



Матричные факторизации

Андрей Зимовнов (Яндекс, ВШЭ)



Уменьшение размерности

- Матрица рейтингов **разреженная** (много пропусков)
- Оценки товаров **коррелируют** (юзерам часто нравятся одни и те же пары товаров)
- **Как сжать матрицу оценок?**



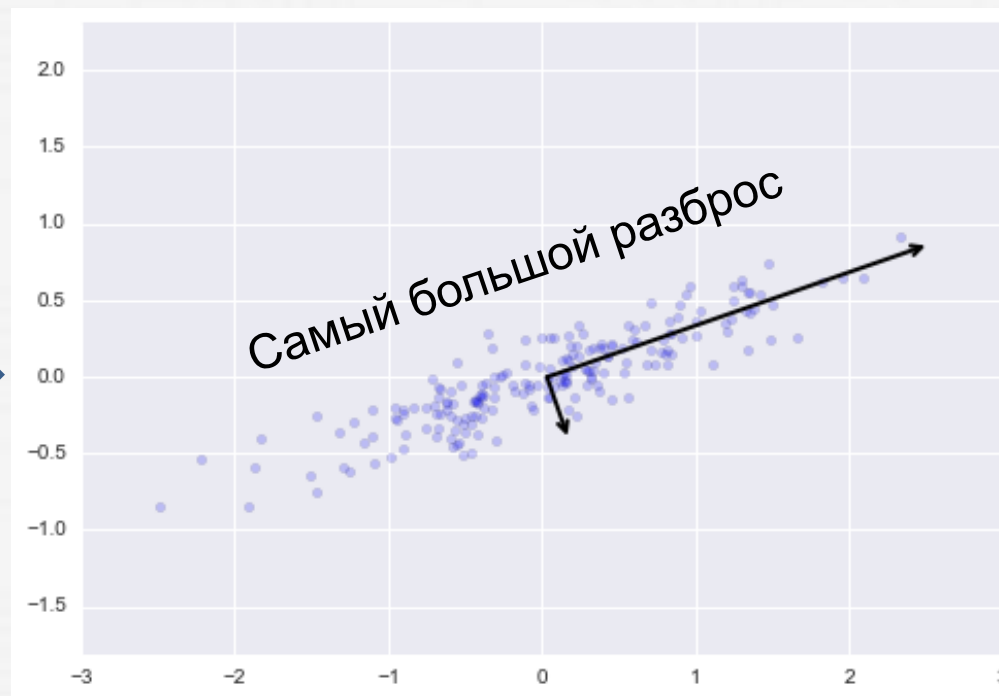
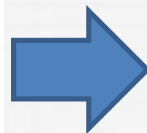
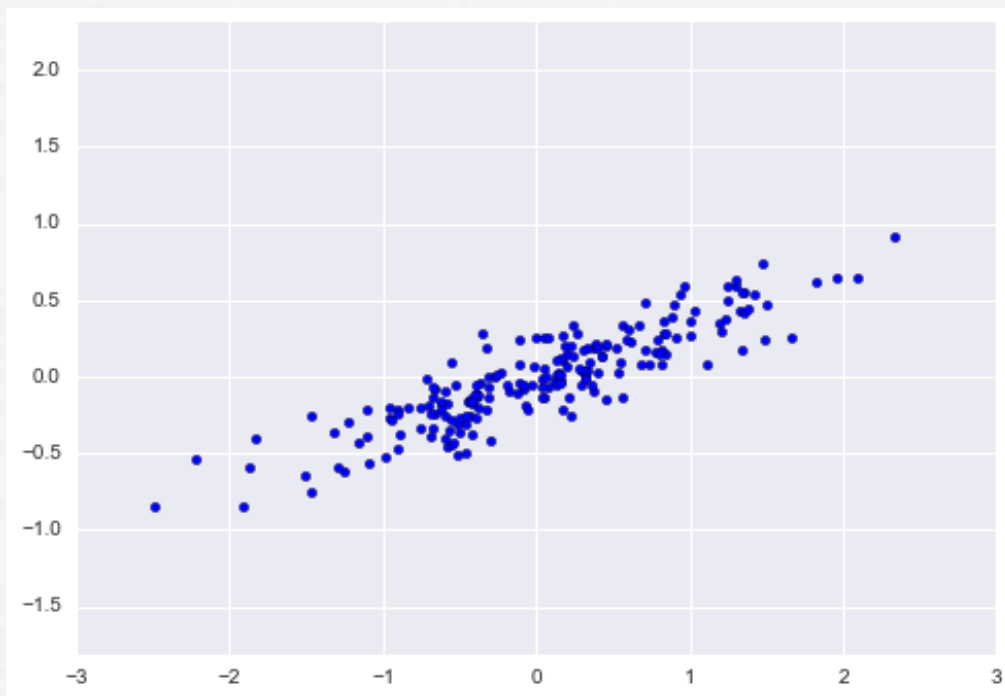
Метод главных компонент

Также известен как PCA (Principal Component Analysis)

Идея: найдем для наших данных новые оси координат, которые лучше объясняют эти данные!

Метод главных компонент

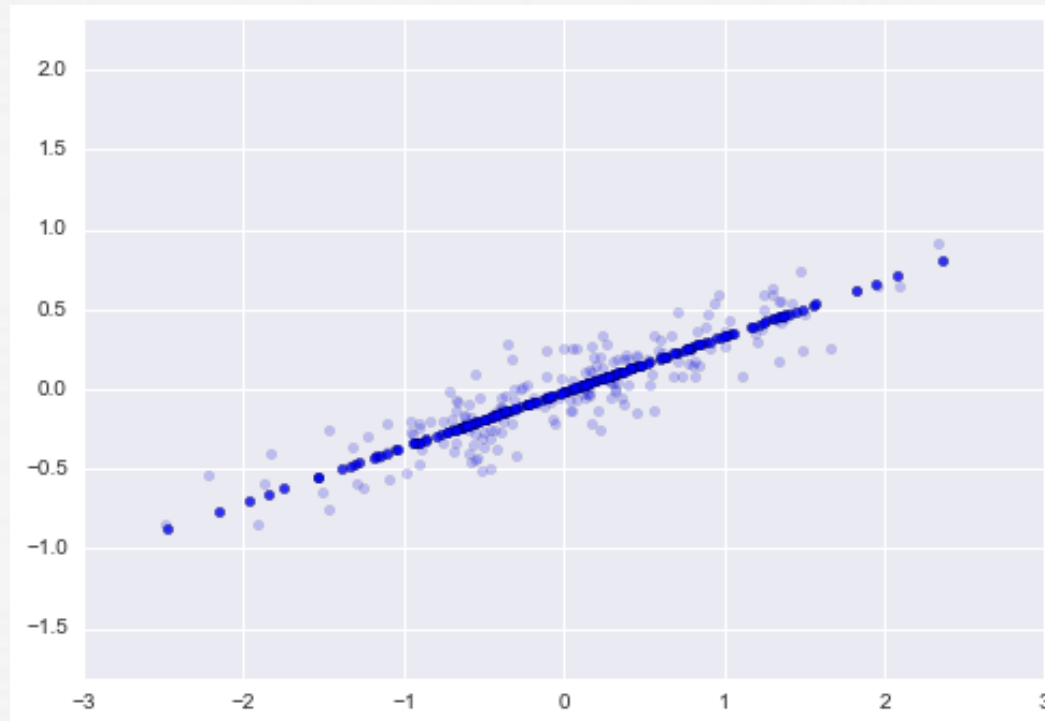
Новые оси координат ортогональны, они же все-таки оси



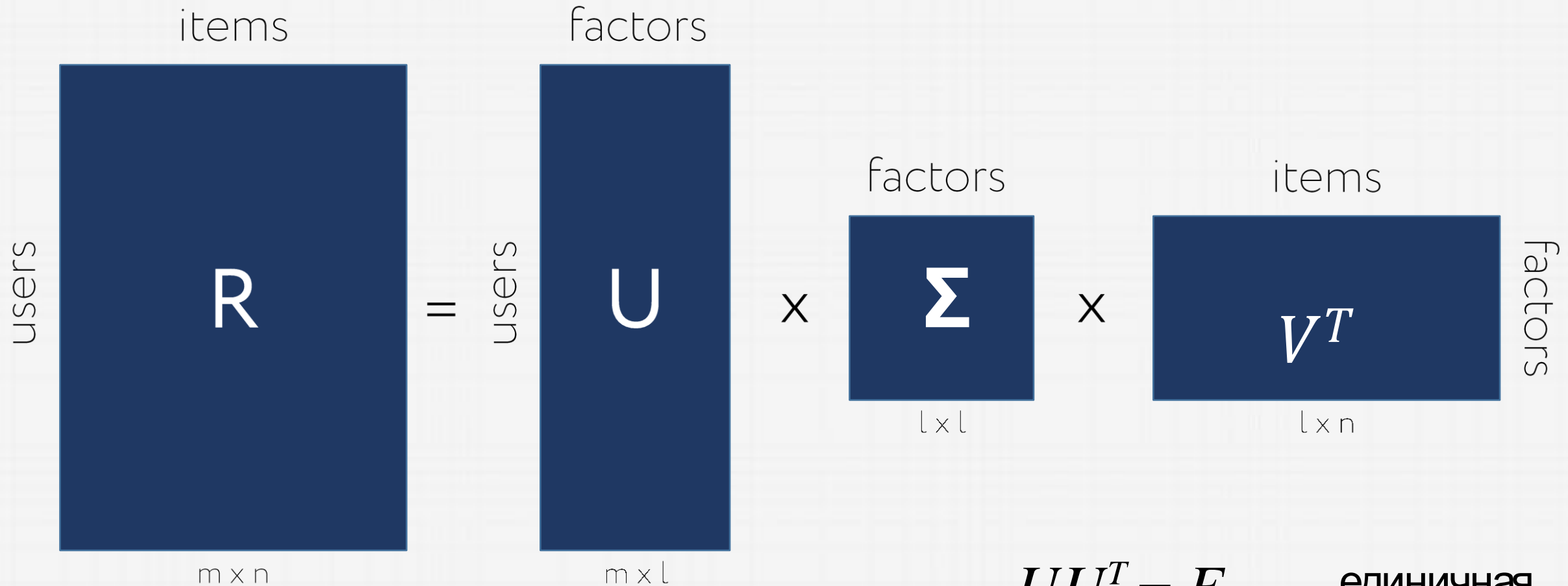
Метод главных компонент

Возьмем проекцию только на первую главную ось (они отсортированы по убыванию объясняемой дисперсии).

Сжали данные, потеряв немного информации!



