Теоремы об отделимости. Задача линейного программирования

(из лекции 4)

Теорема Хана-Банаха

Если X -банахово пространство, Y – его замкнутое подпространство,

на Y задан линейный непрерывный функционал,

то его можно продолжит на пространство X с сохранением нормы

то есть: $Y \subset X,\ g \in Y^*$
тогда существует $f \in X^*,\ \forall y \in Y,\ f(y) = g(y),\ ||f|| = ||g||$

Теорема об отделимости

Пусть M и N – выпуклые множества в банаховом пространстве X, причем M открытое и $M \cap N = \varnothing$, тогда существует линейный непрерывный функционал, разделяющий эти множества

Полученный результат допускает обобщение, важное для приложений в задачах оптимизации.

Теорема (об отделимости групп множеств)

Пусть X – банахово пространство,
$V_0, \dots, V_{n-1}$ – открытые выпуклые подмножества X,
V_n – выпуклое подмножество X,

$$\bigcap_{k=0}^{n} V_k = \varnothing$$

Тогда существуют открытые полупространства $D_0, \dots, D_n$ такие, что

$$V_k \subset D_k,\ r = 0, \dots, n-1,\ V_n \subset \bar{D}_n,$$

$$\bigcap_{k=0}^{n} D_k = \varnothing$$

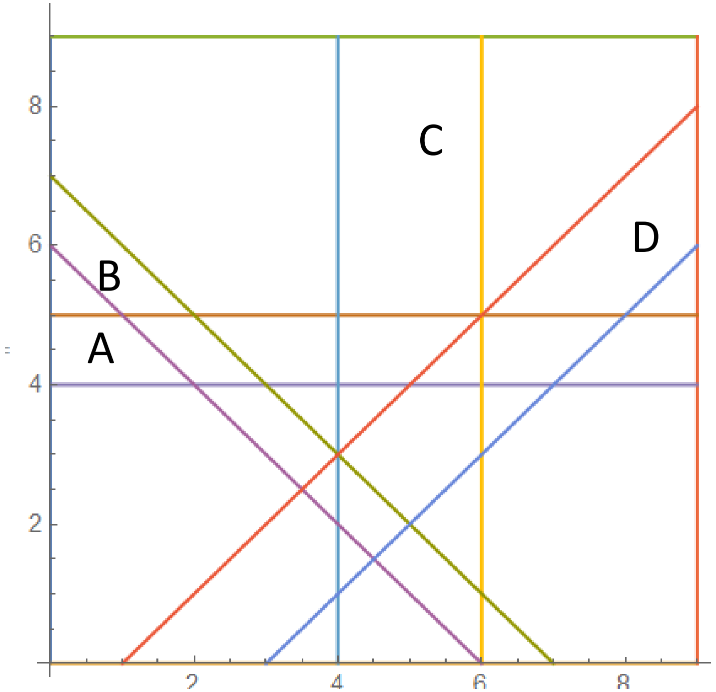
(!)расширение каждого выпуклого множества до полупространства

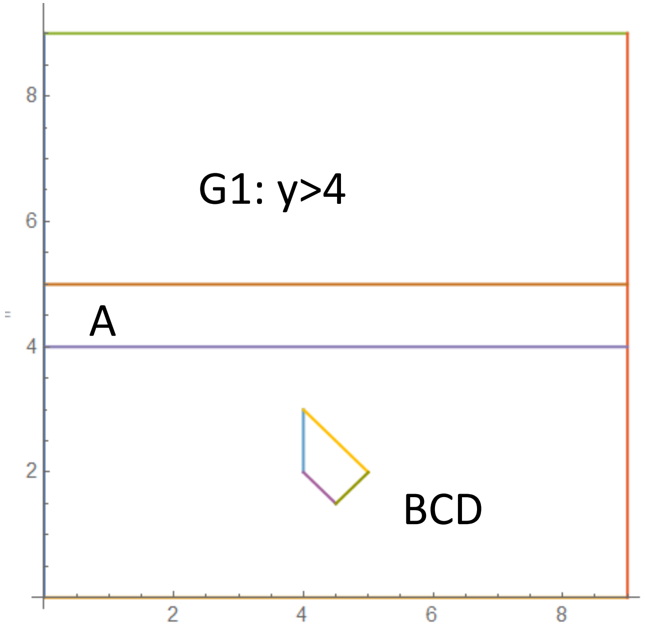
Замечание
Для приложений бывает важно, что одно из множеств не обязано быть открытым. Это обстоятельство не создает принципиальных трудностей, но требует многочисленных технических оговорок в доказательстве. Чтобы избежать этого, здесь будет рассматриваться только случай, когда все множества открыты.

