3 = 1$. $\nabla f(1, -1) = \boxed{(4, 1)}$.

Подробен пример (геометрична интерпретация): Нека $f(x, y) = x^2 + 4y^2$ и точка $P = (1, 1)$.

$$\nabla f(P) = (2, 8).$$

Ако се движим по единичната посока $\boldsymbol{u}_1 = (1, 0)$, тогава $D_{\boldsymbol{u}_1} f(P) = \nabla f(P) \cdot \boldsymbol{u}_1 = 2$. Ако се движим по единичната посока $\boldsymbol{u}_2 = (0, 1)$, тогава $D_{\boldsymbol{u}_2} f(P) = 8$.

Значи около P функцията нараства 4 пъти по-бързо в "y-посока", отколкото в "x-посока". Това обяснява защо нивата са елипси, а не окръжности.

4.6 Насочена производна

Нека $\boldsymbol{u} = (u_1, u_2)$ е **единичен** вектор ($\|\boldsymbol{u}\| = 1$). Насочената производна показва скоростта на промяна на f в посока $\boldsymbol{u}$:

$$D_{\boldsymbol{u}} f(x_0, y_0) = \lim_{t \to 0} \frac{f(x_0 + tu_1,\, y_0 + tu_2) - f(x_0, y_0)}{t}.$$

Твърдение. *Ако f е диференцируема в (x_0, y_0), то*

$$D_{\boldsymbol{u}} f(x_0, y_0) = \nabla f(x_0, y_0) \cdot \boldsymbol{u} = f_x(x_0, y_0)\, u_1 + f_y(x_0, y_0)\, u_2.$$

Доказателство. Разглеждаме $g(t) = f(x_0 + tu_1,\, y_0 + tu_2)$. По верижното правило:

$$g'(t) = f_x(x_0 + tu_1,\, y_0 + tu_2)\, u_1 + f_y(x_0 + tu_1,\, y_0 + tu_2)\, u_2.$$

При $t = 0$: $g'(0) = f_x(x_0, y_0)\, u_1 + f_y(x_0, y_0)\, u_2 = \nabla f(x_0, y_0) \cdot \boldsymbol{u}$. □

Важно! Ако посоката $\boldsymbol{v}$ *не* е единична, трябва първо да я нормираме: $\boldsymbol{u} = \dfrac{\boldsymbol{v}}{\|\boldsymbol{v}\|}$.

Твърдение. *Максималната стойност на $D_{\boldsymbol{u}} f$ е $\|\nabla f\|$, постигана в посока $\boldsymbol{u} = \dfrac{\nabla f}{\|\nabla f\|}$. Минималната е $-\|\nabla f\|$ (в обратна посока).*

Доказателство. $D_{\boldsymbol{u}} f = \nabla f \cdot \boldsymbol{u} = \|\nabla f\| \cos\theta$, където θ е ъгълът между ∇f и $\boldsymbol{u}$. Максимумът е при $\theta = 0$, т.е. $\boldsymbol{u} \parallel \nabla f$. □

Мини-задача. Намерете $D_{\boldsymbol{u}} f(1, 0)$ за $f(x, y) = x^2 + y^2$ и $\boldsymbol{u} = \frac{1}{\sqrt{2}}(1, 1)$.

Отговор: $\nabla f = (2x, 2y)$, в $(1, 0)$: $(2, 0)$. $D_{\boldsymbol{u}} f = \frac{2 \cdot 1 + 0 \cdot 1}{\sqrt{2}} = \boxed{\sqrt{2}}$.

Мини-задача. За $f(x, y) = 3x - 4y$ в каква посока нараства най-бързо? С каква скорост?

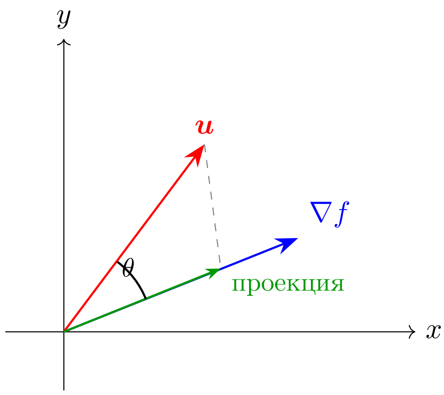
Отговор: $\nabla f = (3, -4)$. Посока: $\frac{1}{5}(3, -4)$. Скорост: $\|\nabla f\| = \sqrt{9 + 16} = \boxed{5}$.

Подробен пример (неединична посока): Нека $f(x, y) = x^2 + y^2$, точка $P = (1, 2)$, посока $\boldsymbol{v} = (3, 4)$. Първо нормализираме:

$$\boldsymbol{u} = \frac{\boldsymbol{v}}{\|\boldsymbol{v}\|} = \left(\frac{3}{5}, \frac{4}{5}\right).$$

$$\nabla f(P) = (2, 4), \qquad D_{\boldsymbol{u}} f(P) = \nabla f(P) \cdot \boldsymbol{u} = 2 \cdot \frac{3}{5} + 4 \cdot \frac{4}{5} = \frac{22}{5}.$$

Ако забравим нормализацията и сметнем със $(3, 4)$, ще получим число, което не е "скорост на единица дължина".



Фигура 6: Насочената производна $D_{\boldsymbol{u}}f = \|\nabla f\| \cos\theta$ е проекцията на ∇f върху $\boldsymbol{u}$.

5 Линейна апроксимация и допирателна равнина

Ясна връзка: Частните производни ни дават локалните наклони по осите, а линейната апроксимация ги комбинира, за да предскаже как се променя f при малко преместване в *произволна* посока.

За функция от една променлива: $f(x_0 + \Delta x) \approx f(x_0) + f'(x_0)\Delta x$.

