OTUS

ОНЛАЙН-ОБРАЗОВАНИЕ

OTUS

Первая нейронная сеть(нет)

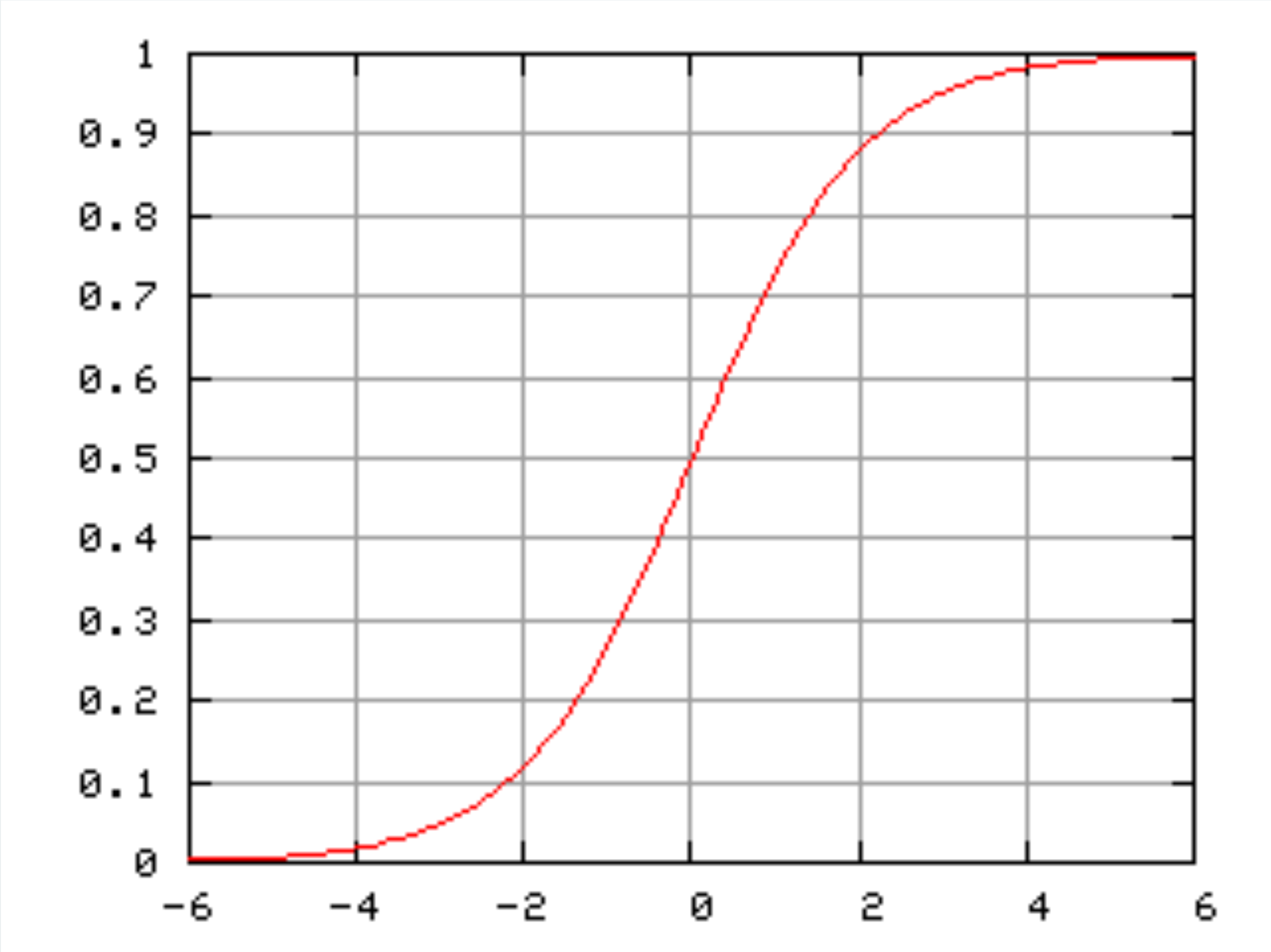
Теоретические основы и разбор кода

Артур Кадурин
Преподаватель

План на сегодня

1. **Логистическая регрессия**
2. Алгоритм обратного распространения
3. Трюк с логистической функцией
4. Практика

