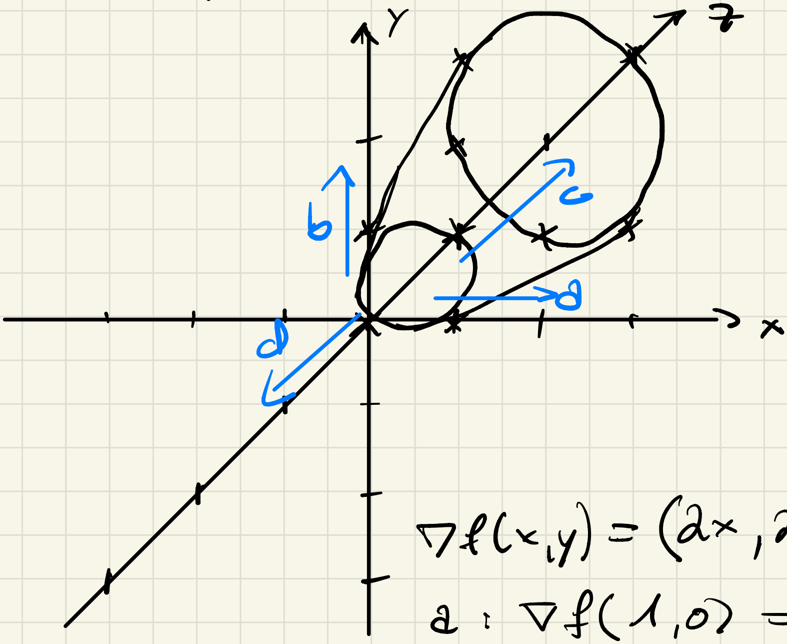
Exercise 4

Malik Mardan

1. $f(x) = x^2 + y^2$



$\nabla f(x,y) = (2x, 2y)$

a : $\nabla f(1,0) = (2,0)$

b : $\nabla f(0,1) = (0,2)$

c : $\nabla f(1,1) = (2,2)$

d : $\nabla f(-1,-1) = (-2,-2)$

each gradien points radially outwards the origin. the gradient shows the direction of the steepness. if you wove along the gradient the height z increases fast and if you move opposite to it, it decreases fast.

2.

a) $\nabla f(y_1, y_2) = (2y_1, 2y_2)$

Jacobian $J_G(x)$ (rows $= \nabla g_i^T$):

$$J_G(x_1, x_2, x_3) = \begin{pmatrix} \frac{dg_1}{dx_1} & \frac{dg_1}{dx_2} & \frac{dg_1}{dx_3} \\ \frac{dg_2}{dx_2} & \frac{dg_2}{dx_2} & \frac{dg_2}{dx_3} \end{pmatrix} =$$

$$\begin{pmatrix} 2x_1 & 3x_2^2 & 1 \\ x_2x_3 & x_1x_3 & x_1x_2 \end{pmatrix}$$

b)

Chain rule for gradients:

$$\nabla_x (f \circ g)(x) = J_G(x)^T \nabla f(g(x))$$

$$g(x) = (x_1^2 + x_2^3 + x_3,\ x_1x_2x_3)$$

$$\nabla f(g(x)) = (2(x^2 + x_2^3 + x_3),\ 2 \cdot (x_1x_2x_3))$$

$$\begin{pmatrix} 2x_1 & x_2x_3 \\ 3x_2^2 & x_1x_3 \\ 1 & x_1x_2 \end{pmatrix} \cdot \begin{pmatrix} 2(x_1^2 + x_2^3 + x_3) \\ 2x_1x_2x_3 \end{pmatrix} =$$

$$\begin{pmatrix} 2x_1 \cdot 2(x_1^2 + x_2^3 + x_3) + x_2x_3 \cdot 2x_1x_2x_3 \\ 3x_2^2 \cdot 2(x_1^2 + x_2^3 + x_3) + x_1x_3 \cdot 2x_1x_2x_3 \\ 2(x_1^2 + x_2^3 + x_3) + x_1x_2 \cdot 2x_1x_2x_3 \end{pmatrix} = \begin{pmatrix} 4x_1(x_1^2 + x_2^3 + x_3) + 2x_1x_2^2x_3^2 \\ 6x_2(x_1^2 + x_2^3 + x_3) + 2x_1^2x_2x_3^2 \\ 2(x_1^2 + x_2^3 + x_3) + 2x_1^2x_2^2x_3 \end{pmatrix}$$

3.

a)

$$\nabla \bar{L} = (0{,}5(\theta_1 - 4), 10(\theta_2 - 3)) = 0$$

$$\Rightarrow \hat{\theta} = (4,3) \quad , \bar{L}(4,3) = 0$$

b)

$$\frac{d^2 L}{d^2\theta_1} = 0{,}5 \quad \frac{d^2 L}{d^2\theta_2} = 10 \quad \frac{d^2 L}{d^2\theta_1\theta_2} = 0$$

$$H = \nabla^2 \bar{L} = \begin{pmatrix} 0{,}5 & 0 \\ 0 & 10 \end{pmatrix}$$ constant, so same at θ

$\sigma_1 = 0{,}5$ $\qquad$ $\sigma_2 = 10$

$v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ $\qquad$ $v_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

c)

Eigenpair: $(\sigma_1, v_1) = (0{,}5, (1,0)^T)$

$(\sigma_2, v_2) = (10, (0,1)^T)$

$$\theta_1(\lambda) = \frac{0{,}5}{0{,}5+\lambda} \cdot 4 = \frac{2}{0{,}5+\lambda}$$

$$\sigma_2(\lambda) = \frac{10}{10+\lambda} \cdot 3 = \frac{30}{10+\lambda}$$

λ	shrink θ_1	$\theta_1(\lambda)$	shrink θ_2	$\theta_2(\lambda)$
0	1	4	1	3
0,1	$0{,}8\bar{3}$	$3{,}\bar{3}$	0,99009	2,9702
1	$0{,}3\bar{3}$	$1{,}3\bar{3}$	0,9091	$2{,}\overline{72}$
10	0,047	0,1904	0,5	1,5

Shrinkage is stronger with smaller eigenvalue, since θ_1 gets pulled faster towards 0 than θ_2.