Отчет по заданию номер 2.

ФИО: Кузьмин Никита.
Группа: 208.
GitHub-репозиторий: тут можно посмотреть код для, с помощью я проверял/искал гипотезы.

Содержание:

Постановка задачи:

Дано:

1. m целочисленных n-мерных векторов: v1, …, vm
2. Матрица попарных расстояний Хэмминга(H), матрица попарного числа совпадений координат для данных векторов(M).

Вопрос:

Верно ли, что у этих матриц пространства линейных комбинаций столбцов совпадают? Заметим, что из-за того, что H и M - симметричны относительно главной диагонали, то эту задачу можно рассматривать и относительно строк.

Ход решения:

1. Сначала подготовим программную часть, с помощью которой можно будет быстро проверять гипотезы, да и просто находить закономерности, от которых можно отталкиваться. Написал функции, которые считают расстояние Хэмминга, число попарных совпадений координат у векторов. Далее написал функции, которые формируют соответствующие матрицы по заданным векторам и функцию, которая генерирует произвольные векторы одинаковой длины(на данный момент они могут совпадать).
2. Нужно определиться с тем, в какую сторону двигаться. Для этого проведем эксперименты, пользуясь соотношением из курса линейной алгебры:

   Теорема 1:

   ] Даны 2 квадратные матрицы порядка n. Соотношения rank(A) = rank(B) = rank(A|B) выполнены <=> пространства линейных комбинаций столбцов/строк этих матриц совпадают.

   С помощью этой теоремы и подготовленной программной части можно предположить, что на вопрос, поставленный в задаче, ответ положительный.

   Добавил функцию **first_hypothesis_check**, которая проверяет на рандомно сгенерированных векторах условия теоремы 1. Проводятся тесты для различных параметров **vectors_amount**, **vectors_dim**, **coordinates_range**. По результатам тестов: условия теоремы нарушаются только для систем векторов, в которых хотя бы два вектора совпадают. Для остальных систем - теорема выполняется, то есть пространства линейных комбинаций столбцов матриц H и M совпадают.

   ## Гипотеза 1:
   ## верно, что пространства линейных комбинаций столбцов данных матриц совпадают.

3. Попробовал найти ранги матриц H и M в общем виде, но это не привело к успеху(пытался методом Гаусса свести к ступенчатой матрице, рассмотрел вариант нахождения ранга с помощью теоремы об окаймляющих минорах, но в общем случае его реализовать тоже тяжело).
4. Можно свести данную задачу к двум системам (m, m). Такого вида:

$$\mathrm{I}\ \begin{cases} \alpha_{11}\cdot\bar{h}_1 + \ldots + \alpha_{1m}\cdot\bar{h}_m = \bar{k}_1 \\ \ldots \\ \alpha_{m1}\bar{h}_1 + \ldots + \alpha_{m,m}\bar{h}_m = \bar{k}_m \end{cases}, \quad \mathrm{II}\ \begin{cases} \beta_{11}\bar{k}_1 + \ldots + \beta_{1m}\bar{k}_m = \bar{h}_1 \\ \ldots \\ \beta_{m1}\bar{k}_1 + \ldots + \beta_{mm}\bar{k}_m = \bar{h}_m \end{cases},$$

   где альфа - неизвестные системы, h1, …, hm и k1,…, km - вектор-столбцы матриц H и M соответственно. Возможно, она поможет решить задачу, воспользовавшись теоремой:

   Теорема 2(Кронекера-Капелли):

Система линейных алгебраических уравнений совместна тогда и только тогда, когда ранг её основной матрицы равен рангу её расширенной матрицы, причём система имеет единственное решение, если ранг равен числу неизвестных, и бесконечное множество решений, если ранг меньше числа неизвестных.

Но: в этой теореме предполагается, что h1, …, hm; k1, …, km - числа, но в системах I и II это не так. Возможно, есть какие-то обобщения теоремы Кронекера-Капелли. Если не останется больше идей, то попробую реализовать эту идею, но пока кажется, что эта идея не приведет к результату.

5. Проведем еще эксперименты с матрицами:
    a. Для определителя: **в экспериментах, когда среди векторов не было совпадающих, определитель был отличен от нуля!**
    b. Для собственных векторов и собственных значений: ничего не было замечено.
    c. Для различных норм: ничего не было замечено.
    d. После проведенных тестов для рангов получилось проследить зависимость максимального ранга(**max_rank)** от размерности векторов(n) и максимального значения координаты(**coordinate_range**) во всех векторах:

$$\text{max\_rank} = \begin{cases} \overbrace{n * (\text{coordinate\_range} - 1) + 1}^{MR}, & \text{если } m \geq MR \\ m, & \text{если } m < MR \end{cases}$$

    НО: это не доказанная закономерность, а просто интересное наблюдение, того, что во всех тестовых случаях это соотношение выполнялось.

