ЛЕКЦИЯ 6

Аннотация. Выпуклые множества.

В этой лекции мы работаем исключительно с аффинными пространствами над полем $\mathbb{R}$ действительных чисел (часто — просто с $\mathbb{R}^n$); прямого аналога излагаемой теории для других полей не существует.

Определение 1. Пусть L — аффинное пространство над $\mathbb{R}$. Множество $X \subset L$ называется выпуклым, если для любых его двух точек $a, b \in X$ и любого числа $t \in [0,1]$ точка $ta + (1-t)b$ также принадлежит X.

Эквивалентное определение:

Определение 1'. Множество $X \subset L$ называется выпуклым, если для любого конечного набора точек $a_1, \dots, a_m \in X$ и любого набора неотрицательных чисел $x_1, \dots, x_m \geq 0$ таких, что $x_1 + \dots + x_m = 1$, точка $x_1 a_1 + \dots + x_m a_m$ также принадлежит X.

Доказательство эквивалентности: очевидно, определение 1 — частный случай ($m = 2$) определения 1', так что если для множества X выполнено определение 1', то выполнено и 1. Обратно, пусть $X \subset L$ обладает свойством 1; докажем индукцией по $m \geq 2$, что и 1' выполнено. Шаг индукции (базу мы предположили выполненной): $a \stackrel{\text{def}}{=} x_1 a_1 + \dots + x_m a_m + x_{m+1} a_{m+1} = (1 - x_{m+1}) b + x_{m+1} a_{m+1}$, где $b = \frac{x_1}{1-x_{m+1}} a_1 + \dots + \frac{x_1}{1-x_m} a_m$ (отметим, что сумма коэффициентов равна 1, так что точка b определена). Поскольку $x_{n+1} = 1 - (x_1 + \dots + x_m) \leq 1$, все коэффициенты в барицентрической комбинации для b неотрицательны, так что по предположению индукции $b \in X$. Тогда по определению 1 также и $a \in X$ — шаг индукции сделан.

Простейшие свойства выпуклых множеств:

Теорема 1. (1) *Пересечение любого семейства выпуклых множеств (в том числе и бесконечного) выпукло или пусто.*
(2) *Образ и прообраз выпуклого множества при аффинном отображении — выпуклое множество.*

Доказательство. Свойство 1 очевидно (точно очевидно?). Докажем свойство 2: пусть $f : K \to L$ — аффинное отображение, а $X \subset K$ выпукло. Пусть $b_1, b_2 \in f(X) \subset L$, то есть $b_1 = f(a_1), b_2 = f(a_2)$, где $a_1, a_2 \in X$, и пусть $t \in [0,1]$. Тогда по свойству аффинных отображений (теорема 4 дополнительной лекции 5А) $tb_1 + (1-t)b_2 = f(ta_1 + (1-t)a_2)$. Поскольку X выпукло, точка $ta_1 + (1-t)a_2 \in X$, откуда $tb_1 + (1-t)b_2 \in f(X)$, то есть образ $f(X)$ — выпуклое множество.

Про прообраз: пусть $Y \subset L$ выпукло, и $a_1, a_2 \in f^{-1}(Y)$, то есть $f(a_1), f(a_2) \in Y$. Для произвольного $t \in [0,1]$ имеем $f(ta_1 + (1-t)a_2) = tf(a_1) + (1-t)f(a_2) \in Y$, поскольку Y выпукло — иными словами, $ta_1 + (1-t)a_2 \in f^{-1}(Y)$, что и означает, что прообраз $f^{-1}(Y)$ — выпуклое множество. □

Примеры выпуклых множеств:

Пример 1. Аффинное подпространство (в произвольном аффинном пространстве) выпукло.