РСА реализуют через SVD



Факторизация матрицы

$$UU^T = E$$

$$VV^T = E$$

единичная
матрица



Интерпретация для рейтингов

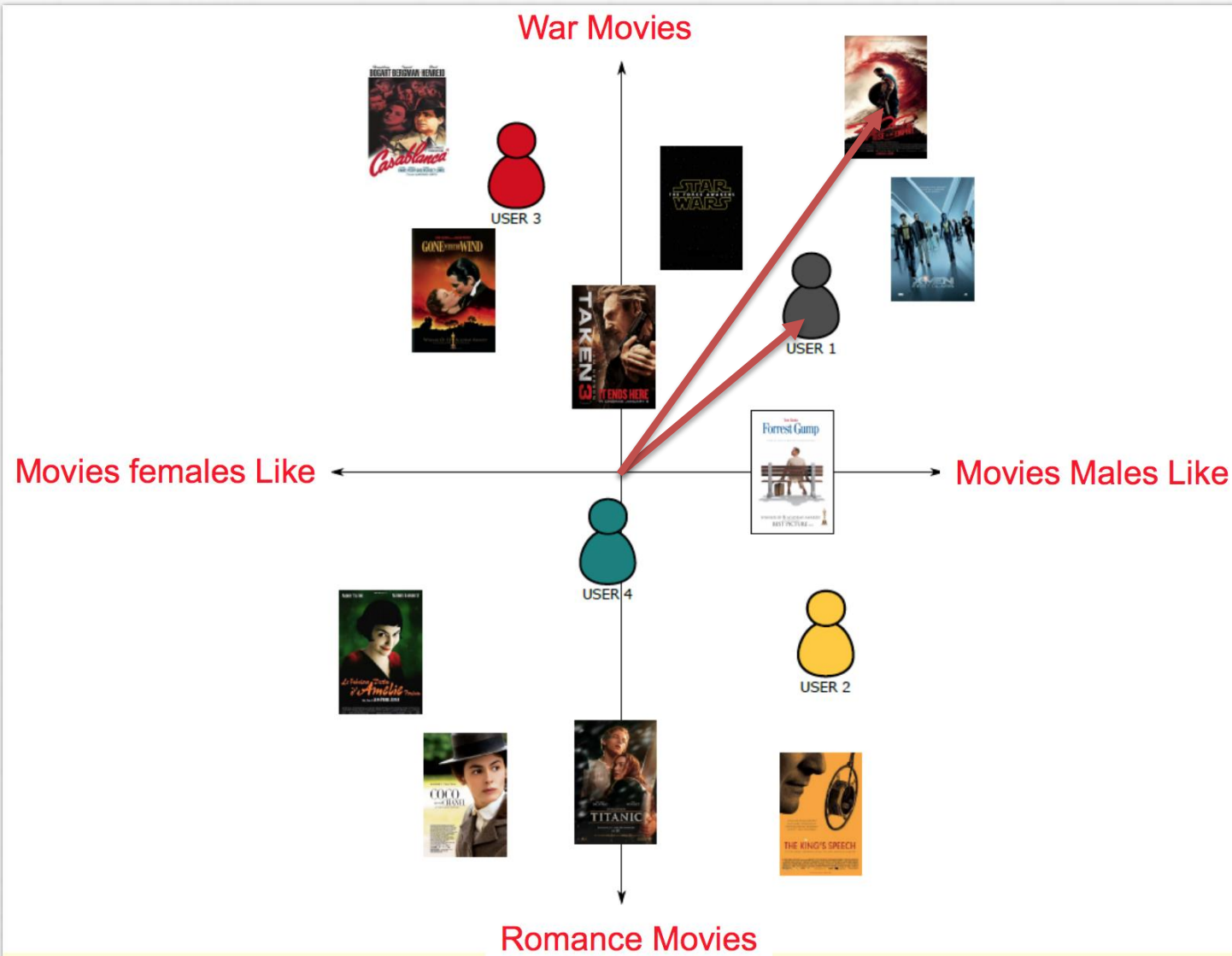
Юзеры и товары погружаются в новое пространство (малой размерности).

Координаты в этом пространстве – «характеристики» товаров и юзеров.

Похожесть в этом пространстве – это скалярное произведение!

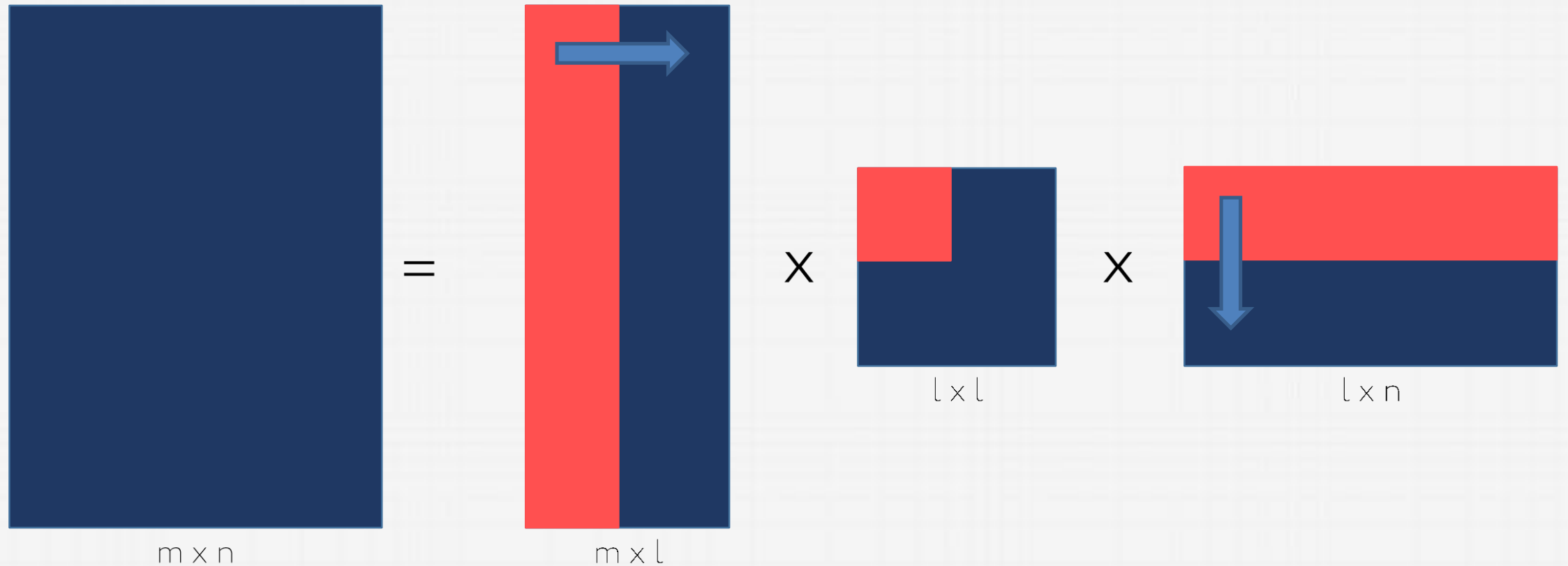
Интерпретация для рейтингов

Взяли 2
компоненты



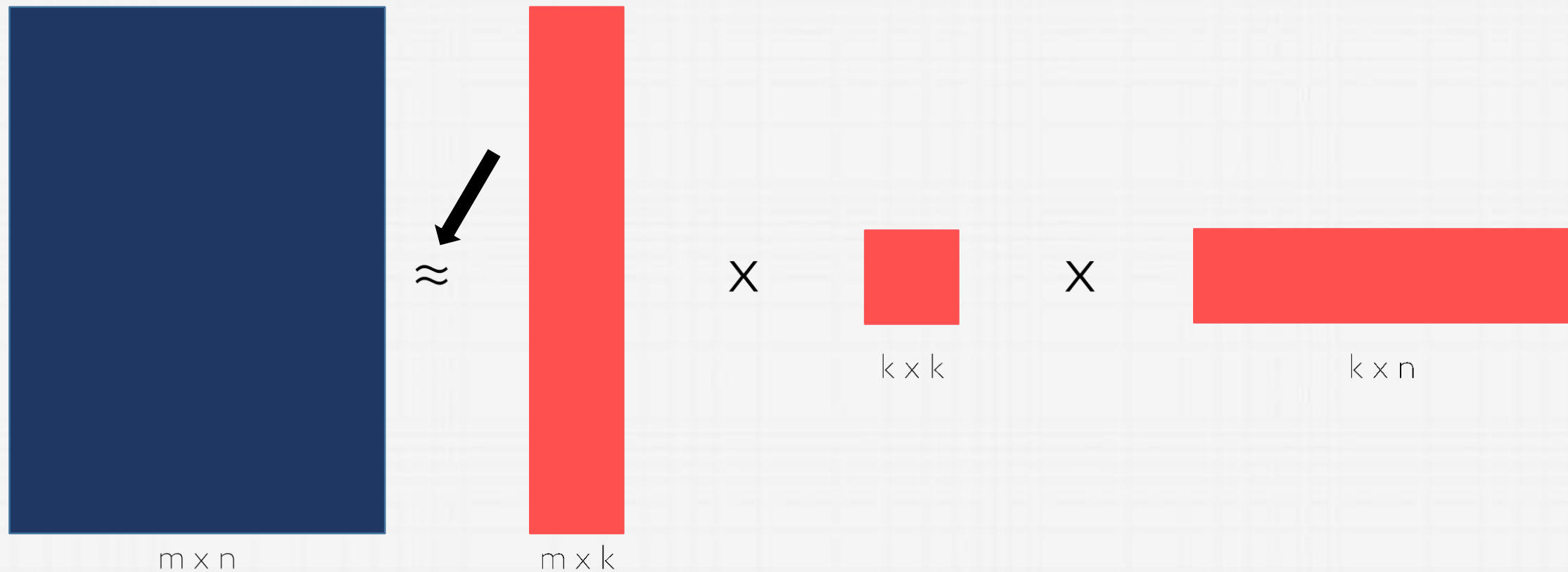
Эти оси модель
может найти
сама!

Свойство SVD



Векторы отсортированы по убыванию вклада в приближение матрицы!

Свойство SVD



Возьмем только часть векторов!

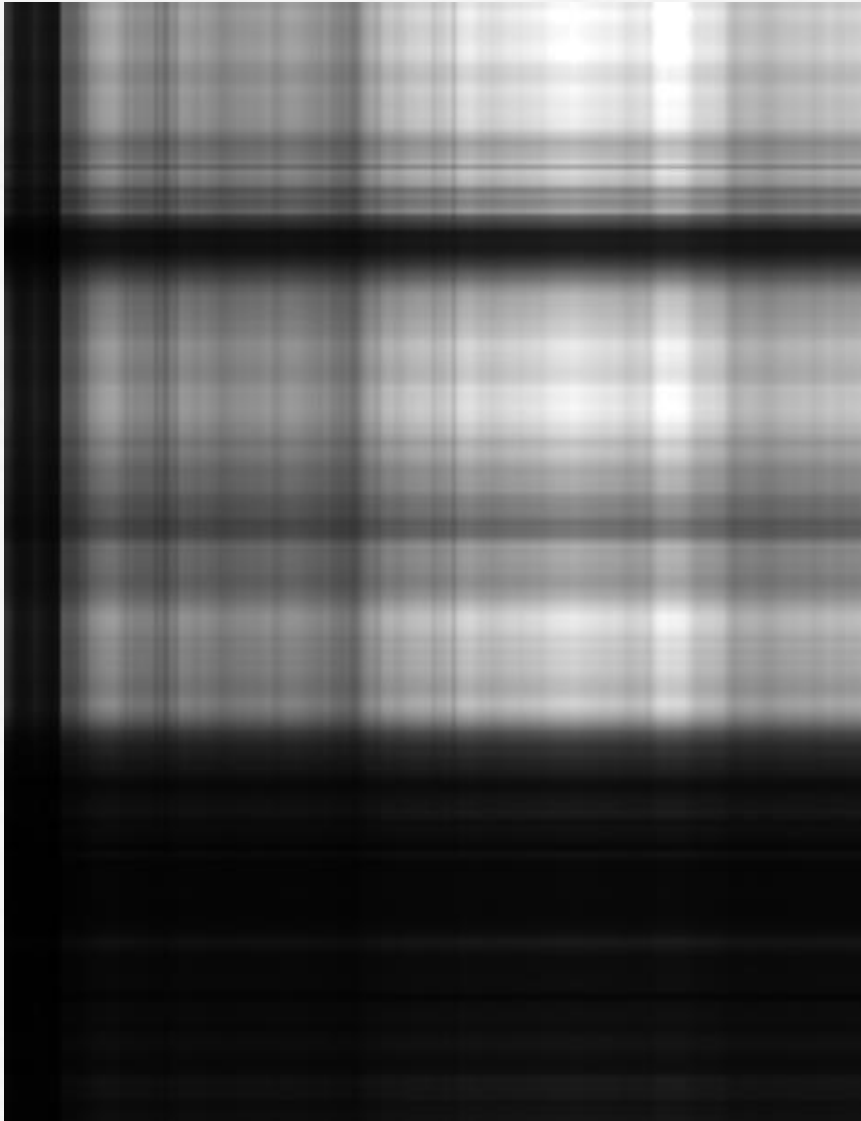
SVD для картинки



Рассмотрим качество приближения для
особенной матрицы – картинки!