6. **Идея:** Так как M = n*E - H, то можно посмотреть, лежит ли единичный столбец в линейной оболочке столбцов H, т.е. совместна ли система Hx = **e**, где **e** = (1, 1, 1, …, 1)^T размерности m. Если это действительно так, то определить, как найти решения x для этой системы. Гипотеза 2 опирается на это.

    Гипотеза 2: если вектор **e**, описанный выше лежит в линейной оболочке вектор-столбцов матрицы H(или M), то линейные оболочки вектор-столбцов матриц H и M совпадают.

    Для проверки идеи написал функцию **second_hypothesis_check** и исправил **generate_vectors** - теперь она генерирует векторы, среди которых нет совпадающих. Как показали тесты, учитывая то, что numpy.linalg.solve не работает с матрицами, у которых det = 0(для них отдельно нужно проверять теорему Кронекера-Капелли - она выполнялась в этих случаях => решений было бесконечно много), то: **вторая гипотеза тоже является верной**(по крайней мере во всех сгенерированных случаях). Пока что никакой закономерности для решений этих систем найти не удалось.

7. **Теорема 3:**
    Гипотеза 1 выполняется <=> вектор e = (1, 1, … 1)^T лежит в пространстве линейных комбинаций столбцов матрицы H(или M).

Доказательство.

▲**Необходимость:**

Учитывая rg(h1, … hm, n***e** - h1, …, n***e** - hm) = rg(h1, …, hm, **e**), предположим, что **e** не лежит в пр-ве линейных комбинаций столбцов матрицы H, тогда: rg(h1, …, hm, **e**) = rg(h1, …, hm) + 1 = rg(h1, … hm, n***e** - h1, …, n***e** - hm), но: так как выполняется гипотеза 1(по условию теоремы) => выполняется теорема 1 => rg(h1, …, hm, n***e** - h1, …, n***e** - hm) = rg(h1, …, hm) = rg(h1, …, hm, **e**) - противоречие!

Достаточность:

] Гипотеза 1 не выполняется => не выполняются условия теоремы 1, то есть хотя бы одно из равенств нарушается: rg(h1, … hm, n***e** - h1, …, n***e** - hm) = rg(h1, …, hm) = rg(n***e** - h1, …, n***e** - hm).

1. Если не выполняется первое равенство, то: так как **e** - лежит в линейной оболочке столбцов матрицы H => rg(h1, …, hm) = rg(h1, …, hm, **e**). То допишем к матрице H m единичных столбцов, умноженных на n. rg(полученной матрицы) = rg(H), так как **e** - лежит в линейной оболочке столбцов H. Затем вычтем из (m+1) столбца 1-й
   из (m+2) столбца 2-й
   …
   из 2m столбца m-й
   Получили матрицу, равную (H|M) => rg(h1, …, hm, n***e** - h1, …, n***e** - hm) = rg(H) => противоречие!
2. Если не выполняется второе равенство, то: так как rg(n***e** - h1, …, n***e** - hm, **e**) = rg(h1, …, hm, **e**) = rg(h1, …, hm) , при этом rg(n***e** - h1, …, n***e** - hm) = rg(n***e** - h1, …, n***e** - hm, **e**), так как через h1, …, hm можно выразить **e** с помощью некоторых коэффициентов альфа, то: складываем (n***e** - hi) с этими коэффициентами => получаем sum(ai *n***e**) - **e** = **e***(ai*n - 1) => **e** лежит в пространстве линейных комбинаций столбцов матрицы M. То есть rg(M) = rg(M|**e**) => rg(h1, …, hm) = rg(n***e** - h1, …, n***e** - hm) => противоречие!
3. Если не выполняется rg(h1, …, hm, n***e** - h1, …, n***e** - hm) = rg(n***e** - h1, …, n***e** - hm), то: т.к. rg(h1, …, hm, n***e** - h1, …, n***e** - hm) = rg(h1, …, hm, **e**) = rg(h1, …, hm), так как **e** - лежит в пр-ве лин. комбинаций h1, …, hm, тогда: припишем к матрице H m столбцов, равных **e.** В таком случае rg(полученной матрицы) = rg(H) и вычтем из m+1, …, 2m столбцов(умноженных на n) полученной матрицы 1, …,m столбцы соответственно, затем удалим первые m столбцов. Так как эти m столбцов являются линейными комбинациями m+1, … 2m столбцов, то rg(полученной матрицы) = rg(H) = rg(h1, …, hm, n***e** - h1, … , n***e** - hm), а полученная матрица тождественно равна M => противоречие. ■

