V06

2. 4)

$\vec{w_0} = (1, -4, 4)$ $\qquad$ $\vec{w_0} = \arg\min\limits_{\vec{w}} E(\vec{w} \mid D)$

$\vec{w_1} = (1, -4, 6)$

$\vec{w_2} = (1, -1, 7)$ $\qquad$ $E(\vec{w_1} \mid D) = E(\vec{w_2} \mid D) = E(\vec{w_3} \mid D)$

$\vec{w_3} = (1, -7, 1)$ $\qquad$ $\Omega(\vec{w}) = \|\vec{w}\|_2^2 \; ; \; \lambda = 100$

$\vec{w_4} = (1, -7, -3)$ $\qquad$ $\frac{\lambda}{2} \Omega(\vec{w}^*) = 400$

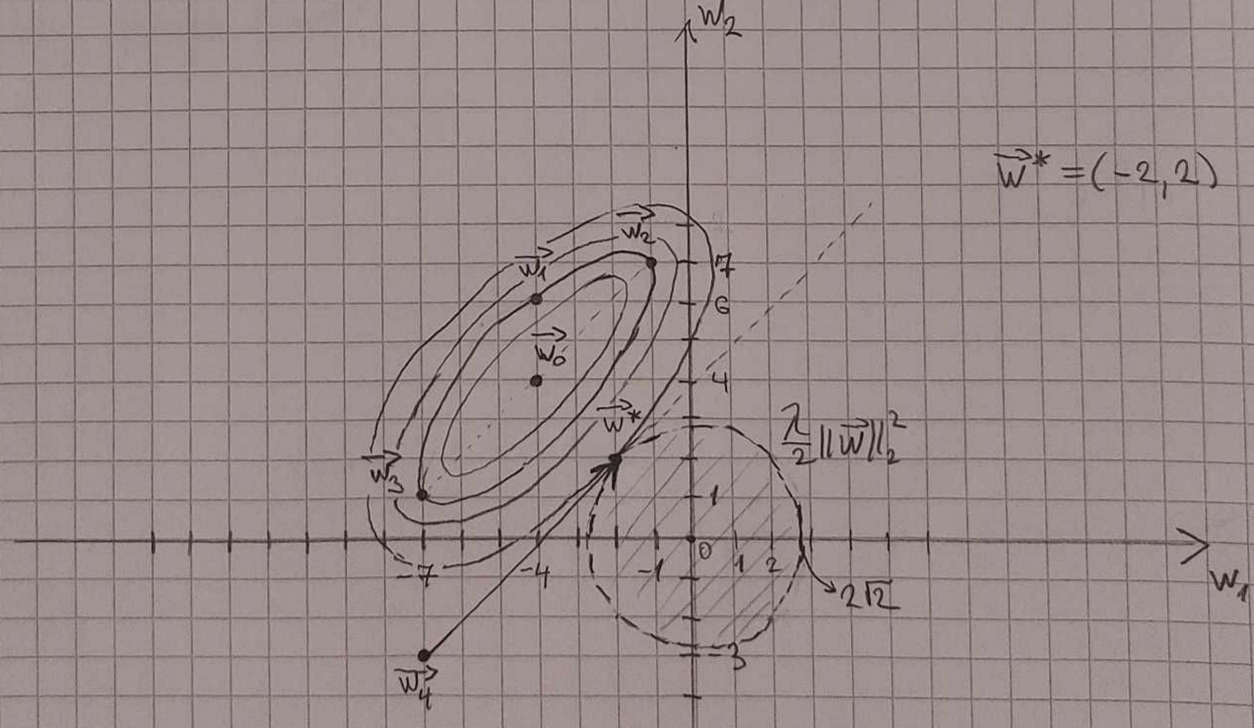
$$E_R(\vec{w} \mid D) = E(\vec{w} \mid D) + \frac{\lambda}{2} \Omega(\vec{w})$$

$$\frac{\lambda}{2} \Omega(\vec{w}^*) = \frac{\lambda}{2} \|\vec{w}^*\|_2^2 = 400$$