Картинка – это набор чисел от 0 до 255
(яркость).

SVD для картинки



Оставили 1 компоненту.

Восстановили наиболее сильные изменения яркости (сверху светло, внизу темно).

SVD для картинки



2 компоненты.

SVD для картинки



10 компонент.

Уже практически виден человек.

SVD для картинки



50 компонент.

Исходная картинка размера 500x600.

Насколько сильно сжали?

SVD для картинки



50 компонент.

Исходная картинка размера 500x600.

Насколько сильно сжали?

$$500*50 + 50*50 + 50*600 = 57500.$$

$$500*600 = 300000.$$

В пять раз!



Пример для рейтингов

| | Yesterday, Beatles | Summertime Sadness, Lana Del Rey | November Rain, Guns 'n Roses | Diamonds, Rihanna | Highway to Hell, AC/DC | What a Wonderful World, Louis Armstrong | Hit the Road Jack!, Ray Charles |
|-------|-----------------------|--|---------------------------------|-------------------|---------------------------|--|------------------------------------|
| Артем | 5 | 2 | 5 | 2 | 5 | 4 | 4 |
| Вася | 4 | 3 | 3 | 3 | 3 | 5 | 5 |
| Маша | 3 | 5 | 2 | 5 | 2 | 4 | 3 |
| Саша | 5 | 2 | 4 | 2 | 5 | 4 | 4 |
| Клара | 5 | 5 | 3 | 4 | 3 | 5 | 5 |



Результат SVD

Профили юзеров:

U

| | f1 | f2 | f3 |
|-------|--------|--------|--------|
| Артем | 0.3262 | 0.5236 | -0.455 |
| Вася | -0.719 | -0.052 | -0.44 |
| Маша | 0.5477 | -0.644 | -0.392 |
| Саша | 0.1567 | 0.4523 | -0.44 |
| Клара | -0.23 | -0.322 | -0.503 |

Важность факторов:

Σ

| | f1 | f2 | f3 |
|----|-------|--------|--------|
| f1 | 22.73 | 0 | 0 |
| f2 | 0 | 5.3211 | 0 |
| f3 | 0 | 0 | 1.7328 |

Профили товаров (по столбцам):

V^T

Интерпретация компонент?

| | Yesterday, Beatles | Summertime Sadness, Lana Del Rey | November Rain, Guns 'n Roses | Diamonds, Rihanna | Highway to Hell, AC/DC | What a Wonderful World, Louis Armstrong | Hit the Road Jack!, Ray Charles |
|----|-----------------------|--|---------------------------------|----------------------|---------------------------|--|---------------------------------------|
| f1 | 0.0205983 | 0.23140511 | 0.29370711 | 0.36389253 | 0.38414103 | -0.35664957 | -0.67274252 |
| f2 | 0.2122569 | -0.57029559 | 0.37908873 | -0.50977121 | 0.46409138 | -0.10192274 | 0.01912943 |
| f3 | -0.43649738 | -0.33353455 | -0.33632132 | -0.31142383 | -0.35567176 | -0.43374969 | -0.41651737 |

Результат SVD

Профили юзеров:

U

| | f1 | f2 | f3 |
|-------|--------|--------|--------|
| Артем | 0.3262 | 0.5236 | -0.455 |
| Вася | -0.719 | -0.052 | -0.44 |
| Маша | 0.5477 | -0.644 | -0.392 |
| Саша | 0.1567 | 0.4523 | -0.44 |
| Клара | -0.23 | -0.322 | -0.503 |

Важность факторов:

Σ

| | f1 | f2 | f3 |
|----|-------|--------|--------|
| f1 | 22.73 | 0 | 0 |
| f2 | 0 | 5.3211 | 0 |
| f3 | 0 | 0 | 1.7328 |

Профили товаров (по столбцам):

V^T

f1 – энергичность

| | Yesterday, Beatles | Summertime Sadness, Lana Del Rey | November Rain, Guns 'n Roses | Diamonds, Rihanna | Highway to Hell, AC/DC | What a Wonderful World, Louis Armstrong | Hit the Road Jack!, Ray Charles |
|----|--------------------|----------------------------------|------------------------------|-------------------|------------------------|---|---------------------------------|
| f1 | 0.0205983 | 0.23140511 | 0.29370711 | 0.36389253 | 0.38414103 | -0.35664957 | -0.67274252 |
| f2 | 0.2122569 | -0.57029559 | 0.37908873 | -0.50977121 | 0.46409138 | -0.10192274 | 0.01912943 |
| f3 | -0.43649738 | -0.33353455 | -0.33632132 | -0.31142383 | -0.35567176 | -0.43374969 | -0.41651737 |



Результат SVD

Профили юзеров:

U

| | f1 | f2 | f3 |
|-------|--------|--------|--------|
| Артем | 0.3262 | 0.5236 | -0.455 |
| Вася | -0.719 | -0.052 | -0.44 |
| Маша | 0.5477 | -0.644 | -0.392 |
| Саша | 0.1567 | 0.4523 | -0.44 |
| Клара | -0.23 | -0.322 | -0.503 |

Важность факторов:

Σ

| | f1 | f2 | f3 |
|----|-------|--------|--------|
| f1 | 22.73 | 0 | 0 |
| f2 | 0 | 5.3211 | 0 |
| f3 | 0 | 0 | 1.7328 |

Профили товаров (по столбцам):

V^T

f2 – рок

| | Yesterday, Beatles | Summertime Sadness, Lana Del Rey | November Rain, Guns 'n Roses | Diamonds, Rihanna | Highway to Hell, AC/DC | What a Wonderful World, Louis Armstrong | Hit the Road Jack!, Ray Charles |
|----|-----------------------|--|---------------------------------|----------------------|---------------------------|--|---------------------------------------|
| f1 | 0.0205983 | 0.23140511 | 0.29370711 | 0.36389253 | 0.38414103 | -0.35664957 | -0.67274252 |
| f2 | 0.2122569 | -0.57029559 | 0.37908873 | -0.50977121 | 0.46409138 | -0.10192274 | 0.01912943 |
| f3 | -0.43649738 | -0.33353455 | -0.33632132 | -0.31142383 | -0.35567176 | -0.43374969 | -0.41651737 |



Результат SVD

Профили юзеров:

U

| | f1 | f2 | f3 |
|-------|--------|--------|--------|
| Артем | 0.3262 | 0.5236 | -0.455 |
| Вася | -0.719 | -0.052 | -0.44 |
| Маша | 0.5477 | -0.644 | -0.392 |
| Саша | 0.1567 | 0.4523 | -0.44 |
| Клара | -0.23 | -0.322 | -0.503 |

Важность факторов:

Σ

| | f1 | f2 | f3 |
|----|-------|--------|--------|
| f1 | 22.73 | 0 | 0 |
| f2 | 0 | 5.3211 | 0 |
| f3 | 0 | 0 | 1.7328 |

Профили товаров (по столбцам):

V^T

f3 – кажется рандом ☺

| | Yesterday, Beatles | Summertime Sadness, Lana Del Rey | November Rain, Guns 'n Roses | Diamonds, Rihanna | Highway to Hell, AC/DC | What a Wonderful World, Louis Armstrong | Hit the Road Jack!, Ray Charles |
|----|--------------------|----------------------------------|------------------------------|-------------------|------------------------|---|---------------------------------|
| f1 | 0.0205983 | 0.23140511 | 0.29370711 | 0.36389253 | 0.38414103 | -0.35664957 | -0.67274252 |
| f2 | 0.2122569 | -0.57029559 | 0.37908873 | -0.50977121 | 0.46409138 | -0.10192274 | 0.01912943 |
| f3 | -0.43649738 | -0.33353455 | -0.33632132 | -0.31142383 | -0.35567176 | -0.43374969 | -0.41651737 |



Хватит 2 компоненты

| | Yesterday, Beatles | Summertime Sadness, Lana Del Rey | November Rain, Guns 'n Roses | Diamonds, Rihanna | Highway to Hell, AC/DC | What a Wonderful World, Louis Armstrong | Hit the Road Jack!, Ray Charles |
|-------|-----------------------|--|---------------------------------|-------------------|---------------------------|--|------------------------------------|
| Артем | 5 | 2 | 5 | 2 | 5 | 4 | 4 |
| Вася | 4 | 3 | 3 | 3 | 3 | 5 | 5 |
| Маша | 3 | 5 | 2 | 5 | 2 | 4 | 3 |
| Саша | 5 | 2 | 4 | 2 | 5 | 4 | 4 |
| Клара | 5 | 5 | 3 | 4 | 3 | 5 | 5 |

| | Yesterday, Beatles | Summertime Sadness, Lana Del Rey | November Rain, Guns 'n Roses | Diamonds, Rihanna | Highway to Hell, AC/DC | What a Wonderful World, Louis Armstrong | Hit the Road Jack!, Ray Charles |
|-------|--------------------|-------------------------------------|---------------------------------|-------------------|---------------------------|--|------------------------------------|
| Артем | 5 | 2 | 5 | 2 | 5 | 4 | 4 |
| Вася | 4 | 3 | 3 | 3 | 3 | 4 | 4 |
| Маша | 3 | 5 | 2 | 5 | 2 | 4 | 4 |
| Саша | 5 | 2 | 4 | 2 | 5 | 4 | 4 |
| Клара | 5 | 5 | 3 | 4 | 3 | 5 | 5 |

Ошибок уже мало!



Заметили подвох?

| | Yesterday, Beatles | Summertime Sadness, Lana Del Rey | November Rain, Guns 'n Roses | Diamonds, Rihanna | Highway to Hell, AC/DC | What a Wonderful World, Louis Armstrong | Hit the Road Jack!, Ray Charles |
|-------|-----------------------|--|---------------------------------|-------------------|---------------------------|--|------------------------------------|
| Артем | 5 | 2 | 5 | 2 | 5 | 4 | 4 |
| Вася | 4 | 3 | 3 | 3 | 3 | 5 | 5 |
| Маша | 3 | 5 | 2 | 5 | 2 | 4 | 3 |
| Саша | 5 | 2 | 4 | 2 | 5 | 4 | 4 |
| Клара | 5 | 5 | 3 | 4 | 3 | 5 | 5 |



Заметили подвох?

| | Yesterday, Beatles | Summertime Sadness, Lana Del Rey | November Rain, Guns 'n Roses | Diamonds, Rihanna | Highway to Hell, AC/DC | What a Wonderful World, Louis Armstrong | Hit the Road Jack!, Ray Charles |
|-------|-----------------------|--|---------------------------------|-------------------|---------------------------|--|------------------------------------|
| Артем | 5 | 2 | 5 | 2 | 5 | 4 | 4 |
| Вася | 4 | 3 | 3 | 3 | 3 | 5 | 5 |
| Маша | 3 | 5 | 2 | 5 | 2 | 4 | 3 |
| Саша | 5 | 2 | 4 | 2 | 5 | 4 | 4 |
| Клара | 5 | 5 | 3 | 4 | 3 | 5 | 5 |

В реальных задачах большинство ячеек пропущены!