Пример 2. Луч $\mathbb{R}_+ \stackrel{\text{def}}{=} \{a \mid a \geq 0\} \subset \mathbb{R}$ — выпуклое множество (проверьте!). Как следствие (выпуклости луча и теоремы 1), если $\ell : L \to \mathbb{R}$ — аффинное отображение (в данном случае его называют также линейной неоднородной функцией: если аффиное пространство L это $\mathbb{R}^n$, то $\ell(x_1, \dots, x_n) = a_1 x_1 + \dots + a_n x_n + b$ для некоторых чисел $a_1, \dots, a_n, b \in \mathbb{R}$), то прообраз луча $\ell^{-1}(\mathbb{R}_+) \stackrel{\text{def}}{=} \{x \in L \mid \ell(x) \geq 0\}$ — выпуклое множество; такое множество называют *полупространством.*

Пример 3. Пусть $B_r \stackrel{\text{def}}{=} \{(x_1, \dots, x_n) \mid x_1^2 + \dots + x_n^2 \leq r^2\}$ — шар радиуса r с центром в начале координат. Докажем лемму:

Лемма 1. $B_r = \{(x_1, \dots, x_n) \mid a_1 x_1 + \dots + a_n x_n \leq r \, \forall (a_1, \dots, a_n) \in B_1\}$.

Доказательство. Выражение $(x_1 - ta_1)^2 + \dots + (x_n - ta_n)^2 = (x_1^2 + \dots + x_n^2) - 2t(a_1 x_1 + \dots + a_n x_n) + t^2(a_1^2 + \dots + a_n^2)$ неотрицательно при всех $t \in \mathbb{R}$ (строго положительно, если векторы $(a_1, \dots, a_n)$ и $(x_1, \dots, x_n)$ не пропорциональны, то есть линейно независимы). Это квадратный трехчлен; таким образом, его дискриминант неотрицателен (строго положителен, если векторы линейно независимы): $(a_1 x_1 + \dots + a_n x_n)^2 \leq (x_1^2 + \dots + x_n^2)(a_1^2 + \dots + a_n^2)$ для всех вещественных a_i и x_i. Отсюда вытекает, что если $(x_1, \dots, x_n) \in B_r$, а $(a_1, \dots, a_n) \in B_1$, то $-r \leq a_1 x_1 + \dots + a_n x_n \leq r$.

Обратно, пусть $(x_1, \dots, x_n)$ таково, что $a_1x_1 + \dots + a_nx_n \le r$ для всех $(a_1, \dots, a_n) \in B_1$; докажем, что $(x_1, \dots, x_n) \in B_r$. Будем считать, что не все x_i равны нулю (если все, то утверждение очевидно). Обозначим $X = x_1^2 + \dots + x_n^2 > 0$ и выберем $a_i = x_i/\sqrt{X}$ для всех $i = 1, \dots, n$ — тогда $a_1^2 + \dots + a_n^2 = (x_1^2 + \dots + x_n^2)/X = 1$, то есть $(a_1, \dots, a_n) \in B_1$. Но тогда $X = x_1^2 + \dots + x_n^2 = (a_1x_1 + \dots + a_nx_n)\sqrt{X} \le r\sqrt{X}$, откуда вытекает, что $\sqrt{X} \le r$, то есть $(x_1, \dots, x_n) \in B_r$. □

На геометрическом языке лемма означает, что шар B_r является пересечением (бесконечного числа) полупространств, заданных неравенствами $a_1x_1 + \dots + a_nx_n \le r$ для всех $(a_1, \dots, a_n) \in B_1$. Из примера 2 и свойства 1 теоремы 1 вытекает, что шар — выпуклое множество. Это относится к шарам с любым центром — они получаются из шаров B_r параллельным переносом (а это аффинное преобразование).

Определение 2. *Выпуклой оболочкой* подмножества $X \subset L$ (обозначение $\operatorname{co} X$) называется пересечение всех выпуклых подмножеств $Y \subset L$ таких, что $X \subset Y$.

Из определения видно, что выпуклая оболочка любого множества X выпукла и содержит множество X как подмножество; если X само выпукло, то его выпуклая оболочка совпадает с ним самим (почему?).

