Матричные разложения и их применения

Андрей Ткачев

НИУ ВШЭ

5 октября, 2018

План

- Что такое матричные разложения
- Для чего нужны матричные разложения
- Матричные разложения для решения СЛАУ
- Разложение Холецкого
- Факторизация неотрицательных матриц

Что такое матричные разложения?

Это представление матрицы в хорошем виде: произведении матриц обладающих хорошими свойствами.

Пример:

$$\begin{pmatrix} \alpha_{1,1} & \dots & \alpha_{1,n} \\ \vdots & \ddots & \vdots \\ \alpha_{n,1} & \dots & \alpha_{n,n} \end{pmatrix} = \begin{pmatrix} \alpha_{1,1} & \dots & \alpha_{1,n} \\ \vdots & \ddots & \vdots \\ \alpha_{n,1} & \dots & \alpha_{n,n} \end{pmatrix} \times \begin{pmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \end{pmatrix}$$

Для чего нужны матричные разложения

- Данные очень часто представляют из себя матрицы: изображения, таблицы сопряженности.
- Хорошее представление матриц позволяет извлекать из данных полезную информацию.
- Некторые матричные разложения также позволяют вычеслительно эффективно решать системы линейных уравнений.

Матричные разложения в решении СЛАУ

Пусть есть СЛАУ, заданная в матричном виде:

$$Ax = b$$

Где A – хорошая. Чтобы получить решение нужно всего-лишь домножить на обратную к A матрицу: $x = A^{-1}b$. Но что делать если обратной не существует? Тогда существует псевдообратная матрица.

Псевдообратная матрица

Матрица A^+ называется псевдо обратной к матрице A если и только если:

- $A \times A^+ \times A = A$
- $A^+ \times A \times A^+ = A^+$
- $(AA^+)^* \times AA^+$
- $(A^+A)^* \times A^+A$

Согласно теореме Мура — Пенроуза псевдообратная существует и единственна.

Псевдообратная матрица

Если A^+ – псевдообратная то решение $Ax = B$ может быть записано в виде:

$$x = A^+b + [E - A^+A]w$$

Где w – решение $Ax = 0$. При этом $x = A^+b$ дает минимум $||Ax - b||^2$.

Ранговая факторизация

Пусть $A \in Mat_{n\times m}$ и ранга r, тогда ранговой факторизация называется:

$$A = B \times C$$

Где $B \in Mat_{n\times r}$, $C \in Mat_{r\times m}$.
С помощью разложения можно вычислить:

$$A^{+} = C^{T}(CC^{T})^{-1}(B^{T}B)^{-1}B^{T}$$

Ранговая факторизация

Способ получить ранговое разложение на примере:

$$A = \begin{pmatrix} 1 & 3 & 1 & 4 \\ 2 & 7 & 3 & 9 \\ 1 & 5 & 3 & 1 \\ 1 & 2 & 0 & 8 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & -2 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$A = \begin{pmatrix} 1 & 3 & 4 \\ 2 & 7 & 9 \\ 1 & 5 & 1 \\ 1 & 2 & 8 \end{pmatrix} \times \begin{pmatrix} 1 & 0 & -2 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

LU-разложение

Само разложение: $A = LU$, где L – нижне треугольная, U – верхнетреугольная матрица.

Для чего? В таком виде удобно решать СЛАУ $Ax = b$:

- Решаем $Ly = b$ прямой подстановкой.
- Решаем $Ux = y$ обратной подстановкой.

LU-разложение

Так же легко находить обтраные матрицы, решая $Ax = E$ или считать определитель:

$$det(A) = det(L)det(U)$$

LU-разложение

LU-разложение

Найти матрицы L и U можно следующим образом (выполнять шаги следует строго по порядку, так как следующие элементы находятся с использованием предыдущих):

$$u_{1j} = a_{1j},\ j = 1 \dots n$$

$$l_{j1} = \frac{a_{j1}}{u_{11}},\ j = 1 \dots n\ (u_{11} \neq 0)$$

Для $i = 2 \dots n$

$$u_{ij} = a_{ij} - \sum_{k=1}^{i-1} l_{ik} u_{kj},\ j = i \dots n$$

$$l_{ji} = \frac{1}{u_{ii}}(a_{ji} - \sum_{k=1}^{i-1} l_{jk} u_{ki}),\ j = i \dots n$$

Разложение холецкого

Это разложение симметричной положительно определенной матрицы A вида:

$$A = L \times L^T$$

где L — нижняя треугольная матрица со строго положительными элементами на диагонали.

Разложение существует и единственно для любой симметричной положительно определенной матрицы.

Разложение холецкого

По аналогии с LU-разложением, разложение холецкого позволяет удобно решать линейные системы $Ax = b$, через решение:

$$\begin{cases} Ly = b \\ L^T x = y \end{cases}$$

Такой метод более вычислительно стаблиный чем метод Гаусса и LU-разложения.

