

Сингулярное разложение и главные компоненты

Обобщение диагонализации

Краткий план:

- Жорданова нормальная форма.

Краткий план:

- Жорданова нормальная форма.
- Сингулярное разложение.

Краткий план:

- Жорданова нормальная форма.
- Сингулярное разложение.
- Доказательство существования.

Не все матрицы диагонализуемы

Утверждение

Квадратная матрица A размера $n \times n$ диагонализуема, если у неё найдётся n линейно независимых собственных векторов.

В этом случае A представима в виде $A = PDP^{-1}$, где P — матрица из собственных векторов, D — диагональная матрица из собственных значений.

Не все матрицы диагонализуемы

Утверждение

Квадратная матрица A размера $n \times n$ диагонализуема, если у неё найдётся n линейно независимых собственных векторов.

В этом случае A представима в виде $A = PDP^{-1}$, где P — матрица из собственных векторов, D — диагональная матрица из собственных значений.

Утверждение

У симметричной матрицы A размера $n \times n$ найдётся n ортогональных собственных векторов единичной длины.

С их помощью матрица A представима в виде

$$A = PDP^T.$$

А если не везёт?

Что делать, если у матрицы A размера $n \times n$ меньше, чем n независимых собственных векторов?

А если не везёт?

Что делать, если у матрицы A размера $n \times n$ меньше, чем n независимых собственных векторов?

Утверждение

Любая квадратная матрица A представима в виде

$$A = PJP^{-1},$$

где **жорданова нормальная форма** J содержит на диагонали **жордановы клетки** J_i :

$$J = \begin{pmatrix} J_1 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & J_2 & \dots & \mathbf{0} \\ \dots & \dots & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \dots & J_k \end{pmatrix}, \quad J_i = \begin{pmatrix} \lambda_i & 1 & \dots & 0 \\ 0 & \lambda_i & \dots & 0 \\ 0 & \dots & \dots & 1 \\ 0 & 0 & \dots & \lambda_i \end{pmatrix}$$

Сингулярное разложение

Утверждение

Любой линейный оператор из \mathbb{R}^k в \mathbb{R}^n представляет собой последовательность действий:

Сингулярное разложение

Утверждение

Любой линейный оператор из \mathbb{R}^k в \mathbb{R}^n представляет собой последовательность действий:

1. Ортогональное преобразование из \mathbb{R}^k в \mathbb{R}^k . Сохраняет длины векторов и углы между ними.

Сингулярное разложение

Утверждение

Любой линейный оператор из \mathbb{R}^k в \mathbb{R}^n представляет собой последовательность действий:

1. Ортогональное преобразование из \mathbb{R}^k в \mathbb{R}^k . Сохраняет длины векторов и углы между ними.
2. Растягивание компонент вектора в \mathbb{R}^k .

Сингулярное разложение

Утверждение

Любой линейный оператор из \mathbb{R}^k в \mathbb{R}^n представляет собой последовательность действий:

1. Ортогональное преобразование из \mathbb{R}^k в \mathbb{R}^k . Сохраняет длины векторов и углы между ними.
2. Растягивание компонент вектора в \mathbb{R}^k .
3. Переход из \mathbb{R}^k в \mathbb{R}^n путём дописывания нулей в вектор, или зачёркивания части координат.

Сингулярное разложение

Утверждение

Любой линейный оператор из \mathbb{R}^k в \mathbb{R}^n представляет собой последовательность действий:

1. Ортогональное преобразование из \mathbb{R}^k в \mathbb{R}^k . Сохраняет длины векторов и углы между ними.
2. Растягивание компонент вектора в \mathbb{R}^k .
3. Переход из \mathbb{R}^k в \mathbb{R}^n путём дописывания нулей в вектор, или зачёркивания части координат.
4. Ортогональное преобразование из \mathbb{R}^n в \mathbb{R}^n . Сохраняет длины векторов и углы между ними.

