Подготовил А. Зеленцов, 201 группа

Для начала хотел бы отметить, что я не то чтобы очень хорош в обозначениях, поэтому заранее извиняюсь за то, что h_{ij} в некоторых местах подразумевает вектор и за конструкции вроде $sign = \geq$. Приятного (надеюсь) прочтения!

В самом начале перепишем некоторые определения.

Определение 1.

Модель алгоритмов классификации называется **полной** (для данной обучающей и конртольной выборок), если для любой матрицы $H \in \mathrm{R}^{q \times l}$ существует оператор модели B, порождающий матрицу оценок H.

Определение 2.

Модель алгоритмов классификации называется **полной относительно решающего правила** C или C**-полной**, если для любой матрицы $G \in \{0, 1\}^{q \times l}$ существует такой оператор модели B, что алгоритм $A = B \cdot C$ реализует матрицу классификации G.

Определение 3.

Модель алгоритмов классификации называется **полной относительно семейства решающих правил** C^* или C^***-полной**, если для любой матрицы $G \in \{0, 1\}^{q \times l}$ существуют такие оператор модели B и решающее правило C из семейства C^*, что алгоритм $A = B \cdot C$ реализует матрицу классификации G.

Так же напишем определение рассматриваемых решающих правил

$$C_{max}(\parallel h_{ij} \parallel) =\parallel \alpha_{ij} \parallel : \alpha_{ij} = \begin{cases} 1 & h_{ij} = \max_t h_{it} \\ 0 & \text{иначе} \end{cases}$$

Если максимумов несколько, условимся выбирать самый левый

$$C_{c_1, c_2}(\parallel h_{ij} \parallel) =\parallel \alpha_{ij} \parallel : \alpha_{ij} = \begin{cases} 1 & c_1 \geq h_{ij} \geq c_2 \\ 0 & \text{иначе} \end{cases}$$

Для C_{c_1, c_2} правила нужно ввести дополнительное ограничение: $c_1 \geq c_2$, потому что иначе матрица G всегда нулевая. Далее везде будем считать, что это ограничение выполняется, дополнительно не оговаривая этого.

Определение 4.

Модель алгоритмов классификации называется **полной относительно решающего правила** C в задаче с непересекающимися классами, если для любой матрицы $G_{\text{н п}} \in \{G \in \{0, 1\}^{q \times l} | G\tilde{1} = \tilde{1}\}$ существует такой оператор модели B, что алгоритм $A = B \cdot C$ реализует матрицу классификации $G_{\text{н п}}$.

Случай линейной модели

Вначале рассмотрим только линейные модели.

Определение 4.

Модель алгоритмов классификации называется линейной, если существуют такие матрицы $H_1, H_2, \ldots, H_l$, матрица H_j имеет размер $q \times k_j, j \in \{1, 2, \ldots, l\}$, что матрица оценок H получается оператором алгоритма модели тогда и только тогда, когда j-й столбец матрицы H имеет вид $H_j x_j, x_j \in \mathrm{R}^{k_j}$ для всех $j \in \{1, 2, \ldots, l\}$.

Утверждение 1.

Докажем теперь, что если ранг матрицы H_j равен q, то, подобрав x_j, мы можем получить любой столбец.

Действительно, по известной теореме Кронекера-Капелли система $H_j x_j = b_j$ совместна тогда и только тогда, когда $rg(H_j) = rg([H_j, b_j])$. Если ранг H_j равен количеству строк, то при приписывании к ней справа любого столбца ранг не поменяется.

Так же, доказательство последующих теорем будет существенно опираться на следующее утверждение, которое было дано в научной работе, поэтому здесь представлено без доказательства

Утверждение 2.

Система неравенств

$$\begin{cases} m_i^T x \geq c & i \in I \\ m_i^T x < c & i \notin I \end{cases}$$

где m_i^T - i-я строка матрицы $M \in \mathrm{R}^{q \times l}$ совместна (относительно неизвестного x) для всех подмножеств $I \subseteq \{1, 2, ..., q\}$, тогда и только тогда, когда $rg(M) = q$.

C_{max} - полнота.

Теорема 1. (Достаточное условие C_{max} - полноты для линейной модели)

Линейная модель C_{max}-полна, если какие-то $l-2$ матрицы из $H_1, H_2, ..., H_l$ имеют ранг q, при этом для остальных двух матриц (пусть их номера i, j) выполняется $rg([H_i, H_j]) = q$.

Доказательство

Для начала рассмотрим случай, когда все l матриц имеют ранг q. Тогда, по **утверждению 1**, мы можем построить любую матрицу оценок H, а следовательно, любую матрицу классификации $G_{\text{н п}}$.