За функция от **две** променливи:

$$f(x_0 + \Delta x,\, y_0 + \Delta y) \approx f(x_0, y_0) + f_x(x_0, y_0)\,\Delta x + f_y(x_0, y_0)\,\Delta y.$$

Допирателната **равнина** към повърхността $z = f(x, y)$ в точка $(x_0, y_0, f(x_0, y_0))$:

$$z = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0).$$

Компактно с градиента:

$$z \approx f(\boldsymbol{a}) + \nabla f(\boldsymbol{a}) \cdot (\boldsymbol{x} - \boldsymbol{a}).$$

Подробен пример (защо формулата е полезна): Нека

$$f(x, y) = \sqrt{4 + x^2 + 2y^2}, \qquad A = (0, 0).$$

Тогава $f(A) = 2$, $f_x(A) = 0$, $f_y(A) = 0$, следователно

$$L(x, y) = 2.$$

Това означава, че близо до $(0, 0)$ повърхността е почти хоризонтална. Например за $(x, y) = (0.1, -0.1)$:

$$f(0.1, -0.1) = \sqrt{4 + 0.01 + 0.02} = \sqrt{4.03} \approx 2.0075,$$

а линеаризацията дава 2 — грешката е малка и това е целта на локалното линейно приближение.

Мини-задача. Намерете допирателната равнина към $z = x^2 + 2y^2$ в $(1, 1)$.

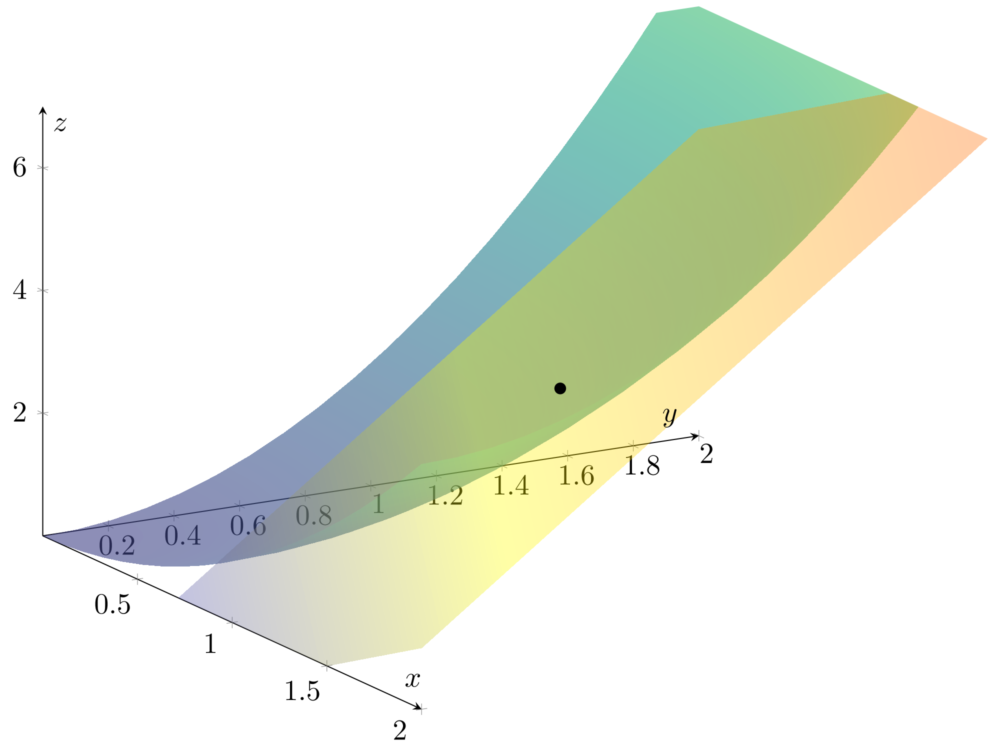
Отговор: $f_x = 2x$, $f_y = 4y$, $f(1, 1) = 3$. Равнина: $z = 3 + 2(x - 1) + 4(y - 1) = \boxed{2x + 4y - 3}$.

Мини-задача. Апроксимирайте $f(1.02,\, 0.97)$ за $f(x, y) = xy + y^2$ около $(1, 1)$.

Отговор: $f(1, 1) = 1 + 1 = 2$, $f_x = y = 1$, $f_y = x + 2y = 3$. $\Delta x = 0.02$, $\Delta y = -0.03$. $f \approx 2 + 1 \cdot 0.02 + 3 \cdot (-0.03) = 2 + 0.02 - 0.09 = \boxed{1.93}$.

Мини-задача. Апроксимирайте $\sqrt{(3.02)^2 + (3.97)^2}$ чрез линеаризация на $f(x,y) = \sqrt{x^2 + y^2}$ около $(3,4)$.

Отговор: $f(3,4) = 5$, $f_x = \frac{x}{\sqrt{x^2+y^2}} = \frac{3}{5}$, $f_y = \frac{y}{\sqrt{x^2+y^2}} = \frac{4}{5}$. $\Delta x = 0.02$, $\Delta y = -0.03$. $f \approx 5 + \frac{3}{5}(0.02) + \frac{4}{5}(-0.03) = 5 + 0.012 - 0.024 = \boxed{4.988}$.



Фигура 7: Повърхност $z = x^2 + 2y^2$ (синьо) и допирателна равнина $z = 2x + 4y - 3$ (червено) в точка $(1,1,3)$.

6 Ниви линии и геометричен смисъл на градиента

Нека $f(x,y) = c$ е линия на ниво.

Тази секция затваря връзката между предишните идеи: частни производни $\to$ градиент $\to$ геометрия на нивата. Ако си представим карта на височините, допирателната към контура е движение "по равно ниво", а градиентът е посоката "нагоре по най-стръмното".