Как быть?



Наивный подход

Заполнить пропуски нулями и применить SVD.

Лучше так не делать:

- Вносим шум в данные, ведь на самом деле там не нули!
- Матрица становится плотной (в противовес разреженной) → занимает терабайты, честный SVD считается годами 😊

Распишем рейтинг

вес фичи

$$\hat{r}_{ai} = \sum_{f=1}^k u_{af} \sigma_f v_{if}$$

интерес пользователя
к данной фиче

релевантность данной
фичи товару

Избавимся от матрицы весов

$$\hat{r}_{ai} = \sum_{f=1}^k u_{af} S_f v_{if} = \sum_{f=1}^k \left(u_{af} \sqrt{S_f} \right) \left(\sqrt{S_f} v_{if} \right) = \underline{q_i^T p_a}$$

Для простоты

Добавим сдвиги

$$\hat{r}_{ui} = \mu + b_i + b_u + q_i^T p_u$$

- μ – глобальное среднее по таблице рейтингов
- b_u – индивидуальный эффект пользователя (оптимизм) u
- b_i – индивидуальный эффект товара (качество) i
- q_i – профиль товара (вектор в пространстве фичей)
- p_u – профиль пользователя (вектор в пространстве фичей)
- q_i^T – транспонированный вектор

Задача оптимизации

$$\min_{q^*, p^*, b^*} \sum_{u,i} (r_{ui} - \alpha - b_u - b_i - q_i^T p_u)^2$$

Сумма только по имеющимся рейтингам!

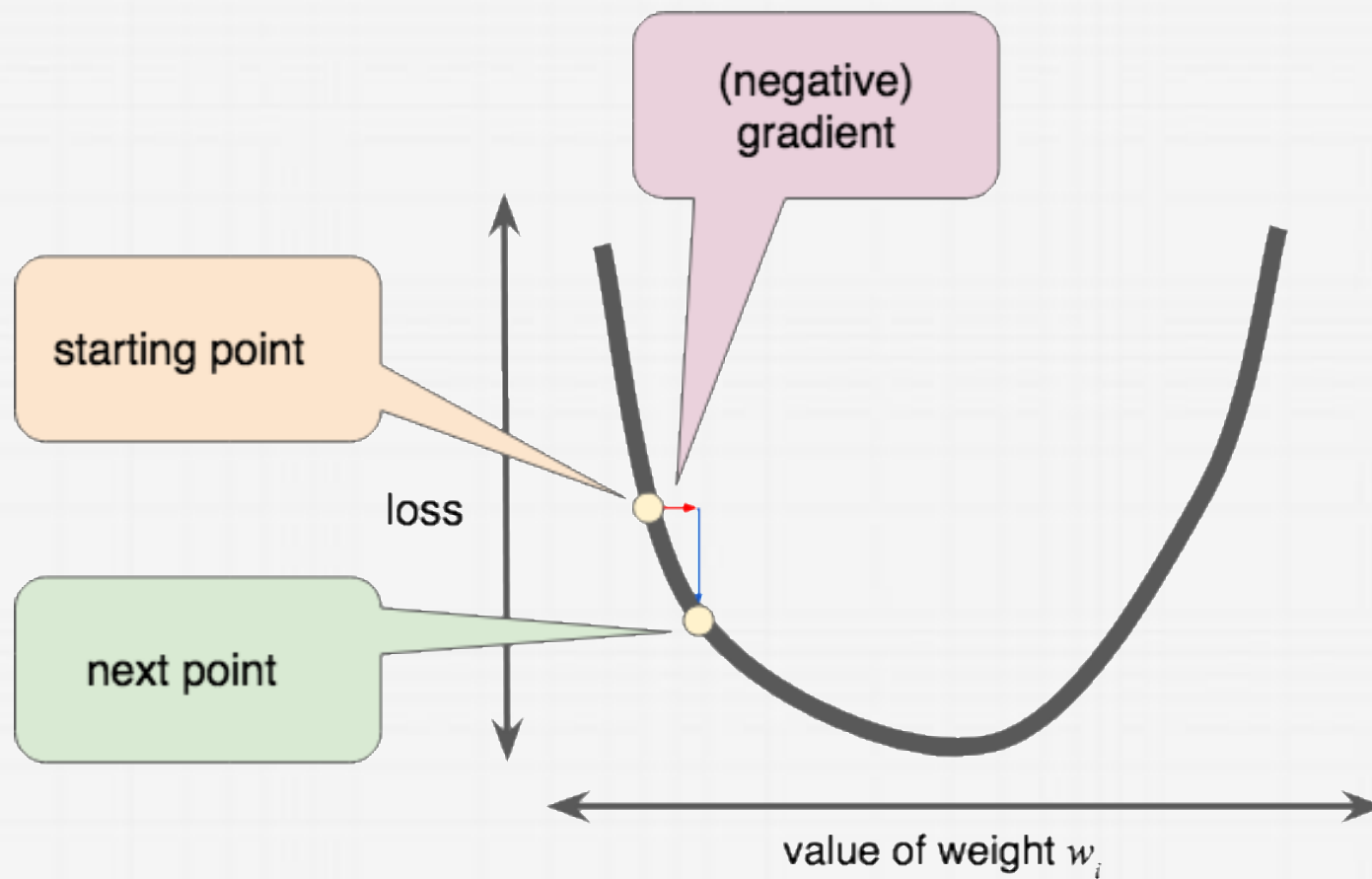
С регуляризацией

$$\min_{q^*, p^*, b^*} \sum_{u,i} (r_{ui} - \alpha - b_u - b_i - q_i^T p_u)^2 + \lambda (b_i^2 + b_u^2 + \|q_i\|^2 + \|p_u\|^2)$$

Сумма только по имеющимся рейтингам!

Как решать?

Градиентный спуск (GD)



Stochastic GD (Funk SVD)

- Начинаем со случайных профилей ($\sim N(0, 0.01)$)
- В цикле:** берем случайный рейтинг, двигаем под него профили товара и юзера

$$q_{if} \leftarrow q_{if} + \gamma (e_{ui} p_{uf} - \lambda q_{if})$$

$$\hat{r}_{ui} = \alpha + b_u + b_i + q_i^T p_u$$

$$p_{uf} \leftarrow p_{uf} + \gamma (e_{ui} q_{if} - \lambda p_{uf})$$

$$e_{ui} = r_{ui} - \hat{r}_{ui}$$

$$b_i \leftarrow b_i + \gamma (e_{ui} - \lambda b_i)$$

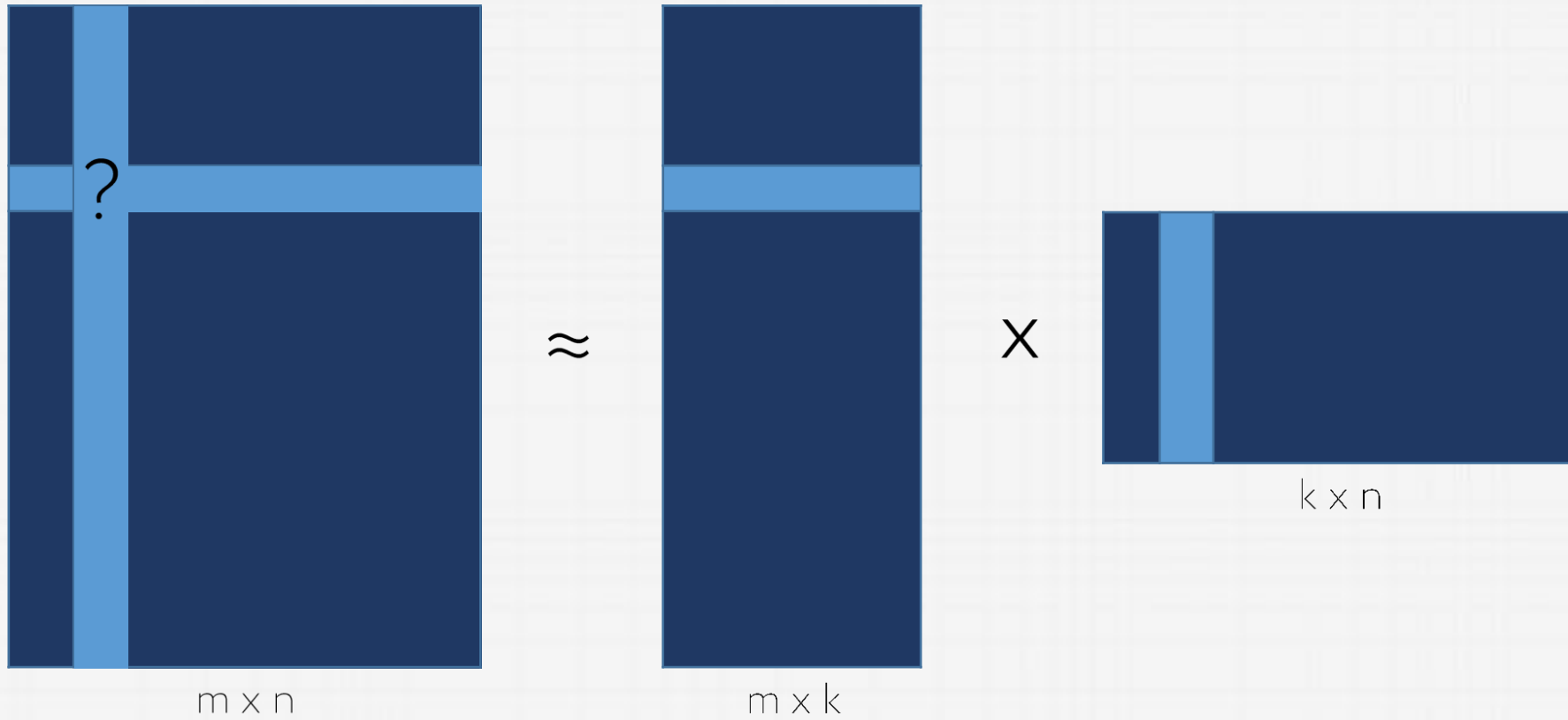
$$b_u \leftarrow b_u + \gamma (e_{ui} - \lambda b_u)$$

Частные производные

- e_{ui} – ошибка
- λ – коэффициент регуляризации
- γ – шаг градиентного спуска

Netflix $\rightarrow \lambda = 0.005, \gamma = 0.02, \gamma \leftarrow 0.9\gamma$

Как делать предсказания





Как обновлять профиль

$$\min_{q^*, p^*, b^*} \sum_{u, i} (r_{ui} - \mu - b_u - b_i - q_i^T p_u)^2$$

Как обновить профиль юзера p_u , если пользователь поставил новую оценку?

Как обновлять профиль

$$\min_{q^*, p^*, b^*} \sum_{u,i} (r_{ui} - \alpha - b_u - b_i - q_i^T p_u)^2$$

Как обновить профиль юзера p_u , если пользователь поставил новую оценку?

Зафиксируем все, кроме его профиля
и решим задачу оптимизации!

Это обычная задача линейной
регрессии!



Резюме: Funk SVD

Работает лучше, чем классическая CF!

Потому что не считаются напрямую попарные похожести (которые могут быть неуверенными)

А ищется такое пространство, в котором все эти неуверенные похожести одновременно объясняются.

