Краткое введение: **Три математические опоры искусственного интеллекта**

Автор: подготовлено для обучающихся в области ИИ

Июль 2025

Содержание

1 Обзор ключевых математических направлений

Искусственный интеллект опирается на три фундаментальных области математики:

1. **Вероятностные модели и статистика**
2. **Оптимизационные методы**
3. **Линейная алгебра**

2 Вероятность и статистика: описание неопределённости

2.1 Случайные величины и их распределения

Случайная величина — это числовая модель случайного эксперимента. Примеры:

- **Равномерное распределение:** $p(x) = 1, \quad x \in [0, 1]$
- **Биномиальное распределение:**

$$p_k(x) = \frac{(l_k + w_k + 1)!}{l_k!\, w_k!} x^{w_k} (1 - x)^{l_k}$$

Ожидание выигрыша:

$$\mathbb{E}[P_k] = \frac{w_k + 1}{l_k + w_k + 2}$$

2.2 Математическое ожидание и дисперсия

Основные понятия:

$$\mathbb{E}[\xi] = \sum \xi_i p_i$$

$$\mathrm{Var}(\xi) = \mathbb{E}[(\xi - \mathbb{E}[\xi])^2]$$

Для суммы независимых случайных величин верно:

$$\mathbb{E}\left[\sum \xi_k\right] = \sum \mathbb{E}[\xi_k], \quad \mathrm{Var}\left(\sum \xi_k\right) = \sum \mathrm{Var}(\xi_k)$$

2.3 Центральная предельная теорема (ЦПТ)

ЦПТ утверждает, что:

$$\frac{1}{\sqrt{n}} \sum_{k=1}^{n} \xi_k \xrightarrow{d} \mathcal{N}(0, \sigma^2)$$

Для среднего значения:

$$\frac{1}{n} \sum_{k=1}^{n} \xi_k = m + \frac{\sigma}{\sqrt{n}} \eta, \quad \eta \sim \mathcal{N}(0, 1)$$

Доверительная граница:

$$\mathbb{P}\left(\left|\frac{1}{n} \sum \xi_k - h\right| \leq \frac{\sigma}{\sqrt{n}} \sqrt{\ln(1/\delta)}\right) \geq 1 - \delta$$

3 Оптимизация: поиск наилучших решений

3.1 Максимальное правдоподобие

Оценка параметров θ:

$$\hat{\theta} = \arg\max_{\theta} \prod_{k=1}^{n} p(\xi_k | \theta)$$

Для монетки:

$$L = x^S (1 - x)^{n - S}, \quad \hat{x} = \frac{S}{n}$$

3.2 Метод наименьших квадратов

Минимизация среднеквадратичной ошибки:

$$\hat{h} = \arg\min_h \sum_{k=1}^{n} (v_k - hr_k)^2$$

3.3 Байесовская оценка и регуляризация

При наличии априорной информации:

$$\hat{x} = \arg\min \left[\frac{1}{2\sigma_\xi^2} \sum (\xi_k - x)^2 + \frac{x^2}{2\sigma_x^2} \right]$$

В случае LASSO-регуляризации:

$$\hat{x} = \arg\min \left[\frac{1}{2\sigma_\xi^2} \sum (\xi_k - x)^2 + \lambda |x| \right]$$

4 Линейная алгебра: язык многомерных данных

4.1 Ключевые инструменты

- **SVD-разложение** (сингулярное разложение)
- **Метод главных компонент (PCA)**

Эти методы позволяют выделять важные признаки и сокращать размерность данных.

5 Машинное обучение: от теории к практике

5.1 Формулировка задачи

Статистическая модель:

$$y = f(a, x) + \xi, \quad \xi \sim \mathcal{N}(0, \sigma^2)$$

Цель — минимизировать:

$$\min_x \mathbb{E}_{(a,y)}[(y - f(a, x))^2]$$

5.2 Эмпирический подход

В практическом ML минимизируется эмпирический риск:

$$\min_x \frac{1}{n} \sum_{k=1}^{n} (y_k - f(a_k, x))^2$$

5.3 Переобучение и проверка моделей

Ключевые моменты:

- Разделение данных на обучающую и тестовую выборки
- Регуляризация для предотвращения переобучения

6 Выводы

- **Теория вероятностей** объясняет природу случайности и формирует статистические основы.
- **Методы оптимизации** позволяют эффективно обучать модели.
- **Линейная алгебра** обеспечивает обработку и анализ многомерных данных.
- **Понимание статистики необходимо для корректного применения машинного обучения.**