При условиях теоремы, мы гарантированно можем построить произвольные хотя бы $l-2$ столбцов матрицы H.

Не ограничивая общности, будем считать, что это $l-2$ последних столбцов. Тогда, в условиях нашей теоремы, выходит, что H_1, H_2 имеют ранг $\leq q$ и выполнено $rg([H_1, H_2]) = q$.

Наша задача - уметь подбирать $x_1, x_2, ..., x_l$ таким образом, чтобы в каждой строке максимум был на требуемой позиции.

Для начала докажем, что мы можем обеспечить(здесь и далее "обеспечить" значит - возможно подобрать такие $x_1, x_2, ..., x_l$) любое соотношение $sign \in \{\geq, <\}^q$ для первых двух столбцов в каждой строке одновременно. Если мы докажем это, то в силу того, что мы можем выбрать остальные столбцы произвольными, теорема доказана.

Почему так? Рассмотрим $i-$ю строку. Допустим, нужно обеспечить максимум в позиции k_i. Если $k_i > 2$, то нам без разницы, что стоит в $(i, 1)$ и $(i, 2)$ позициях матрицы H так как мы можем сделать нужный нам элемент сколь угодно большим, а остальные $l-3$ элемента - сколь угодно маленькими в силу **утверждения 1**. Иначе если, к примеру, $k = 2$, то нужно требовать, чтобы элемент в $(i, 2)$ позиции был больше чем в $(i, 1)$ позиции, а остальные элементы строки можно сделать сколь угодно малыми. И соотвествующие соотношения должны выполняться для всех строк одновременно. Если для любого i, $k_i \leq 2$, то нас волнуют все соотношения одновременно.

Перепишем это математически. Для всех подмножеств $I \subseteq \{1, 2, ..., q\}$ должно быть возможно подобрать такие x_1, x_2, чтобы выполнялось:

$$\begin{cases} h_{1i}^T x_1 \geq h_{2i}^T x_2 & i \in I \\ h_{1i}^T x_1 < h_{2i}^T x_2 & i \notin I \end{cases}$$

Где h_{ji} - i-я строка матрицы H_j

Введем $x_2' = -x_2$ и получим:

$$\begin{cases} h_{1i}^T x_1 \geq -h_{2i}^T x_2' & i \in I \\ h_{1i}^T x_1 < -h_{2i}^T x_2' & i \notin I \end{cases}$$

Перепишем по другому:

$$\begin{cases} h_{1i}^T x_1 + h_{2i}^T x_2^{'} \geq 0 & i \in I \\ h_{1i}^T x_1 + h_{2i}^T x_2^{'} < 0 & i \notin I \end{cases}$$

Теперь пусть $M = [H_1, H_2]$, $x = [x_1, x_2^{'}]$, тогда систему выше можно переписать как:

$$\begin{cases} m_i^T x \geq 0 & i \in I \\ m_i^T x < 0 & i \notin I \end{cases}$$

Эта система совместна в силу $rg(M) = q$ и **утверждения 2**.

Таким образом, доказали, что мы всегда можем найти x_1, x_2, чтобы обеспечить любое соотношение $sign \in \{\geq, <\}^q$ для первых двух столбцов в каждой строке одновременно.

Остальные столбцы можно выбрать произвольными $\implies$ можно подобрать $x_1, x_2, \ldots, x_l$ таким образом, чтобы в каждой строке максимум был на требуемой позиции для любого требуемого набора позицией $\implies$ можно построить любую матрицу классификации $G_{\text{н п}}$.

Теорема доказана.

Утверждение 3.

Если линейная модель C_{max} полна, то мы можем обеспечить любое соотношение $sign \in \{\geq, <\}^q$ для любых двух столбцов в каждой строке одновременно.

Допустим линейная модель C_{max} полна и для x-того и y-того столбцов существует такой набор знаков $sign \in \{\geq, <\}^q$, который мы не можем обеспечить. Тогда мы не можем получить матрицу $G_{\text{н п}}$, в которой максимальный элемент для i-й строки:

$$k_i = \begin{cases} x & sign[i] = \geq \\ y & sign[i] = < \end{cases}$$

Противоречие. Утверждение доказано.

Теорема 2. (Необходимое условие C_{max}-полноты для линейной модели)

Если линейная модель C_{max}-полна, то для любых $i, j \in \{1, 2, \ldots, l\} : i < j$ выполнено $rg([H_i, H_j]) = q$.