Разложение холецкого

- Есть вектор из независимых стандартных случайныйх нормальных величин x.
- Хотим получить вектор корреллированных случайных нормальных величин.
- Берем желаемую матрицу Σ ковариации итогового вектора.
- Раскладываем $\Sigma = LL^T$.
- Вектор $y = Lx$ искомый.

Факторизация неотрицательных матриц (NMF)

Она же Non-negative Matrix Factorization (NMF): $V = WH$, где $W \in Mat_{n\times r}$ и $H \in Mat_{r\times m}$ положительные матрицы. Точное разложение – NP задача, поэтому обычно

используют приближенное решение, минимизируя фунцию потерь. В качестве функции потерь используют:

- Дивергенцию Кульбака-Лейблера $D(V, WH)$, $D(A, B) = \sum_{i,j} a_{ij} log(\frac{a_i}{b_j}) - a_i j + b_i j$.
- Норму Фробениуса $||V - WH||_F^2$.

NMF кластреризация

Ценное свойство NMF – кластеризация столбцов.

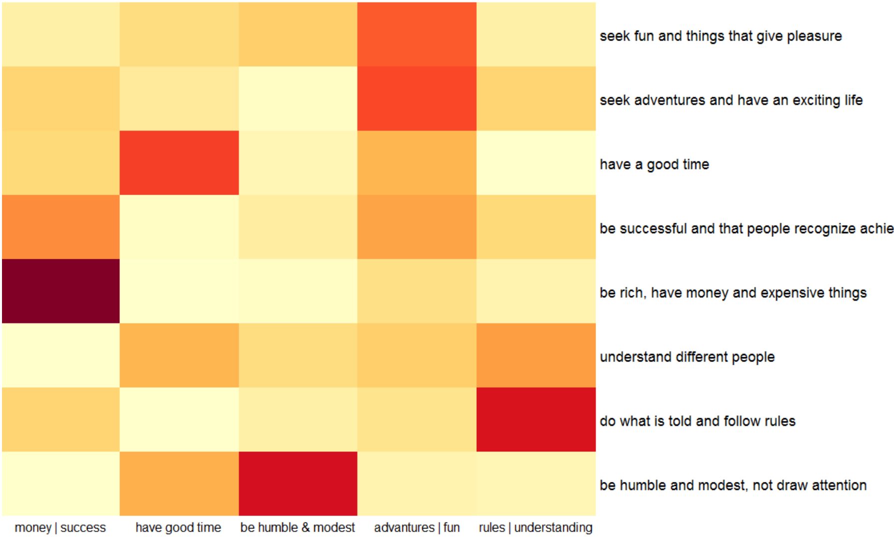
Пример:

- Данных 6 раунда европейского социального исследования (ESS).
- 29 стран.
- Анкета с 21 вопросом с ответаот «Very much like me» до «Not like me at all».
- Составим таблицу частот ответа на вопрос в каждой стране.

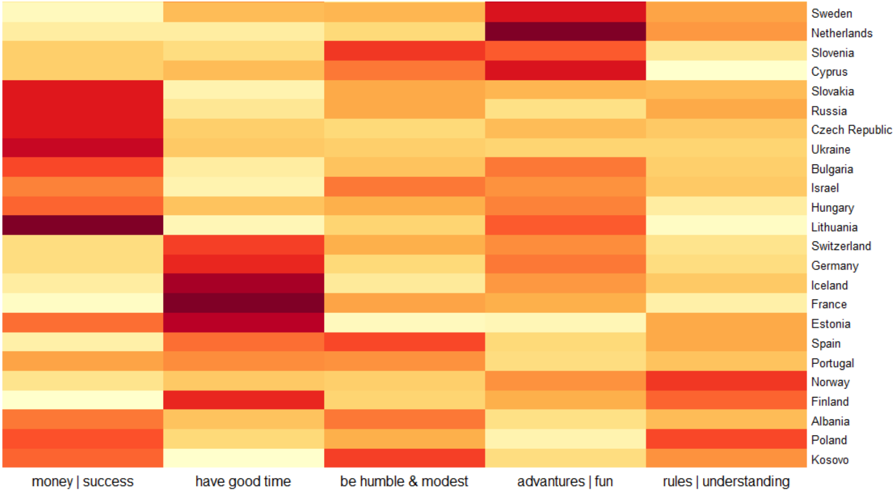
Percentage of human values support in ESS countries

	think new ideas and being creative	be rich, have money and expensive things	people are treated equally and have equal opportunities	show abilities and be admired
Albania	65.4	49.3	80.1	66.1
Belgium	53.5	17.1	73.2	50.9
Bulgaria	50.6	29	61.6	71.2
Switzerland	69.6	11.7	81.9	56.7
Cyprus	82	27.6	89.5	60.3

