Линейные отображения векторных пространств. Матричные разложения.

Детерминант матрицы

Определение. *Детерминантом* или *определителем* квадратной матрицы A размера n называется число $\det A = \sum_{j=1}^{n} a_{ij}(-1)^{i+j} d_{ij}$, где d_{ij} — детерминант матрицы, полученной вычеркиванием из матрицы A i-й строки и j-го столбца.

Утверждение 1. ***(а)*** $\det A = \det A^T$.

(б) *Если столбцы матрицы линейно зависимы, то детерминант матрицы равен 0.*

(в) *Если переставить местами два столбца, то детерминант умножится на (-1).*

(г) *Если прибавить к одному столбцу другой, умноженный на некоторое число, то детерминант не изменится.*

(д) $\det AB = \det A \det B$.

Задача 1

Выразите $\det \alpha A$ через $\det A$.

Задача 2

Вычислите $\begin{vmatrix} 1 & 0 & 8 & 1 \\ 1 & 4 & 0 & 3 \\ 2 & 0 & 9 & 3 \\ 1 & 5 & 4 & 1 \end{vmatrix}$

Связь с системой линейных уравнений

Систему линейных уравнений

$$
\begin{aligned}
a_1^1x_1 + \ldots + a_n^1x_n &= b_1 \\
a_1^2x_1 + \ldots + a_n^2x_xn &= b_2 \\
&\ldots \\
a_1^mx_1 + \ldots + a_n^mx_n &= b_m
\end{aligned}
$$

можно переписать в виде $Ax = b$, где A — матрица из коэффициентов при неизвестных, x – столбец из неизвестных, а b — столбец из чисел в правых частях уравнений.

Линейные отображения

Определение. *Отображением* A линейного пространства L в линейное пространство $\overline{L}$ называется закон, по которому каждому вектору из L сопоставлен ровно один вектор из $\overline{L}$. Будем писать $A\colon L \to \overline{L}$.

Определение. Отображение $A\colon L \to \overline{L}$ называется *линейным*, если для любых векторов x, y из L и любого числа α выполнены равенства $A(x) + A(y) = A(x + y)$ и $A(\alpha x) = \alpha A(x)$.

Определение. Если пространства L и $\overline{L}$ совпадают, то линейное отображение называется *линейным преобразованием.*

Утверждение 2. *При линейном отображении*

(а) *линейная комбинация векторов переходит в такую же линейную комбинацию своих образов;*

(б) *нулевой вектор переходит в нулевой;*

(в) *линейно зависимые векторы переходят в в линейно зависимые;*

(г) *линейное подпространство переходит в линейное подпространство, не большей размерности.*

Определение. Множество образов всех векторов из L является линейным подпространством $A(L)$ в $\overline{L}$ и называется *множеством значений* отображения A. Оно обозначается $\mathrm{Im}A$.

Определение. Размерность множества значений отображения называется *рангом* отображения.

Определение. Множество векторов, отображающихся в нулевой вектор называется *ядром* отображения и обозначается $\mathrm{Ker}A$.

Матрица линейного отображения

Определение. *Матрицей линейного отображения* $A\colon L \to \overline{L}$ в паре базисов e и f называется матрица, столбцы которой — координатные столбцы векторов $A(e_1), \ldots, A(e_n)$ в базисе f.

Утверждение 3. *Верны следующие утверждения*

(а) *Ранг матрицы линейного отображения равен рангу этого отображения.*

(б) *Сумма ранга отображения и размерности его ядра равна размерности отображаемого пространства.*

Определение. *Матрицей линейного преобразования* A в базисе e называется матрица, столбцы которой — координатные столбцы векторов $A(e_1), \ldots, A(e_n)$ в базисе e.

Задача 3

Найдите матрицу линейного преобразования, соответствующего повороту плоскости на угол α против часовой стрелки вокруг точки $(0, 0)$.

Собственные подпространства

Определение. Если для числа λ подпространство $\mathrm{Ker}(A - \lambda E)$ ненулевое, то λ называется *собственным значением* преобразования A, а подпространство — *собственным подпространством*, соответствующим собственному значению λ.