Теорема 2.
(1) *Выпуклая оболочка $\operatorname{co} X$ множества $X \subset L$ состоит из всех барицентрических комбинаций $x_1a_1 + \dots + x_ma_m$ (для всевозможных $m \ge 1$), где точки $a_1, \dots, a_m \in X$, а веса $x_1, \dots, x_m > 0$, $x_1 + \dots + x_m = 1$.*
(2) *Если $\dim L = n$, то для любой точки $b \in \operatorname{co} X$ найдется $r \le n + 1$ точек $a_1, \dots, a_r \in X$ таких, что $b \in \operatorname{co}\{a_1, \dots, a_r\}$.*

Утверждение 2 называется теоремой Каратеодори.

Доказательство. Обозначим Y множество всевозможных барицентрических комбинаций $x_1a_1 + \dots + x_ma_m$ (для всевозможных $m \ge 1$), где точки $a_1, \dots, a_m \in X$, а веса $x_1, \dots, x_m > 0, x_1 + \dots + x_m = 1$ (это называется выпуклой комбинацией). Утверждение 1 заключается в том, что $Y = \operatorname{co} X$. По определению выпуклой оболочки $X \subset \operatorname{co} X$, так что все точки $a_1, \dots, a_m \in \operatorname{co} X$. Но $\operatorname{co} X$ выпукло (как пересечение выпуклых множеств), так что по определению 1' любая выпуклая комбинация этих точек лежит в $\operatorname{co} X$ — это означает, что $Y \subseteq \operatorname{co} X$.

Пусть теперь $a = x_1a_1 + \dots + x_ma_m$ и $b = y_1b_1 + \dots + y_kb_k$ — произвольные точки Y; здесь имеется в виду, что $a_1, \dots, a_m, b_1, \dots, b_k \in X$, $x_1, \dots, x_m, y_1, \dots, y_k > 0$ и $x_1 + \dots + x_m = 1 = y_1 + \dots + y_k$. Возьмем $t \in [0, 1]$, тогда $c \stackrel{\text{def}}{=} ta + (1-t)b = tx_1a_1 + \dots + tx_ma_m + (1-t)y_1b_1 + \dots + (1-t)y_kb_k$. Это выпуклая комбинация точек из X: все коэффициенты $ta_i \ge 0$, $(1-t)b_j \ge 0$, и $tx_1 + \dots + tx_m + (1-t)y_1 + \dots + (1-t)y_k = t + (1-t) = 1$, то есть $c \in Y$. Тем самым доказано, что Y — выпуклое множество; поскольку оно содержит X, получается $\operatorname{co} X \subseteq Y$ и, значит, $Y = \operatorname{co} X$.

Для доказательства утверждения 2 (теоремы Каратеодори) отметим начальную точку $o \in L$, что дает возможность (см. начало лекции 5) отождествить L с векторным пространством V размерности n. Возьмем теперь произвольную точку $b \in \operatorname{co} X$; по доказанному выше $b = x_1a_1 + \dots + x_ma_m$ — выпуклая комбинация точек $a_1, \dots, a_m \in X$. Пусть $m \ge n + 2$; тогда векторы $a_2 - a_1, \dots, a_m - a_1 \in V$ линейно зависимы, т.е. существуют числа $\mu_2, \dots, \mu_m$, не все равные нулю, для которых $\mu_2(a_2 - a_1) + \dots + \mu_m(a_m - a_1) = 0$. Полагая $\mu_1 \stackrel{\text{def}}{=} -\mu_2 - \dots - \mu_m$, получим $\mu_1a_1 + \dots + \mu_ma_m = 0$ и $\mu_1 + \dots + \mu_m = 0$.

