Математическое обоснование (Задание (Expert Level))

Постановка задачи

Пусть $X \in \mathbb{R}^{n\times m}$ — матрица данных. Требуется найти ортонормированные направления $\mathbf{w}_1, \dots, \mathbf{w}_k$, и доказать, что они максимизируют дисперсию проекций.

Матрица ковариаций

$$\Sigma = \frac{1}{n-1} X^\top X \quad \text{(симметричная, положительно определённая).}$$

Максимизация дисперсии

Дисперсия проекции на вектор $\mathbf{w}$:

$$\mathrm{Var}(\mathbf{w}) = \mathbf{w}^\top \Sigma \mathbf{w}.$$

Задача оптимизации с ограничением:

$$\max_{\mathbf{w}} \mathbf{w}^\top \Sigma \mathbf{w} \quad \text{при} \quad \|\mathbf{w}\| = 1.$$

Решение через множители Лагранжа

Введём функцию Лагранжа:

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

Где:
λ — множитель Лагранжа,
$\mathbf{w}^\top \mathbf{w} - 1 = 0$ — условие нормировки.
Для поиска экстремума вычислим градиент фунции Лагранжа по $\mathbf{w}$ и приравняем его к нулю:

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

Таким образом, $\mathbf{w}$ — собственный вектор Σ, а λ — собственное значение.

Связь дисперсии и собственных значений

Из $\Sigma\mathbf{w} = \lambda\mathbf{w}$ следует:

$$\mathrm{Var}(\mathbf{w}) = \mathbf{w}^\top\Sigma\mathbf{w} = \lambda\mathbf{w}^\top\mathbf{w} = \lambda.$$

т.к. $\mathbf{w}^\top\mathbf{w} = 1$
Следовательно, максимальная дисперсия соответствует наибольшему λ.

Доказательство отсутствия других решений

По теореме о спектральном разложении:

- Собственные векторы $\mathbf{w}_i$ симметричной матрицы Σ ортонормированы и образуют базис.
- Соответствующие λ_i упорядочены: $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m \geq 0$.

Любой вектор $\mathbf{v}$ можно разложить по этому базису:
$\mathbf{v} = \sum_{i=1}^m c_i\mathbf{w}_i, \quad$ где $\sum_{i=1}^m c_i^2 = 1$.

Дисперсия проекции: $\mathbf{v}^\top S\mathbf{v} = \sum_{i=1}^m \lambda_i c_i^2$.

Максимум достигается при $c_1 = 1$, $c_2 = c_3 = ... = 0$, то есть $\mathbf{v} = \mathbf{w}_1$.

Обобщение на несколько компонент

: 1. Вторая компонента: Ищется в подпространстве, ортогональном $\mathbf{w}_1$:
$\max_{\mathbf{u}\perp\mathbf{w}_1} \mathbf{u}^\top\Sigma\mathbf{u}, \quad \|\mathbf{u}\| = 1$.
Решение — $\mathbf{w}_2$ (собственный вектор с λ_2).

2. Ортогональность компонент:
Следует из ортогональности собственных векторов симметричной матрицы.

Итог

- Главные компоненты PCA - собственные векторы матрицы Σ, упорядоченные по убыванию собственных значений
- Метод множителей Лагранжа доказывает, что экстремумы достигаются на собственных векторах, а собственные значения задают "силу"направлений.