Доказательство

Используем **утверждение 3** и перепишем его вывод в виде системы:

Для любых $i, j \in \{1, 2, \ldots, l\} : i < j$ и всех подмножеств $K \subseteq \{1, 2, \ldots, q\}$ возможно подобрать такие x_i, x_j, чтобы выполнялось:

$$\begin{cases} h_{ik}^T x_i \geq h_{jk}^T x_j & k \in K \\ h_{ik}^T x_i < h_{jk}^T x_j & k \notin K \end{cases}$$

Где h_{mk} - k-я строка матрицы H_m

Делаем такие же преобразования, как в предыдущей теореме, вводим $M = [H_i, H_j]$, $x = [x_i, x_j^{'}]$ и получаем, что система

$$\begin{cases} m_k^T x \geq 0 & k \in K \\ m_k^T x < 0 & k \notin K \end{cases}$$

совместна для всех подмножеств $K \subseteq \{1, 2, \ldots, q\}$.

Следовательно, по **утверждению 2**, $rg(M) = rg([H_i, H_j]) = q$ для любых $i, j \in \{1, 2, \ldots, l\}$

Теорема доказана.

Замечания к теоремам 1, 2

В случае $l = 2$ ($l < 2$ в данном случае рассматривать не имеет смысла), из необходимого условия начинает следовать достаточное и они "схлопываются" в одно необходимое и достаточное условие.

Почему необходимое условие не является достаточным в общем случае?

Рассмотрим три столбца с номерами i, j, k и два соотношения $sign_{ij}, sign_{ik} \in \{\geq, <\}^q$ между i, j и i, k соответсвенно. Необходимое условие гарантирует нам нахождение таких x_{i1}, x_j и таких x_{i2}, x_k, чтобы соотношения выполнялись. Однако x_{i1} и x_{i2} должны совпадать, а это не гарантировано. В этом месте доказательство достаточности "ломается".

Почему достаточное условие не является необходимым?

Существует контрпример.

Если $l = 3$, $q = 2$,

$$H_1 = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} H_2 = \begin{pmatrix} 2 & 2 \\ 3 & 3 \end{pmatrix} H_3 = \begin{pmatrix} 3 & 3 \\ 4 & 4 \end{pmatrix}$$

то подбирая x_1, x_2, x_3, мы можем построить любую матрицу классификации, однако $rg(H_j) = q - 1$ для любого $j = 1, 2, 3$.

Почему мы можем построить любую матрицу классификации? Потому что один столбец в ней всегда нулевой, соответсвующий x мы подбираем таким, чтобы получились очень маленькие числа, остается случай с $l = 2$ и любые две матрицы удовлетворяют необходимому условию, которое в случае $l = 2$ является достаточным

C_{c_1, c_2} - полнота

Столбцы матрицы классификации G в случае решающего правила C_{c_1, c_2}, вообще говоря, никак не зависят друг от друга. Поэтому зависимости между матрицами H_i, H_j искать бессмысленно. Поэтому мы можем каждую матрицу H_j рассматривать отдельно от остальных.

Сформулируем сначала несколько очевидных условий.

Лемма 1. (Достаточное условие C_{c_1, c_2}-полноты)

Линейная модель C_{c_1, c_2}-полна если для любого $j \in \{1, 2, \ldots, l\}$ $rg(H_j) = q$.

Доказательство

Из **утверждения 1** и рангов матриц $H_1, H_2, \ldots, H_l$ следует, что мы можем построить любую матрицу оценок $H \implies$ любую матрицу классификации G.

Лемма доказана.

Замечание

Необходимым условие выше не является.

Пример $l = 1$

$$H_1 = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}$$

$c_1 = 20, c_2 = 10$

$rg(H_1) = q - 1$

В примере модель полна, но условие не выполняется.

Лемма 2 (необходимое условие C_{c_1, c_2}-полноты)

Если линейная модель C_{c_1, c_2}-полна, то для любого $j \in \{1, 2, \ldots, l\}$ в матрице H_j нет нулевых и одинаковых строк.

Доказательство

Если в матрице H_j есть нулевая строка, то в j-м столбце матрицы оценок всегда стоит 0 (пусть на i-й позиции) $\implies$ если система **S** совместна для подможеств $I: i \in I$, то она не совместна для подможеств $I: i \notin I$ и наоборот $\implies$ система **S** совместна не для всех подмножеств $I \subseteq \{1, 2, \ldots, q\}$.

Если в матрице H_j есть одинаковые строки, то в j-м столбце матрицы оценок два элемента всегда одинаковые(пусть это элементы i и k), $\implies$ система **S** не совместна для случая $i \in I$, $k \notin I$.

Лемма доказана.

Отметим, что **Лемма 1** рассматривает фиксированные c_1, c_2, поэтому и не является необходимым условием. Сформулируем также следующую интересную теорему.

Утверждение 4

Если строки матрицы A линейно зависимы, то если $Ax = b$, то компоненты столбца b линейно зависимы, причем с такими же коэффициентами.