Музыкальная пауза

Visualizing high-dimensional data

This allows data matrices of high-dimensionality to be approximated optimally by one of rank 2:



so that the data can be visualized in a two-dimensional space for ease of interpretation



Alternating Least Squares (ALS)

Еще один метод настройки

$$\begin{matrix} m \times n \\ \text{Yellow Box} \end{matrix} = \begin{matrix} m \times k \\ \text{Red Box } P \end{matrix} \times \begin{matrix} k \times n \\ \text{Red Box } Q^t \end{matrix}$$

одновременно находить P и Q сложно - невыпуклая задача

$$\min_{q_*, p_*, b_*} \sum_{u,i} (r_{ui} - q_i^T p_u)^2 + \lambda (\|q_i\|^2 + \|p_u\|^2)$$

$$\min_{q_*, p_*, b_*} \sum_{u,i} (r_{ui} - q_i^T p_u)^2 + \lambda (\|q_i\|^2 + \|p_u\|^2)$$

В цикле:

- Фиксируем матрицу Q
- Находим оптимальную матрицу P (аналитическое решение)
- И наоборот

Для неявных оценок:

$$p_{ui} = \begin{cases} 1 & r_{ui} > 0 \\ 0 & r_{ui} = 0 \end{cases}$$

Был клик или не было

$$c_{ui} = 1 + \alpha r_{ui}$$

Если был, то вес побольше

$$\min_{y_*, y_*} \sum_{u,i} c_{ui} (p_{ui} - x_u^T y_i)^2 + \lambda \left(\sum_u \|x_u\|^2 + \sum_i \|y_i\|^2 \right)$$

Взвешенные потери

Сумма по всем ячейкам!

Как обучать?

Implicit ALS: хитрость

Формулы для ALS шага:

Диагональ с весами оценок юзера

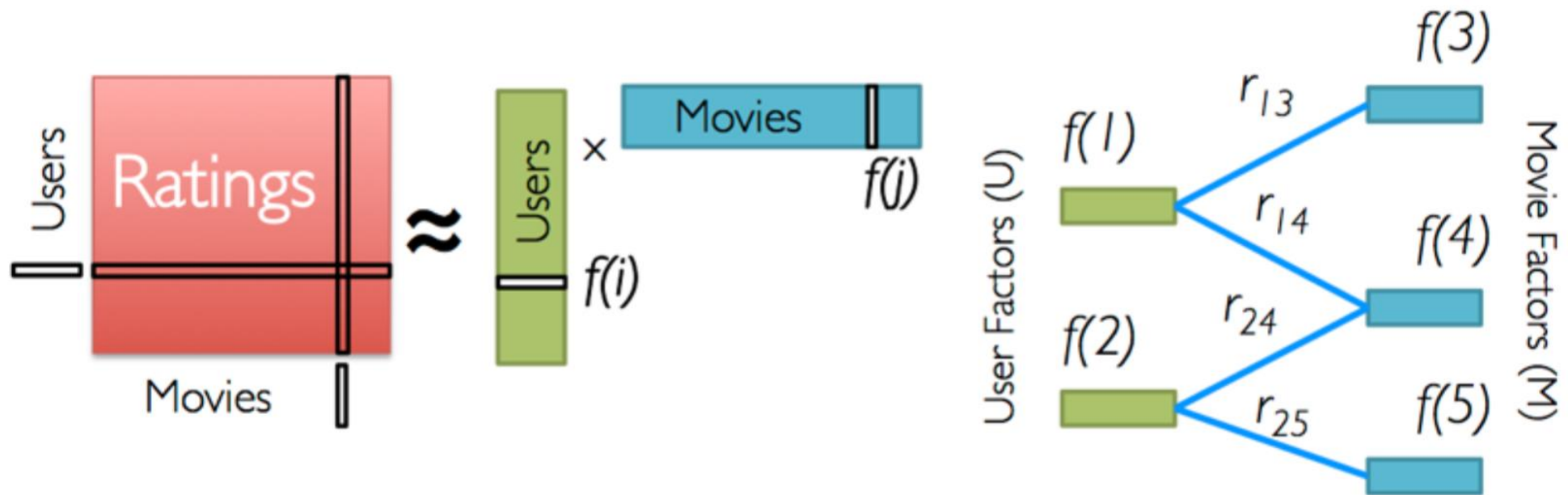
$$x_u = (Y^T C^u Y + \lambda I)^{-1} Y^T C^u p(u)$$

$$Y^T C_u Y = Y^T Y + Y^T (C_u - I) Y$$

Считается один раз!

Не ноль только на кликнутых!

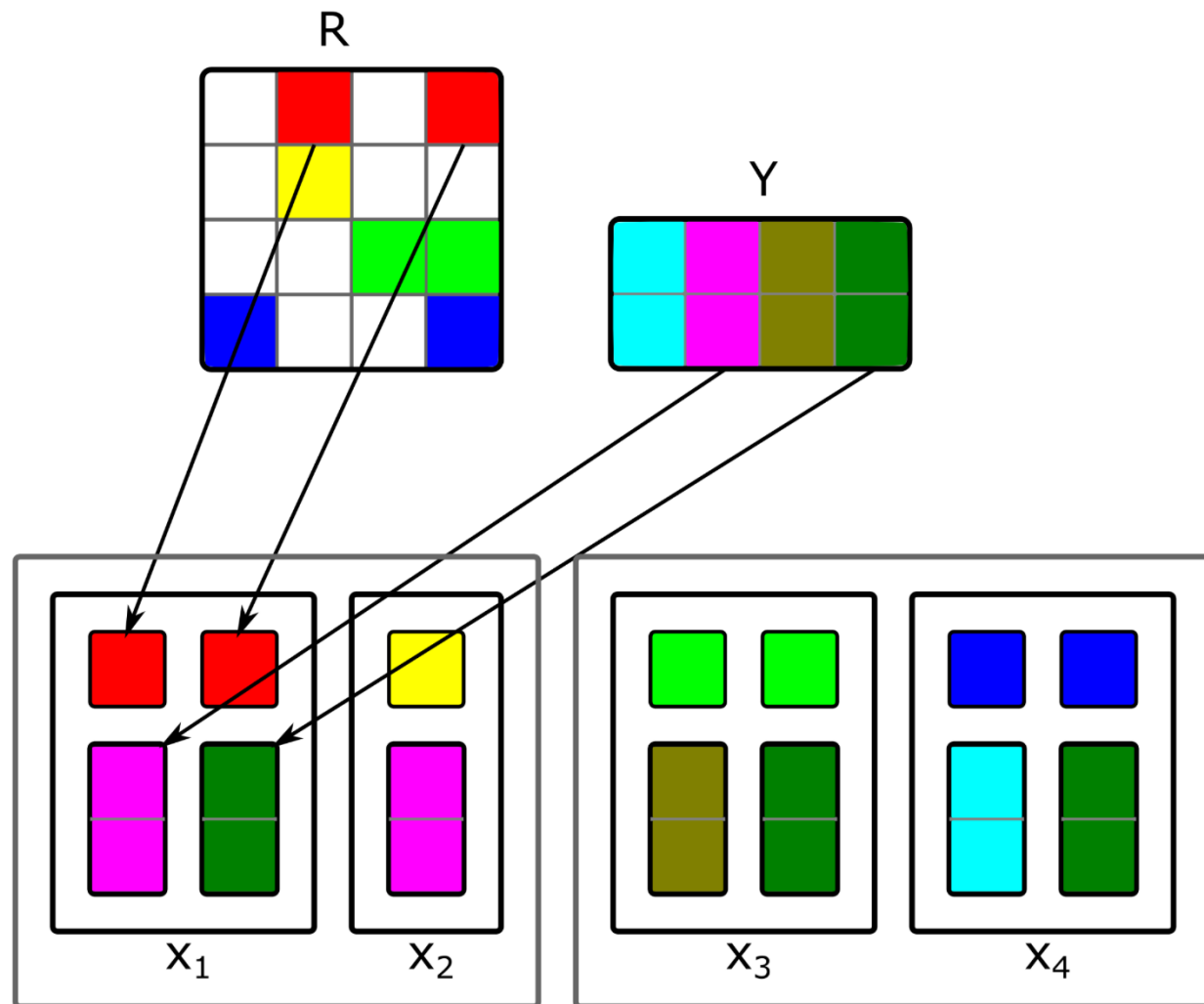
Параллельный ALS



Iterate:

$$f[i] = \arg \min_{w \in \mathbb{R}^d} \sum_{j \in \text{Nbrs}(i)} (r_{ij} - w^T f[j])^2 + \lambda ||w||_2^2$$

Наивный параллельный ALS

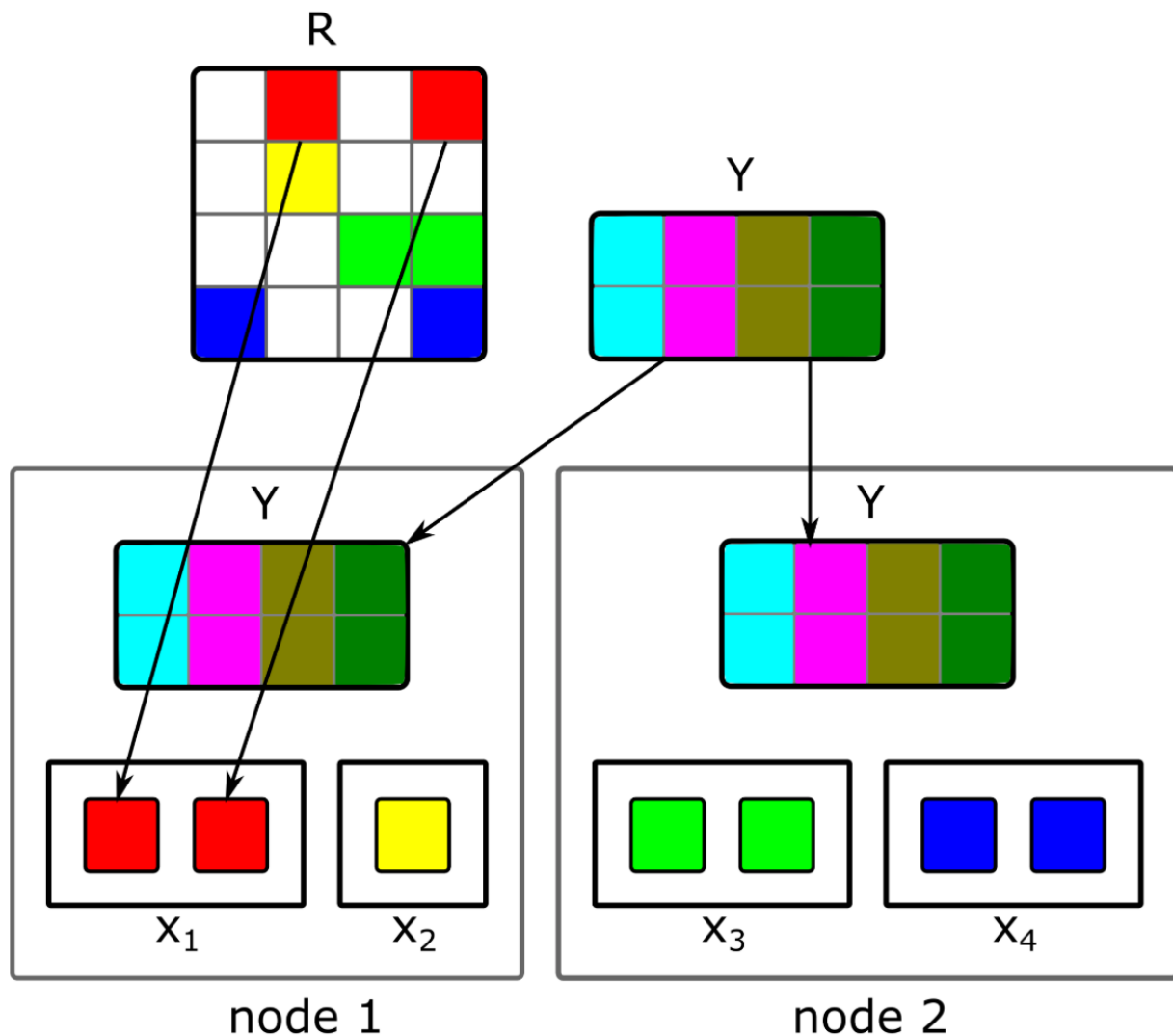


$R = XY$, обновляем X

- Join R и Y по товарам $\rightarrow (u, r_{ui}, y_i)$
- Группируем по u
- На каждой машине пересчитываем x_u