Логистическая функция



$$\sigma(z) = \frac{1}{1 + e^{-z}}$$

Логистическая регрессия

Логистическая регрессия — это модель применяющаяся для предсказания вероятности наступления события в зависимости от значений набора признаков.

$$\mathbb{P}\left\{y = 1 \middle| \boldsymbol{x}\right\} = \sigma\left(\boldsymbol{\theta}^T \boldsymbol{x}\right) \qquad \sigma(\boldsymbol{z}) = \frac{1}{1 + e^{-\boldsymbol{z}}}$$

Логистическая регрессия

Логистическая регрессия — это модель применяющаяся для предсказания вероятности наступления события в зависимости от значений набора признаков.

$$\mathbb{P}\left\{y = 1 \middle| \boldsymbol{x}\right\} = \sigma\left(\boldsymbol{\theta}^T \boldsymbol{x}\right) \qquad \sigma(\boldsymbol{z}) = \frac{1}{1 + e^{-\boldsymbol{z}}}$$

$$\mathbb{P}\left\{y \middle| \boldsymbol{x}\right\} = \sigma\left(\boldsymbol{\theta}^T \boldsymbol{x}\right)^y \left(1 - \sigma\left(\boldsymbol{\theta}^T \boldsymbol{x}\right)\right)^{1-y}$$

$$L\left(\boldsymbol{x},\ y \middle| \boldsymbol{\theta}\right) = \prod_{i=0}^{N} \mathbb{P}\left\{y = y^{(i)} \middle| \boldsymbol{x} = \boldsymbol{x}^{(i)}\right\}$$

Логистическая регрессия

$$L\left(\boldsymbol{x},\ y\middle|\boldsymbol{\theta}\right) = \prod_{i=0}^{N} \mathbb{P}\left\{y = y^{(i)}\middle|\boldsymbol{x} = \boldsymbol{x}^{(i)}\right\}$$

$$\mathbb{P}\left\{y\middle|\boldsymbol{x}\right\} = \sigma\left(\boldsymbol{\theta}^T\boldsymbol{x}\right)^y\left(1 - \sigma\left(\boldsymbol{\theta}^T\boldsymbol{x}\right)\right)^{1-y}$$

$$p^{(i)} = \sigma\left(\boldsymbol{\theta}^T\boldsymbol{x}^{(i)}\right)$$

$$\mathscr{L}\left(\boldsymbol{x}, y\middle|\boldsymbol{\theta}\right) = -\frac{1}{N}\log L\left(\boldsymbol{x},\ y\middle|\boldsymbol{\theta}\right) =$$

$$-\frac{1}{N}\sum_{i}\left[y^{(i)}\log p^{(i)} + \left(1 - y^{(i)}\right)\log\left(1 - p^{(i)}\right)\right]$$

Логистическая регрессия

$$L\left(\boldsymbol{x},\ y \middle| \boldsymbol{\theta}\right) = \prod_{i=0}^{N} \mathbb{P}\left\{y = y^{(i)} \middle| \boldsymbol{x} = \boldsymbol{x}^{(i)}\right\}$$

$$\mathbb{P}\left\{y \middle| \boldsymbol{x}\right\} = \sigma\left(\boldsymbol{\theta}^T \boldsymbol{x}\right)^y \left(1 - \sigma\left(\boldsymbol{\theta}^T \boldsymbol{x}\right)\right)^{1-y}$$

