УНИВЕРСИТЕТ ИТМО

Основы линейной алгебры для анализа данных и машинного обучения

Воркшоп 4

Масляев Михаил, аспирант,
младший научный сотрудник

Ключевые пункты

1. Базис векторного пространства, смена базиса
2. Линейные операторы
3. Собственные вектора и числа, спектр оператора
4. Матричные разложения

Вопросы на helpdesk

УНИВЕРСИТЕТ ИТМО

https://docs.google.com/spreadsheets/d/1hcsgVuWwZAPzxvYz97gPdTvesnCBlaPnj-y98Hqdy1A/edit?usp=sharing

ITsMOre than a UNIVERSITY

Линейное пространство и его базис

Напоминание: линейное пространство

Определение того, с чем будем работать:

Векторное пространство - множество V, элементы которого удовлетворяют следующим аксиомам:

А) (Абелева группа по сложению) Для каждой пары элементов множества V: x и y вводится операция суммы элементов: $x + y$, такая что:

1) сумма коммутативна: $x + y = y + x$;
2) сумма ассоциативна: $x + (y + z) = (x + y) + z$;
3) существует 0: $x + 0 = x$ для любого x из V;
4) для каждого x из V существует уникальный элемент "-x": $x + (-x) = 0$.

Б) Для каждой пары a - скаляра и x - элемента из V вводится операция умножения ax из V со следующими свойствами:

1) произведение ассоциативно по скалярным множителям: $a(bx) = (ab)x$;
2) $1x = x$ для любого x из V;
3) Умножение на скаляр - дистрибутивно относительно векторов: $a(x + y) = ax + ay$;
4) Умножение на скаляр - дистрибутивно относительно скаляров: $(a + b)x = ax + bx$;

Напоминание: линейное пространство

Примеры: не стоит ограничиваться R^2

- Пространство матриц фиксированной размерности (и в целом тензоры произвольных рангов):

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \ \& \ B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \ : \gamma(A+B) = \begin{bmatrix} \gamma a_{11} + \gamma b_{11} & \gamma a_{12} + \gamma b_{12} \\ \gamma a_{21} + \gamma b_{21} & \gamma a_{22} + \gamma b_{22} \end{bmatrix}$$

- Множество непрерывных на некотором интервале функций *C[(a, b)]*;
- Множество решений дифференциального уравнения $y'' + py' + qy = 0$ на определённом интервале с различными начальными условиями;
- Множество остовных подграфов некоторого конечного простого ненаправленного графа *G*. Для *H* и *H'* подграфов говорим, что вершины смежны в *H + H'*, если они смежны в *H* или *H'*.

Покажите, что подобные множества определяют векторные пространства. Что выступает в роли нулевого элемента?

Двойственные пространства

Определение: линейный функционал на векторном пространстве *V*: отображение *f* из векторного пространства в множество действительных чисел, удовлетворяющее свойству:

$$f : V \longrightarrow \mathbb{R},\ f(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 f(x_1) + \alpha_2 f(x_2)$$

Определение: двойственное пространство *V'*: множество всех возможных линейных функционалов на векторном пространстве *V*. Двойственное пространство - также является векторным пространством:

В нём можно ввести нуль: тождественное-нулевую функцию: $f(x) \equiv 0$

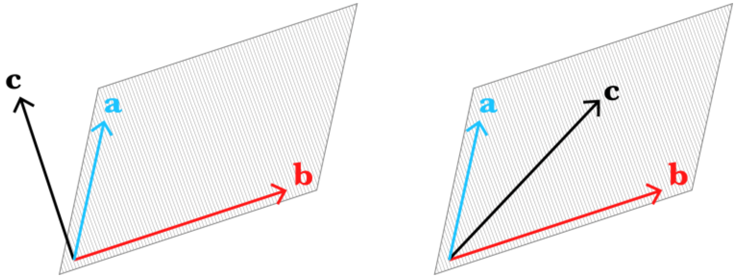
По свойствам линейного функционала: $f(x) = \alpha_1 f_1(x) + \alpha_2 f_2(x)$

Напоминание: базис линейного пространства

Что такое базис линейного пространства?

Определение: линейная зависимость: вектора $(v_1, \ldots, v_n)$ называют линейно-зависимыми, если существуют коэффициенты $(a_1, \ldots, a_n)$:

$$\sum_{i=1}^{n} a_i v_i = 0, \ \exists a_j \neq 0$$



Линейно-независимые

Линейно-зависимые

Линейно-независимые векторы - равенство комбинации только при (0, …, 0).

