Batch Normalization

Batch Norm at test time

deeplearning.ai

Batch Norm at test time

$$\mu = \frac{1}{m} \sum_i z^{(i)}$$

$$\sigma^2 = \frac{1}{m} \sum_i (z^{(i)} - \mu)^2$$

$$z_{\text{norm}}^{(i)} = \frac{z^{(i)} - \mu}{\sqrt{\sigma^2 + \varepsilon}}$$

$$\tilde{z}^{(i)} = \gamma z_{\text{norm}}^{(i)} + \beta$$

μ, σ^2: estimate using exponentially weighted average (across mini-batches).

$X^{\{1\}}, X^{\{2\}}, X^{\{3\}}, \dots$

$\mu^{\{1\}[l]}$ $\quad \mu^{\{2\}[l]}$ $\quad \mu^{\{3\}[l]} \longrightarrow \mu$

$\theta_1 \qquad \theta_2 \qquad \theta_3 \qquad \qquad \sigma^2$

$\sigma^{2\{1\}[l]} \qquad \sigma^{\{2\}[l]} \qquad \dots$

$z_{\text{norm}} = \frac{z - \mu}{\sqrt{\sigma^2 + \varepsilon}} \qquad \tilde{z} = \gamma z_{\text{norm}} + \beta$

Andrew Ng