Сингулярное разложение

Утверждение

Любой линейный оператор из \mathbb{R}^k в \mathbb{R}^n представляет собой последовательность действий:

1. Ортогональное преобразование из \mathbb{R}^k в \mathbb{R}^k . Сохраняет длины векторов и углы между ними.
2. Растягивание компонент вектора в \mathbb{R}^k .
3. Переход из \mathbb{R}^k в \mathbb{R}^n путём дописывания нулей в вектор, или зачёркивания части координат.
4. Ортогональное преобразование из \mathbb{R}^n в \mathbb{R}^n . Сохраняет длины векторов и углы между ними.

Все эти действия мы рассмотрели на первой лекции!

Сингулярное разложение проекции

Оператор $H : \mathbb{R}^n \rightarrow \mathbb{R}^n$ проецирует векторы на линейную оболочку M .

Сингулярное разложение проекции

Оператор $H : \mathbb{R}^n \rightarrow \mathbb{R}^n$ проецирует векторы на линейную оболочку M .

Выберем ортогональный базис $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ в M .

Сингулярное разложение проекции

Оператор $H : \mathbb{R}^n \rightarrow \mathbb{R}^n$ проецирует векторы на линейную оболочку M .

Выберем ортогональный базис $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ в M .

Дополним его до ортогонального базиса в \mathbb{R}^n векторами $\mathbf{v}_{k+1}, \dots, \mathbf{v}_n$.

Сингулярное разложение проекции

Оператор $H : \mathbb{R}^n \rightarrow \mathbb{R}^n$ проецирует векторы на линейную оболочку M .

Выберем ортогональный базис $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ в M .

Дополним его до ортогонального базиса в \mathbb{R}^n векторами $\mathbf{v}_{k+1}, \dots, \mathbf{v}_n$.

1. Повернём-отразим пространство, чтобы $\mathbf{v}_1, \dots, \mathbf{v}_n$ перешли в $\mathbf{e}_1, \dots, \mathbf{e}_n$.

Сингулярное разложение проекции

Оператор $H : \mathbb{R}^n \rightarrow \mathbb{R}^n$ проецирует векторы на линейную оболочку M .

Выберем ортогональный базис $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ в M .

Дополним его до ортогонального базиса в \mathbb{R}^n векторами $\mathbf{v}_{k+1}, \dots, \mathbf{v}_n$.

1. Повернём-отразим пространство, чтобы $\mathbf{v}_1, \dots, \mathbf{v}_n$ перешли в $\mathbf{e}_1, \dots, \mathbf{e}_n$.
2. Домножим первые k компонент вектора на 1, а остальные на 0.

Сингулярное разложение проекции

Оператор $H : \mathbb{R}^n \rightarrow \mathbb{R}^n$ проецирует векторы на линейную оболочку M .

Выберем ортогональный базис $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ в M .

Дополним его до ортогонального базиса в \mathbb{R}^n векторами $\mathbf{v}_{k+1}, \dots, \mathbf{v}_n$.

1. Повернём-отразим пространство, чтобы $\mathbf{v}_1, \dots, \mathbf{v}_n$ перешли в $\mathbf{e}_1, \dots, \mathbf{e}_n$.
2. Домножим первые k компонент вектора на 1, а остальные на 0.
3. Смены размерности нет.

Сингулярное разложение проекции

Оператор $H : \mathbb{R}^n \rightarrow \mathbb{R}^n$ проецирует векторы на линейную оболочку M .

Выберем ортогональный базис $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ в M .

Дополним его до ортогонального базиса в \mathbb{R}^n векторами $\mathbf{v}_{k+1}, \dots, \mathbf{v}_n$.

1. Повернём-отразим пространство, чтобы $\mathbf{v}_1, \dots, \mathbf{v}_n$ перешли в $\mathbf{e}_1, \dots, \mathbf{e}_n$.
2. Домножим первые k компонент вектора на 1, а остальные на 0.
3. Смены размерности нет.
4. Повернём-отразим пространство, чтобы $\mathbf{e}_1, \dots, \mathbf{e}_n$ перешли в $\mathbf{v}_1, \dots, \mathbf{v}_n$.