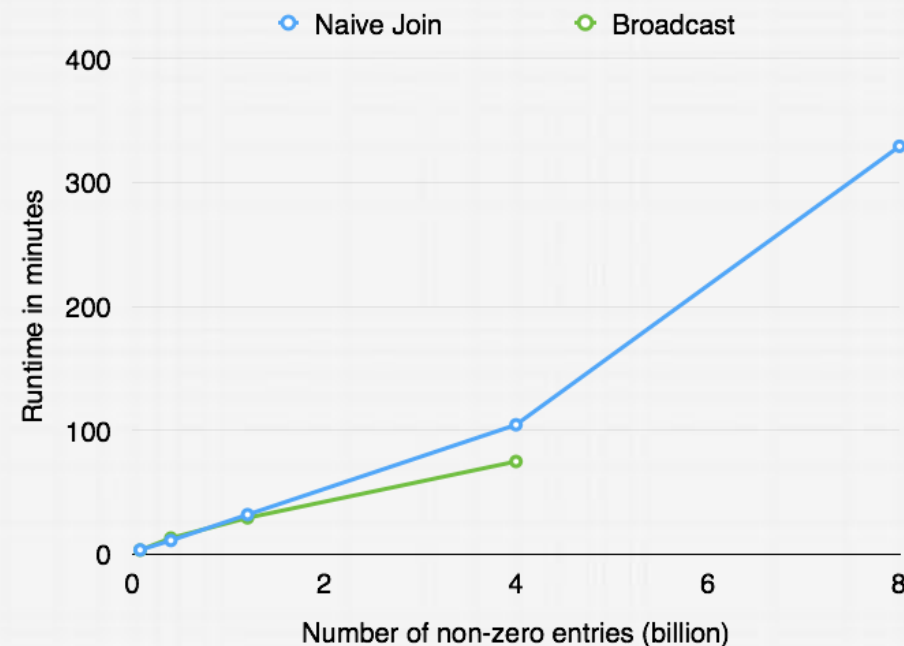
Один и тот же вектор y_i может понадобится нескольким юзерам на машине, и будет отправлен несколько раз!

Может быть спасет broadcast Y?

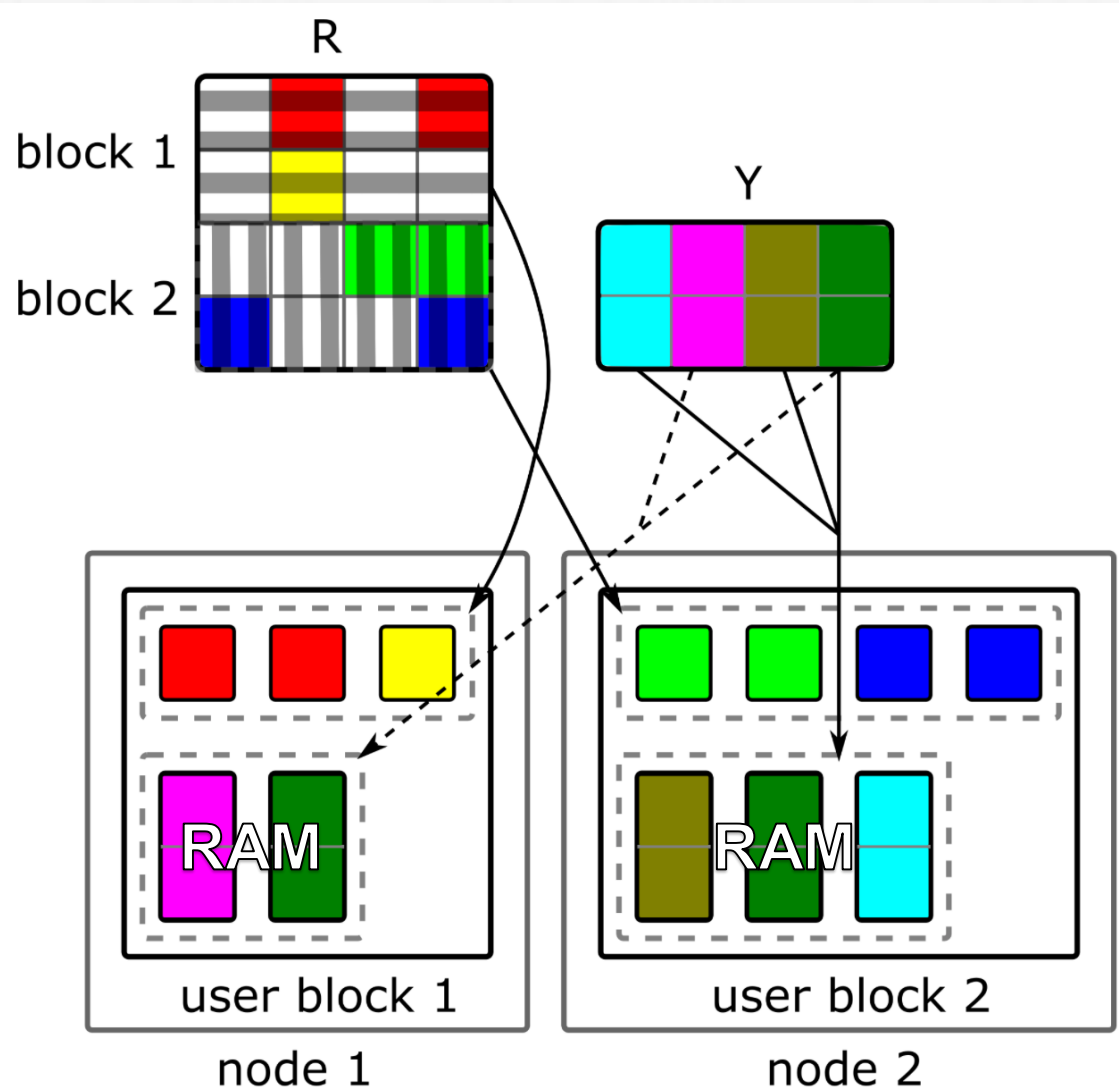


Чуть быстрее!

Имеет предел масштабирования
(Y должен влезать в память)

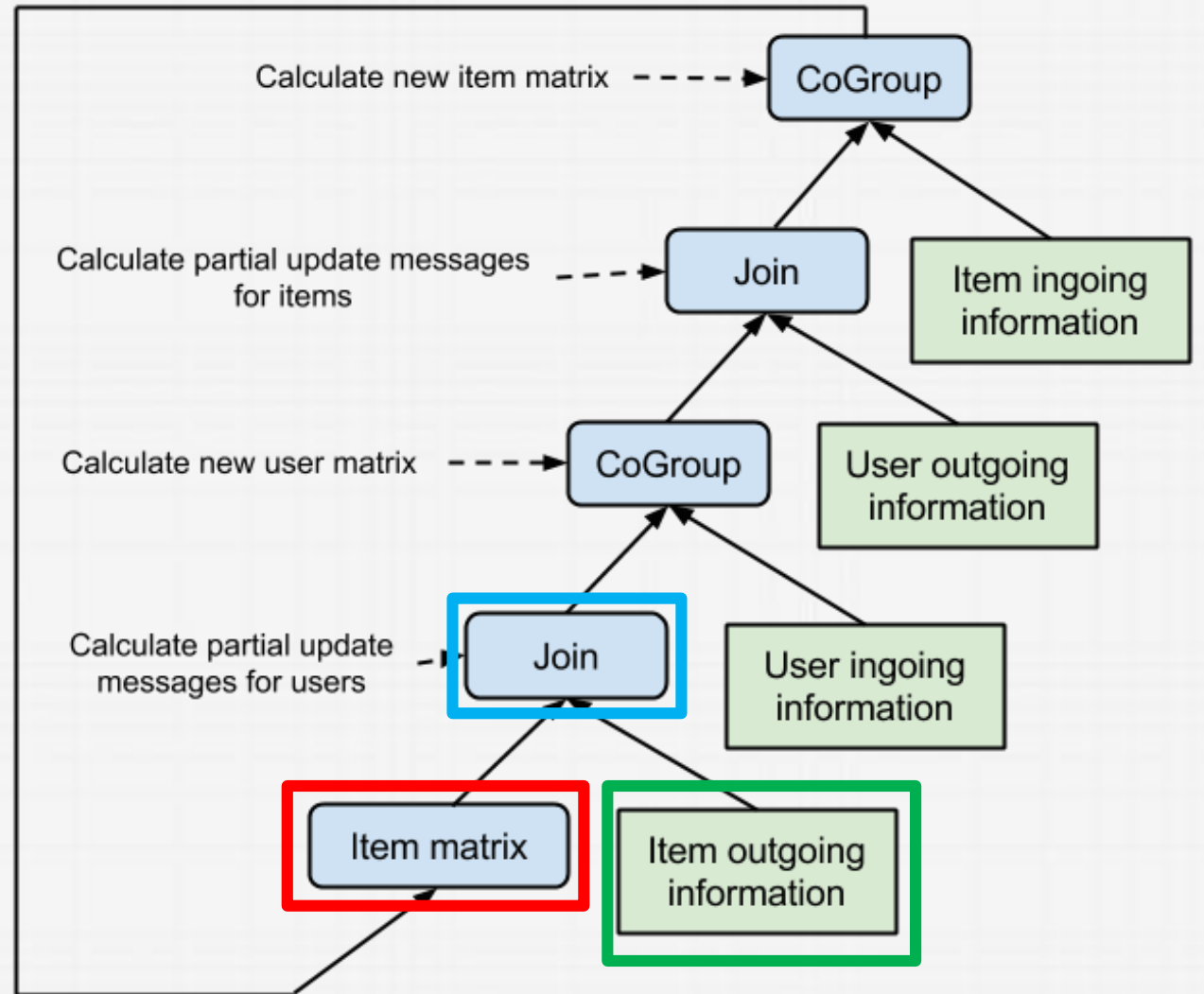
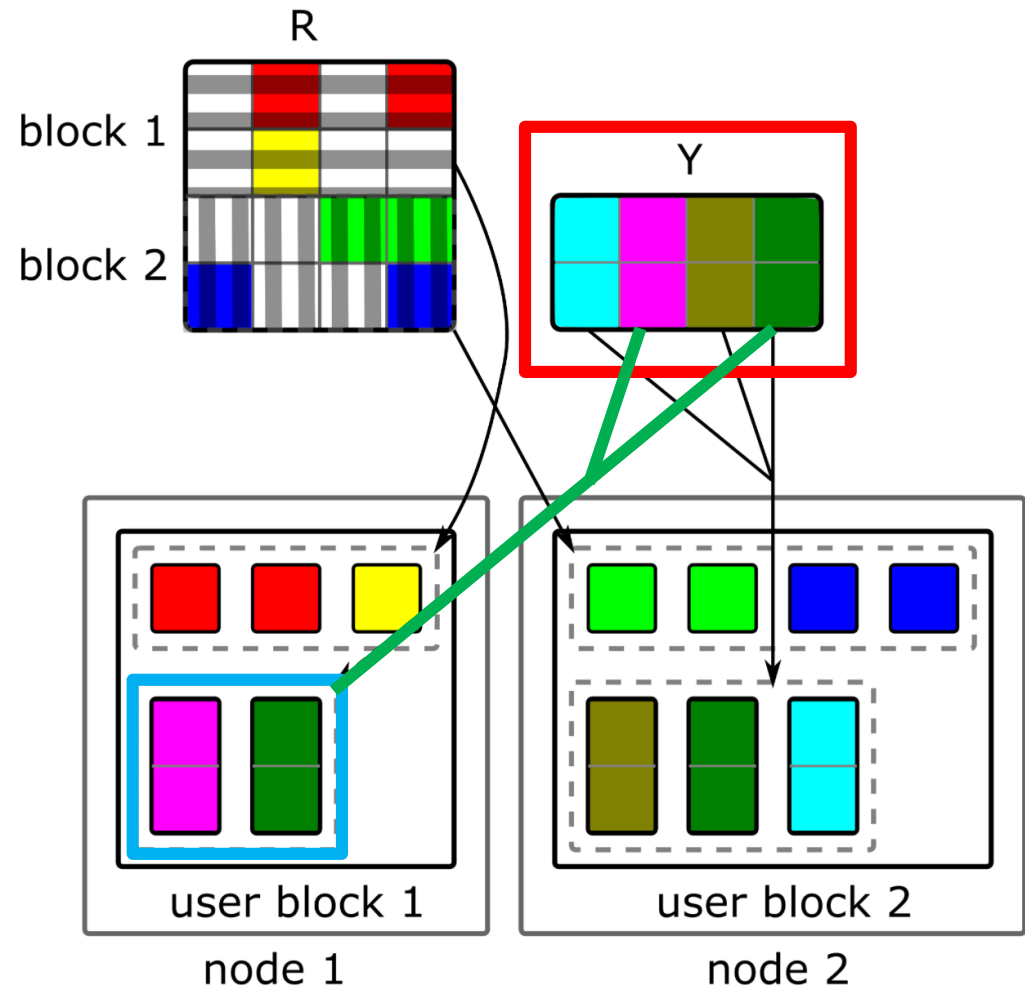