$$\|\vec{w}^*\|_2^2 = \frac{800}{\lambda} = \frac{800}{100} = 8 \; / \sqrt{\ }$$

$$\|\vec{w}^*\|_2 = \sqrt{8} = 2\sqrt{2}$$

$\vec{w}^* = (-2, 2)$



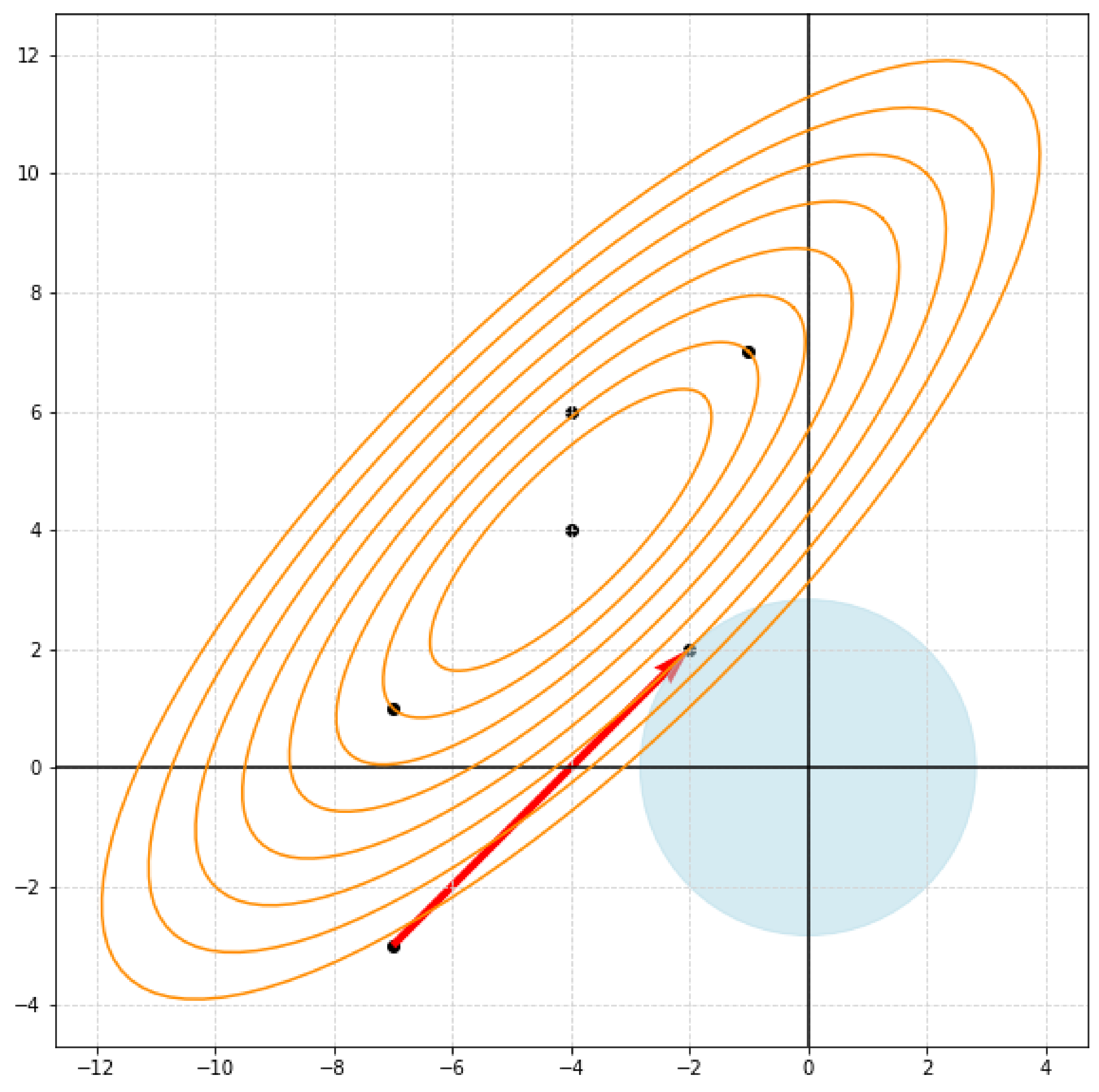
Vektor $(\vec{w}^* - \vec{w_4})$ ortogonalan je na normale izokontura $E(\vec{w} \mid D)$ i $\frac{\lambda}{2} \Omega(\vec{w})$.

Polazeći iz točke koja odgovara vektoru $\vec{w_4}$, vektor negativnog gradijenta <u>ukupne</u> pogreške E_R s obzirom na težine $-\nabla_{\vec{w}} E_R(\vec{w}|D)\,\big|_{\vec{w}=\vec{w_4}}$ leži na istom pravcu kao i vektor $(\vec{w^*}-\vec{w_4})$. Budući da koristimo linijsko pretraživanje, s inicijalnim parametrima $\vec{w_4}$ ćemo u <u>jednom koraku</u> doći do minimizatora $\vec{w^*}$.

Preostali vektori $\vec{w_1}-\vec{w_3}$ nemaju to svojstvo, stoga će algoritam sigurno trebati više od jednog koraka ako su inicijalni parametri $\vec{w_1}-\vec{w_3}$.

Zaključujemo da će algoritam najbrže konvergirati s inicijalnim parametrima $\vec{w_4}$.

U nastavku je dana računalna skica zadatka radi bolje preglednosti.


12
10
8
6
4
2
0
−2
−4
−12
−10
−8
−6
−4
−2
0
2
4