8. **Рассмотрим частный случай:** n = 1.
   Тогда векторы становятся числами(а так как по условию задачи сказано, что все векторы попарно различны => матрица H будет иметь вид: нули на главной диагонали, остальные элементы - единицы. Аналогично для M: единицы на главной диагонали, а остальные элементы - нули. С помощью линейных преобразований матрицы H можно показать, что rg(H) = rg(M)(H сводится к M) => rg(H) = rg(M) = rg(H|M), то есть **Гипотеза 1 верна** для n = 1.
   Посмотрел на тесты в программе для тех случаев, когда среди векторов могут быть равные(но не все равны между собой!), то - результат тоже положительный. Думаю, ничего страшного не будет, если попытаюсь расширить условия задачи при n = 1. Пусть среди векторов могут быть равные. Покажем, что Гипотеза 1 верна. Проделаем это более аккуратно:
   1) Перенумеруем векторы так, чтобы равные шли подряд, т.е.

$$\underbrace{v_1, v_1, \dots, v_1}_{l_1}, \underbrace{v_2, \dots, v_2}_{l_2}, \dots, \underbrace{v_t, \dots, v_t}_{l_t}$$

2) Тогда матрицы H и M станут блочными для такой нумерации векторов:
Обозначения : O^(l1*l1) - квадратная матрица размерности l1, состоящая из нулей
E^(l1*l1) - квадратная матрица размерности l1, состоящая из единиц

$$H = \begin{array}{c} l_1\left\{\begin{matrix} \\ \\ \\ \\ \end{matrix}\right. \\ \\ \\ \end{array} \begin{bmatrix} \overbrace{0 & 0 & 0 & 0 & \cdots & 0}^{l_1} & 1 & \cdots & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 & \cdots & 1 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 & \cdots & 1 \\ 1 & 1 & 1 & 1 & \cdots & 1 & 0 & \cdots & \\ & & & & \cdots & & & & \end{bmatrix} = \begin{bmatrix} \Theta^{l_1\times l_1} & & & \mathbb{1} \\ & \boxed{\Theta^{l_2\times l_2}} & & \\ & & \ddots & \\ \mathbb{1} & & & \boxed{\Theta^{l_t\times l_t}} \end{bmatrix}$$

$$M = \begin{bmatrix} E^{l_1\times l_1} & & & \Theta \\ & \boxed{E^{l_2\times l_2}} & & \\ & & \ddots & \\ \Theta & & & E^{l_t\times l_t} \end{bmatrix}$$

то: $rg(M) = rg(E^{l_1\times l_1} + \dots + E^{l_t\times l_t}) = t$

3) Разберемся с rg(H). Докажем, что rg(H) = rg(M) = t, так как H можно свести к M с помощью линейных преобразований:

$$H=\left[\begin{array}{c|cccc}
\Theta^{l_1\times l_1} & & & \mathbb{1} & \\ \hline
 & \boxed{\Theta^{l_2\times l_2}} & & & \\
\mathbb{1} & & \ddots & & \\
 & & & & \boxed{\Theta^{l_t\times l_t}}
\end{array}\right]$$

3.1) Вычитаем 1-й столбец из (l1+1)-го, (l1+2)-го, …, m-го.

$$\left[\begin{array}{c|c|c}
\Theta^{l_1\times l_1} & E & E \\ \hline
E & \boxed{-E^{l_2\times l_2}} & \Theta \\
 & \ddots & \\ \hline
E & \Theta & \boxed{-E^{l_t\times l_t}}
\end{array}\right]$$

3.2) Прибавляем к l1 первым строкам (l1+1)-ю,

(l1+2)-ю,

…,

m-ю строки.

$$\left[\begin{array}{c|ccc}
(m-l_1)E^{l_1\times l_1} & & \Theta & \\ \hline
 & \boxed{-E^{l_2\times l_2}} & & \Theta \\
E & & \ddots & \\
 & \Theta & & \boxed{-E^{l_t\times l_t}}
\end{array}\right]$$

3.3) Вычитаем из первых l1 столбцов (l1+1)-й,
(l1+l2+1)-й,
…,
(l1+l2+…+lt+1)-й столбцы
(то есть первые столбцы из каждого блока).

$$\begin{bmatrix} (m-l_1)E^{l_1\times l_1} & & & \Theta \\ & -E^{l_2\times l_2} & & \\ & & \ddots & \\ \Theta & & & -E^{l_t\times l_t} \end{bmatrix}$$

3.4) Делим первые l1 столбцов на (m-l1) != 0, так как не все векторы равны между собой.
Умножаем остальные столбцы на (-1).

$$\begin{bmatrix} E^{l_1\times l_1} & & & \Theta \\ & E^{l_2\times l_2} & & \\ & & \ddots & \\ \Theta & & & E^{l_t\times l_t} \end{bmatrix}$$

Получили матрицу M из H с помощью линейных преобразований => rg(H)=rg(M)=t.