Доказательство По теореме Хана-Банаха найдется полупространство D_0, содержащее множество V_0 и не пересекающееся с множеством $\bigcap\limits_{k=1}^{n} V_k$, так как множества V_0 и $\bigcap\limits_{k=1}^{n} V_k$ выпуклы, открыты и не пересекаются. Здесь и далее используется то обстоятельство, что пересечение открытых выпуклых множеств тоже открыто и выпукло.

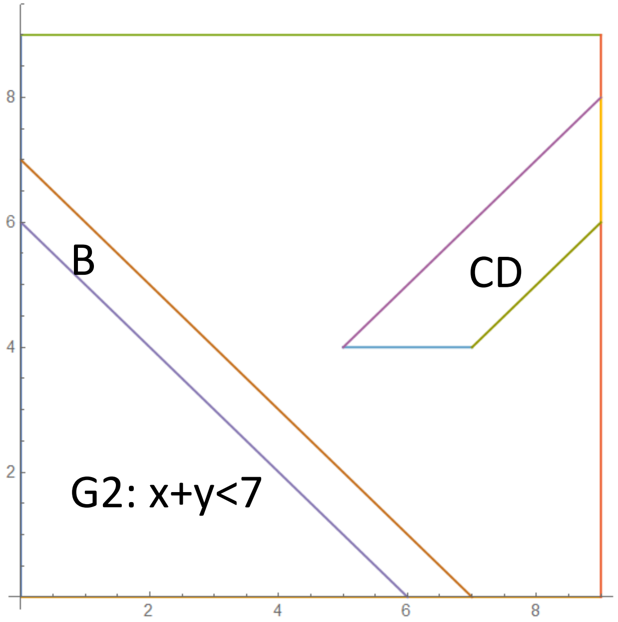
Рассмотрим теперь пару множеств V_1 и $D_0 \bigcap \left(\bigcap\limits_{k=2}^{n} V_k \right)$. По построению эти множества не пересекаются, оба множества открыты и выпуклы. Они не пересекаются, так как если бы существовала точка x, принадлежащая и тому и другому множеству, то эта точка попала бы в пересечение множеств D_0 и $\bigcap\limits_{k=1}^{n} V_k$, что противоречит определению множества D_0.

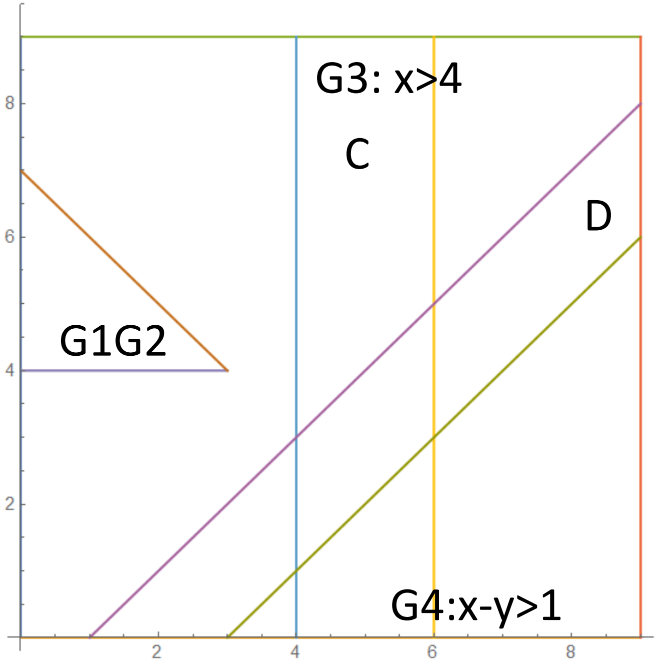
Снова воспользуемся теоремой Хана-Банаха и построим полупространство D_1, содержащее множество V_1 и не пересекающееся с множеством $(D_0 \bigcap D_1) \bigcap \left(\bigcap\limits_{k=2}^{n} V_k \right)$. В ходе этого процесса «вытеснения» на n-м шаге будет построено полупространство D_n, содержащее множество V_n и не пересекающееся с множеством $D_0 \bigcap \left(\bigcap\limits_{k=1}^{n-1} D_k \right)$. Остается заметить, что построенная система полупространств искомая.