Действительно, распишем столбец b как столбец из скалярных произведений вида $\langle a_i, x \rangle$ и заметим, что по свойствам линейности скалярного произведения утверждение выше выполнено.

Теорема 3. (Критерий полноты относительно любого правила из семейства C_{c_1, c_2})

Линейная модель полна относительно любого решающего правила из семейства C_{c_1, c_2} тогда и только тогда, когда для любого $j \in \{1, 2, \ldots, l\}$ $rg(H_j) = q$

Доказательство достаточности

Лемма 1 работает для любых фиксированных c_1, c_2, поэтому достаточность доказана.

Доказательство необходимости.

Нужно доказать, что если $\exists j \in \{1, 2, \ldots, l\} : rg(H_j) < q$, то $\exists c_1, c_2$: линейная модель не полна относительно C_{c_1, c_2}.

Будем рассматривать случай $c_1 = c_2 \neq 0$ и j-ый столбец матрицы оценок H, $rg(H_j) < q$.

То есть чтобы получить j-ый столбец матрицы классификации G, состоящим только из единиц, нужно получить j-ый столбец мматрицы оценок H состоящим только из c_1.

Если это невозможно, то теорема доказана.

Если это возможно, то заметим, что

1. Так как $rg(H_j) < q$, в матрице есть линейно зависимые строчки.
2. Если, например, k-ая строчка линейно зависит от всех остальных, то k-ая компонента j-го столбца матрицы H линейно зависит от остальных компонент столбца, причем с такими же коэффициентами. Это следует из **утверждения 4**

Рассмотрим k-ую строчку, которая линейно зависит от остальных с коэффициентами $\alpha_1, \ldots, \alpha_{q-1}$. Следовательно $(\alpha_1 + \ldots + \alpha_{q-1}) * c_1 = c_1$. Значит, мы не можем построить j-ый столбец матрицы H, где на всех позициях стоит c_1, а на k-ой позиции - другое число. Значит мы не можем построить j-ый столбец матрицы G, в котором везде единицы, кроме k-ой позиции. Теорема доказана.

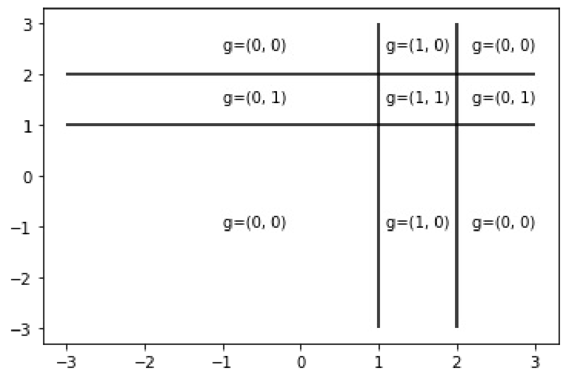
Рассуждения

Переформулируем условие. Рассмотрим матрицу H_j отдельно от остальных, обозначим j-ый столбец матрицы оценок H как b и j-ый столбец матрицы G как g для сокращения записей.

Таким образом, b можно представить как точку в пространстве R^q. И всё пространство можно поделить за зоны в зависимости от того, какой g получается применением рещающего правила к точке из этой зоны.

В случае $q = 2$, если провести ортонормированные оси классическим образом $(0, 1)(1, 0)$ (горизонтальная ось - 1ая компонента, вертикальная - вторая), это выглядит так. Прямые соответствуют значениям c_1, c_2.

In [33]:



Таким образом, для каждой зоны у нас появляется метка - какой вектор g соответсвует этой зоне.

Очевидно, эту картинку можно расширить на q-мерное пространство. В общем случае, нужно провести ортонормированные оси и отметить на каждой оси перпендикулярными к этой оси прямыми точки c_1, c_2. Тогда в центре будет q-мерный куб и будет 2^q меток для зон.

Идея такая: у каждой матрицы есть образ $\{b \mid \exists x_j : H_j * x_j = b\}$: размерность образа равна рангу матрицы. Таким образом переформулировка такая: Линейная модель C_{c_1, c_2}-полна тогда и только тогда, когда образ каждой матрицы H_j содержит хотя бы по одной точке из каждой из групп зон с одинаковыми метками.

Отметим так же, что на картинке выше границы квадрата принадлежат квадрату, а границы полосок - полоскам.