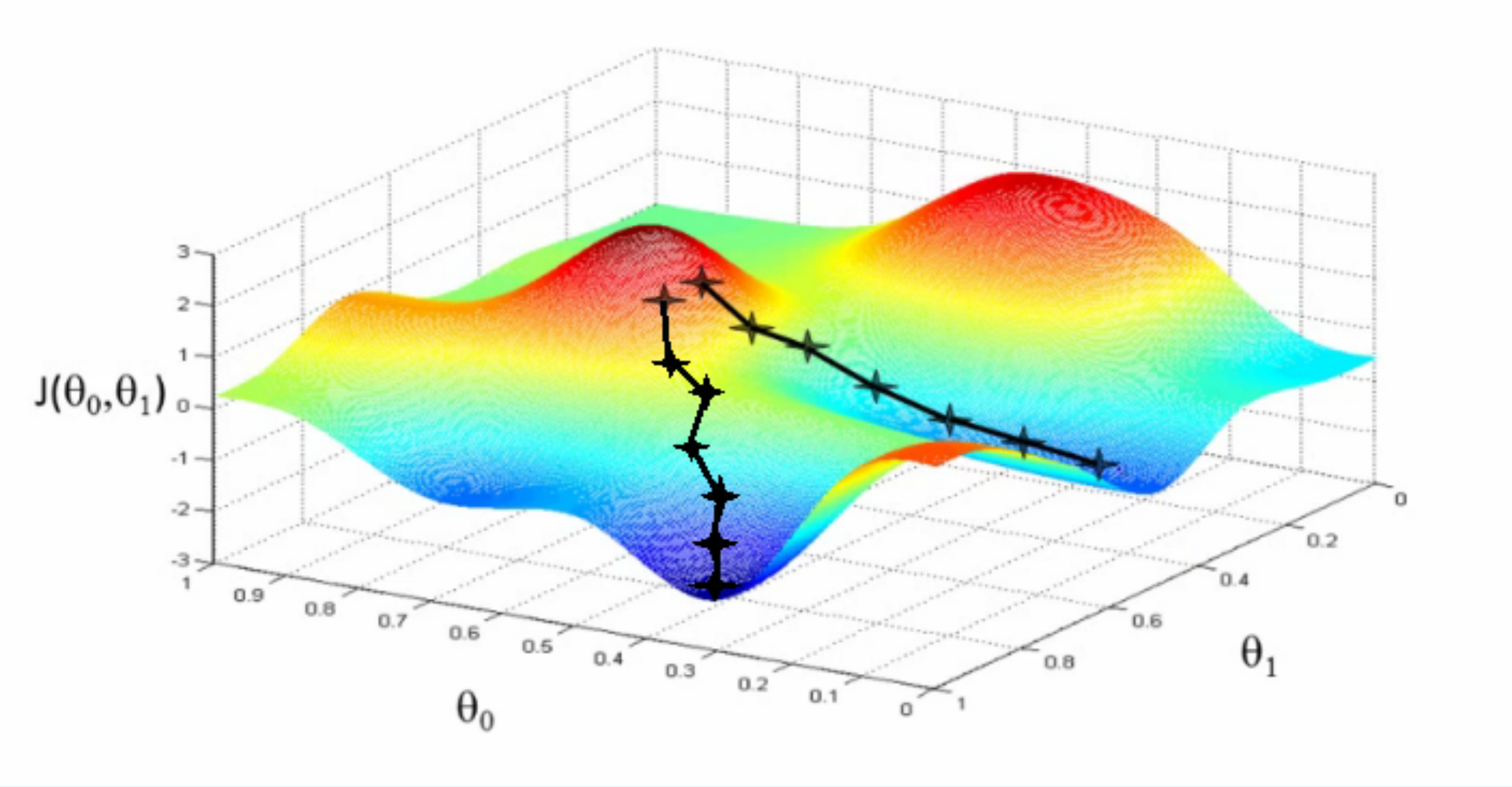
$$p^{(i)} = \sigma\left(\boldsymbol{\theta}^T \boldsymbol{x}^{(i)}\right)$$

$$\mathscr{L}\left(\boldsymbol{x}, y \middle| \boldsymbol{\theta}\right) = -\frac{1}{N} \log L\left(\boldsymbol{x},\ y \middle| \boldsymbol{\theta}\right) =$$

$$-\frac{1}{N} \sum_{i} \left[y^{(i)} \log p^{(i)} + \left(1 - y^{(i)}\right) \log\left(1 - p^{(i)}\right)\right]$$

Как минимизировать функцию потерь?