Пресечение
G1,G2,G3,G4
пусто





Следующая теорема составляет основу для «перевода» абстрактных алгоритмов функционального анализа в вычислительные процедуры в пространстве R^n.

Теорема (о численном описании пустоты пересечения полупространств)

Пусть $D_k = \{x : f_k(x) < c_k,\ f_k \in X^*,\ c_k \in \mathbb{R},\ k = 1, \dots, n\}$ – произвольное множество полупространств в банаховом пространстве X.

Пересечение этих полупространств пусто тогда и только тогда, когда

существуют числа $\lambda_k \geq 0$ не все равные нулю такие, что

$$\sum_{k=1}^{n} \lambda_k f_k = 0 \text{ и } \sum_{k=1}^{n} \lambda_k c_k \leq 0$$

Доказательство
Необходимость.

Рассмотрим отображение пространства X в $\mathbb{R}^n$,
сопоставляющее $x \in X$ точку $(f_1(x) - c_1, \dots, f_n(x) - c_n) \in \mathbb{R}^n$.

Заметим, что все координаты такой точки не могут быть отрицательны, иначе пересечение полупространств было бы не пусто.

$x : f_k(x) - c_k < 0 \rightarrow x$в пересечении ?!!

Обозначим через A образ всего пространства X при этом отображении. A является сдвигом линейного пространства
$\{y = (y_1, \ldots, y_n) : y_k = f_k(x),\ x \in X\}$ -(c_1,..,c_n)

Как отмечено выше, в образе пространства X не могут находиться точки, все координаты которых отрицательны.
Другими словами, конус $B = \{(y_1, \ldots, y_n) \in R^n : y_k < 0\}$ не пересекается с множеством A.

Поскольку множество A является сдвигом линейного пространства, то существует, содержащая его плоскость $P = \{y \in \mathbb{R}^n : \sum_{k=1}^{n} \lambda_k y_k = \lambda\}$
которая не пересекается с множеством B.

Таким образом, $x \in A \to \sum_{k=1}^{n} \lambda_k (f_k(x) - c_k) = \lambda$
и $\sum_{k=1}^{n} \lambda_k y_k < \lambda$ для всех $y \in B$ $(A \cap B = \emptyset)$.

Проведем анализ коэффициентов.
Числа λ_k неотрицательны, так как наличие $\lambda_k < 0$ сделало бы невозможным неравенство $\sum_{k=1}^{n} \lambda_k y_k < \lambda$.

Число λ не может быть отрицательным, иначе сумма $\sum_{k=1}^{n} \lambda_k y_k$ при подходящих значениях $y \in B$ могла бы принимать значения сколь угодно близкие к 0 и в пределе неравенство было бы нарушено.

Из первого условия следует, что
$\sum\limits_{k=1}^{n} \lambda_k f_k(x) = \sum\limits_{k=1}^{n} c_k \lambda_k + \lambda$
для всех $x \in X$.

Поскольку правая часть равенства не зависит от x, то левая может быть только константой, но при $x = 0$ левая часть равна нулю $(f_k(0) = 0)$, значит,
$\sum\limits_{k=1}^{n} \lambda_k f_k = 0$.

Таким образом, $\sum\limits_{k=1}^{n} c_k \lambda_k + \lambda = 0$ и $\sum\limits_{k=1}^{n} c_k \lambda_k \leq 0$.

Достаточность.
Если выполнены условия
$\sum\limits_{k=1}^{n} \lambda_k f_k = 0$,
все $\lambda_k \geq 0$, $\sum\limits_{k=1}^{n} \lambda_k c_k \leq 0$
и пересечение полупространств не пусто,

то для точки x_* из этого пересечения были бы выполнены соотношения:
$f_k(x_*) < c_k$ и $\sum\limits_{k=1}^{n} \lambda_k f_k(x_*) < \sum\limits_{k=1}^{n} c_k \lambda_k \leq 0$.