Твърдение. *Ако f е диференцируема и (x_0, y_0) лежи на нивото $f = c$, то $\nabla f(x_0, y_0)$ е **перпендикулярен** на допирателната към линията на ниво.*

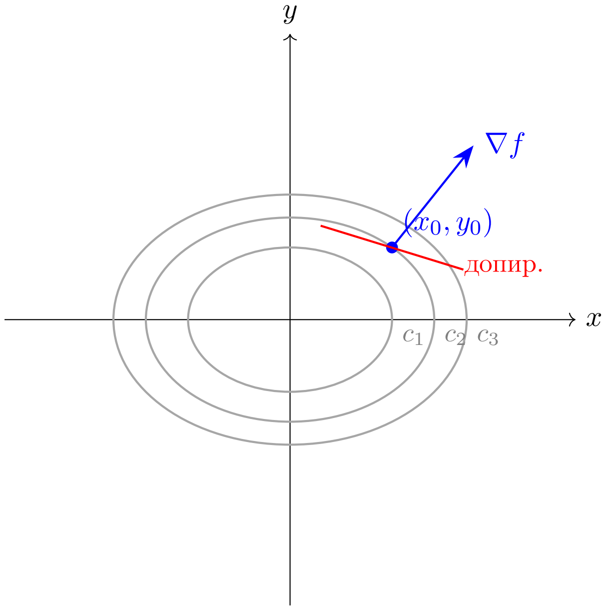
Доказателство. Нека $\boldsymbol{r}(t) = (x(t), y(t))$ е параметризация на нивото и $\boldsymbol{r}(0) = (x_0, y_0)$. Тогава $f(x(t), y(t)) \equiv c$. Диференцираме по t:

$$f_x\, x'(0) + f_y\, y'(0) = 0 \quad \Longrightarrow \quad \nabla f \cdot \boldsymbol{r}'(0) = 0.$$

Следователно ∇f е перпендикулярен на $\boldsymbol{r}'(0)$ (допирателния вектор на нивото). □

Приложение: Уравнение на *допирателната права* към нивото $f(x,y) = c$ в точка (x_0, y_0):

$$f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) = 0.$$



Фигура 8: Елиптични нива линии на $f(x,y) = x^2 + 2y^2$. Градиентът е перпендикулярен на всяка нивова крива.

Мини-задача. За $f(x,y) = x^2 + y^2$ намерете допирателната права към нивото $f = 5$ в точка $(1,2)$.

Отговор: $\nabla f = (2x, 2y) = (2,4)$. Нормала: $(2,4)$, или $(1,2)$. Уравнение: $1 \cdot (x-1) + 2 \cdot (y-2) = 0$, т.е. $\boxed{x + 2y - 5 = 0}$.

7 Верижно правило за функции на няколко променливи

В едномерния случай: $(f(g(x)))' = f'(g(x))\,g'(x)$.

Многомерната версия казва същото: промяната "по пътя" се получава като сума от *(чувствителност по всяка променлива)* $\times$ *(скорост на тази променлива)*.

Нека сега $f(x,y)$ и $x = x(t)$, $y = y(t)$. Тогава $F(t) = f(x(t), y(t))$ и:

$$\frac{dF}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} = \nabla f \cdot \boldsymbol{r}'(t).$$

Ако $x = x(s,t)$, $y = y(s,t)$:

$$\frac{\partial f}{\partial s} = f_x \frac{\partial x}{\partial s} + f_y \frac{\partial y}{\partial s}, \qquad \frac{\partial f}{\partial t} = f_x \frac{\partial x}{\partial t} + f_y \frac{\partial y}{\partial t}.$$

Мини-задача. Нека $f(x,y) = x^2 y$ и $x = \cos t$, $y = \sin t$. Намерете $\frac{dF}{dt}$ при $t = \frac{\pi}{4}$.

Отговор: $f_x = 2xy$, $f_y = x^2$, $x' = -\sin t$, $y' = \cos t$.
$\frac{dF}{dt} = 2xy(-\sin t) + x^2 \cos t$. При $t = \frac{\pi}{4}$: $x = y = \frac{1}{\sqrt{2}}$.
$= 2 \cdot \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} \cdot \left(-\frac{1}{\sqrt{2}}\right) + \frac{1}{2} \cdot \frac{1}{\sqrt{2}} = -\frac{1}{\sqrt{2}} + \frac{1}{2\sqrt{2}} = \boxed{-\dfrac{1}{2\sqrt{2}}}$.

Подробен пример (практическо тълкуване): Нека $T(x,y) = x^2 + xy$ е температура, а траекторията е $x(t) = 1 + t$, $y(t) = 2 - t$. Тогава

$$\frac{dT}{dt} = T_x\,x'(t) + T_y\,y'(t), \qquad T_x = 2x + y, \quad T_y = x.$$

При $t = 0$: $x = 1$, $y = 2$, $x' = 1$, $y' = -1$.

$$\left.\frac{dT}{dt}\right|_{t=0} = (2 \cdot 1 + 2) \cdot 1 + 1 \cdot (-1) = 4 - 1 = \boxed{3}.$$

Тоест по тази траектория температурата нараства със скорост 3 в момента $t = 0$.

8 Втори частни производни и тест за екстремум

8.1 Втори производни

Ако f_x и f_y са сами диференцируеми, можем да вземем техните частни производни:

$$f_{xx} = \frac{\partial^2 f}{\partial x^2}, \quad f_{xy} = \frac{\partial^2 f}{\partial y\, \partial x}, \quad f_{yx} = \frac{\partial^2 f}{\partial x\, \partial y}, \quad f_{yy} = \frac{\partial^2 f}{\partial y^2}.$$

Твърдение (Шварц / Клеро)**.** *Ако f_{xy} и f_{yx} са непрекъснати, то $f_{xy} = f_{yx}$ (смесените производни са равни, реда не е важен).*

Мини-задача. За $f(x, y) = x^3 + x^2y^3 - 2y^2$ проверете, че $f_{xy} = f_{yx}$.

Отговор: $f_x = 3x^2 + 2xy^3$, $f_{xy} = 6xy^2$. $f_y = 3x^2y^2 - 4y$, $f_{yx} = 6xy^2$. Наистина $f_{xy} = f_{yx} = \boxed{6xy^2}$.

