ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ

«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ ИТМО»

Университет ИТМО

Отчет по лабе № 1

Выполнили работу:

Чекалин Макар Максимович (ИСУ: 467982)

Академические группы:

№ j3111

Санкт-Петербург 2025

Задача:

Доказать, что оптимальные направления PCA совпадают с собственными векторами матрицы ковариаций

Пусть дан центрированный набор данных X∈Rn×m, где:
n — количество объектов,
m — количество признаков.

Ковариационная матрица:

$$\Sigma = \frac{1}{n-1} X^T X.$$

Цель PCA

Найти такое направление, что при проекции данных дисперсия будет максимальной.
Задача формализуется как:

$$\max_w \operatorname{Var}(Xw) = \frac{1}{n-1} \|Xw\|^2 = w^T \Sigma w,$$

Решение через метод Лагранжа

Вводим функцию Лагранжа:

$$\mathcal{L}(w, \lambda) = w^T \Sigma w - \lambda(w^T w - 1)$$

Условие стационарности:

$$\frac{\partial \mathcal{L}}{\partial w} = 2\Sigma w - 2\lambda w = 0 \implies \Sigma w = \lambda w.$$

Получаем собственное уравнение, где:

w — собственный вектор ковариационной матрицы Σ

λ — собственное значение.

Спектральное разложение

Поскольку Σ — симметричная и положительно полуопределённая, она допускает разложение:

$$\Sigma = \sum_{i=1}^{m} \lambda_i v_i v_i^T,$$

$\lambda 1 \geq \lambda 2 \geq \cdots \geq \lambda m \geq 0$— собственные значения,

vi — соответствующие ортонормированные собственные векторы.

Максимизация дисперсии

Любой единичный вектор w можно разложить по базису собственных векторов:

$$w = \sum_{i=1}^{m} a_i v_i, \quad \text{где } \sum_{i=1}^{m} a_i^2 = 1.$$

Подставляя в дисперсию:

$$w^T \Sigma w = \sum_{i=1}^{m} \lambda_i a_i^2.$$

Максимум достигается, когда $a_1 = 1$ (для $\lambda_{\max}$), а остальные $a_j = 0$. Таким образом, оптимальное направление:

$$w = v_1,$$

где v_1 — собственный вектор, соответствующий наибольшему собственному значению λ_1.

Последующие компоненты:
Для поиска остальных направлений вводим условие ортогональности:
Вторая компонента: $w2 = v2$ (собственный вектор для $\lambda 2$)
И так далее: $wk = vk$

Вывод:
Главные компоненты PCA — это собственные векторы ковариационной матрицы, отсортированные по убыванию собственных значений.

Первая компонента $v1$ максимизирует дисперсию.
Остальные $v2, v3 \dots$ добавляют максимум возможной информации, сохраняя ортогональность.

Этот результат логически следует из задачи оптимизации и свойств спектрального разложения.