

θ_1 GD $\theta^{(0)}$

$$\textcircled{1} \quad \vec{g}^{(t)} = \nabla_{\theta} L(\theta^{(t)}, \mathcal{T})$$

$$\textcircled{2} \quad \theta^{(t+1)} = \theta^{(t)} - \eta \vec{g}^{(t)}$$

$$\textcircled{3} \quad c? \Rightarrow s \text{ top}$$

$$\mathcal{T} = \{x, y\}$$

SGD

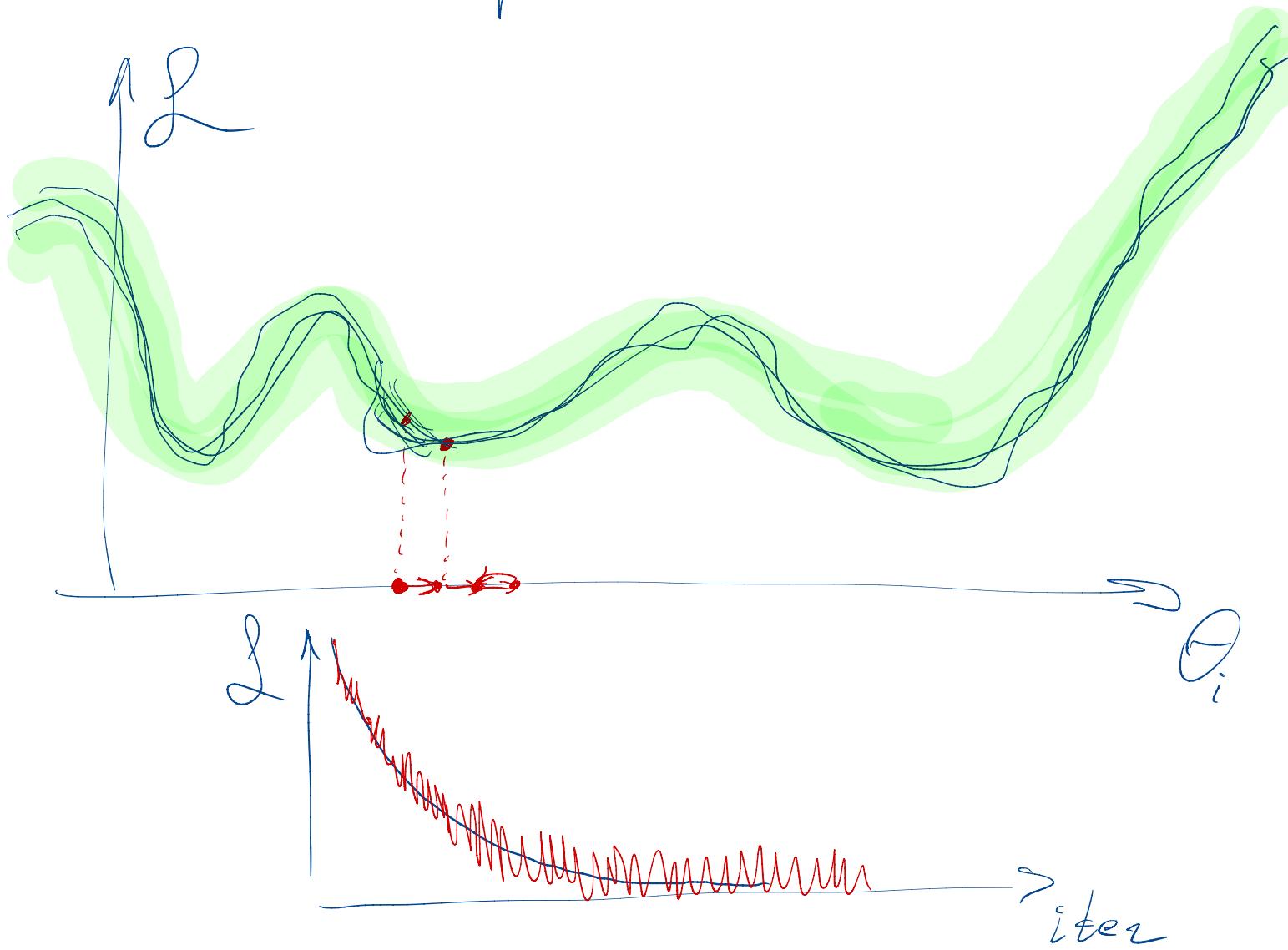
