

**ЛИНЕЙНАЯ АЛГЕБРА И ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ****LINEAR ALGEBRA AND ARTIFICIAL INTELLIGENCE****Н. М. Гордеева, Е. С. Попущина**

Московский государственный технический университет им. Н.Э. Баумана  
(национальный исследовательский университет), г. Москва, Россия

**N. M. Gordeeva, E. S. Popushina**

Bauman Moscow State Technical University (National Research University),  
Moscow, Russia

**Аннотация.** В работе рассматриваются особенности преподавания линейной алгебры тем студентам, которые планируют работать в сфере информационных технологий. Отмечается, что все темы, изучаемые при традиционном подходе в большинстве вузов, необходимы, но не достаточны. С развитием искусственного интеллекта выпускники встречаются на работе с машинным обучением и понятием кластеризации множеств. Для решения поставленных задач будущий специалист должен хорошо понимать, как устроена структура пространств и подпространств, что такое расстояние от объекта до подпространства, косинусная мера близости, какие бывают способы задания пространств, какие бывают метрики. В статье предлагается ряд задач с методическими комментариями и описываются способы создания задач для большого количества вариантов.

**Ключевые слова:** линейная алгебра; метрики; кластеризация; машинное обучение.

**Annotation.** The paper discusses the features of teaching linear algebra to those students who plan to work in the field of information technology. It is noted that all the topics studied with the traditional approach in most universities are necessary, but insufficient. With the development of artificial intelligence, graduates meet at work with machine learning and the concept of clustering sets. To solve the tasks set, a future specialist should understand well how the structure of spaces and subspaces is arranged, what is the distance from an object to a subspace, the cosine measure of proximity, what are the ways of defining spaces, what are the metrics. The article offers a number of tasks with methodological comments and describes ways to create tasks for a large number of options.

**Keywords:** linear algebra, metrics; clustering; machine learning.

**Введение и описание проблемы.** С развитием цифровизации всех отраслей народного хозяйства растет потребность в специалистах в области информационных технологий и все больше выпускников вузов самых разных специальностей нуждаются хотя бы в минималь-

ном объеме знаний, касающихся искусственного интеллекта, позволяющим им разобраться с поставленными задачами. Одно из самых популярных направлений – машинное обучение.

Машинное обучение – целый класс методов искусственного интеллекта, использующих средства разных математических дисциплин и реализующих различные техники работы с данными в цифровой форме. Оно делится на два вида: обучение с учителем и обучение без учителя. В первом случае требуется сгруппировать изучаемые объекты в группы – кластеры, используя данные о попарном сходстве объектов. Это так называемый кластерный анализ. Главное отличие кластеризации от классификации состоит в том, что перечень групп (кластеров) изначально не задан и определяется в процессе работы алгоритма.

В общем виде этапы кластерного анализа можно описать так:

1. Отбор объектов для кластеризации.
2. Определение множества переменных, по которым будут оцениваться объекты в выборке. При необходимости (почти всегда) – нормализация значений переменных или установление «весов».
3. Выбор метрики – т.е. меры сходства между объектами. Выбор метрики полностью лежит на исследователе, поскольку результаты кластеризации могут существенно отличаться при использовании разных метрик.
4. Вычисление меры сходства для каждой пары объектов.
5. Применение кластерного анализа для создания групп сходных объектов, т.е. кластеров.
6. Представление результатов анализа.
7. После получения и анализа результатов возможна корректировка выбранной метрики и метода кластеризации до получения оптимального результата, т.е. повторения этапов 3 – 6.

Данные в цифровой форме – это моделирование любых изучаемых объектов путем записи их основных свойств в виде набора характеристик, измеряемых по некоторой утвержденной шкале, т.е. набор векторов или матриц.

Практически в любом учебном плане (даже у гуманитарных направлений, у которых не обязательно отдельно выделена линейная алгебра) есть дисциплина, в которой студенты знакомятся с такими объектами, как арифметические векторы и матрицы, и осваивают основные действия с ними. Большинство учебных программ содержит раздел, посвященный решениям матричных уравнений и систем линейных алгебраических уравнений. Инженерные и математические специальности предполагают изучение в курсе линейной алгебры знакомство с линейными пространствами, подпространствами и линейными операторами.

Таким образом, первые два шага не вызывают трудностей, а третьему надо уделить особое внимание. В курсе линейной алгебры в большинстве случаев вводится скалярное произведение, с помощью которого строятся метрика и норма. И не обязательно даже вводятся эти понятия, часто просто говорят о длине вектора.