4) Осталось рассмотреть rg(H|M). Тут достаточно занулить первую квадратную матрицу(H) размера m с помощью линейных преобразований.
5) То есть получили: rg(H) = rg(M) = rg(H|M) => **условие задачи выполняется**.

Есть большое желание попробовать доказать по индукции для n. Попробуем!

9. Для n=1, т.е. база индукции доказана.
] для n = k - верно, т.е. пусть даны векторы v1, …, vm и для них выполняется условие задачи. Рассмотрим n = k+1, то есть размерность векторов v1, …, vm увеличивается на 1. Можно заметить, что H = H1 + H2 + … + Hk+1 и M = M1 + M2 + … + Mk+1, где Hi, Mi - матрицы расстояний Хэмминга и попарных совпадений координат соответственно для i-й компоненты векторов. То есть по предположению индукции для H' = H1 + … + Hk и M' = M1 + … + Mk выполняется условие задачи. Причем, для матриц Hk+1 и Mk+1 тоже выполняется условие, так как они основаны на векторах размерности n = 1(т.е. на числах), среди которых могут быть совпадающие, а для такого случая в **пункте 8** было доказано утверждение задачи.
Нужно посмотреть на то, как меняются ранги матриц H' и M' при добавлении Hk+1 и Mk+1. Известно, что столбцы H', M' линейно выражаются через друг друга.

Известно, что столбцы Hk+1 и Mk+1 линейной выражаются через друг друга.
Пусть:

$$H' = \left[\bar{h}'_1 \,\middle|\, \bar{h}'_2 \,\middle|\, \dots \,\middle|\, \bar{h}'_m\right]$$

$$M' = \left[\bar{r}'_1 \,\middle|\, \bar{r}'_2 \,\middle|\, \dots \,\middle|\, \bar{r}'_m\right]$$

Тогда:

$$\begin{cases} \bar{r}'_1 = \sum_{i=1}^{m} \alpha_{1i} \bar{h}'_i \\ \dots\dots\dots \\ \bar{r}'_m = \sum_{i=1}^{m} \alpha_{mi} \bar{h}'_i \end{cases}$$

То есть матрица M’ будет иметь вид:

$$M' = \left[\sum_{i=1}^{m} \alpha_{1i} \bar{h}'_i \,\middle|\, \dots \,\middle|\, \sum_{i=1}^{m} \alpha_{mi} \bar{h}'_i\right]$$

Из доказанного утверждения аналогичный вид будет у матрицы Mk+1. Если h1, h2, …, hm; r1, r2, …, rm - векторы матриц Hk+1 и Mk+1 соответственно, то:
можно выписать вид H’ + Hk+1 и M’ + Mk+1:

$$H' + H_{k+1} = \left[\bar{h}'_1 + \bar{h}_1 \,\middle|\, \dots \,\middle|\, \bar{h}'_m + \bar{h}_m\right]$$

$$M' + M_{k+1} = \left[\sum_{i=1}^{m} (\alpha_{1i} \bar{h}'_i + \beta_{1i} \bar{h}_i) \,\middle|\, \dots \,\middle|\, \sum_{i=1}^{m} (\alpha_{mi} \bar{h}'_i + \beta_{mi} \bar{h}_i)\right]$$

Пока не получается завершить доказательство по индукции.