NMF кластеризация



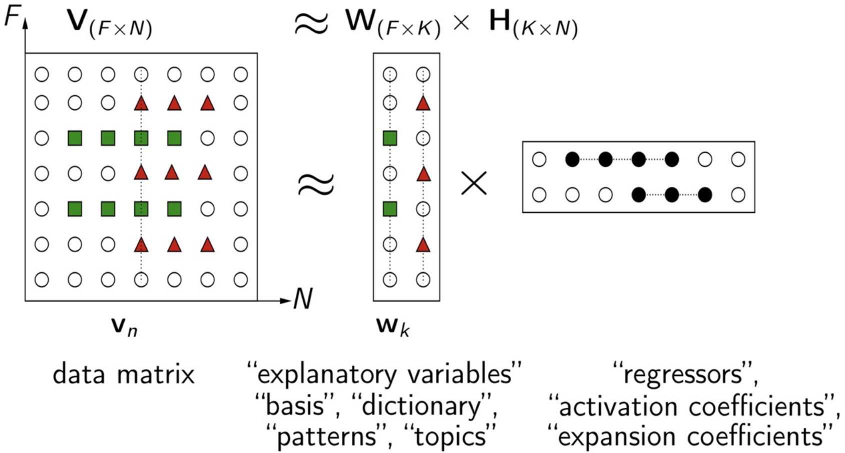
NMF кластеризация



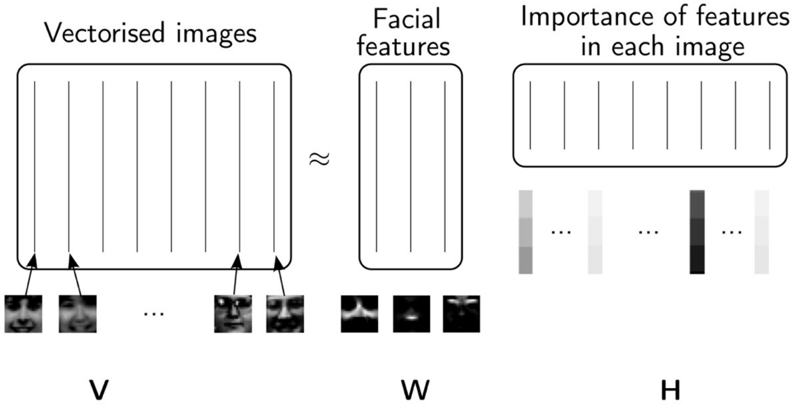
NMF интуиция

Интуитивное объяснение: положительность матриц влечет их разряженность и как следствие происходит выделение кластеров.

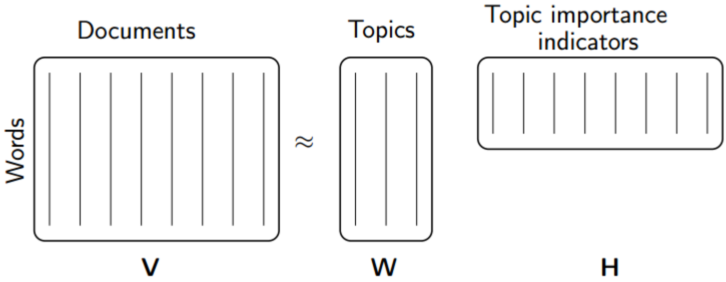
NMF интуиция



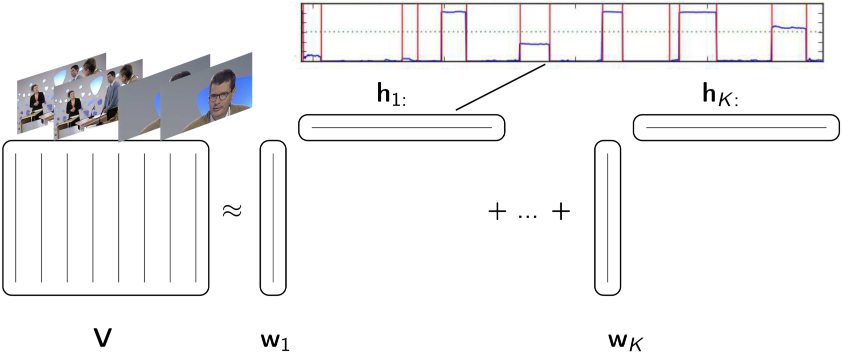
NMF интуиция



NMF анализ текстов



NMF временнтая сегментация



Сравнение NMF и SVD

Отличия от SVD:

- Для NMF не существует алгоритма быстро вычисляющего точное решение.
- Легко интерпретировать результаты из-за встроенной кластеризации и положительности значений. По матрице W можно востановить исходные данные в NMF.

Сравнение NMF и SVD

Evaluation for face recognition:

- **Dataset**: Olivetti faces, 40 classes
- **Classifiers**: LDA (Linear Discriminant Analysis)
- **Cross-validated results**:

	Accuracy
PCA	93%
ICA	93%
NMF	**96%**

Выводы

Матричные разложения бывают полезны:

- Для решения СЛАУ точных и преблизительных
- Для оптимизации численного решения СЛАУ
- Генерации псевдо случайных корреллерированных величин.

Не SVD единым. NMF может пригодиться:

- Когда важна неотрицательная природа данных.
- В анализе таблиц соответствий.
- В извлечении признаков.

Литература и дополнительная информация

- The least-squares method for matrices dependent on parameters
- Topic Modeling with NMF and SVD
- Contingency tables and NMF
- NMF with applications