Но это означало бы, что $\sum\limits_{k=1}^{n} \lambda_k f_k(x_*) \neq 0$, а это противоречит сделанному предположению.

Задача

$C[0,1], 0 < x_1 < x_2 < x_3 < 1$
$f \in C[0,1],\ F_j(f) = a_{j1} f(x_1) + a_{j1} f(x_2) + a_{j3} f(x_3),\ j = 1, 2, 3$

Докажите, что F_j линейно независимы тогда и только тогда, когда $\det A \neq 0$

При каких коэффициентах $c_k,\ k = 1, 2, 3$ полупространства $D_k = \{x : F_k(x) < c_k\}$ имеют непустое пересечение

Для описания решения задачи линейной оптимизации последнюю теорему надо модифицировать, выделив в ней один из функционалов. При этом в доказательстве будет использовано следующее простое утверждение.

Предложение

Если два открытых множества имеют непустое пересечение, то замыкание их пересечения равно пересечению замыканий.

Теорема

Пусть $g, f_1, \ldots, f_n \in X^*$, $c_k \in R$, $D_k = \{x \in X : f_k(x) < c_k\}$.

Если $\bigcap\limits_{k=1}^{n} D_k \neq \varnothing$,

$g \in X^*$ и $g(x) \geq 0$ при $x \in \overline{\bigcap\limits_{k=1}^{n} D_k}$,

то существуют неотрицательные числа λ_k, такие что $g + \sum\limits_{k=1}^{n} \lambda_k f_k = 0$.

Доказательство

Обозначим $D_0 = \{x : g(x) < 0\}$. Из условия теоремы следует, что

$D_0 \bigcap \left(\bigcap\limits_{k=1}^{n} \bar{D}_k \right) = \varnothing.$

можно поменять порядок замыкания и объединения

$\bigcap\limits_{k=1}^{n} \bar{D}_k = \overline{\bigcap\limits_{k=1}^{n} D_k}.$

Следовательно,

$D_0 \bigcap \left(\bigcap\limits_{k=1}^{n} D_k \right) \subset D_0 \bigcap \left(\overline{\bigcap\limits_{k=1}^{n} D_k} \right) = \varnothing.$

По доказанной выше теореме найдутся неотрицательные коэффициенты λ_k, такие что $\lambda_0 g + \sum\limits_{k=1}^{n} \lambda_k f_k = 0.$

Остается заметить,что $\lambda_0 \neq 0$ так как из той же теоремы следовало, что $\bigcap\limits_{k=1}^{n} D_k = \varnothing$

Еще одна вспомогательная теорема показывает, каким может быть функционал, обращающийся в ноль на пересечении ядер нескольких других функционалов.

Теорема
Предположим, что $g, f_1, \ldots, f_n \in X^*$.
Если $g(x) = 0$ для $x \in \bigcap\limits_{k=1}^{n} \ker f_k$,

то существуют вещественные числа α_k, такие что $g = \sum\limits_{k=1}^{n} \alpha_k f_k$.

Доказательство
Рассмотрим отображение $\phi(x) = (f_1(x), \ldots, f_n(x))$ банахова пространства X в пространство $\mathbb{R}^n$.

На образе пространства X в $\mathbb{R}^n$ построим еще одно отображение ψ по формуле $\psi(y) = g(x) \in \mathbb{R}$, если $y = \phi(x)$.

Поскольку одному y может соответствовать несколько x, необходимо проверить корректность определения ψ, то есть убедиться, что функционал g принимает на них одинаковые значения. Пусть $\phi(x_1) = \phi(x_2)$, тогда $x_1 - x_2 \in \bigcap\limits_{k=1}^{n} \ker f_k$ и по условию $x_1 - x_2 \in \ker g$. Следовательно, $g(x_1) = g(x_2)$.
Очевидно, что ψ – линейное отображение из $\mathbb{R}^n$ в $\mathbb{R}$. Всякое такое отображение имеет вид $\psi(y) = \sum\limits_{k=1}^{n} \alpha_k y_k$. Остается заметить, что $g(x) = \psi(\phi(x)) = \sum\limits_{k=1}^{n} \alpha_k f_k$.