10. Попробуем рассмотреть задачу с другой стороны: так как все элементы матриц H и M неотрицательны, то можно рассмотреть теорию матриц с неотрицательными элементами. В курсе линейной алгебры подробно эта тема не рассматривалась, поэтому придется гуглить. Для начала надо определить, является ли разложимой каждая из матриц H и M. **UPD:** Тоже не вышло...
11. Вернемся к замечанию **пункта 5.a).** В нем проверялось значение определителя матриц H и M на разных тестах, которые говорят о невырожденности данных матриц, если исходные векторы попарно различны.
    Теорема 4:
    Если det(H) != 0 => **e** = (1, …, 1)^T лежит в линейной оболочке столбцов матрицы H.
    Доказательство:
    ▲Так как det(H) != 0 => существует обратная матрица H^(-1). Домножим H на обратную матрицу справа => H*H^(-1) = I. По свойству произведения матриц: i-й столбец произведения равен линейной комбинации столбцов матрицы H с соответствующими коэффициентами из i-го столбца матрицы H^(-1). То есть с помощью линейной комбинации **только столбцов** матрицы H получили I. А вектор **e** принадлежит линейной оболочке столбцов матрицы I => вектор **e** принадлежит линейной оболочке матрицы H. ■
    UPD: Получилось доказать более лаконично данную теорему:
    ▲Так как матрица H невырождена, то её столбцы образуют базис m-мерного пространства => можно получить любой вектор этого пространства с помощью линейной комбинации базисных столбцов, в частности и вектор **e**. ■)
    То есть из det(H) != 0 => (**теорема 3**) => доказательство **гипотезы 1**.

12. Для матрицы H при m = 3 показать, что det(H) != 0 - легко. Но при увеличении m сложность растет в разы(по крайней мере, для моего подхода). Покажем, что для m=3 => det(H) != 0.
    Выше было замечено, что H = H1 + H2+ … + Hm, где Hi - матрица попарных расстояний Хэмминга, составленная для i-ой координаты исходных векторов. То есть при m=3 H = H1 + H2 + H3. Видов компонент будет не так много(4) и их можно перебрать:

1) $\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ 2) $\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$

3) $\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$ 4) $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$

И рассмотрим всевозможные их суммы: H = a*H1 + b*H2 + c*H3 + d*H4, где a, b, c, d целочисленные коэффициенты, удовлетворяющие следующим соотношениям:

1) a + b + c + d = 3.
2) a, b, c, d >= 0.

получим матрицу:

$$\begin{bmatrix} 0 & a+b+c & b+c+d \\ a+b+c & 0 & a+c+d \\ b+c+d & a+c+d & 0 \end{bmatrix}$$

Определитель которой равен: 2*(b+c+d)*(a+b+c)*(a+c+d) > 0, так как хотя бы один из коэффициентов будет отличен от нуля.
Но для m=4 это уже будет сделать не так просто.

Что можно было сделать еще:

1) Матрица попарных расстояний Хэмминга очень похожа на матрицу расстояний между векторами(||vi - vj||), а матрица попарных совпадений координат - напоминает матрицу Грама для системы векторов. Можно проверить, есть ли какая-то связь между ними.
2) Можно заметить, что |rank(H) - rank(M)| <= 1. Это следует из определения матрицы M: M = n*E - H и того, что rank(H|**e**) = rank(n*E - H), а ранг матрицы при приписывании к ней одного столбца может измениться не более, чем на единицу. Возможно, получилось бы доказать, что |rank(H) - rank(M)| = 0.
3) Как я уже писал в **пункте 4:** можно исследовать две системы:

$$\mathrm{I}\ \begin{cases} \alpha_{11}\cdot\bar{h}_1 + \ldots + \alpha_{1m}\cdot\bar{h}_m = \bar{k}_1 \\ \ldots\ldots\ldots\ldots\ldots \\ \alpha_{m1}\bar{h}_1 + \ldots + \alpha_{m,m}\bar{h}_m = \bar{k}_m \end{cases} ; \quad \mathrm{II}\ \begin{cases} \beta_{11}\bar{k}_1 + \ldots + \beta_{1m}\bar{k}_m = \bar{h}_1 \\ \ldots\ldots\ldots\ldots\ldots \\ \beta_{m1}\bar{k}_1 + \ldots + \beta_{mm}\bar{k}_m = \bar{h}_m \end{cases}$$

4) Более подробно изучить теорию неотрицательных матриц, так как найти её применение к этой задаче за это время не получилось.
5) Поработать еще индукцией в **пункте 9.** Как мне кажется, замечательное свойство матриц H и M (H = H1 + … Hn) идеально подходит для доказательства с помощью индукции по n.
6) Совсем не рассматривал(кроме **пункта 5**) ограничение на координаты векторов. Возможно, для каких-то частных случаев удалось бы решить задачу.
7) При доказательствах не удалось нигде воспользоваться свойством матрицы попарных расстояний Хэмминга - неравенством треугольника.
8) Можно было загуглить решение и исходить от него, но в этом задании не стал пользоваться такой возможностью, а то не смог бы придумать чего-то более-менее отличающегося от предложенного решения. (Похожая ситуация была в задании №1, когда был бэнчмарк, от которого я сильно не смог отойти(в плане предобработки данных). Нужно было анализировать его только после проверки своих идей).