Логистическая регрессия

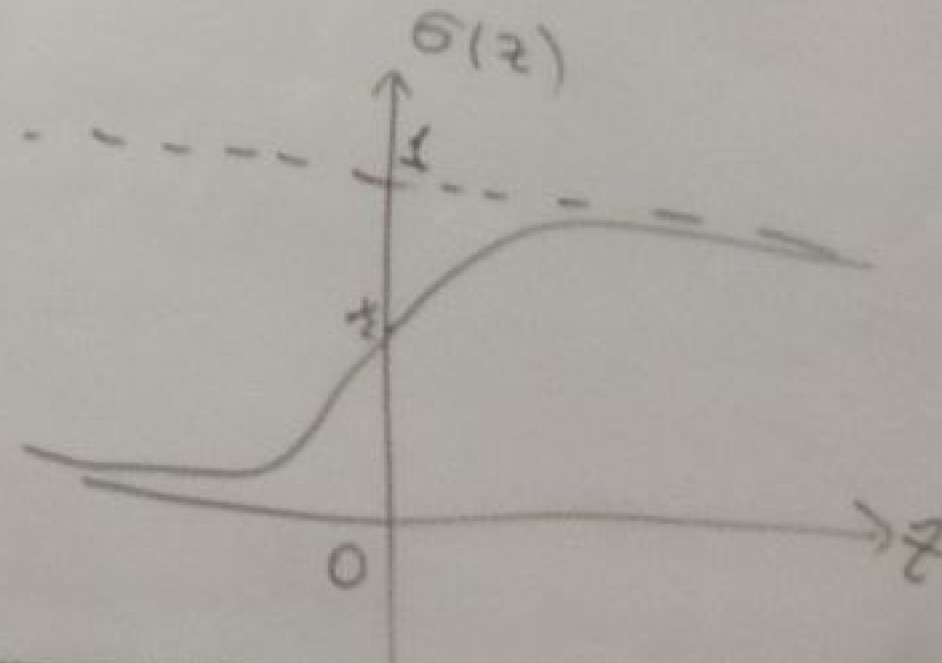
$y = \{0, 1\}$ - двухклассовая регр.

объекты почти выборка

$y_1, \dots, y_n$ - выборка, где $y_i \sim Bern(\mu_\theta(x_i))$

$\mu_\theta(x_i) = \sigma(\theta^T x_i)$

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



Сво…

$\{ \sigma($…


JEANS
CALVIN KLEIN

Логистическая регрессия

$$\frac{\partial L}{\partial \theta} = \sum_{i=1}^{n} \Big[y_i \left(1 - \sigma(\theta^T x_i)\right) + (1 - y_i) \cdot \left(-\sigma(\theta^T x_i)\right) \Big] x_i =$$

$$= \sum_{i=1}^{n} \left(y_i - \sigma(\theta^T x_i) \right) x_i = X^T \left(y - S(\theta) \right)$$

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \quad S(\theta) = \begin{pmatrix} \sigma(\theta^T x_1) \\ \vdots \\ \sigma(\theta^T x_n) \end{pmatrix}$$

Свойства лог.ре

$\{ \sigma(\theta^T x) = \frac{1}{2} \}$

σ – растёт –

Свойства сиг

1. $\sigma(-z)$
2. $\sigma^{-1}(z$
3. $\sigma'(z)$

Классификация

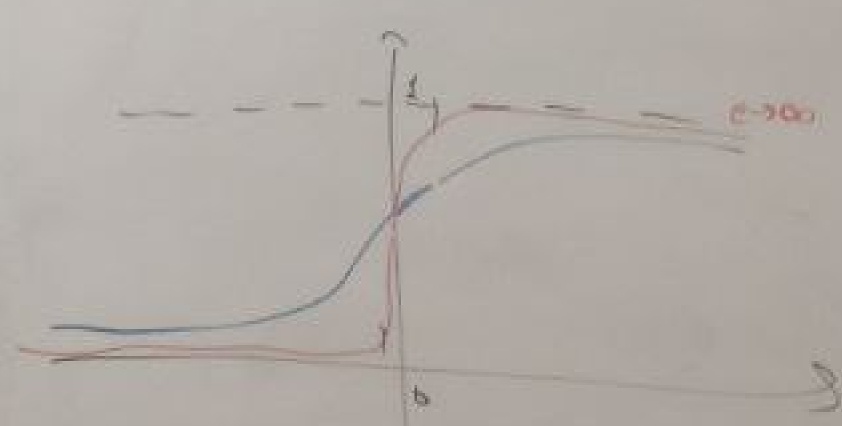
Переобучение

Рассмотрим случай.
- классы линейно раздел.
- среди признаков есть един.
- $\exists\theta$, т.ч. $\{\theta^T x = 0\}$ — разделяет классы

Логистическая регрессия

$\forall c > 0 \in \mathbb{R}: \{ c\,\theta^T x = 0 \} = \{ \theta^T x = 0 \}$ – разделяющая гипер плоскость

$\sigma(c\,\underbrace{\theta^T x}_{>0}) \xrightarrow[c \to \infty]{} 1$, повышается вер.



Проблемы

- мы п.и. либо ~0
  либо ~1
- $\theta^T X \approx 0 \to$

$\Rightarrow$ сложно классифицировать

Градиентный спуск(п

$\theta_{k+1} = \theta_k + \eta \nabla F(\theta_k)$

$\theta_{k+1} = \theta_k + \eta \; X^T($


JEANS


JEANS
CALVIN KLEIN

Логистическая регрессия

$$DDF(\theta) = \left(\sum_{i=1}^{n} \sigma(\theta^T x_i) x_i\right)' =$$

$$= \sum_{i=1}^{n} \left[x_i^T \sigma(\theta^T x_i) \cdot (1 - \sigma(\theta^T x_i)) x_i \right]$$

$$= \underset{d\times n}{X^T} \underset{n\times n}{V(\theta)} \underset{n\times d}{X} \qquad V(\theta) = \mathrm{diag}\left(\sigma(\theta^T x_i)(1 - \sigma(\theta^T x_i))\right)$$

$$\theta_{k+1} = \theta_k - (X^T V(\theta_k) X)^{-1} X^T (S(\theta_k) - y) =$$

$$= (X^T V(\theta) X)^{-1} \left[X^T V(\theta) X \theta_k - X^T (S(\theta_k) - y) \right] =$$

IF

D

DD

IRLS (Itterative Reweighted Least Squ

$$= (X^T V(\theta_k) X)^{-1} X^T V(\theta_k)\, z(\theta_k)$$

$$z(\theta_k) = \left[X\theta_k - V(\theta_k)^{-1}(S(\theta_k) - y)\right]$$