

8. Adam Optimization Algorithm

Tags

check

Adam optimization algorithm

Q. What is Adam?

Adam optimization algorithm

Adam optimization algorithm

$$V_{dw}=0, S_{dw}=0, V_{db}=0, S_{db}=0$$

On iteration t :

Compute dw, db using current mini-batch

$$V_{dw} = \beta_1 V_{dw} + (1-\beta_1) dw, \quad V_{db} = \beta_1 V_{db} + (1-\beta_1) db \quad \leftarrow \text{"momentum"} \beta_1$$

$$S_{dw} = \beta_2 S_{dw} + (1-\beta_2) dw^2, \quad S_{db} = \beta_2 S_{db} + (1-\beta_2) db^2 \quad \leftarrow \text{"RMSprop"} \beta_2$$

$$V_{dw}^{\text{corrected}} = V_{dw} / (1-\beta_1^t), \quad V_{db}^{\text{corrected}} = V_{db} / (1-\beta_1^t)$$

$$S_{dw}^{\text{corrected}} = S_{dw} / (1-\beta_2^t), \quad S_{db}^{\text{corrected}} = S_{db} / (1-\beta_2^t)$$

$$w := w - \alpha \frac{V_{dw}^{\text{corrected}}}{\sqrt{S_{dw}^{\text{corrected}} + \epsilon}}, \quad b := b - \alpha \frac{V_{db}^{\text{corrected}}}{\sqrt{S_{db}^{\text{corrected}} + \epsilon}}$$

Andrew Ng

Q. What is Adam?