8.2 Хесиева матрица

Втората производна се събира в **Хесиева матрица**:

$$H_f(x, y) = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix}.$$

8.3 Тест за екстремум (за две променливи)

Нека $\nabla f(x_0, y_0) = \mathbf{0}$ (критична точка). Дефинираме **дискриминанта**:

$$D = f_{xx}(x_0, y_0)\, f_{yy}(x_0, y_0) - \big(f_{xy}(x_0, y_0)\big)^2 = \det H_f(x_0, y_0).$$

Условие	**Извод**
$D > 0$ и $f_{xx} > 0$	локален **минимум**
$D > 0$ и $f_{xx} < 0$	локален **максимум**
$D < 0$	**седлова точка**
$D = 0$	тестът е **неубедителен**

Мини-задача. Намерете и класифицирайте критичните точки на $f(x, y) = x^2 + y^2 - 4x - 2y$.

Отговор: $f_x = 2x - 4 = 0 \Rightarrow x = 2$, $f_y = 2y - 2 = 0 \Rightarrow y = 1$. Критична точка: $(2, 1)$. $f_{xx} = 2$, $f_{yy} = 2$, $f_{xy} = 0$, $D = 4 - 0 = 4 > 0$ и $f_{xx} = 2 > 0$. $\Rightarrow$ **локален минимум** в $\boxed{(2, 1)}$.

Подробен пример (с пълна класификация): Нека $f(x, y) = x^2 + xy + y^2 - 6x$.

1) Критична точка:

$$f_x = 2x + y - 6 = 0, \qquad f_y = x + 2y = 0.$$

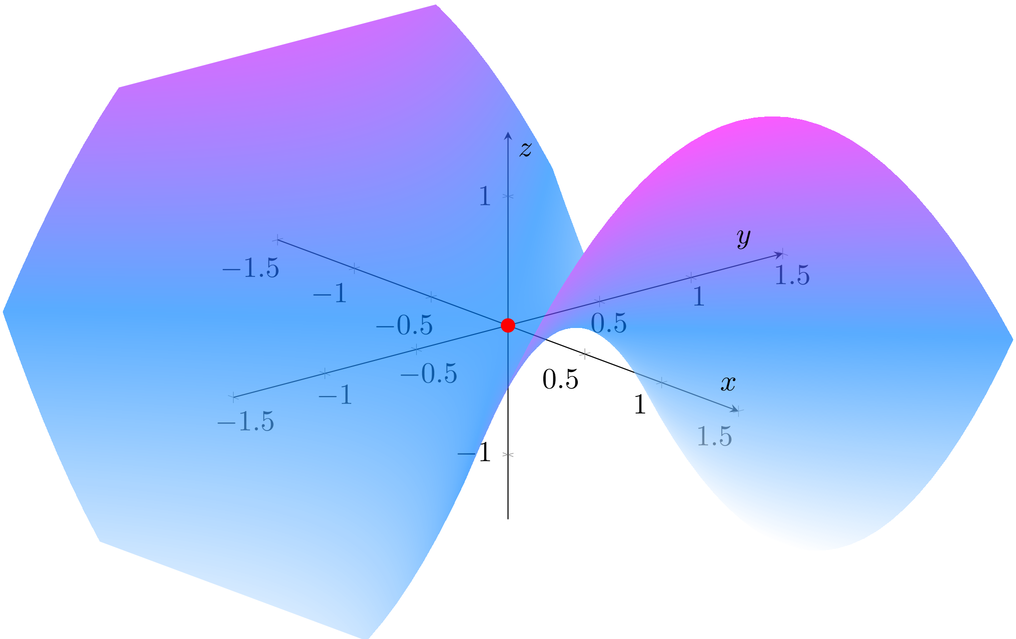
От второто $x = -2y$. Замяна в първото: $2(-2y) + y - 6 = 0 \Rightarrow -3y = 6 \Rightarrow y = -2$, $x = 4$.

2) Хесиева матрица:

$$f_{xx} = 2, \qquad f_{yy} = 2, \qquad f_{xy} = 1.$$

$$D = 2 \cdot 2 - 1^2 = 3 > 0, \qquad f_{xx} = 2 > 0.$$

Следователно в $(4, -2)$ имаме **локален минимум**.



Фигура 9: Седлова точка на $z = x^2 - y^2$ в началото. Повърхността се "качва" по x и "слиза" по y.

9 Кога частните производни не са достатъчни

Съществуват функции, за които частните производни съществуват в дадена точка, но функцията **не е диференцируема** там.

Пример:

$$f(x, y) = \begin{cases} \dfrac{x^2 y}{x^2 + y^2}, & (x, y) \neq (0, 0), \\ 0, & (x, y) = (0, 0). \end{cases}$$

Частни производни в $(0, 0)$**:**

$$f_x(0, 0) = \lim_{h \to 0} \frac{f(h, 0) - f(0, 0)}{h} = \lim_{h \to 0} \frac{0}{h} = 0,$$

$$f_y(0, 0) = \lim_{h \to 0} \frac{f(0, h) - f(0, 0)}{h} = \lim_{h \to 0} \frac{0}{h} = 0.$$

Но по правата $y = x$: $f(t, t) = \frac{t^3}{2t^2} = \frac{t}{2} \to 0$, а по правата $y = x^2$: $f(t, t^2) = \frac{t^4}{t^2 + t^4} = \frac{t^2}{1 + t^2} \to 0$.

Линейната апроксимация $L = 0 + 0 \cdot \Delta x + 0 \cdot \Delta y = 0$ не приближава f добре по всички пътища (грешката не е $o(\|\Delta\|)$), следователно f **не е диференцируема** в $(0, 0)$.

Извод: Диференцируемостта е *по-силно* условие от съществуването на частни производни. Достатъчно условие: f_x *и* f_y *са непрекъснати* $\Rightarrow f$ е диференцируема.

10 Приложение: градиентен спуск (Gradient Descent)

В машинното обучение искаме да **минимизираме** функция на загуба $L(\boldsymbol{w})$, зависеща от тегла $\boldsymbol{w} = (w_1, \ldots, w_n)$.