Определение: линейный базис векторного пространства *V*: множество *X* линейно-независимых векторов из *V*, что любой вектор v из *V* можно представить через линейную комбинацию элементов *X*:

$$v = \sum_{i=1}^{N} \alpha_i x_i, \ X = \{x_1, \ldots, x_N\}.$$

Коэффициенты вектора V в базисе X

Выбор базиса линейного пространства

- Пространство матриц фиксированной размерности (и в целом тензоры произвольных рангов):

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \ \& \ B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} : \gamma(A+B) = \begin{bmatrix} \gamma a_{11} + \gamma b_{11} & \gamma a_{12} + \gamma b_{12} \\ \gamma a_{21} + \gamma b_{21} & \gamma a_{22} + \gamma b_{22} \end{bmatrix}$$

- Множество непрерывных на некотором интервале функций $C[(a, b)]$;
- Множество решений дифференциального уравнения $y'' + py' + qy = 0$ на определённом интервале с различными начальными условиями;
- Множество остовных подграфов некоторого конечного простого ненаправленного графа G. Для H и H' подграфов говорим, что вершины смежны в $H + H'$, если они смежны в H или H'.

Какие из этих пространств - конечномерные? Как на этих векторных пространствах можно задать базис? Приведите примеры подобных базисов.

Местами выбор базиса неочевиден, но может быть совершён по аксиоме выбора / лемме Цорна. Базис Гамеля.

Замена базиса векторного пространства

Операция по замене базиса: пусть в пространстве *V* введён фиксированный базис $\{e_1, \ldots, e_n\}$, любая система из n-векторов $\{e'_1, \ldots, e'_n\}$ может быть определена через квадратную матрицу перехода $C = (c_{ij})$:

Матрица перехода

$$e'_j = \sum_i e_i c_{ij}$$

Запись одной координаты

$$(e'_1, \ \ldots \ , e'_n) = (e_1, \ \ldots \ , e_n)C$$

Запись в векторной форме

$$\begin{cases} e'_1 = e_1 c_{11} + \ldots e_n c_{1n} \\ \ldots \\ e'_n = e_1 c_{n1} + \ldots e_n c_{nn} \end{cases}$$

$$det(C) \neq 0$$

$\{e'_1, \ldots, e'_n\}$ - линейно-независимые, то есть может быть базисом *V*.

Замена базиса векторного пространства

Рассмотрим преобразование произвольного вектора x из *V* в базисах *X* и *X'*:

$$x = x_1 e_1 + \ \dots \ + x_n e_n = x'_1 e'_1 + \ \dots \ x'_n e'_n$$

Координатные векторы в старом и новом базисе: $X = (x_1, \ \dots \ , x_n)^T \qquad X' = (x'_1, \ \dots \ , x'_n)^T$

$$x = (e'_1, \ \dots \ , e'_n) X' = (e_1, \ \dots \ , e_n) C X' \Longrightarrow X = CX'$$

Зачем нам менять базисы? Ортогональность? Разреженность? Нормированность?

Удобство представления пространства:

- Ортогональность: из просто *n*-линейно-независимых в *V* (должно быть предгильбертовым пространством) хотим получить, что $\langle e_1, e_2 \rangle = 0$;
- Часто удобно выбрать длину (норму) каждого из базисных векторов единичной, такой базис называется нормированным;
- Допустимый подход - ортогонализация Грама-Шмидта.

Линейные преобразования

Определение: линейный оператор/преобразование - линейное отображение векторного пространства *V* в самого себя, т.е. удовлетворяет условиям:

1) *A*(x + y) = *A*x + *A*y, для любых x и y из *V*;
2) *A*(kx) = k*A*x, для любого x из *V* и k из *R*;

Если в V задан базис $\{e_1, \ \dots \ , e_n\}$, то допускается матричная форма по аналогии с матрицей замены базиса: $\mathcal{A}e_i = \sum_j a_{ij}e_j$

Получаем для образа: $y = \sum_j x_j \mathcal{A}e_j = \sum_{i,j} x_j a_{ij} e_i = \sum_i y_i e_i$