Перейдем к рассмотрению задачи **линейной оптимизации**.

Пусть в банаховом пространстве X задан «многоугольник»
$A = \bigcap\limits_{j=1}^{n} \{x : f_j(x) < c_j\}$ (здесь $f_j \in X^*$, $c_j \in \mathbb{R}$),

многоугольник рассечен сдвинутым линейным пространством
$Y = \{y : g_k(y) = d_k\}$, здесь $g_k \in X^*,\ d_k \in \mathbb{R},\ k = 1, \ldots, m$.

Задан функционал $h \in X^*$,

требуется найти его максимум на $A \cap Y$.

Поставленная задача будет решена в несколько расширенной постановке, а именно, будет дано описание точек $\bar{x} \in A \cap Y$, для которых

$h(\bar{x}) = \max\{h(x) : x \in A \cap Y\}$

Теорема (решение задачи линейного программирования)
Элемент $\bar{x}$ является решением поставленной выше задачи оптимизации тогда и только тогда, когда существуют числа

$\lambda_j \geq 0$ ($j \in J(\bar{x})$ и μ_k ($k = 1, \ldots, m$), такие что

$$h = \sum_{j \in J} \lambda_j f_j + \sum_{k=1}^{m} \mu_k g_k$$

здесь J – множество индексов «выхода на грань»: $J(\bar{x}) = \{j : f_j(\bar{x}) = c_k\}$.

Замечание
Появление в формулировке множества индексов J важно для приложений. Это можно понять на простом примере вычисления наибольшего значения линейной функции на треугольнике. Без упоминания множества J теорема 8.6 утверждала бы очевидный факт: наибольшее значение достигается в одной из вершин, а в существующей формулировке указывается вершина, в которой достигается наибольшее значение. Формулировка теоремы может создавать ощущение о трудности ее применения в практических задачах, но это не так. Выяснив, по какой минимальной системе функционалов из числа f_j можно разложить функционал h так, чтобы коэффициенты были положительны, (*тоже что ДЗ 1*)
можно точно описать то множество точек, на котором функционал принимает наибольшее значение.

В ДЗ 3 обсуждается именно это обстоятельство

Для доказательства теоремы важно следующее свойство совокупности индексов J. Оно имеет простой геометрический смысл. Функционал h достигает наибольшего значения в вершине, на ребре или грани многогранника A.
Следующее предложение показывает, что рядом с этим множеством принадлежность точки многограннику зависит только от функционалов с номерами из J.

Предложение
Пусть $C = \{x \in A : h(x) = h(\bar{x})\}$.
Тогда для любой достаточно малой окрестности V множества C выполняется равенство

$$V \bigcap \left(\bigcap_{j \in J} \{x : f_j(x) < c_j\} \right) = V \cap A.$$

Доказательство теоремы

1) Рассмотрим сначала случай, когда нет ограничений вида $g_k(y) = d_k$. В этом случае условие теоремы можно записать так: множества (полупространства)
$D_j = \{x : f_j(x) < c_j\},\ j = 1, \ldots, n$, таковы, что
$A = = \bigcap\limits_{j=1}^{n} \{x : f_j(x) < c_j\} \neq \varnothing$, и
$D_0 = \{x : h(x) > h(\bar{x})\}$ таково, что $D_0 \cap A = \varnothing$.

По доказанной выше теореме это утверждение равносильно существованию чисел $\lambda_j \geq 0$ таких, что $h = \sum\limits_{j=1}^{n} \lambda_j f_j$.

Для завершения доказательства достаточно показать, что сумму можно «сократить», то есть ограничить сумму индексами, попадающим в множество J.

Та же теорема гарантирует, что для этого достаточно доказать, что функционал h достигает в точке $\bar{x}$ своего наибольшего значения на множестве $\bigcap\limits_{j\in J} D_j$.

Ранее доказанное предложение утверждает, что существует отрытое множество V, содержащее точку $\bar{x}$ такое, что

$$A \cap V = V \bigcap \Big(\bigcap_{j\in J} D_j \Big).$$

По условию на этом множестве $h(x) < h(\bar{x})$.
Возьмем произвольную точку из множества $x_1 \in \bigcap\limits_{j \in J} D_j$

Поскольку это пересечение выпукло и точка $\bar{x}$ лежит на его границе, то отрезок $\{x = x_1 + t(\bar{x} - x_1) : 0 \le t < 1\}$ содержится в этом множестве.