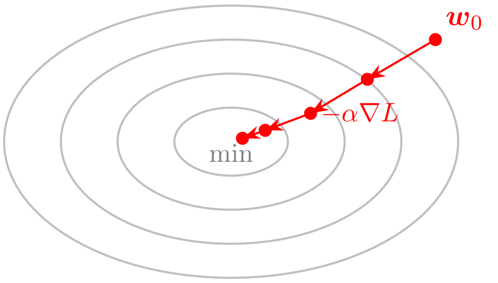
Алгоритъм:

$$\boldsymbol{w}_{\text{ново}} = \boldsymbol{w}_{\text{старо}} - \alpha \nabla L(\boldsymbol{w}_{\text{старо}}),$$

където $\alpha > 0$ е *скорост на обучение* (learning rate).

Защо работи? Градиентът сочи "нагоре" (най-бързо нарастване). Вървим в обратна посока, за да намалим L.

Ясна връзка с насочена производна: ако вземем единична посока $\boldsymbol{u}$, скоростта на промяна е $D_{\boldsymbol{u}}L = \nabla L \cdot \boldsymbol{u}$. Минималната възможна стойност се получава при $\boldsymbol{u} = -\dfrac{\nabla L}{\|\nabla L\|}$, затова gradient descent естествено избира тази посока.



Фигура 10: Градиентен спуск: стъпваме в обратна посока на градиента. Нивата са елипси на L.

Мини-задача. Нека $L(w_1, w_2) = w_1^2 + 4w_2^2$. Старти от $\boldsymbol{w} = (2, 1)$ с $\alpha = 0.1$. Намерете $\boldsymbol{w}_{\text{ново}}$.
Отговор: $\nabla L = (2w_1, 8w_2) = (4, 8)$. $\boldsymbol{w}_{\text{ново}} = (2, 1) - 0.1 \cdot (4, 8) = (2 - 0.4,\ 1 - 0.8) = \boxed{(1.6,\ 0.2)}$.

Подробен пример (две поредни стъпки): Същата функция $L(w_1, w_2) = w_1^2 + 4w_2^2$, същото начало $\boldsymbol{w}_0 = (2, 1)$ и $\alpha = 0.1$.

Първа стъпка вече имаме:

$$\boldsymbol{w}_1 = (1.6, 0.2).$$

Втора стъпка:

$$\nabla L(\boldsymbol{w}_1) = (2 \cdot 1.6,\ 8 \cdot 0.2) = (3.2, 1.6).$$

$$\boldsymbol{w}_2 = \boldsymbol{w}_1 - 0.1\nabla L(\boldsymbol{w}_1) = (1.6, 0.2) - 0.1(3.2, 1.6) = (1.28, 0.04).$$

Виждаме как координатите се приближават към $(0, 0)$, където L има минимум.

11 Теория-куиз (без смятане)

1. Какво означава $\dfrac{\partial f}{\partial x}$ с думи?

2. Ако $\nabla f(a,b) = \mathbf{0}$, какво можем да кажем за точката (a,b)?
3. Защо градиентът е перпендикулярен на нивата?
4. Каква е разликата между f_x и $D_{\boldsymbol{u}} f$?
5. Ако $D > 0$ и $f_{xx} > 0$, какъв тип екстремум имаме?
6. Защо в градиентния спуск вървим в посока $-\nabla L$?
7. Може ли функция с нулеви частни производни в точка да не е диференцируема там?
8. Каква е геометричната интерпретация на допирателната равнина?

Кратки отговори:

1. Скоростта на промяна на f, когато се движим само по оста x (с фиксирано y).
2. Критична точка — кандидат за минимум, максимум или седлова точка.
3. Защото по нивото $f = c$ функцията е постоянна, т.е. промяната по допирателния вектор е нула: $\nabla f \cdot \boldsymbol{r}' = 0$.
4. f_x е частна производна (промяна по оста x), а $D_{\boldsymbol{u}} f = \nabla f \cdot \boldsymbol{u}$ е промяна в произволна посока.
5. Локален минимум.
6. Защото ∇L сочи "нагоре" (нарастване), а ние искаме намаляване.
7. Да — примерът $f(x,y) = \frac{x^2 y}{x^2+y^2}$.
8. Най-доброто линейно приближение на повърхността в дадена точка.

12 Задачи за самостоятелна работа

Задача 1. Намерете f_x и f_y за $f(x,y) = x^2 y - 3x + y^3$.

Задача 2. Намерете $\nabla f(1,-2)$ за функцията от задача 1.

Задача 3. Намерете f_x и f_y за $f(x,y) = e^{x^2+y}$.

Задача 4. Намерете f_x и f_y за $f(x,y) = \ln(x^2 + y^2)$.

Задача 5. Намерете $\nabla f(1,1)$ за $f(x,y) = \dfrac{x}{x+y}$.

Задача 6. Намерете насочената производна на $f(x,y) = x^2 + xy + y^2$ в точка $(0,1)$ по посока на $(1,1)$.

Задача 7. Намерете $D_{\boldsymbol{u}} f(1,1)$ за $f(x,y) = x^2 y$ и посока $\boldsymbol{u} = \frac{1}{\sqrt{5}}(2,-1)$.

Задача 8. За $f(x,y) = xe^y$ намерете посоката на най-бързо нарастване и скоростта в точка $(0,0)$.

Задача 9. За $f(x,y) = x^2 + y^2$ намерете допирателната права към нивото $f = 5$ в точка $(1,2)$.

Задача 10. Намерете уравнението на допирателната равнина към $z = x^2 + 2y^2$ в точка $(1,1)$.

Задача 11. Намерете допирателната равнина към $z = \sin x \cos y$ в $(0,0)$.

Задача 12. Използвайте линейна апроксимация, за да оцените $f(1.02, 0.97)$ за $f(x,y) = xy + y^2$, около $(1,1)$.

Задача 13. Използвайте линейна апроксимация, за да оцените $f(1.02, 1.98)$ за $f(x,y) = \sqrt{x^2 + y^2}$, около $(1,2)$.

Задача 14. Намерете критичните точки на $f(x,y) = x^2 + y^2 - 4x - 2y$ и ги класифицирайте.