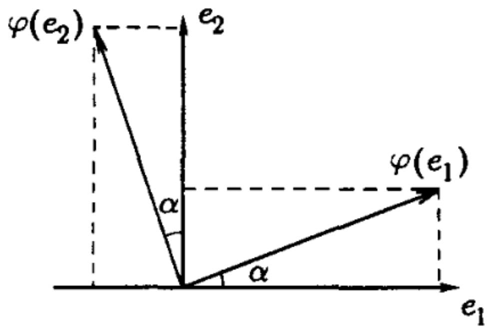
Линейные преобразования. Примеры

Классические примеры линейных преобразований:

Вращение векторов линейного пространства на заданный угол α:

В пространстве с ортонормированным Базисом $\{e_1, e_2\}$:

$$\begin{cases} \psi(e_1) = e_1 \cos(\alpha) + e_2 \sin(\alpha) \\ \psi(e_2) = -e_1 \sin(\alpha) + e_2 \cos(\alpha) \end{cases}$$



По этой записи базисных векторов: матрица оператора:

$$A = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}.$$

Линейные преобразования. Образ и ядро

Определение: для линейного отображения $\mathcal{A} : V \longrightarrow U$:

Образ: $Im\ \mathcal{A} = \{\mathcal{A}a : a \in V\} \subset U$

Ядро: $Ker\ \mathcal{A} = \{a \in V : \mathcal{A}a = 0\} \subset V$

Ядро и образ линейного оператора образуют подпространства;

Пример для векторного пространства дифференцируемых функций и отображения дифференцирования.

- Почему оно линейное?
- Что в данное случае является образом?
- Что ядром?

Линейные преобразования. Примеры

Рассмотрим пример P_2 - векторного пространства полиномов степеней не более 2:

$$P_2 = \{a_0 + a_1 x + a_2 x^2 \mid a_0, a_1, a_2 \in \mathbb{R}\}$$

Базис в пространстве полиномов: $\mathcal{B} = (1, x, x^2)$

Используем линейное преобразование: $T(p)(x) = p'(x) - p(x)$ $T(p)(x) : P_2 \longrightarrow P_2$

Исследуем результат преобразования:

$$T(a_0 + a_1 x + a_2 x^2) = (a_0 + a_1 x + a_2 x^2)' - (a_0 + a_1 x + a_2 x^2) =$$
$$= a_1 - a_0 + (2a_2 - a_1)x - a_2 x^2$$

В итоге матричная форма линейного преобразования:

$$\mathbf{T}_{\mathcal{B}} = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 2 \\ 0 & 0 & -1 \end{bmatrix}.$$

Линейные преобразования

Изменение формы линейного оператора при замене базиса:

$$(e'_1, \ \dots, e'_n) = (e_1, \ \dots, e_n)C$$

В силу линейности оператора получаем:

$$(\mathcal{A}e'_1, \ \dots, \mathcal{A}e'_n) = (\mathcal{A}e_1, \ \dots, \mathcal{A}e_n)C =$$
$$= (e_1, \ \dots, e_n)AC = (e'_1, \ \dots, e'_n)C^{-1}AC$$

В итоге матрица оператора в новом базисе: $B = C^{-1}AC$

Собственные вектора и числа

Собственные вектора и числа

Что такое собственный вектор?

- По определению: ненулевой вектор *x* из *V*, удовлетворяющий условию $Ax = \lambda x$
- В таком случае, λ - собственное число оператора *A*.

Полезное свойство линейных операторов:

Если представить основное пространство V через прямую сумму инвариантных подпространств (где $U \in V : AU \subset U$), то матрица превращается в блочную матрицу

$$A = \begin{bmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & \dots & 0 \\ & \cdots & & \\ 0 & 0 & \dots & A_k \end{bmatrix}$$

Собственные вектора и числа

Что такое собственный вектор?
- По определению: ненулевой вектор *x* из *V*, удовлетворяющий условию $Ax = \lambda x$
- В таком случае, λ - собственное число оператора *A*.

Каждый собственный вектор - отдельное инвариантное подпространство!
Размерность каждого блока блочной матрицы оператора - 1, то есть скаляр, получаем диагонализацию линейного оператора!

$$A = \begin{bmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & \dots & 0 \\ & \cdots & & \\ 0 & 0 & \dots & A_k \end{bmatrix} \Rightarrow A' = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ & \cdots & & \\ 0 & 0 & \dots & \lambda_k \end{bmatrix}$$

Собственные вектора и числа. Определение

Что такое собственный вектор?