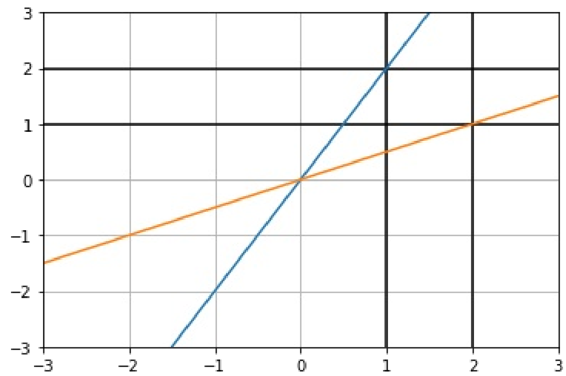
Заметим, что для случая $q = 2$ уже можно сформулировать необходимое и достаточное условие. Образ матрицы в данном случае либо вся плоскость, либо прямая, либо точка (0, 0). Последний случай очевидно не подходит. Первый случай очевидно подходит.

Во втором случае все сильно зависит от c_1, c_2.

Отметим, что если одна из строчек матрицы нулевая, то прямая получается строго вертикальная или горизонтальная, это не подходит, поэтому отбросим этот случай.

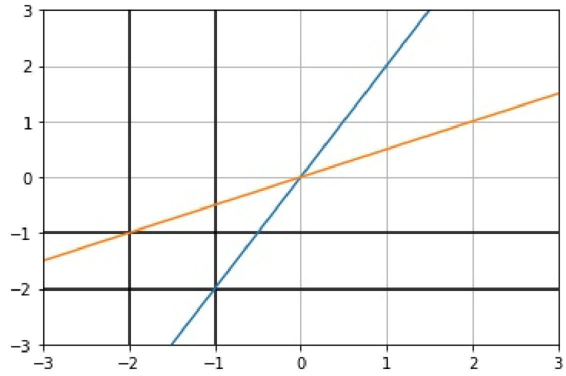
Если c_1, c_2 положительные, то получается, что нам подходят все прямые, лежащие между прямыми с рисунка кроме прямой $x = y$.

In [35]:



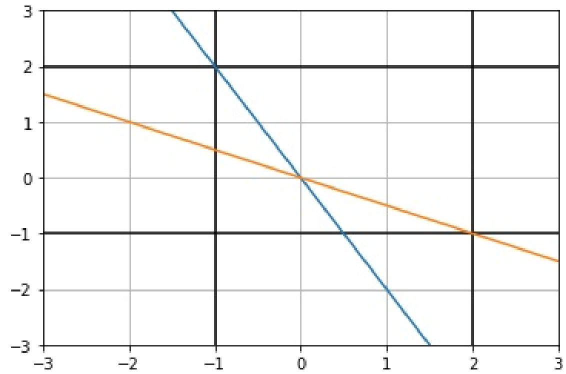
Если c_1, c_2 - оба отрицательные, то по сути ничего не меняется

In [36]:



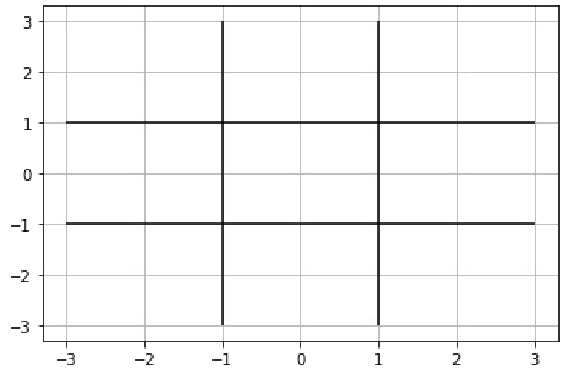
Если разных знаков, то подходят все прямые между изображенными на рисунке

In [43]:



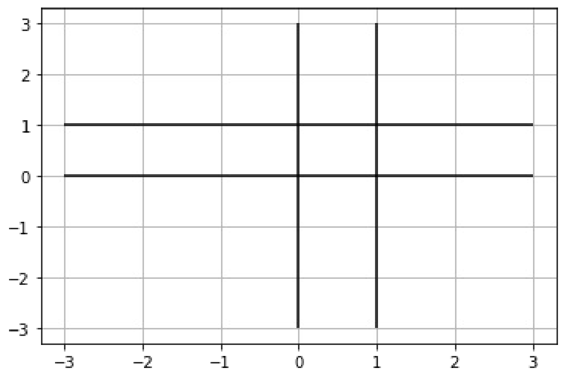
Отдельный неприятный случай $c_1 = -c_2$. В этом случае не подходит ни одна прямая.

In [8]:



Рассмотрим еще крайние случаи. Пусть $c_2 = 0, c_1 > 0$. В этом случае подходит любое $\alpha < 0$. Если наоборот, $c_2 < 0, c_1 = 0$, то подходят все $\alpha > 0$ кроме $\alpha = 1$.

In [9]:



Последний случай $c_1 = c_2$. В этом случае не подходит ни одна прямая.