Задача 15. Намерете и класифицирайте критичните точки на $f(x,y) = x^2 - y^2 + 2x - 4y$.

Задача 16. Намерете и класифицирайте критичните точки на $f(x,y) = x^3 + y^3 - 3xy$.

Задача 17. Намерете f_{xx}, f_{xy}, f_{yy} за $f(x,y) = x^3y + e^{xy}$.

Задача 18. Нека $f(x,y) = x^2y$ и $x = \cos t$, $y = \sin t$. Намерете $\frac{dF}{dt}$.

Задача 19. Нека $f(x,y) = \ln(x^2 + y^2)$ и $x = e^t$, $y = e^{-t}$. Намерете $\frac{dF}{dt}$.

Задача 20. За $f(x,y) = \ln(x^2 + y^2)$ намерете $\nabla f(1,1)$ и насочената производна по посока на $(1,0)$.

Задача 21. Проверете диференцируемостта на

$$f(x,y) = \begin{cases} \dfrac{x^2y}{x^2+y^2}, & (x,y) \neq (0,0), \\ 0, & (x,y) = (0,0), \end{cases}$$

в точка $(0,0)$.

Задача 22. Нека $L(w_1, w_2) = (w_1 - 3)^2 + (w_2 + 1)^2$. Намерете минимума и направете една стъпка градиентен спуск от $(0,0)$ с $\alpha = 0.5$.

Задача 23. За $f(x,y) = 4x^2 + y^2$ намерете посоката, в която $D_{\boldsymbol{u}}f(1,2) = 0$ (перпендикулярна на градиента).

Задача 24. Намерете и класифицирайте критичните точки на $f(x,y) = xy - x^3 - y^3$.

Задача 25. За $f(x,y,z) = x^2 + 2y^2 + 3z^2 - 2xy$ намерете ∇f и $\nabla f(1,1,1)$.

13 Подробни решения (стъпка по стъпка)

Решение на задача 1

$f(x,y) = x^2y - 3x + y^3$.

Диференцираме по x (y е конст.):

$$f_x = 2xy - 3.$$

Диференцираме по y (x е конст.):

$$f_y = x^2 + 3y^2.$$

$$\boxed{f_x = 2xy - 3, \quad f_y = x^2 + 3y^2}.$$

Решение на задача 2

От задача 1: $f_x = 2xy - 3,\ f_y = x^2 + 3y^2$.

В точка $(1, -2)$:

$$f_x(1,-2) = 2 \cdot 1 \cdot (-2) - 3 = -7, \quad f_y(1,-2) = 1^2 + 3 \cdot 4 = 13.$$

$$\boxed{\nabla f(1,-2) = (-7,\, 13)}.$$

Решение на задача 3

$f(x,y) = e^{x^2+y}$.

По верижно правило:

$$f_x = e^{x^2+y} \cdot 2x = 2x\, e^{x^2+y}, \quad f_y = e^{x^2+y} \cdot 1 = e^{x^2+y}.$$

$$\boxed{f_x = 2x\, e^{x^2+y}, \quad f_y = e^{x^2+y}}.$$

Решение на задача 4

$f(x,y) = \ln(x^2 + y^2)$.

$$f_x = \frac{2x}{x^2+y^2}, \qquad f_y = \frac{2y}{x^2+y^2}.$$

$$\boxed{f_x = \frac{2x}{x^2+y^2}, \quad f_y = \frac{2y}{x^2+y^2}}.$$

Решение на задача 5

$f(x,y) = \dfrac{x}{x+y}$.

По правило за частно (по x):

$$f_x = \frac{1 \cdot (x+y) - x \cdot 1}{(x+y)^2} = \frac{y}{(x+y)^2}.$$

По y:

$$f_y = \frac{0 \cdot (x+y) - x \cdot 1}{(x+y)^2} = -\frac{x}{(x+y)^2}.$$

В $(1,1)$:

$$f_x(1,1) = \frac{1}{4}, \quad f_y(1,1) = -\frac{1}{4}.$$

$$\boxed{\nabla f(1,1) = \left(\frac{1}{4},\, -\frac{1}{4}\right)}.$$

Решение на задача 6

$f(x, y) = x^2 + xy + y^2$, точка $(0, 1)$, посока $(1, 1)$.

Първо нормираме: $\boldsymbol{u} = \frac{1}{\sqrt{2}}(1, 1)$.

$\nabla f = (2x + y,\ x + 2y)$. В $(0, 1)$: $\nabla f = (1, 2)$.

$$D_{\boldsymbol{u}} f = \nabla f \cdot \boldsymbol{u} = \frac{1}{\sqrt{2}}(1 \cdot 1 + 2 \cdot 1) = \boxed{\frac{3}{\sqrt{2}}}.$$

Решение на задача 7

$f(x, y) = x^2 y$, точка $(1, 1)$, $\boldsymbol{u} = \frac{1}{\sqrt{5}}(2, -1)$.

$\nabla f = (2xy, x^2)$. В $(1, 1)$: $\nabla f = (2, 1)$.

$$D_{\boldsymbol{u}} f = \frac{1}{\sqrt{5}}(2 \cdot 2 + 1 \cdot (-1)) = \boxed{\frac{3}{\sqrt{5}}}.$$

Решение на задача 8

$f(x, y) = xe^y$.

$f_x = e^y$, $f_y = xe^y$. В $(0, 0)$: $\nabla f = (e^0,\ 0 \cdot e^0) = (1, 0)$.

Посока на най-бързо нарастване: $(1, 0)$ (по положителната ос x). Скорост: $\|\nabla f\| = 1$.

$$\boxed{\text{Посока } (1, 0), \quad \text{скорост } 1}.$$

Решение на задача 9

$f(x, y) = x^2 + y^2$, ниво $f = 5$, точка $(1, 2)$.

$\nabla f = (2x, 2y) = (2, 4)$. Нормален вектор: $(2, 4)$, или $(1, 2)$.

Уравнение: $1 \cdot (x - 1) + 2 \cdot (y - 2) = 0$, т.е.:

$$\boxed{x + 2y - 5 = 0}.$$

Решение на задача 10

$z = x^2 + 2y^2$, точка $(1, 1)$.

