Derivation of the Logistic Loss from the Power-of-h Form

1. Setup

We consider data points (x, y) with $x \in \mathbb{R}^n$ and labels $y \in \{-1, +1\}$. Define the sigmoid function:

$$h_\theta(x) = \sigma\big(\theta^\top x\big) = \frac{1}{1 + e^{-\theta^\top x}}.$$

2. Likelihood via Powers of h

For a single example,

$$p\big(y \mid x; \theta\big) = h_\theta(x)^{\frac{y+1}{2}} \big(1 - h_\theta(x)\big)^{\frac{1-y}{2}}.$$

Over a dataset of m examples, the *negative log-likelihood* (loss) is

$$\mathcal{L}(\theta) = -\sum_{i=1}^{m} \log p\big(y^{(i)} \mid x^{(i)}; \theta\big).$$

3. Expressing Logs

Let $z^{(i)} = \theta^\top x^{(i)}$. Then

$$\log h_\theta(x^{(i)}) = -\log\big(1 + e^{-z^{(i)}}\big), \quad \log\big(1 - h_\theta(x^{(i)})\big) = -z^{(i)} - \log\big(1 + e^{-z^{(i)}}\big).$$

Substitute into the loss:

$$\mathcal{L}(\theta) = -\sum_{i=1}^{m} \Big[\tfrac{y^{(i)}+1}{2}\big(-\log(1 + e^{-z^{(i)}})\big) + \tfrac{1-y^{(i)}}{2}\big(-z^{(i)} - \log(1 + e^{-z^{(i)}})\big)\Big]$$

$$= \sum_{i=1}^{m} \big[\log(1 + e^{-z^{(i)}}) + \tfrac{1-y^{(i)}}{2}\, z^{(i)}\big].$$

4. Final Simplification

Using the fact that $y^{(i)} \in \{-1, +1\}$, one shows

$$\log(1 + e^{-z^{(i)}}) + \tfrac{1-y^{(i)}}{2}\, z^{(i)} = \log\big(1 + e^{-y^{(i)} z^{(i)}}\big).$$

Hence the loss admits the compact form

$$\boxed{\mathcal{L}(\theta) = \sum_{i=1}^{m} \log\big(1 + e^{-y^{(i)}\, \theta^\top x^{(i)}}\big).}$$