Таким образом, скажем, что $h_2 = \alpha * h_1$, где h_1, h_2 - строчки матрицы H_j. Нужные нам **ограничения**:

1. $c_1, c_2 \neq 0$, одного знака, $c_1 \neq c_2$ и $\alpha \in [\frac{c_2}{c_1}, \frac{c_1}{c_2}]$ (или наоборот, в зависимости от того, какая из дробей больше) за исключением случая $\alpha = 1$.
2. $c_1, c_2 \neq 0$, разных знаков, $c_1 \neq -c_2$ и $\alpha \in [\frac{c_2}{c_1}, \frac{c_1}{c_2}]$ (или наоборот, в зависимости от того, какая из дробей больше).
3. $c_2 = 0$, $c_1 > 0, \alpha < 0$
4. $c_2 < 0, c_1 = 0, \alpha > 0, \alpha \neq 1$

Теорема 4. (Критерий C_{c_1,c_2}-полноты при $q = 2$)

Если $q = 2$, то

Линейная модель C_{c_1,c_2}-полна тогда и только тогда, когда для каждой матрицы $H_j, j \in \{1, 2, \ldots, l\}$выполнено одно из двух условий:

1. $rg(H_j) = q$
2. $rg(H_j) = q - 1$, матрица не имеет нулевых строчек и выполнено одно из **ограничений** выше.

Отметим что доказательство этой теоремы в обе стороны подробно описано выше.

Таким образом, так как для любого $\alpha \neq 0, \exists c_1, c_2 : \alpha \in [\frac{c_2}{c_1}, \frac{c_1}{c_2}]$, можно сформулировать еще одну теорему:

Теорема 5. (Критерий полноты относительно семейства C_{c_1,c_2} при $q = 2$)

Если $q = 2$, то

Линейная модель полна относительно семейства решающих правил C_{c_1,c_2} тогда и только тогда, когда для каждая матрица $H_j, j \in \{1, 2, \ldots, l\}$ не имеет нулевых и одинаковых строчек.

Доказательство

Матрица не имеет нулевых и одинаковых строчек тогда и только тогда, когда её образ - либо вся плоскость, либо не вертикальная и не горизонтальная прямая, проходящая через начало координат, но не $x = y$. В этом и только в этом случае можно найти подходящие c_1, c_2 и построить нужные полоски и квадрат.

Сформулируем теперь критерий для произвольного q и линейной модели. Следующее утверждение активно используется в доказательстве **теоремы 6**.

Утверждение 5

Имеется линейное уравнение: $\alpha_1 * x_1 + \ldots + \alpha_n * x_n = 0$. Тогда, если $x_i \in [a, b], x_i \neq a, b$. А $x_j = a$ или $x_j = b$, то можно выбрать новые $x_i^{'}, x_j^{'}$ такими, чтобы $x_i^{'} \in [a, b], x_j^{'} \notin [a, b]$ и линейное уравнение оставалось верным.

Действительно, обозначим все остальные слагаемые за $-C$, тогда $\alpha_i * x_i + \alpha_j * x_j = C$, теперь рассмотрим случай $x_j = a$ (для $x_j = b$ аналогично). Скажем, что $x_j^{'} = a - \delta, \delta > 0$. Тогда, чтобы линейное уравнение оставалось верным, нужно $x_i^{'} = x_i + \frac{\alpha_j}{\alpha_i} * \delta$. Подберем $\delta > 0$ столь малым, чтобы $x_i^{'} \in [a, b]$ и утверждение доказано.

Теорема 6 (Критерий C_{c_1,c_2}-полноты)

Линейная модель C_{c_1,c_2}-полна тогда и только тогда, когда для каждой матрицы $H_j, j \in \{1, 2, \ldots, l\}$выполнено одно из двух условий:

1. $rg(H_j) = q$
2. $rg(H_j) = q - 1$ и гиперплоскость образа матрицы "отсекает" одну из вершин q-мерного куба от ее соседних вершин т.е. вершина лежит по разные стороны от гиперплоскости со своими соседями. (В крайнем случае "отсекаемая" вершина может лежать и на гиперплоскости)

Доказательство достаточности

Рассмотрим отдельно какую-то из матриц $H_1, \ldots, H_l$.

1ое условие рассмотрено в **Лемме 1**.

Поэтому рассмтрим 2ое условие. Очевидно, строчки матрицы линейно зависимы. То есть $\exists\{\alpha_1, \ldots, \alpha_q\} : \alpha_1 * h_1 + \ldots + \alpha_q * h_q = 0$, где $h_1, \ldots, h_q$ - строчки матрицы.

Заметим при этом, что в таком случае вектор $\alpha = (\alpha_1, \ldots, \alpha_q)$ - нормаль к гиперплоскости, так как по **утверждению 4** компоненты стобца матрицы оценок линейно зависимы с такими же коэффициентами, а это и есть уравнение гиперплоскости. Напомним, что точки гиперплоскости - столбцы матрицы оценок, которые можно получить.