① $\theta^{(0)}, \gamma, |B|, C$

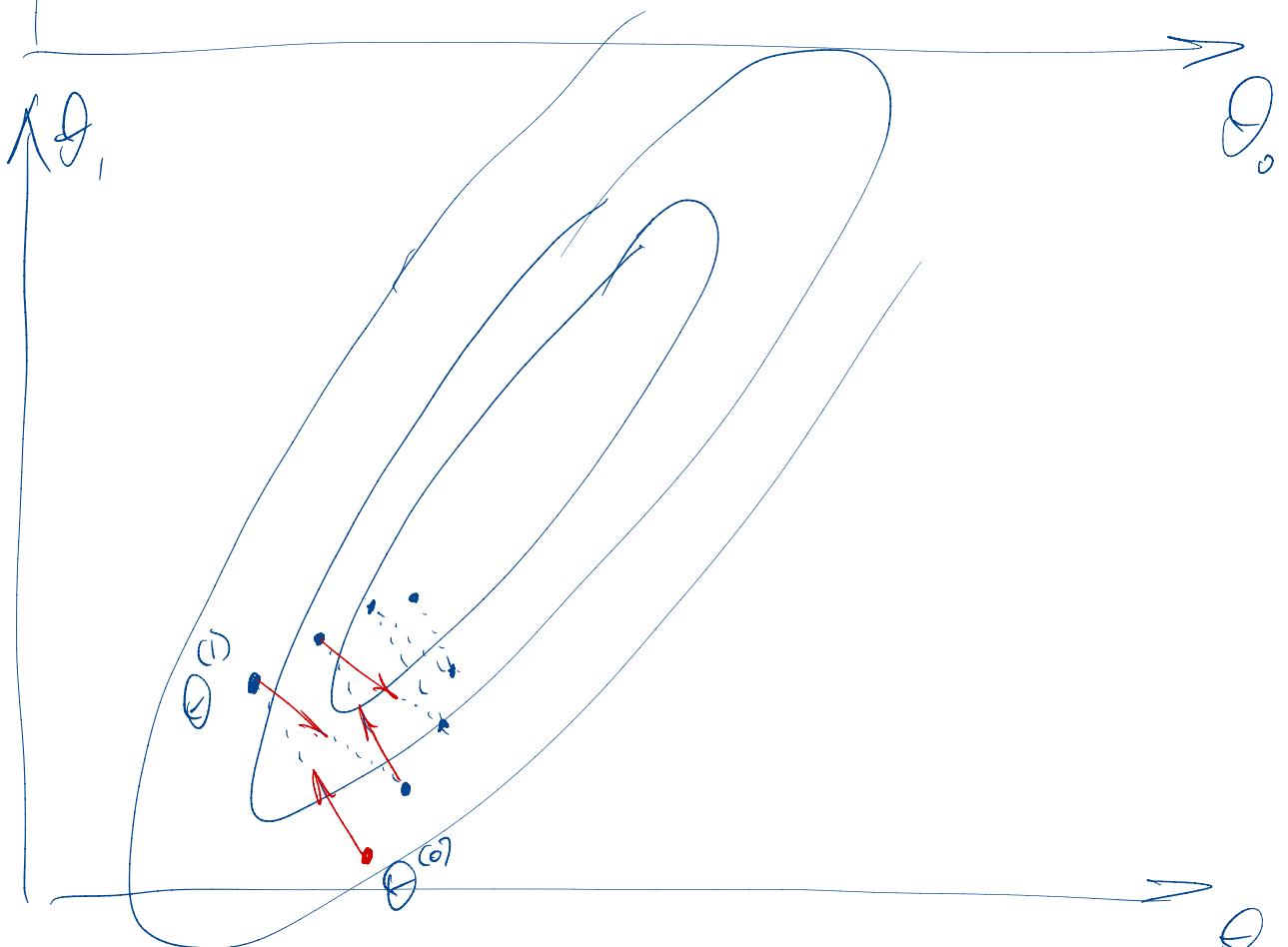
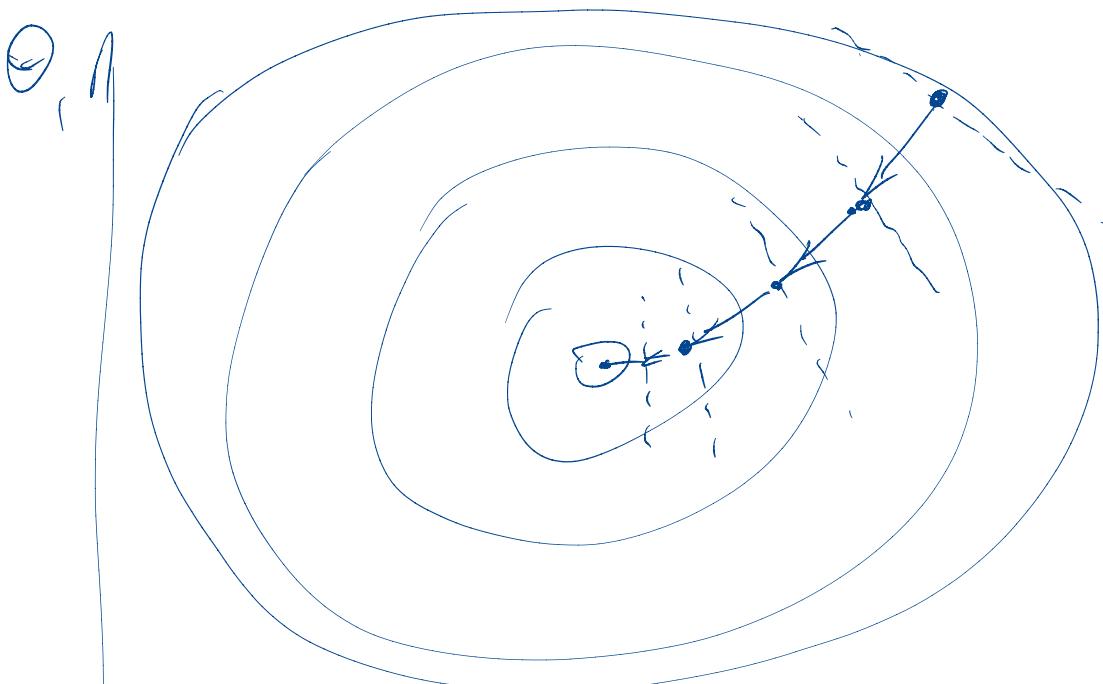
$$\text{② } \vec{g}^{(t)} = \vec{\nabla}_{\theta} \mathcal{L}(B, \theta^{(t)})$$

$$\text{③ } \theta^{(t+1)} = \theta^{(t)} - \gamma \vec{g}^{(t)}$$

$$B = \{x, y\}^{\text{[Obs]}}$$

④ $C \Rightarrow \text{stop}$





$$\vec{g} = \nabla_{\theta} L(\theta^{(t)}, \beta)$$

$$\theta^{(t+1)} = \theta^{(t)} - \eta g^{(t)}$$

Унгілбенсікі метод

$$\vec{g} = \nabla_{\theta} L(\theta^{(t)}, B)$$

$$\vec{m}^{(t)} = \beta \vec{m}^{(t-1)} + (1 - \beta) \vec{g}^{(t)}$$

$$\theta^{(t+1)} = \theta^{(t)} - h \vec{m}^{(t)}$$

