

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

Решение задачи

Пусть $\mathbf{X} \in \mathbb{R}^{n \times m}$ центрированная матрица данных, где n — число наблюдений, m — число признаков, и каждый столбец имеет нулевое среднее ($\sum_{i=1}^n x_{ij} = 0$). Матрица ковариаций:

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

Цель RSA — найти единичные векторы $\mathbf{v} \in \mathbb{R}^m$ ($\mathbf{v}^T \mathbf{v} = 1$), максимизирующие дисперсию проекций данных на \mathbf{v} .

Проекция i -го наблюдения $\mathbf{x}_i \in \mathbb{R}^m$ (строка матрицы \mathbf{X}) на вектор \mathbf{v} :

$$z_i = \mathbf{x}_i^T \mathbf{v}.$$

Так как данные центрированы, среднее проекций равно нулю:

$$\frac{1}{n} \sum_{i=1}^n z_i = \mathbf{v}^T \left(\frac{1}{n} \sum_{i=1}^n \mathbf{x}_i \right) = 0.$$

Дисперсия проекций:

$$\text{Var}(z) = \frac{1}{n-1} \sum_{i=1}^n z_i^2 = \frac{1}{n-1} (\mathbf{X}\mathbf{v})^T (\mathbf{X}\mathbf{v}) = \frac{1}{n-1} \mathbf{v}^T \mathbf{X}^T \mathbf{X} \mathbf{v} = \mathbf{v}^T \Sigma \mathbf{v}.$$

Задача PCA:

$$\max_{\mathbf{v}} \mathbf{v}^T \Sigma \mathbf{v} \quad \text{при} \quad \mathbf{v}^T \mathbf{v} = 1.$$

Составим функцию Лагранжа:

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

Дифференцируем по \mathbf{v} :

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

Таким образом, \mathbf{v} — собственный вектор Σ , а λ — собственное значение. Дисперсия:

$$\mathbf{v}^T \Sigma \mathbf{v} = \mathbf{v}^T (\lambda \mathbf{v}) = \lambda.$$

Для максимизации дисперсии выбираем $\mathbf{v} = \mathbf{v}_1$, соответствующий наибольшему λ_1 , что даёт дисперсию λ_1 .

Вторая компонента максимизирует дисперсию при $\mathbf{v}^T \mathbf{v}_1 = 0$. Функция Лагранжа:

$$\mathcal{L}(\mathbf{v}, \lambda, \mu) = \mathbf{v}^T \Sigma \mathbf{v} - \lambda(\mathbf{v}^T \mathbf{v} - 1) - \mu \mathbf{v}^T \mathbf{v}_1.$$

Дифференцируем:

$$2\Sigma \mathbf{v} - 2\lambda \mathbf{v} - \mu \mathbf{v}_1 = 0.$$

Умножим на \mathbf{v}_1^T :

$$2\mathbf{v}_1^T \Sigma \mathbf{v} - 2\lambda \mathbf{v}_1^T \mathbf{v} - \mu = 0.$$

Так как $\mathbf{v}_1^T \mathbf{v} = 0$ и $\mathbf{v}_1^T \Sigma \mathbf{v} = \mathbf{v}^T \Sigma \mathbf{v}_1 = \lambda_1 \mathbf{v}^T \mathbf{v}_1 = 0$, то $\mu = 0$. Получаем:

$$\Sigma \mathbf{v} = \lambda \mathbf{v}.$$

Представим $\mathbf{v} = \sum_{i=2}^m \alpha_i \mathbf{v}_i$. Дисперсия:

$$\mathbf{v}^T \Sigma \mathbf{v} = \sum_{i=2}^m \alpha_i^2 \lambda_i \leq \lambda_2 \sum_{i=2}^m \alpha_i^2 = \lambda_2,$$

максимум при $\mathbf{v} = \mathbf{v}_2$. Аналогично, k -я компонента — \mathbf{v}_k , собственный вектор для λ_k , ортогональный $\mathbf{v}_1, \dots, \mathbf{v}_{k-1}$.

Таким образом, оптимальные направления PCA — собственные векторы Σ , где k -я компонента соответствует λ_k , максимизируя дисперсию в ортогональном подпространстве.

Выполнили:

Абаянцева Евгения Юрьевна J3110, ИСУ 329004
Смольникова Юлия Юрьевна, J3114, ИСУ 474363