При этом, если $\exists i : \alpha_i = 0$, то гиперплоскость параллельна одной из осей либо содержит эту ось (т.к. нормаль перпендикулярна этой оси). Такая гиперплоскость не может "отсекать" одну из вершин т.к. параллельна некоторым сторонам куба или содержит их. Поэтому $\forall i : \alpha_i \neq 0$. Значит любая строчка линейно выражается через остальные, причем с ненулевыми коэффициентами.

В таком случае по **утверждению 1**, мы можем выбрать $q - 1$ элементов столбца матрицы оценок произвольными (так как $q - 1$ строчки матрицы линейно независимы), а оставшийся (причем "оставшимся" мы можем выбрать какой угодно элемент) будет их линейной комбинацией с ненулевыми коэффициентами.

Таким образом рассмотрим конкретный столбец матрицы классификаци G (назовём этот столбец g) и докажем, что мы можем построить таковой.

Если в столбце g есть хотя бы два нуля (допустим в позициях i, j), то мы гарантированно можем его построить следующим образом:

1. Подбираем для всех элементов соотвествующего столбца матрицы оценок H кроме i-того и j-того подходящие значения, т.е. такие значения, чтобы после применения решающего правила в столбце g были нужные нам числа.
2. Таким образом, можем переписать уравнение линейной зависимости строк по другому: $\alpha_i * x_i + \alpha_j * x_j = C$, где x_i, x_j - соответствующие элементы столбца матрицы оценок, а C - все остальные слагаемые уравнения гиперплоскости (мы их зафиксировали в пункте 1).
3. На x_i, x_j есть условия: они не должны попадать в ограниченный отрезок $[c_2, c_1]$, т.к. на этих позициях в столбце g стоят 0. То есть точка (x_i, x_j) не входит в квадрат с вершинами $(c_1, c_1), (c_1, c_2), (c_2, c_1), (c_2, c_2)$
4. Такая система уравнений и неравентсв имеет решения x_i, x_j, так как решение равенства $\alpha_i * x_i + \alpha_j * x_j = C$ - прямая, а на прямой в любом случае есть точки, не входящие в ограниченный квадрат, так как она бесконечна.

Далее, если переформулировать 2ое условие теоремы математически,то: Существует такой вектор $c \in \{c_1, c_2\}^q$, что $\langle \alpha, c \rangle \geq 0$ либо ≤ 0, но для любого вектора v, получающегося из c заменой одной компоненты (c_1 на c_2 и наоборот): $\langle \alpha, v \rangle < 0$ либо соответственно > 0. $\langle x, y \rangle$ - скалярное произведение.

Значит, существуют точки, лежащие на всех ребрах куба, выходящих из "отсеченной" вершины, принадлежащие гиперплоскости. Назовем множество этих точек A.Они принадлежат границе куба, а следовательно, мы можем построить столбец g, состоящий только из единиц.

Осталось рассмотреть случай, когда в столбце g ровно один ноль. Допустим он должен быть на i-ой позиции.

Тогда если гиперплоскость не проходит через "отсекаемую" вершину куба, то возьмем точку из множества A таким образом, чтобы на i-ой координате было одно из c_1, c_2. Тогда по свойству множества A, $\exists j:$ на j-ой позиции будет стоять элемент $v \in [c_2, c_1], v \neq c_1, c_2$. Тогда, в силу **утверждения 5**, можем "подвинуть" i-ую и j-ую компоненты таким образом, чтобы i-ая компонента $\notin [c_2, c_1]$, а j-ая $\in [c_2, c_1]$ и линейная зависимость не была нарушена (точка принадлежит гиперплоскости). Тогда получили нужный вектор и достаточность доказана.

А вот если гиперплоскость проходит через "отсекаемую" точку, то множество A "схлопывается" к одной точке - "отсекаемой". Получается что гиперплоскость проходит "покасательно" к кубу, имея с ним только одну общую точку - "отсекаемую" и не имея точек внутри куба. В этом случае будем подобно **утверждению 5** будем "двигать" i-ую и j-ую компоненту не нарушая уравнения линейной зависимости на $0 < \delta < \frac{c_1 - c_2}{2}$ таким образом, чтобы i-ая компонента $\notin [c_2, c_1]$, а j-ая $\in [c_2, c_1]$. Если у нас не получилось это сделать, то выходит, что обе компоненты "двигаются" вместе либо вглубь отрезка $[c_2, c_1]$, либо из него(эквивалентно движению вглубь отрезка со знаком "-"), не нарушая уравнения гиперплоскости. Тогда получается, что гиперплоскость имеет точки внутри куба. Противоречие. Следовательно, мы точно можем "подвинуть" компоненты нужным образом не нарушая линейной зависимости. Значит, мы можем получить нужный стоблец g.