$f(1, 1) = 1 + 2 = 3$, $f_x = 2x = 2$, $f_y = 4y = 4$.

$$z = 3 + 2(x - 1) + 4(y - 1) = \boxed{2x + 4y - 3}.$$

Решение на задача 11

$z = \sin x \cos y$, точка $(0, 0)$.

$f(0, 0) = \sin 0 \cdot \cos 0 = 0$. $f_x = \cos x \cos y$, в $(0, 0)$: 1. $f_y = -\sin x \sin y$, в $(0, 0)$: 0.

$$z = 0 + 1 \cdot (x - 0) + 0 \cdot (y - 0) = \boxed{x}.$$

Решение на задача 12

$f(x,y) = xy + y^2$, около $(1,1)$.

$f(1,1) = 1 + 1 = 2$, $f_x = y = 1$, $f_y = x + 2y = 3$.

$\Delta x = 0.02$, $\Delta y = -0.03$:

$$f(1.02, 0.97) \approx 2 + 1 \cdot 0.02 + 3 \cdot (-0.03) = 2 + 0.02 - 0.09 = \boxed{1.93}.$$

Решение на задача 13

$f(x,y) = \sqrt{x^2 + y^2}$, около $(1,2)$.

$f(1,2) = \sqrt{5}$. $f_x = \frac{x}{\sqrt{x^2+y^2}} = \frac{1}{\sqrt{5}}$, $f_y = \frac{y}{\sqrt{x^2+y^2}} = \frac{2}{\sqrt{5}}$.

$\Delta x = 0.02$, $\Delta y = -0.02$:

$$f(1.02, 1.98) \approx \sqrt{5} + \frac{1}{\sqrt{5}} \cdot 0.02 + \frac{2}{\sqrt{5}} \cdot (-0.02) = \sqrt{5} - \frac{0.02}{\sqrt{5}}.$$

$$\boxed{f(1.02, 1.98) \approx \sqrt{5} - \frac{0.02}{\sqrt{5}} \approx 2.2316}.$$

Решение на задача 14

$f(x,y) = x^2 + y^2 - 4x - 2y$.

$f_x = 2x - 4 = 0 \Rightarrow x = 2$, $f_y = 2y - 2 = 0 \Rightarrow y = 1$.

Критична точка: $(2,1)$.

$f_{xx} = 2$, $f_{yy} = 2$, $f_{xy} = 0$. $D = 2 \cdot 2 - 0^2 = 4 > 0$ и $f_{xx} = 2 > 0$.

$$\boxed{\text{Локален минимум в } (2,1), \quad f(2,1) = -5}.$$

Решение на задача 15

$f(x,y) = x^2 - y^2 + 2x - 4y$.

$f_x = 2x + 2 = 0 \Rightarrow x = -1$, $f_y = -2y - 4 = 0 \Rightarrow y = -2$.

Критична точка: $(-1,-2)$.

$f_{xx} = 2$, $f_{yy} = -2$, $f_{xy} = 0$. $D = 2 \cdot (-2) - 0 = -4 < 0$.

$$\boxed{\text{Седлова точка в } (-1,-2)}.$$

Решение на задача 16

$f(x,y) = x^3 + y^3 - 3xy$.

$f_x = 3x^2 - 3y = 0 \Rightarrow y = x^2$. $f_y = 3y^2 - 3x = 0 \Rightarrow x = y^2$.

Заместваме $y = x^2$ в $x = y^2$: $x = (x^2)^2 = x^4$. $x^4 - x = 0 \Rightarrow x(x^3 - 1) = 0 \Rightarrow x = 0$ или $x = 1$.

Критични точки: $(0,0)$ с $y = 0$ и $(1,1)$ с $y = 1$.

$f_{xx} = 6x$, $f_{yy} = 6y$, $f_{xy} = -3$.

В $(0,0)$**:** $D = 0\cdot 0 - 9 = -9 < 0$. $\Rightarrow$ Седлова точка.

В $(1,1)$**:** $D = 6\cdot 6 - 9 = 27 > 0$, $f_{xx} = 6 > 0$. $\Rightarrow$ Локален минимум, $f(1,1) = 1 + 1 - 3 = -1$.

$$\boxed{(0,0) \text{ — седлова}, \quad (1,1) \text{ — лок. мин. с } f = -1}.$$

Решение на задача 17

$f(x,y) = x^3y + e^{xy}$.

$f_x = 3x^2y + ye^{xy}$.

$f_{xx} = 6xy + y^2e^{xy}$.

$f_{xy} = 3x^2 + e^{xy} + xye^{xy}$.

$f_y = x^3 + xe^{xy}$.

$f_{yy} = x^2e^{xy}$.

$$\boxed{f_{xx} = 6xy + y^2e^{xy}, \quad f_{xy} = 3x^2 + e^{xy} + xye^{xy}, \quad f_{yy} = x^2e^{xy}}.$$

Решение на задача 18

$f(x,y) = x^2y$, $x = \cos t$, $y = \sin t$.

$f_x = 2xy$, $f_y = x^2$, $x' = -\sin t$, $y' = \cos t$.

$$\frac{dF}{dt} = 2xy(-\sin t) + x^2\cos t = 2\cos t\sin t(-\sin t) + \cos^2 t\cos t.$$

$$\boxed{\frac{dF}{dt} = -2\cos t\sin^2 t + \cos^3 t = \cos t(\cos^2 t - 2\sin^2 t)}.$$

Решение на задача 19

$f(x,y) = \ln(x^2 + y^2)$, $x = e^t$, $y = e^{-t}$.

