

SVD Метод Simon Funk'a

Вещественная матрица X размера $n \times m$, ее элементы x_{ij} .

Приближаем ее матрицей \tilde{X} размера $n \times m$, ее элементы \tilde{x}_{ij} .
 $\tilde{X} = U \cdot V$,

Размер матрицы U равен $n \times k$, ее элементы u_{ij} .

Размер матрицы V равен $k \times m$, ее элементы v_{ij} .

Тогда справедливо $\tilde{x}_{ij} = u_{i \cdot} \cdot v_{\cdot j}$

Качество приближения измеряем

$$L_1 = \sum e_{ij}^2$$

где $e_{ij} = x_{ij} - \tilde{x}_{ij}$

и суммируем только по элементам матрицы X , отличным от пропусков.

Не путаем разреженную матрицу (sparse matrix) и матрицу с большой долей пропусков!

В первой много нулей, не пропусков!

Матрицы U и V находим методом скорейшего спуска, итеративно.

https://surprise.readthedocs.io/en/stable/matrix_factorization.html#surprise.prediction_algorithms.matrix_factorization.SVDpp

Матрица триплетов получается из исходной матрицы X .

Номер строки * номер столбца * элемент матрицы

Перед каждой эпохой строки матрицы триплетов переставляем.