- По определению: ненулевой вектор *x* из *V*, удовлетворяющий условию $\mathcal{A}x = \lambda x$
- В таком случае, λ - собственное число оператора *A*.

Условие существования собственных векторов и значений для линейного оператора: вырожденность оператора $\mathcal{A} - \lambda E$, то есть равенство нулю определителя:

$$det|\mathcal{A} - \lambda E| = 0$$

$$det|\mathcal{A} - \lambda E| = \begin{bmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ & \cdots & & \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{bmatrix}$$

Собственные вектора и числа. Определение

Характеристический многочлен линейного оператора:

$$f_{\mathcal{A}}(t) = (-1)^n \, det(A - tE)$$

Существуют теоремы:

- Собственные значения линейного оператора - корни его характеристического многочлена;
- Любой линейный оператор в **комплексном пространстве** имеет собственный вектор;

Для существования базиса из собственных векторов необходимо (и достаточно):

1) Чтобы характеристический многочлен $f_{\mathcal{A}}(t)$ раскладывался на линейные множители;

2) Размерность каждого собственного подпространства $V_\lambda(\mathcal{A}) = Ker(A - \lambda E)$ должна соответствовать кратности корней многочлена.

Как выполнять проверку на практике?

Собственные вектора и числа

Проверка: через характеристический многочлен. Если у него n - различных корней, то существует базис из собственных векторов оператора *A*.

Нахождение собственных векторов и чисел:
1. Составление матрицы оператора $A - \lambda E$, по ней - характеристический полином;
2. Поиск корней характеристического полинома, оценка их числа / того, действительные ли это числа;
3. Для каждого собственного числа: поиск собственных векторов через решение системы линейных алгебраических уравнений:

$$\begin{cases} (a_{11} - \lambda)x_1 + a_{12}x_2 + \ \dots \ + a_{1n}x_n = 0 \\ a_{21}x_1 + (a_{22} - \lambda)x_2 + \ \dots \ + a_{2n}x_n = 0 \\ \dots \\ a_{n1}x_1 + a_{n2}x_2 + \ \dots \ + (a_{nn} - \lambda)x_n = 0 \end{cases}$$

Собственные вектора и числа

Проблема: задача поиска корней произвольного полинома, для высоких степеней отсутствует аналитическое решение.

Решение: использовать специальные алгоритмы, в общем случае - итеративные. Большинство исполнений - не универсальны, применимы лишь к конкретным матрицам, или находят лишь одно собственное значение.

Пример для вещественной симметричной матрицы: метод Якоби
Основная идея - диагонализация матрицы при помощи матрицы поворота Гивенса - обобщения матрицы поворота:

Заменяем матрицу на близкую к ней: $S' = GSG^T$ последовательно, пока матрица на станет близкой к диагональной (по аналогии с методом решения СЛАУ):

$$G_{ij}: \qquad G_{24} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \cos(\vartheta) & 0 & \sin(\vartheta) & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & -\sin(\vartheta) & 0 & \cos(\vartheta) & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Матричные разложения. Спектральное разложение

Представление данной матрицы *A* как произведения матриц с заданными свойствами: диагональностью, симметричностью и пр. для удобств последующих вычислений или анализа.

Практическая задача в анализе данных (PCA): спектральное разложение матрицы ковариаций: представление матрицы *A* через произведение матриц $A = V\Lambda V^{-1}$

В этом разложении V - матрица со столбцами - собственными векторами оператора *A*; матрица Λ - диагональная матрица, где на главных диагоналях стоят собственные значения *A*.

$$\mathbf{M} = \underbrace{\begin{pmatrix} \uparrow & \uparrow & & \uparrow \\ \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_d \\ \downarrow & \downarrow & & \downarrow \end{pmatrix}}_{V} \underbrace{\begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_d \end{pmatrix}}_{\mathbf{\Lambda}} \underbrace{\begin{pmatrix} \longleftarrow & \mathbf{u}_1 & \longrightarrow \\ \longleftarrow & \mathbf{u}_2 & \longrightarrow \\ & \vdots & \\ \longleftarrow & \mathbf{u}_d & \longrightarrow \end{pmatrix}}_{V^T}$$

Дальнейшее развитие вопроса - в лабораторной работе

УНИВЕРСИТЕТ ИТМО

Вопросы по воркшопу?

ITsMOre than a UNIVERSITY