Блочный ALS



- Пользователи и товары бьются на группы, каждая группа обрабатывается на своей машине
- Перед началом итераций вычисляем на какие машины отправлять новые y_i и x_u (хранится в памяти, линейно относительно измерений)
- Пересылаются только нужные части матриц

Блочный ALS



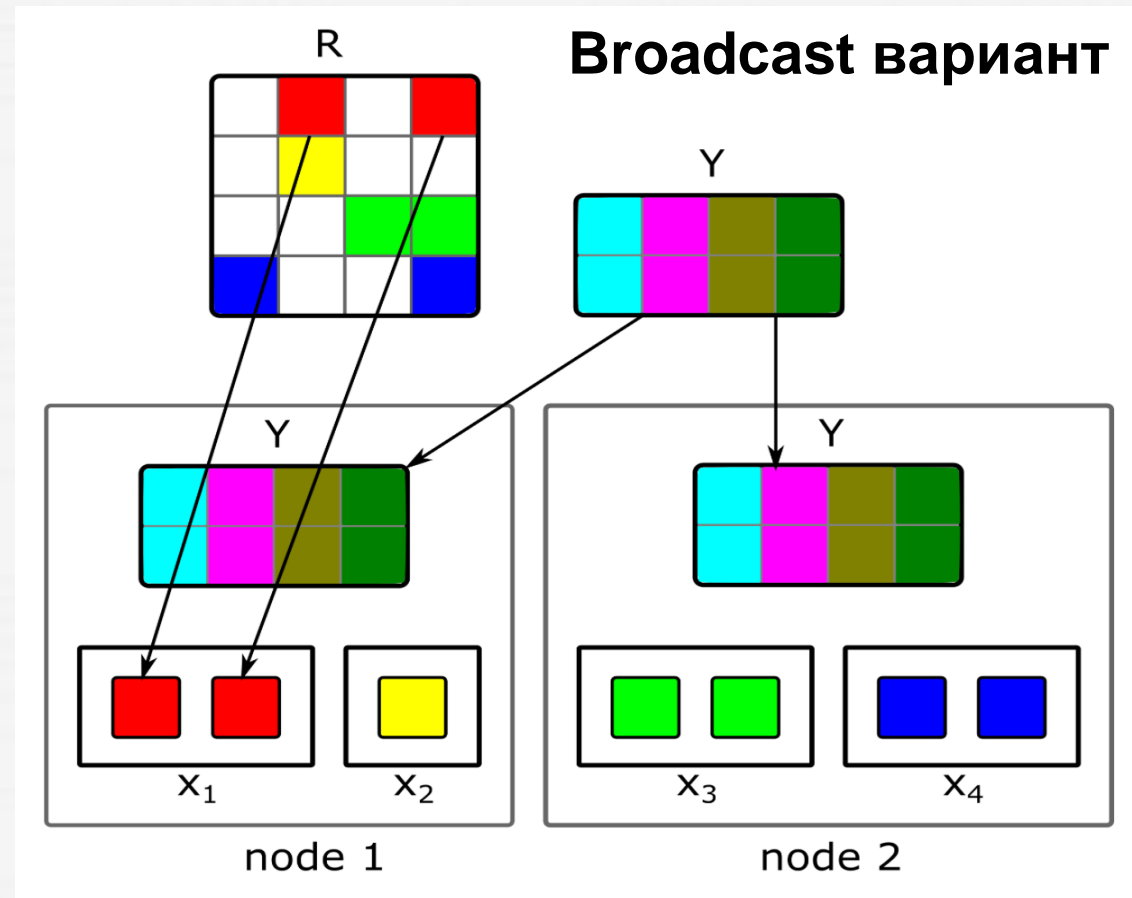
Так реализовано в Spark!

Пример в Spark

```
from pyspark.mllib.recommendation import ALS, Rating
sc.setCheckpointDir("/checkpoints/")
model = ALS.train(train_scores_for_als, 100, 30, seed=42)
```

```
import math
RMSE = math.sqrt(
    model
    .predictAll(test_scores_for_als
                 .map(lambda x: (x.user, x.product)))
    .map(lambda x: ((x.user, x.product), x.rating))
    .join(test_scores_for_als
          .map(lambda x: ((x.user, x.product), x.rating)))
    .map(lambda x: (x[1][0] - x[1][1]) ** 2)
    .mean()
)
print RMSE
```

- 50М пользователей, 5М товаров, **300ГБ** оценок
- Профили размера 100 в памяти занимают **20ГБ** и **2ГБ**
- Разошлем профили на машины и легко сделаем шаг ALS на каждой
- 20 итераций, обучается за день





Для новых товаров

- Накапливаем оценки для новых товаров за последний час
- Профили пользователей в памяти (устарели на день)
- Делаем шаг ALS для обучения профилей новых товаров (в офлайне)



Почему это работает?

- По сути мы учим для каждого товара регрессию в пространстве пользователей (кому понравится, а кому нет)
- Вкусы пользователей меняются не так быстро, устаревание профиля на день не критично (наверное)



Новые пользователи

- Шаг ALS в онлайн для всех пользователей делается всегда!
- Если пользователь кликал в товары с профилем, то у него тоже считается профиль за **20 мс.**

Factorization Machines (FM)

| Feature vector \mathbf{x} | | | | | | | | | | | | | | | | | Target y | | | | | |
|-----------------------------|------|---|---|-----|-------|----|----|----|-----|--------------------|-----|-----|-----|-----|------|------------------|------------|----|----|-----|---|-----------|
| $\mathbf{x}^{(1)}$ | 1 | 0 | 0 | ... | 1 | 0 | 0 | 0 | ... | 0.3 | 0.3 | 0.3 | 0 | ... | 13 | 0 | 0 | 0 | 0 | ... | 5 | $y^{(1)}$ |
| $\mathbf{x}^{(2)}$ | 1 | 0 | 0 | ... | 0 | 1 | 0 | 0 | ... | 0.3 | 0.3 | 0.3 | 0 | ... | 14 | 1 | 0 | 0 | 0 | ... | 3 | $y^{(2)}$ |
| $\mathbf{x}^{(3)}$ | 1 | 0 | 0 | ... | 0 | 0 | 1 | 0 | ... | 0.3 | 0.3 | 0.3 | 0 | ... | 16 | 0 | 1 | 0 | 0 | ... | 1 | $y^{(2)}$ |
| $\mathbf{x}^{(4)}$ | 0 | 1 | 0 | ... | 0 | 0 | 1 | 0 | ... | 0 | 0 | 0.5 | 0.5 | ... | 5 | 0 | 0 | 0 | 0 | ... | 4 | $y^{(3)}$ |
| $\mathbf{x}^{(5)}$ | 0 | 1 | 0 | ... | 0 | 0 | 0 | 1 | ... | 0 | 0 | 0.5 | 0.5 | ... | 8 | 0 | 0 | 1 | 0 | ... | 5 | $y^{(4)}$ |
| $\mathbf{x}^{(6)}$ | 0 | 0 | 1 | ... | 1 | 0 | 0 | 0 | ... | 0.5 | 0 | 0.5 | 0 | ... | 9 | 0 | 0 | 0 | 0 | ... | 1 | $y^{(5)}$ |
| $\mathbf{x}^{(7)}$ | 0 | 0 | 1 | ... | 0 | 0 | 1 | 0 | ... | 0.5 | 0 | 0.5 | 0 | ... | 12 | 1 | 0 | 0 | 0 | ... | 5 | $y^{(6)}$ |
| | A | B | C | ... | TI | NH | SW | ST | ... | TI | NH | SW | ST | ... | Time | TI | NH | SW | ST | ... | | |
| | User | | | | Movie | | | | | Other Movies rated | | | | | | Last Movie rated | | | | | | |

$$\hat{y}(\mathbf{x}) := w_0 + \sum_{i=1}^n w_i x_i + \sum_{i=1}^n \sum_{j=i+1}^n \langle \mathbf{v}_i, \mathbf{v}_j \rangle x_i x_j \quad \rightarrow \text{SGD}$$

Funk SVD как частный случай

| | User | | | | Item | | | |
|----------|----------|----------|----------|----------|----------|----------|----------|----------|
| x^1 | 1 | 0 | 0 | ... | 1 | 0 | 0 | ... |
| x^2 | 1 | 0 | 0 | ... | 0 | 1 | 0 | ... |
| x^3 | 0 | 1 | 0 | ... | 1 | 0 | 0 | ... |
| \vdots | \vdots | \vdots | \vdots | \vdots | \vdots | \vdots | \vdots | \vdots |
| x^n | 0 | 0 | 0 | ... | 0 | 0 | 1 | ... |

$$\hat{y}(\mathbf{x}) := w_0 + \sum_{i=1}^n w_i x_i + \sum_{i=1}^n \sum_{j=i+1}^n \langle \mathbf{v}_i, \mathbf{v}_j \rangle x_i x_j$$



$$\hat{y}(\mathbf{x}) = w_0 + w_u + w_i + \langle \mathbf{v}_u, \mathbf{v}_i \rangle$$

SVD++ как частный случай

$$\hat{r}_{ui} = \mu + b_i + b_u + q_i^T \left(p_u + \frac{1}{\sqrt{|R(u)|}} \sum_{j \in R(u)} y_j \right)$$

- $R(u)$ - множество товаров, по которым пользователь проставил рейтинг
- y_j - дополнительный вектор в пространстве признаков для товара j

Учитывается неявный фидбек – проставил рейтинг.

