

Корреляционные меры сходства и меры ассоциативности

Введение в кластерный анализ

Коэффициент корреляции Пирсона

$$d_{ij} = \frac{\sum_{k=1}^m (x_{ik} - \bar{x}_i)(x_{jk} - \bar{x}_j)}{\sqrt{\sum_{k=1}^m (x_{ik} - \bar{x}_i)^2 (x_{jk} - \bar{x}_j)^2}}$$

Расстояние Махаланобиса (обобщённое расстояние)

$$d_{ij} = \sqrt{(x_i - x_j)^T \Sigma^{-1} (x_i - x_j)}$$

x_i, x_j — объекты

Σ — ковариационная матрица объектов

Коэффициенты ассоциативности

$$\varphi = \sqrt{\frac{\sum (f_0 - f)^2}{f}}$$

$$f = \frac{n_r n_c}{n}$$

f_0 — значение в ячейке

n_r — итоговое число по строке

n_c — итоговое число по столбцу

n — итоговое число по строкам и столбцам