

Abschnitt 2 - Gradientenabstieg

Freitag, 22. März 2024 14:40

2.1 Gradientenabstieg 1d

- $J_s(\theta)$: Mittlerer Verlust auf der Stichprobe
- $J'(\theta)$: Steigung der Verlustfunktion bei konkreten Modellparametern
- $\theta(t) := \theta(t-1) - \lambda \cdot J'(\theta(t-1))$: Neubestimmung der Modellparameter auf Basis der Steigung der Verlustfunktion mit Lernrate

Annahme hier: nur ein Modellparameter (1d)

- Gegeben ist die Funktion $L_b(\theta) = \begin{cases} 2 \cdot \theta^2 + \frac{3}{16} & \theta \leq 0 \\ \frac{\theta}{2} + \frac{3}{16} & \theta > 0 \end{cases}$ Berechnen Sie mit Stift und Papier:

I Bei Startwert $\theta = -1$ und Lernrate $\lambda = 0.1$, was sind die nächsten beiden Schritte des Gradientenabstiegs?

II Bei Startwert $\theta = 4$ und unbekannter Lernrate benötigt der Gradientenabstieg 12 Schritte, bis $L_b(\theta) = 0$ erreicht wird. Bestimme die Lernrate.

$$\theta = 0$$

$$L_b'(\theta) = \begin{cases} 4\theta & \theta \leq 0 \\ \frac{1}{2} & \theta > 0 \end{cases}$$

I $\theta(0) = -1 \quad \lambda = 0.1$

$$\theta(1) = -1 - 0.1 \cdot 4(-1) = -0.6$$

$$\theta(2) = -0.6 - 0.1 \cdot 4(-0.6) = -0.36$$

II $\theta(0) = 4 \Rightarrow L_b'(4) = \frac{1}{2}$

$$\theta(1) = 4 - \frac{1}{2} \lambda$$

$$\theta(2) = 4 - \frac{1}{2} \lambda - \frac{1}{2} \lambda = 4 - 2\left(\frac{1}{2} \lambda\right)$$

\vdots

$$\theta(12) = 4 - 12\left(\frac{1}{2} \lambda\right) = 0$$

$$4 = 12 \cdot \frac{1}{2} \lambda = 6 \lambda$$

$$\lambda = \frac{4}{6} = \frac{2}{3}$$