Пользователь не случайно выбирает чему поставить оценку.

SVD++ как частный случай

| | | | | | | | | | | | | | | |
|--------------------|------|---|---|-----|-------|----|----|----|-----|--------------------|-----|-----|-----|-----|
| $\mathbf{x}^{(1)}$ | 1 | 0 | 0 | ... | 1 | 0 | 0 | 0 | ... | 0.3 | 0.3 | 0.3 | 0 | ... |
| $\mathbf{x}^{(2)}$ | 1 | 0 | 0 | ... | 0 | 1 | 0 | 0 | ... | 0.3 | 0.3 | 0.3 | 0 | ... |
| $\mathbf{x}^{(3)}$ | 1 | 0 | 0 | ... | 0 | 0 | 1 | 0 | ... | 0.3 | 0.3 | 0.3 | 0 | ... |
| $\mathbf{x}^{(4)}$ | 0 | 1 | 0 | ... | 0 | 0 | 1 | 0 | ... | 0 | 0 | 0.5 | 0.5 | ... |
| $\mathbf{x}^{(5)}$ | 0 | 1 | 0 | ... | 0 | 0 | 0 | 1 | ... | 0 | 0 | 0.5 | 0.5 | ... |
| $\mathbf{x}^{(6)}$ | 0 | 0 | 1 | ... | 1 | 0 | 0 | 0 | ... | 0.5 | 0 | 0.5 | 0 | ... |
| $\mathbf{x}^{(7)}$ | 0 | 0 | 1 | ... | 0 | 0 | 1 | 0 | ... | 0.5 | 0 | 0.5 | 0 | ... |
| | A | B | C | ... | TI | NH | SW | ST | ... | TI | NH | SW | ST | ... |
| | User | | | | Movie | | | | | Other Movies rated | | | | |

$$x_j := \begin{cases} 1, & \text{if } j = i \vee j = u \\ \frac{1}{\sqrt{|N_u|}}, & \text{if } j \in N_u \\ 0, & \text{else} \end{cases}$$

$$\hat{y}(\mathbf{x}) = \overbrace{w_0 + w_u + w_i + \langle \mathbf{v}_u, \mathbf{v}_i \rangle}^{\text{SVD++}} + \frac{1}{\sqrt{|N_u|}} \sum_{l \in N_u} \langle \mathbf{v}_i, \mathbf{v}_l \rangle + \frac{1}{\sqrt{|N_u|}} \sum_{l \in N_u} \left(w_l + \langle \mathbf{v}_u, \mathbf{v}_l \rangle + \frac{1}{\sqrt{|N_u|}} \sum_{l' \in N_u, l' > l} \langle \mathbf{v}_l, \mathbf{v}_{l'} \rangle \right)$$



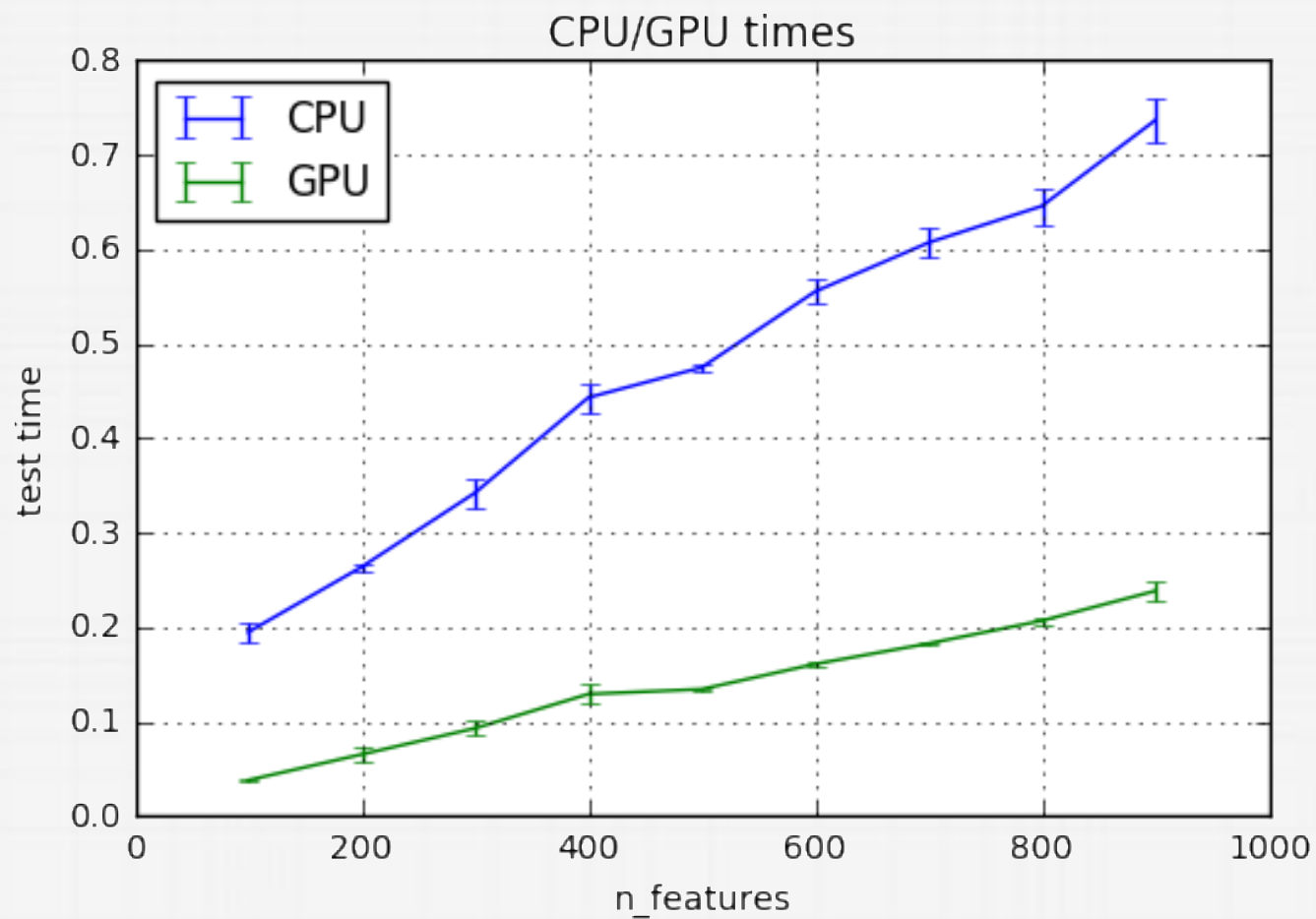
FM универсальны

Можно придумывать новые разновидности факторизаций, меняя только описание признаков.

Каноничная реализация – libFM, <http://www.libfm.org/>

Реализация на TensorFlow (GPU) – tffm, <https://github.com/geffy/tffm>

На GPU быстрее

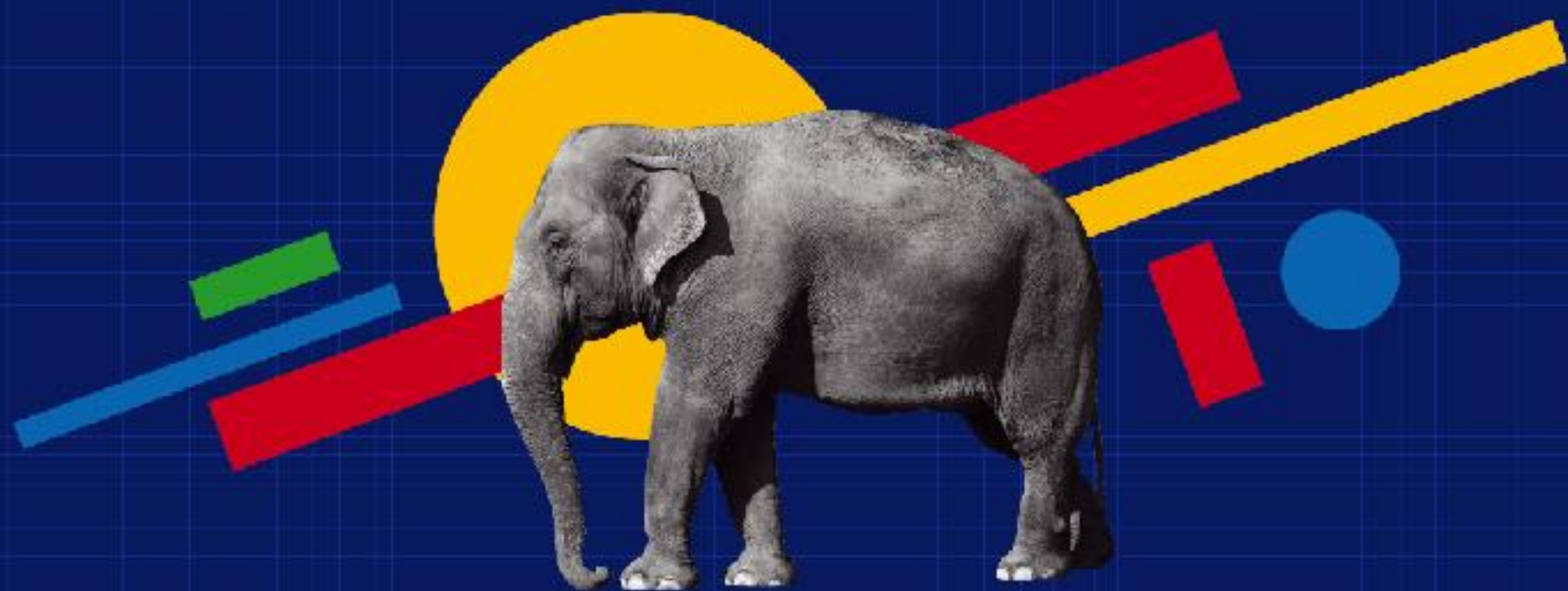


https://github.com/geffy/tffm/blob/master/gpu_benchmark.ipynb

FM: линейное время работы

Можем усложнять
сколько угодно, а
время все равно
линейно!

$$\begin{aligned}
 & \sum_{i=1}^n \sum_{j=i+1}^n \langle \mathbf{v}_i, \mathbf{v}_j \rangle x_i x_j \\
 &= \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \langle \mathbf{v}_i, \mathbf{v}_j \rangle x_i x_j - \frac{1}{2} \sum_{i=1}^n \langle \mathbf{v}_i, \mathbf{v}_i \rangle x_i x_i \\
 &= \frac{1}{2} \left(\sum_{i=1}^n \sum_{j=1}^n \sum_{f=1}^k v_{i,f} v_{j,f} x_i x_j - \sum_{i=1}^n \sum_{f=1}^k v_{i,f} v_{i,f} x_i x_i \right) \\
 &= \frac{1}{2} \sum_{f=1}^k \left(\left(\sum_{i=1}^n v_{i,f} x_i \right) \left(\sum_{j=1}^n v_{j,f} x_j \right) - \sum_{i=1}^n v_{i,f}^2 x_i^2 \right) \\
 &= \frac{1}{2} \sum_{f=1}^k \left(\left(\sum_{i=1}^n v_{i,f} x_i \right)^2 - \sum_{i=1}^n v_{i,f}^2 x_i^2 \right)
 \end{aligned}$$



BIG DATA IS LOVE

NEWPROLAB.COM