Из последнего равенства следует, что среди чисел $\mu_1, \dots, \mu_m$ есть положительные. Пусть $t \stackrel{\text{def}}{=} \frac{x_i}{\mu_i} \stackrel{\text{def}}{=} \min\{\frac{x_j}{\mu_j} \mid j : \mu_j > 0\}$. Тогда все числа $y_j \stackrel{\text{def}}{=} x_j - t\mu_j \ge 0$ (включая и те j, для которых $\mu_j < 0$!), а $y_i = 0$. С другой стороны, $b = (x_1 - t\mu_1)a_1 + \dots + (x_m - t\mu_k)a_m = y_1a_1 + \dots + \widehat{y_ia_i} + \dots + y_ma_m$ ("крышка" означает пропуск соответствующего слагаемого), причем $y_1 + \dots + y_m = (x_1 + \dots + x_m) - t(\mu_1 + \dots + \mu_m) = 1$, — таким образом, b это выпуклая комбинация $(m-1)$ точек $a_1, \dots, \widehat{a_i}, \dots, a_m \in X$. Тем самым количество точек в выпуклой комбинации можно уменьшать, пока оно не станет $m \le n + 1$. □

Пример 4. Пусть L — плоскость, $X = \{a_1, a_2, a_3, a_4, a_6\}$ — множество вершин правильного шестиугольника. По утверждению 1 теоремы 2 выпуклая оболочка $\operatorname{co} X$ содержит все точки вида $ta_i + (1-t)a_j$, $0 \le t \le 1$, то есть все отрезки и все диагонали шестиугольника. Если b — точка на таком отрезке или диагонали, то, поскольку $\operatorname{co} X$ выпукло, оно содержит все точки вида $ta_i + (1-t)b$, то есть все точки всех отрезков, соединяющих вершины со всеми точками на сторонах и диагоналях. Нетрудно видеть, что объединение всех таких отрезков — сам шестиугольник S. С другой стороны, шестиугольник S выпуклый и содержит точки $a_1, \dots, a_6$ — следовательно, $\operatorname{co} X \subseteq S$. Тем самым $\operatorname{co} X = S$.

Утверждение 2 означает, что каждая точка $b \in S$ лежит внутри какого-нибудь треугольника (здесь $n = 2$, так что $n+1 = 3$) — это действительно так. При этом треугольник для разных точек $b \in S$ может быть разный (ни один треугольник не совпадает с S) и не обязан быть единственным (в данном конкретном примере он никогда не единственный).

Еще свойства выпуклых оболочек в конечномерных аффинных пространствах. Везде ниже L — аффинное пространство и $\dim L = n$.

Теорема 3 (Радона). *Любые $n+2$ точки $a_1, \ldots, a_{n+2} \in L$ можно разбить на два непустых непересекающихся подмножества, выпуклые оболочки которых пересекаются.*

Пример 5. Среди трех точек на прямой одна лежит между двумя другими и, следовательно, является их выпуклой комбинацей.

Если три из 4 точек на плоскости лежат на одной прямой, то одна из них является выпуклой комбинацией двух других, что доказывает в этом случае теорему (четвертую точку можно присоединить к любому из подмножеств). Если таких трех точек нет, то 4 точки являются вершинами либо выпуклого четырехугльника, либо невыпуклого. В первом случае каждое из двух подмножеств содержит по 2 точки — концы диагоналей; диагонали (выпуклые оболочки пар своих концов) пересекаются. Во втором случае одна из точек — вершина угла, большего развернутого — составляет одно подмножество и лежит внутри треугольника, образованного тремя оставшимися точками, которые образуют второе подмножество.