Достаточность доказана.

Доказательство необходимости

Кроме указанных в теореме для произвольной матрицы H_j возможны 3 случая:

1. $rg(H_j) < q - 1$.
2. $rg(H_j) = q - 1$ и гиперплоскость не имеет точек внутри куба или на его границе
3. $rg(H_j) = q - 1$ и гиперплоскость "отсекает" более чем одну вершину куба.

В первом случае мы не можем построить подпространство размерности $q - 2$, чтобы оно содержало в себе точки из зон со всеми метками.

Назовём "рукавами" области с меткой, содержащей только один 0. (В двумерном случае это продолжения полосок, в трехмерном эти объекты напоминают, извините, бесконечные квадратные вентиляционные рукава, выходящие из куба). Таких рукавов всего q.

Отметим, что через любые $q - 1$ точек можем построить линейное многообразие размерности $q - 2$ и притом только одно. Выберем любые $q - 1$ рукавов и отметим в них любые $q - 1$ точек соответственно. Через эти $q - 1$ можно провести одно линейное многообразие размерности $q - 2$. Таким образом, по построению, такое многообразие не будет иметь общих точек с оставшимся рукавом, так как рукав перпендикулярен всем остальным рукавам.

Значит, в этом случае, модель не полна.

Во втором случае не можем построить столбец $(1, \dots, 1)$, так как гиперплоскость не имеет точек внутри куба или на его границе. Модель не полна.

В третьем случае среди отрезанных вершин найдутся те, между которыми есть ребро (соседние). Иначе отрезать более одной вершины нельзя.

Таким образом, $\exists c \in \{c_1, c_2\}^q$ и $\exists i$ такие что если поменять в c значение на i-ой позиции (c_1 на c_2 и наоборот) и если назвать этот вектор $c_{changed}$, то $\langle \alpha, c_{changed} \rangle$ и $\langle \alpha, c \rangle$ будут одного знака, но не равны 0. А значит, $\nexists v : v \in [c_2, c_1]$, такого что если поставить в векторе c на i-ую позицию v и назвать такой новый вектор c_v, то $\langle \alpha, c_v \rangle = 0$. Значит, гиперплоскости не принадлежит ни одной точки, такой, что i-ая координата $\in [c_2, c_1]$. Значит, в столбце g на i-ой позиции не может стоять 1. Значит модель не полна.

Таким образом мы рассмотрели все случаи кроме указанных в теореме и доказали, что в этих случаях модель не полна.

Необходимость доказана.

Замечание

Отметим, что проверять 2ое условие **Теоремы 6** может оказаться вычислительно сложно, так как при наивном подходе нужно перебрать 2^q скалярных произведений. Поэтому на практике гораздо проще пользоваться **Леммами 1 и 2**. Либо придумать не наивный подход, но это еще одна не менее сложная теорема.

Тем не менее, при фиксированном q возможно вычислить конкретные критерии, как мы сделали для случая $q = 2$, введя ограничения на коэффициенты линейной зависимости строк матрицы. Такие критерии проверяются за константное время.

Случай произвольной модели

В общем случае, достаточное условие C_{max}-полноты можно сформулировать так:

Модель классификации C_{max}-полна, если мы можем подобрать параметры модели таким образом, чтобы хотя бы $l - 2$ столбца матрицы оценок H были произвольными, а для остальных двух можно было обеспечить любое соотношение $sign \in \{\geq, <\}^q$

Поставленная выше гипотеза доказывается предположительно по индукции.

Необходимое условие:

Если модель классификации C_{max}-полна, то можно подобрать параметры модели таким образом, чтобы для любых двух столбцов матрицы оценок H выполнялись соотшения $sign \in \{\geq, <\}^q$

Почему это верно - подробно указано в доказательстве для линейной модели.

Достаточное условие C_{c_1, c_2}-полноты:

Если мы можем подобрать параметры модели, таким образом, чтобы матрица оценок H была произвольной, то модель C_{c_1, c_2}-полна.

Условие выше является достаточным для любого полного решающего правила. Решающее правило C будем называть полным, если для любой матрицы классификации G существует матрица оценок H: $\mathrm{C}(H) = G$

Критерий C_{c_1,c_2}-полноты:

Назовем образом модели все возможные матрицы оценок, которые могут получиться подбором параметров. Модель C_{c_1,c_2}-полна тогда и только тогда, когда образ содержит хотя бы по одной точке из каждой из групп зон с одинаковыми метками.

Данный критерий описывается выше, в рассуждениях для линейной модели.