

Обзор: Три математических опоры искусственного интеллекта

Автор: Альтернативная редакция лекции по ИИ

Июль 2025

Ключевые разделы математики в ИИ

1. **Теория вероятностей и статистика:** Байесовский подход, Центральная предельная теорема, Оценка максимального правдоподобия
2. **Методы оптимизации:** Идеи Ферма и Лагранжа, стохастический градиентный спуск
3. **Линейная алгебра:** SVD, PCA, спектральные методы

1 Вероятностные модели и статистические методы

1.1 Байесовский взгляд и задача многорукого бандита

- Рассматривается ситуация: у каждой "ручки" вероятность выигрыша P_k
- После l_k неудач и w_k удачных попыток апостериорное распределение:

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

- Ожидаемая вероятность выигрыша:

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

1.2 Центральная предельная теорема (ЦПТ)

- Сумма независимых одинаково распределённых величин стремится к нормальному распределению:

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

- Плотность Гаусса:

$$\rho(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2}{2\sigma^2}\right)$$

1.3 Статистические оценки и доверительные интервалы

- Оценка среднего:

$$\hat{h} = \frac{1}{n} \sum \xi_k$$

- Доверительный интервал:

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

- Метод максимального правдоподобия:

$$\hat{x} = \arg \max_x \prod p(\xi_k, x)$$

- Пример линейной регрессии:

$$v_k = H \cdot r_k + \varepsilon_k, \quad \hat{H} = \arg \min_H \sum (v_k - H r_k)^2$$

2 Оптимизация и машинное обучение

2.1 Байесовская регуляризация и LASSO

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

Для L1-регуляризации:

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

2.2 Машинное обучение и статистика: различия

$$\text{ML: } \min_x \frac{1}{n} \sum (y_k - f(a_k, x))^2$$

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

2.3 Переобучение и валидация

- Необходимость разделения на обучающую и тестовую выборки
- Проблема переобучения при увеличении числа параметров

3 Фундаментальные идеи и комментарии

- Теория вероятностей объясняет случайность и неопределённость
- Статистика восстанавливает скрытые зависимости из выборок
- Оптимизация обучает модели и минимизирует ошибки
- Линейная алгебра даёт инструменты для анализа и сжатия данных

«Если множество случайных факторов взаимодействует, результат — гауссовское распределение.»