Доказательство теоремы Радона. Как и при доказательстве теоремы Каратеодори, отождествим L и V (паралллельное L векторное пространство разиерности n). Векторы $a_1 - a_{n+2}, \ldots, a_{n+1} - a_{n+2} \in V$ линейно зависимы: существуют числа $x_1, \ldots, x_{n+1}$, не все равные нулю и такие, что $\sum_{i=1}^{n+1} x_i(a_i - a_{n+2}) = 0$. Полагая $x_{n+2} \overset{\text{def}}{=} -\sum_{i=1}^{n+1} x_i$, получим $\sum_{i=1}^{n+2} x_i a_i = 0$ и $\sum_{i=1}^{n+2} x_i = 0$. Поскольку не все x_i равны нулю, множества $I_+ \overset{\text{def}}{=} \{i \mid x_i > 0\}$ и $I_- \overset{\text{def}}{=} \{i \mid x_i < 0\}$ непусты. Положим $c \overset{\text{def}}{=} \sum_{i \in I_+} x_i > 0$, тогда $\sum_{i \in I_-} x_i = -c$. Тогда точка $a \overset{\text{def}}{=} \sum_{i \in I_+} \frac{x_i}{c} a_i = -\sum_{i \in I_-} \frac{x_i}{c} a_i$ принадлежит выпуклой оболочке множества $\{a_i \mid i \in I_+\}$ и выпуклой оболочке множества $\{a_i \mid i \in I_-\}$. Точки a_i, для которых $x_i = 0$ (если такие существуют), можно присоединить к любому из двух подмножеств. $\square$

Теорема 4 (теорема Хелли). *Если $X_1, \ldots, X_N \subset L$ — конечный набор выпуклых множеств такой, что каждые $n+1$ из них имеют общую точку, то все они имеют общую точку.*

Доказательство. Если $N \le n+1$, то теорема тривиальна. Если $N \ge n+2$, то применим индукцию по N; база $N = n+2$. Для каждого $i = 1, \ldots, n+2$ пусть a_i — общая точка всех $X_1, \ldots, X_{n+2}$, кроме, возможно, X_i (она существует по условию). Согласно теореме Радона, существует разбиение $\{1, \ldots, n+2\} = I_+ \sqcup I_-$ на два непустых множества и существует точка $a \in \mathrm{co}(\{a_i \mid i \in I_+\}) \cap \mathrm{co}(\{a_i \mid i \in I_-\})$. Докажем, что $a \in \bigcap_{j=1}^{n+2} X_j$. Рассмотрим произвольное число $j \in \{1, \ldots, n+2\}$; без ограничения общности $j \in I_+$. Точки $a_i, i \ne j$, все принадлежат X_j; в частности, все точки $a_i, i \in I_-$, принадлежат X_j (поскольку $i \in I_+$). Следовательно, $\mathrm{co}(\{a_i \mid i \in I_-\}) \subset X_j$ (напомним, что X_j выпукло!) и, таким образом, $a \in X_j$. Поскольку j произвольно, $a \in \bigcap_{j=1}^{n+2} X_j$, и тем самым пересечение непусто.

Пусть теперь $N \ge n+3$ (шаг индукции) и для $N-1$ множеств теорема уже известна. Рассмотрим $N-1$ множеств: $X_1, \ldots, X_{N-2}$ и $Y \overset{\text{def}}{=} X_{N-1} \cap X_N$, и выберем любые $n+1$ из них. Если все это множества X_i, то они пересекаются по условию теоремы. Если же одно из множеств Y, а остальные $X_{i_1}, \ldots, X_{i_n}$, то в наборе из $n+2$ множеств $X_{i_1}, \ldots, X_{i_n}, X_{N-1}, X_N$ любые $n+1$ имеют общую точку — следовательно, все $n+2$ также имеют общую точку (база индукции!) и, таким образом, общую точку имеют $X_{i_1}, \ldots, X_{i_n}, Y$. Значит, набор $N-1$ множеств $X_1, \ldots, X_{N-2}, Y$ удовлетворяет условию теоремы и, по предположению индукции, имеет общую точку, которая является общей точкой всех множеств $X_1, \ldots, X_N$. $\square$

Пример 6. Рассмотрим конечную систему линейных уравнений и неравенств (строгих и нестрогих) от n переменных $x_1, \ldots, x_n$. Из теоремы Хелли и примера 2 вытекает такое утверждение: если любые $n+1$ из этих уравнений и неравенств имеют общее решение, то и вся система имеет решение.