Сингулярное разложение

Утверждение

Любую матрицу A размера $n \times k$ можно представить в виде

$$A = U\Sigma V^T,$$

где матрица U размера $n \times n$ — ортогональная, $U^T U = I$,

матрица V размера $k \times k$ — ортогональная, $V^T V = I$,

матрица Σ размера $n \times k$ — диагональная.

Сингулярное разложение

Утверждение

Любую матрицу A размера $n \times k$ можно представить в виде

$$A = U\Sigma V^T,$$

где матрица U размера $n \times n$ — ортогональная, $U^T U = I$,

матрица V размера $k \times k$ — ортогональная, $V^T V = I$,

матрица Σ размера $n \times k$ — диагональная.

Данное разложение также называется *SV D-разложением*,
singular value decomposition.

Присмотримся к матрицам

Если $n \geq k$, то SVD -разложение примет вид

$$A = \begin{pmatrix} | & & | \\ \mathbf{u}_1 & \dots & \mathbf{u}_n \\ | & & | \end{pmatrix} \cdot \begin{pmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \sigma_k \\ 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 \end{pmatrix} \begin{pmatrix} - & \mathbf{v}_1^T & - \\ \vdots & & \\ - & \mathbf{v}_k^T & - \end{pmatrix}$$

Зачем нужно SVD -разложение?

Сильно упрощает многие вычисления.

Зачем нужно SVD -разложение?

Сильно упрощает многие вычисления.

Показывает внутренний мир матрицы.

Зачем нужно $SV D$ -разложение?

Сильно упрощает многие вычисления.

Показывает внутренний мир матрицы.

Существует быстрая и устойчивая итеративная процедура нахождения $SV D$ -разложения.

Раскладываем все матрицы!

Сингулярное разложение для A^T :

$$A^T = (U\Sigma V^T)^T = V\Sigma^T U^T$$

Раскладываем все матрицы!

Сингулярное разложение для A^T :

$$A^T = (U\Sigma V^T)^T = V\Sigma^T U^T$$

Сингулярное разложение для $A^T A$:

$$A^T A = V\Sigma^T U^T \cdot U\Sigma V^T = V\Sigma^T \Sigma V^T$$

Раскладываем все матрицы!

Сингулярное разложение для A^T :

$$A^T = (U\Sigma V^T)^T = V\Sigma^T U^T$$

Сингулярное разложение для $A^T A$:

$$A^T A = V\Sigma^T U^T \cdot U\Sigma V^T = V\Sigma^T \Sigma V^T$$

Сингулярное разложение для AA^T :

$$AA^T = U\Sigma V^T \cdot V\Sigma^T U^T = U\Sigma\Sigma^T U^T$$

Существование разложения

Наша задача предъявить разложение $A = U\Sigma V^T$.

Для удобства будем считать, что $n \geq k$.

Существование разложения

Наша задача предъявить разложение $A = U\Sigma V^T$.

Для удобства будем считать, что $n \geq k$.

Доказательство

1. Матрица $A^T A$ является матрицей Грама. Она положительно определена и симметрична. А потому представима в виде $A^T A = V D V^T$.

Существование разложения

Наша задача предъявить разложение $A = U\Sigma V^T$.

Для удобства будем считать, что $n \geq k$.

Доказательство

1. Матрица $A^T A$ является матрицей Грама. Она положительно определена и симметрична. А потому представима в виде $A^T A = V D V^T$.
2. Диагональные элементы D неотрицательны. Поэтому D представима в виде $\Sigma^T \Sigma$.

Существование разложения

Наша задача предъявить разложение $A = U\Sigma V^T$.

Для удобства будем считать, что $n \geq k$.