$f_x = \frac{2x}{x^2+y^2}$, $f_y = \frac{2y}{x^2+y^2}$. $x' = e^t$, $y' = -e^{-t}$.

$$\frac{dF}{dt} = \frac{2x}{x^2+y^2}\cdot e^t + \frac{2y}{x^2+y^2}\cdot(-e^{-t}) = \frac{2e^{2t} - 2e^{-2t}}{e^{2t} + e^{-2t}}.$$

$$\boxed{\frac{dF}{dt} = \frac{2(e^{2t} - e^{-2t})}{e^{2t} + e^{-2t}} = 2\tanh(2t)}.$$

Решение на задача 20

$f(x,y) = \ln(x^2 + y^2)$.

$\nabla f = \left(\frac{2x}{x^2+y^2}, \frac{2y}{x^2+y^2}\right)$.

В $(1,1)$: $\nabla f = \left(\frac{2}{2}, \frac{2}{2}\right) = (1,1)$.

Насочена производна по $(1,0)$ (вече единичен): $D_{\boldsymbol{u}}f = (1,1)\cdot(1,0) = \boxed{1}$.

Решение на задача 21

$f(x,y) = \dfrac{x^2 y}{x^2+y^2}$ за $(x,y) \neq (0,0)$, $f(0,0) = 0$.

Показахме, че $f_x(0,0) = 0$ и $f_y(0,0) = 0$ (в секция 8).

Ако f е диференцируема в $(0,0)$, то $f(h,k) \approx 0 + 0 \cdot h + 0 \cdot k = 0$.

Грешката: $\frac{|f(h,k)-0|}{\sqrt{h^2+k^2}} = \frac{h^2|k|}{(h^2+k^2)^{3/2}}$.

По правата $k = h$: $\frac{h^3}{(2h^2)^{3/2}} = \frac{1}{2\sqrt{2}} \neq 0$.

Следователно грешката не клони към 0 по всички пътища.

$$\boxed{f \text{ не е диференцируема в } (0,0)}.$$

Решение на задача 22

$L(w_1, w_2) = (w_1 - 3)^2 + (w_2 + 1)^2$.

$\nabla L = (2(w_1 - 3),\ 2(w_2 + 1))$. $\nabla L = \mathbf{0} \Rightarrow w_1 = 3,\ w_2 = -1$. Минимум в $(3, -1)$, $L = 0$.

Градиентен спуск от $(0,0)$ с $\alpha = 0.5$: $\nabla L(0,0) = (2(-3), 2(1)) = (-6, 2)$. $\boldsymbol{w}_1 = (0,0) - 0.5(-6,2) = (3,-1)$.

$$\boxed{\text{Минимум } (3,-1); \quad \text{след 1 стъпка: } \boldsymbol{w}_1 = (3,-1) \text{ (точно минимумът!)}}.$$

Решение на задача 23

$f(x,y) = 4x^2 + y^2$, точка $(1,2)$.

$\nabla f = (8x, 2y) = (8, 4)$.

Насочената производна е нула перпендикулярно на градиента. $(8,4)$ е перпендикулярен на (a,b) ако $8a + 4b = 0$, т.е. $b = -2a$.

Единичен вектор: $\boldsymbol{u} = \frac{1}{\sqrt{5}}(1,-2)$.

$$\boxed{\boldsymbol{u} = \frac{1}{\sqrt{5}}(1,-2) \text{ (или } \frac{1}{\sqrt{5}}(-1,2))}.$$

Решение на задача 24

$f(x,y) = xy - x^3 - y^3$.

$f_x = y - 3x^2 = 0$, $f_y = x - 3y^2 = 0$.

От първото: $y = 3x^2$. Заместваме: $x - 3(3x^2)^2 = 0 \Rightarrow x - 27x^4 = 0 \Rightarrow x(1 - 27x^3) = 0$. $x = 0 \Rightarrow y = 0$, или $x^3 = \frac{1}{27} \Rightarrow x = \frac{1}{3} \Rightarrow y = 3 \cdot \frac{1}{9} = \frac{1}{3}$.

$f_{xx} = -6x$, $f_{yy} = -6y$, $f_{xy} = 1$.

В $(0,0)$: $D = 0 \cdot 0 - 1 = -1 < 0$. Седлова.

В $\left(\frac{1}{3}, \frac{1}{3}\right)$: $f_{xx} = -2$, $f_{yy} = -2$, $D = 4 - 1 = 3 > 0$, $f_{xx} < 0$. Локален максимум. $f\left(\frac{1}{3}, \frac{1}{3}\right) = \frac{1}{9} - \frac{1}{27} - \frac{1}{27} = \frac{3-1-1}{27} = \frac{1}{27}$.

$$\boxed{(0,0) \text{ — седлова}, \quad \left(\frac{1}{3}, \frac{1}{3}\right) \text{ — лок. макс. с } f = \frac{1}{27}}.$$

Решение на задача 25

$f(x, y, z) = x^2 + 2y^2 + 3z^2 - 2xy$.

$$f_x = 2x - 2y, \quad f_y = 4y - 2x, \quad f_z = 6z.$$

$$\nabla f = (2x - 2y,\ 4y - 2x,\ 6z).$$

В $(1, 1, 1)$:

$$\nabla f(1, 1, 1) = (2 - 2,\ 4 - 2,\ 6) = \boxed{(0, 2, 6)}.$$

14 Обобщение: какво научихме

- **Частни производни:** f_x и f_y — производни по отделните променливи с фиксиране на останалите.
- **Градиент:** ∇f — вектор от частните производни. Сочи в посоката на най-бързо нарастване.
- **Насочена производна:** $D_{\boldsymbol{u}} f = \nabla f \cdot \boldsymbol{u}$ — промяна в произволна посока.
- **Линейна апроксимация:** $f \approx f(\boldsymbol{a}) + \nabla f(\boldsymbol{a}) \cdot (\boldsymbol{x} - \boldsymbol{a})$.
- **Допирателна равнина:** обобщение на допирателната права.
- **Ниви линии:** ∇f е перпендикулярен на тях.
- **Хесиева матрица и тест:** класификация на критични точки чрез $D = f_{xx} f_{yy} - f_{xy}^2$.
- **Градиентен спуск:** $\boldsymbol{w}_{\text{ново}} = \boldsymbol{w} - \alpha \nabla L$ — основа на оптимизацията в AI.