ДЗ №4: Свойства логистической регрессии

Вес = 0,5

Дедлайн - 30.04.2025

В лекциях была рассмотрена модель логистической регрессии для задачи бинарной классификации.

Напомним, в этой модели с практической точки зрения есть два ключевых свойства:

- Вероятность для объекта быть объектом класса $c = 1$ в бинарном случае моделируется как математическое ожидание случайной величины y_i:

  $$y_i \sim \mathcal{B}(p_i), \tag{1}$$

  где $\mathcal{B}(p_i)$ - обозначение распределения Бернулли с параметром p_i, i - индекс, нумерующий объекты выборки. При этом параметр p_i этого распределения предполагается связанным с признаковым описанием x_i объекта следующей обобщенной линейной функцией (сигмоидой):

  $$\begin{gathered} p = \frac{1}{1 + e^{-z}}, \\ z = \theta \cdot x \end{gathered} \tag{2}$$

  где θ - параметры модели, оптимизируемые в подходе максимального правдоподобия (MLE, Maximum Likelihood Estimation).

- Подход оптимизации (максимизации) правдоподобия в случае бинарной задачи приводит к следующей функции потерь:

  $$\mathcal{L}(\theta, X) = -\sum_{i=1}^{N} \left(y_i * \ln p_i + (1 - y_i) * \ln(1 - p_i)\right) \tag{3}$$

На лекции без доказательства были произнесены следующие свойства логистической регрессии:

1.

У функции потерь (3) есть минимум, он единственный и потому глобальный. **Покажите это.**

> Подсказка: для этого необходимо показать, что:
>
> - функции $f_1 = -\ln(\sigma(z))$ и $f_2 = -\ln(1 - \sigma(z))$ выпуклые на одномерном пространстве z. А именно: производная этих функций по z везде возрастает, или, что эквивалентно с допущением о двойном дифференцировании, вторая производная по z везде положительна;
> - Дважды дифференцируемая функция аффинной функции $(\theta \cdot x)$ выпукла на пространстве θ.

2.

Разделяющая гиперповерхность в решении с применением модели логистической регрессии линейна. То есть, в случае двумерного признакового описания это прямая, в случае трехмерного - плоскость, в случае многомерного - гиперплоскость. **Покажите это.**

> Подсказка: для этого имеет смысл воспользоваться смыслом разделяющей поверхности: это поверхность, во всех точках которой выполняется условие $p = C$, где C - некоторая константа, например, $C = 0,5$.