Градиентный спуск



План на сегодня

1. Логистическая регрессия
2. **Алгоритм обратного распространения**
3. Трюк с логистической функцией
4. Практика

Обратное распространение

$$\mathscr{L}\left(y,\ p\middle|\boldsymbol{\theta}\right) = -\frac{1}{N}\sum_i \left[y^{(i)}\log p^{(i)} + \ \left(1 - y^{(i)}\right)\log\left(1 - p^{(i)}\right)\right]$$

$$p = \sigma\left(\boldsymbol{\theta x} + b\right) = \sigma(\boldsymbol{z}) \qquad z = \boldsymbol{\theta x} + b \qquad \sigma(\boldsymbol{z}) = \frac{1}{1 + e^{-\boldsymbol{z}}}$$

Обратное распространение

$$\mathcal{L}\left(y,\ p\middle|\boldsymbol{\theta}\right) = -\frac{1}{N}\sum_i\left[y^{(i)}\log p^{(i)} + \ \left(1 - y^{(i)}\right)\log\left(1 - p^{(i)}\right)\right]$$

$$p = \sigma\left(\boldsymbol{\theta x} + b\right) = \sigma(\boldsymbol{z}) \qquad z = \boldsymbol{\theta x} + b \qquad \sigma(\boldsymbol{z}) = \frac{1}{1 + e^{-\boldsymbol{z}}}$$

$$\frac{\partial\mathcal{L}}{\partial\theta} = \frac{\partial\mathcal{L}}{\partial p}\frac{\partial p}{dz}\frac{\partial z}{\partial\theta}$$

$$\frac{\partial\mathcal{L}}{\partial p} = \frac{y}{p} - \frac{1-y}{1-p} \qquad \frac{\partial p}{\partial z} = \sigma(z)\left(1 - \sigma(z)\right) \qquad \frac{\partial z}{\partial\theta} = x$$

Обратное распространение

$$\mathscr{L}\left(y,\ p\middle|\theta\right) = -\frac{1}{N}\sum_i \left[y^{(i)}\log p^{(i)} + \left(1 - y^{(i)}\right)\log\left(1 - p^{(i)}\right)\right]$$

$$p = \sigma\left(\boldsymbol{\theta x} + b\right) = \sigma(\boldsymbol{z}) \qquad z = \boldsymbol{\theta x} + b \qquad \sigma(\boldsymbol{z}) = \frac{1}{1 + e^{-\boldsymbol{z}}}$$

$$p = \sigma_1\left(z_1\left(\sigma_0\left(z_0(x)\right)\right)\right) \qquad \frac{\partial \mathscr{L}}{\partial \theta_0} = \frac{\partial \mathscr{L}}{\partial \sigma_1}\frac{\partial \sigma_1}{\partial z_1}\frac{\partial z_1}{\partial \sigma_0}\frac{\partial \sigma_0}{\partial z_0}\frac{\partial z_0}{\partial \theta_0}$$

$$\frac{\partial \mathscr{L}}{\partial p} = \frac{y}{p} - \frac{1 - y}{1 - p} \qquad \frac{\partial p}{\partial z} = \sigma(z)\left(1 - \sigma(z)\right) \qquad \frac{\partial z}{\partial \theta} = x$$

План на сегодня

1. Логистическая регрессия
2. Алгоритм обратного распространения
3. **Трюк с логистической функцией**
4. Практика

Логистическая функция

$$\sigma(z) = \frac{1}{1 + e^{-z}}$$

Какие могут быть проблемы?

Логистическая функция

$$\sigma(z) = \frac{1}{1 + e^{-z}}$$

Какие могут быть проблемы?

При относительно небольших абсолютных значениях z e^{-z} может оказаться слишком большим.

Логистическая функция

$$\sigma(z) = \frac{1}{1 + e^{-z}}$$

Какие могут быть проблемы?

При относительно небольших абсолютных значениях z $\mathrm{e}^{-\mathrm{z}}$ может оказаться слишком большим.

При вычислении функции ошибки мы считаем логарифм от сигмоиды, если сигмоида равна нулю, то логарифм $-\mathrm{inf}$.

Логистическая функция

$$-\log\sigma(z) = -\log\frac{1}{1+e^{-z}} = \log\left(1+e^{-z}\right) =$$

$$\log\frac{1+e^{z}}{e^{z}} = -z + \log\left(1+e^{z}\right) =$$

$$\log\left(1+e^{-|z|}\right) - min(0, z)$$

План на сегодня

1. Логистическая регрессия
2. Алгоритм обратного распространения
3. Трюк с логистической функцией
4. **Практика**

Спасибо
за внимание!