Доказательство

1. Матрица $A^T A$ является матрицей Грама. Она положительно определена и симметрична. А потому представима в виде $A^T A = V D V^T$.
2. Диагональные элементы D неотрицательны. Поэтому D представима в виде $\Sigma^T \Sigma$.
3. Осталось найти U из целевого разложения:

$$A = U\Sigma V^T$$

Существование разложения

Наша задача предъявить разложение $A = U\Sigma V^T$.

Для удобства будем считать, что $n \geq k$.

Доказательство

1. Матрица $A^T A$ является матрицей Грама. Она положительно определена и симметрична. А потому представима в виде $A^T A = V D V^T$.
2. Диагональные элементы D неотрицательны. Поэтому D представима в виде $\Sigma^T \Sigma$.
3. Осталось найти U из целевого разложения:

$$A = U\Sigma V^T \text{ или } AV = U\Sigma$$

Существование разложения

Уже нашли Σ и V . Осталось найти U из $AV = U\Sigma$.

Существование разложения

Уже нашли Σ и V . Осталось найти U из $AV = U\Sigma$.

Окончание доказательства

4. Находим вектора u_i по очереди:

$$u_1 = Av_1/\sigma_1,$$

Существование разложения

Уже нашли Σ и V . Осталось найти U из $AV = U\Sigma$.

Окончание доказательства

4. Находим вектора u_i по очереди:

$$u_1 = Av_1/\sigma_1, \quad u_2 = Av_2/\sigma_2, \dots$$

Существование разложения

Уже нашли Σ и V . Осталось найти U из $AV = U\Sigma$.

Окончание доказательства

4. Находим вектора u_i по очереди:

$$u_1 = Av_1/\sigma_1, u_2 = Av_2/\sigma_2, \dots$$

5. Вектора v_i кончатся раньше u_i .

Существование разложения

Уже нашли Σ и V . Осталось найти U из $AV = U\Sigma$.

Окончание доказательства

4. Находим вектора u_i по очереди:

$$u_1 = Av_1/\sigma_1, \quad u_2 = Av_2/\sigma_2, \dots$$

5. Вектора v_i кончатся раньше u_i . Оставшиеся u_{k+1}, \dots, u_n выберем произвольными, чтобы U была ортогональной матрицей.

Поиск SVD разложения

Это видеофрагмент с доской, слайдов здесь нет :)

Нахождение проекции при известном SVD

Это видеофрагмент с доской, слайдов здесь нет :)

РСА: максимизация разброса

Краткий план:

- Немного статистики.

Краткий план:

- Немного статистики.
- Метод главных компонент.

Центрирование

Определение

Для вектора чисел $x = (x_1, x_2, \dots, x_n)$ **выборочным средним** называют величину

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}.$$

Центрирование

Определение

Для вектора чисел $x = (x_1, x_2, \dots, x_n)$ **выборочным средним** называют величину

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}.$$

Определение

Центрирование переменной — переход от набора чисел (x_1, x_2, \dots, x_n) к набору чисел $(x_1 - \bar{x}, x_2 - \bar{x}, \dots, x_n - \bar{x})$.

Центрирование

Определение

Для вектора чисел $x = (x_1, x_2, \dots, x_n)$ **выборочным средним** называют величину

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}.$$

Определение

Центрирование переменной — переход от набора чисел (x_1, x_2, \dots, x_n) к набору чисел $(x_1 - \bar{x}, x_2 - \bar{x}, \dots, x_n - \bar{x})$.

Пример. $(4, 2, 3, 3) \rightarrow (1, -1, 0, 0)$.

Свойства центрирования

Утверждение

Если x' — это центрированный вектор x , $y_i = x_i - \bar{x}$, то $\bar{x}' = 0$.