Поэтому для большинства выпускников этап «выбор метрики» вообще выглядит очень туманным. А от этого выбора зависит эффективность применения метода. Более того, при вычислении меры сходства и определении кластеров приходится оперировать расстояниями между множествами. И надо очень хорошо понимать структуру линейных пространств, под-

пространств, что такое расстояние от вектора до подпространства, что такое косинусная близость (схожесть), которую используют в решении задач обработки естественного языка и т.п.

После не всегда бесплатных курсов, которые проходят желающие работать в IT-сфере, выпускники с гордостью оперируют понятием «манхэттенское расстояние», хотя это всего лишь метрика, предложенная Г. Минковским более ста лет назад. Она является одним из многообразных примеров метрики, с которыми можно познакомить студентов уже на первом курсе без ущерба для остальных тем, и можно рассмотреть целесообразность применения разных метрик в зависимости от решаемой проблемы.

Поскольку цифровизация данных представляет из себя запись объекта в виде вектора, то дальнейшая работа любыми объемами данных происходит как работа с линейными подпространствами. Линейные подпространства можно задавать двумя основными способами. Первый способ – явное задание, как линейная оболочка системы векторов, второй способ – неявное задание, как решение однородной системы линейных алгебраических уравнений. Это очень важные понятия, необходимые для работы с большими данными, но, к сожалению, редко встречающиеся в задачниках. Способы задания, переход от одного способа к другому, нахождение пересечения и суммы подпространств в разных случаях явного и неявного задания, описаны, например, в пособии [1].

Далее в работе предлагается рассмотреть ряд задач, которые полезно включить в программу, учитывая вышесказанное.

**Задача 1.** Проверьте, задает ли функция  $f(x, y) = x_1y_1 + 2x_2y_3 + 5x_3y_2 + x_1y_2 - 3x_2y_1$  скалярное произведение в пространстве  $\mathbb{R}^3$ .

**Задача 2.** Покажите, что норму в  $n$  – мерном линейном пространстве  $X$  можно определить равенством:  $\|x\|_1 = |\alpha_1| + |\alpha_2| + \dots + |\alpha_n|$ , где  $\alpha_i$  – координаты вектора  $x$ .

**Задача 3.** Покажите, что пространство ограниченных числовых последовательностей имеет метрику  $\rho(x, y) = \sup_i |x_i - y_i|$ .

**Задача 4.** Даны векторы  $a_1, a_2, a_3, q$ . Пусть пространство  $L$  – это линейная оболочка векторов  $a_1, a_2, a_3$ . Найдите проекцию вектора  $q$  на пространство  $L$  и угол между вектором  $q$  и пространством  $L$ .

**Задача 5.** Даны векторы  $a_1, a_2, q$ . Постройте наилучшее приближение вектора  $q$  в виде линейной комбинации векторов  $a_1$  и  $a_2$ . Вычислите абсолютную и относительную ошибку приближения.

**Заключение.** Курс, включающий данные темы, читается в МГТУ им. Н.Э. Баумана. Варианты задач 2 и 3 берутся из классических учебников, так как в них достаточно показать студентам способ доказательства. Освоение тем из задач 1, 4, 5 требует большого количества вариантов, чтобы гарантированно отработать навык. Для составления задач этих типов написана программа, позволяющая делать неповторяющиеся варианты для каждого студента, а также ответы для быстрой проверки преподавателем. С помощью этой программы можно легко менять сложность задачи, например, изменением размерности векторов; отсеивать «некрасивые ответы», а именно содержащие дроби, корни, громоздкие промежуточные выкладки и т.д. Такой подход исключает простое заучивание студентом десятка классических

задач, позволяя подбирать индивидуальную траекторию для каждого студента, подобно тому как это делается при работе со студентами, имеющими проблемы со здоровьем в методике укрупнения дидактических единиц [2]. В дальнейшем предполагается добавить в программу задачи на другие темы из курса линейной алгебры.

### **Библиографический список**

1. Власов П.А., Кавинов А.В., Канатников А.Н. Алгебра: Метод. указания к выполнению домашнего задания – М.: Изд-во МГТУ им. Н.Э. Баумана, 2007. – 80 с.
2. Семакин А. Н., Емгушева Г. П. Повышение эффективности преподавания линейной алгебры посредством методики укрупнения дидактических единиц для студентов с ограниченными возможностями здоровья // *Alma Mater* (Вестник высшей школы). 2021. № 10. С. 65-71. DOI 10.20339/AM.10-21.065.

#### **Сведения об авторах:**

Надежда Михайловна Гордеева

E-mail: nmgordeeva@bmstu.ru; SPIN-код 8406-7531, ORCID: 0009-0005-1907-9884.

Екатерина Сергеевна Попушина

E-mail: popushinaes@bmstu.ru; SPIN-код: 3108-1062, ORCID: 0009-0008-6830-7443.