Так как множество V открыто и точка $\bar{x}$ принадлежит границе этого множества, то на отрезке найдется точка $x_2 = x_1 + t_2(\bar{x} - x_1) \in V$. Как уже было отмечено, в этом случае $h(\bar{x}) > h(x_2)$.

Покажем теперь, что $h(x_2) > h(x_1)$
$(1 - t_2)x_1 = x_2 - t_2\bar{x}$, откуда
$(1 - t_2)(x_1 - x_2) = t_2x_2 - t_2\bar{x}$, следовательно,
$(1 - t_2)(h(x_1) - h(x_2)) = t_2(h(x_2) - h(\bar{x}))$ и $h(x_2) > h(x_1)$.

2) Рассмотрим общий случай, когда присутствуют ограничения вида $g_k(y) = d_k$. Чтобы перейти к линейным пространствам проведем сдвиг $x \to x - \bar{x}$, то есть переместим точку, где достигается наибольшее значение, в 0. При этом в ограничениях изменятся постоянные: c_j на $c_j - f_j(\bar{x})$, d_k на $d_k - g_k(\bar{x}) = 0$. Не меняя обозначений для функционалов, запишем эквивалентную задачу: выяснить необходимые и достаточные условия для того, чтобы 0 доставлял решение экстремальной задачи $h(\bar{x}) = \max\{h(x) : x \in A \cap Y\}$.

Сузим задачу на линейное пространство Y, которое после сдвига стало линейным пространством $\{x : g_k(x) = 0,\ k = 1, \ldots, m\}$.

Обозначим через $\tilde{f}_j$ и $\tilde{h}$ сужения функционалов f_j и h на линейное пространство Y. Задача на наибольшее значение «наследуется» пространством Y, но при этом теряет ограничения, связанные с равенствами. Такая задача решена в первой части доказательства.

Функционал $\tilde{h}$ является ее решением тогда и только тогда, когда существуют числа

$\lambda_j \geq 0$ такие, что $\tilde{h} = \sum_{j \in \tilde{J}} \lambda_j \tilde{f}_j$
(здесь $\tilde{J} = \{j : \tilde{f}_j = c_j - f_j(\bar{x})\}$).
Заметим, что множества $\tilde{J}$ и J совпадают.

Чтобы завершить доказательство, надо вернуться в пространство X. Это легко сделать, поскольку функционалы f_j являются продолжениями $\tilde{f}_j$ с пространства Y на X. Чтобы получить все решения, надо добавить к имеющемуся любой из функционалов, обращающийся в 0 на пространстве Y. Из теорем, доказанных ранее следует, что 0 является решением задачи оптимизации тогда и только тогда, когда

$$h + \sum_{j \in J} \lambda_j f_j + \sum_k \mu_k g_k = 0.$$

Для завершения доказательства остается сделать обратный сдвиг, переводящий точку 0 в $\bar{x}$.

Интересна история доказательства этой теоремы. Впервые оно было получено советским математиком С. В. Канторовичем в 1938 г., как решение практической задачи оптимизации раскроя материала на реальном производстве. Оно было внедрено на производстве и дало хороший эффект, но не прижилось, поскольку предприятие перестало выполнять план по сдаче отходов производства. Тем не менее Канторович опубликовал построенный алгоритм, остававшийся решением теоретической задачи [5]. Общая задача линейного программирования была впервые поставлена в 1947 г. Данцигом и Вудом. Первое успешное решение задачи линейного программирования на ЭВМ было проведено в 1952 г. Алгоритм стал очень популярным в экономических моделях. В итоге в 1975 г. Канторовичу была присуждена Нобелевская премия по экономике. В дальнейшем оказалось, что идеи, заложенные в алгоритм решения задачи, можно перенести на задачу выпуклой оптимизации (поиск наибольшего значения выпуклого функционала на выпуклом множестве). Первый шаг в этом направлении приведен выше – это теорема об отделимости группы выпуклых множеств. Следует отметить, что математический аппарат необходимый для решения этой задачи в общем виде значительно сложнее. Подробное описание этих вопросов имеется в книге [6].