Свойства центрирования

Утверждение

Если x' — это центрированный вектор x , $y_i = x_i - \bar{x}$, то $\bar{x}' = 0$.

$$\sum_{i=1}^n x'_i = \sum_{i=1}^n (x_i - \bar{x}) = \sum_{i=1}^n x_i - n\bar{x} = 0$$

Свойства центрирования

Утверждение

Если x' — это центрированный вектор x , $y_i = x_i - \bar{x}$, то $\bar{x}' = 0$.

$$\sum_{i=1}^n x'_i = \sum_{i=1}^n (x_i - \bar{x}) = \sum_{i=1}^n x_i - n\bar{x} = 0$$

С геометрической точки зрения:

$(\bar{x}, \bar{x}, \dots, \bar{x})$ — проекция вектора x на $\text{Span } \mathbf{v}$, где $\mathbf{v} = (1, 1, \dots, 1)$.

Свойства центрирования

Утверждение

Если x' — это центрированный вектор x , $y_i = x_i - \bar{x}$, то $\bar{x}' = 0$.

$$\sum_{i=1}^n x'_i = \sum_{i=1}^n (x_i - \bar{x}) = \sum_{i=1}^n x_i - n\bar{x} = 0$$

С геометрической точки зрения:

$(\bar{x}, \bar{x}, \dots, \bar{x})$ — проекция вектора x на $\text{Span } \mathbf{v}$, где $\mathbf{v} = (1, 1, \dots, 1)$.

$(x_1 - \bar{x}, x_2 - \bar{x}, \dots, x_n - \bar{x})$ — проекция вектора x на $\text{Span}^\perp \mathbf{v}$.

Выборочная дисперсия

Определение

Выборочной дисперсией набора чисел (x_1, x_2, \dots, x_n) называют величину $\|\mathbf{x}'\|^2 / (n - 1)$, где вектор \mathbf{x}' — это центрированный вектор \mathbf{x} .

Выборочная дисперсия

Определение

Выборочной дисперсией набора чисел (x_1, x_2, \dots, x_n) называют величину $\|\mathbf{x}'\|^2 / (n - 1)$, где вектор \mathbf{x}' — это центрированный вектор \mathbf{x} .

Выборочная дисперсия вектора показывает «разброс» x_i , насколько далеки x_i от своего среднего \bar{x} .

Выборочная корреляция

Определение

Выборочной корреляцией двух наборов чисел x и y называют косинус угла между их центрированными версиями:

$$\rho(\mathbf{x}, \mathbf{y}) = \cos \angle(\mathbf{x}', \mathbf{y}') = \frac{\langle \mathbf{x}', \mathbf{y}' \rangle}{\|\mathbf{x}'\| \|\mathbf{y}'\|},$$

где $x'_i = x_i - \bar{x}$ и $y'_i = y_i - \bar{y}$.

Выборочная корреляция

Определение

Выборочной корреляцией двух наборов чисел x и y называют косинус угла между их центрированными версиями:

$$\rho(\mathbf{x}, \mathbf{y}) = \cos \angle(\mathbf{x}', \mathbf{y}') = \frac{\langle \mathbf{x}', \mathbf{y}' \rangle}{\|\mathbf{x}'\| \|\mathbf{y}'\|},$$

где $x'_i = x_i - \bar{x}$ и $y'_i = y_i - \bar{y}$.

Если после центрирования вектора ортогональны, то выборочная корреляция равна 0.

Выборочная корреляция

Определение

Выборочной корреляцией двух наборов чисел x и y называют косинус угла между их центрированными версиями:

$$\rho(\mathbf{x}, \mathbf{y}) = \cos \angle(\mathbf{x}', \mathbf{y}') = \frac{\langle \mathbf{x}', \mathbf{y}' \rangle}{\|\mathbf{x}'\| \|\mathbf{y}'\|},$$

где $x'_i = x_i - \bar{x}$ и $y'_i = y_i - \bar{y}$.

Если после центрирования вектора ортогональны, то выборочная корреляция равна 0.

Выборочная корреляция лежит в диапазоне от -1 до 1 .

Скринкаст: *SVD* для снижения размерности

Это видеофрагмент с доской, слайдов здесь нет :)

Бонус: геометрическая алгебра

Это видеофрагмент с доской, слайдов здесь нет :)