Утверждение 4. *Ограничение преобразования на собственном подпространстве является или нулевым преобразованием, или гомотетией: оно умножает каждый вектор этого подпространства на собственное значение.*

Определение. Ненулевой вектор x называется *собственным вектором* преобразования A, соответствующим собственному значению λ, если $A(x) = \lambda x$.

Определение. Равенство $\det(A - \lambda E) = 0$ называется *характеристическим уравнением* матрицы A. Его корни называются *характеристическими числами* матрицы A.

Теорема 1. *Все корни характеристического уравнения и только они являются собственными значениями.*

Теорема 2. *Пусть собственное значение λ_0 преобразования A — корень характеристического уравнения кратности s. Тогда размерность соответствующего собственного подпространства не превосходит s.*

Задача 4

Докажите, что ненулевое линейное преобразование, для которого все ненулевые векторы собственные, является гомотетией.

Задача 5

Найдите собственные значения матрицы $\begin{pmatrix} 3 & -2 & 6 \\ -2 & 6 & 3 \\ 6 & 3 & -2 \end{pmatrix}$

Разложение по сингулярному значению (SVD-разложение)

Определение. Матрица A называется *ортогональной*, если $AA^T = E$.

Теорема 3. *Для любой матрицы A существуют две ортогональные матрицы U и V такие, что U^TAV — диагональная матрица Σ.*

Матрицы U и V можно выбрать так, чтобы диагональные элементы матрицы σ удовлетворяли условию $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r > \sigma_{r+1} = \ldots = \sigma_n = 0$, где r — ранг матрицы A.

Определение. Столбцы матриц U и V называются *левыми* и *правыми сингулярными векторами*, а числа $\sigma_1, \ldots, \sigma_n$ — *сингулярными* числами матрицы A.

Определение. Запись из теоремы можно переписать в виде $A = U\Sigma V^T$. Такое разложение матрицы A называют *разложением по сингулярному значению* или *SVD-разложением.*

Сингулярные числа связаны с собственными числами. С одной стороны,

$$AA^T = U\Sigma V^T V\Sigma U^T = U\Sigma^2 U^T$$

С другой стороны,

$$A^T A = V\Sigma U^T U\Sigma V^T = V\Sigma^2 V^T$$

Тогда столбцы матрицы U — собственные векторы матрицы AA^T, матрицы V — матрицы $A^T A$, а квадраты сингулярных чисел — собственными числами этих матриц.

Рекомендательные системы

У нас есть матрица, состоящая из рейтингов (лайков, фактов покупки и т.п.), которые пользователи (строки матрицы) присвоили продуктам (столбцы матрицы). Как правило, такие матрицы разрежены, потому что лишь незначительная доля продуктов оценена большим количеством пользователей.

Тогда можно представить каждого пользователя вектором из r факторов u_i. Точно также каждый продукт будет представлен вектором из r факторов v_j. Тогда, чтобы рассчитать рейтинг пользователя для некоторого продукта достаточно взять скалярное произведение векторов $(u_i, v_j) = u_i^T v_j$.

Метод главных компонент (PCA)

Пусть у нас есть матрица «объект-признак» $X \in \mathbb{R}^{n\times k}$, то есть есть n объектов и значения их k признаков. Наша цель — уменьшить размерность пространства до d. Считаем, что данные центрированы, то есть среднее в каждом столбце равно 0.

Будем искать главные компоненты $u_1, \ldots, u_d \in \mathbb{R}^k$, которые удовлетворяют следующим свойствам:

• Они ортогональны, то есть $(u_i, u_j) = 0$ для любых $i \neq j$

• Они нормированы, то есть $|u_i| = 1$

• При проецировании выборки на компоненты $u_1, \ldots, u_d$ получается максимальная дисперсия среди всех возможных способов выбрать d компонент.

Утверждение 5. *Главная компонента* u_i *равна собственному вектору, который соответствует* i*-му по величине собственному значению.*