

1 凸集

1.1 仿射集合和凸集

直线和线段

设 $x_1 \neq x_2$ 为 R^n 空间中的两个点，具有如下形式：

$$y = \theta x_1 + (1 - \theta)x_2, \theta \in R$$

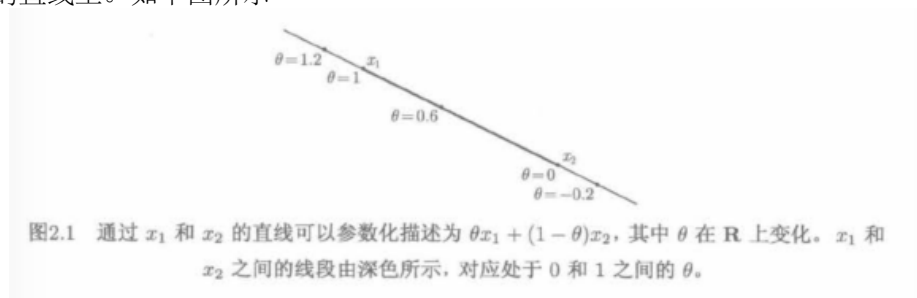
组成一条穿越 x_1, x_2 的直线。参数 $\theta = 0$ 对应 $y = x_2$ ，而 $\theta = 1$ 表明 $y = x_1$ 。

参数 θ 的值在 0 和 1 之间波动，构成了 x_1, x_2 之间的（闭）线段。

y 的表示形式：

$$y = x_2 + \theta(x_1 - x_2)$$

另一种解释： y 是基点 x_2 （对应 $\theta = 0$ ）和方向 $(x_1 - x_2)$ （由 x_2 指向 x_1 ）乘以参数 θ 的和。因此 θ 给出了 y 在由 x_2 通向 x_1 的路上的位置。当 θ 由 0 增加到 1，点 y 相应地由 x_2 移动到 x_1 。如果 $\theta > 1$ ，点 y 在超越了 x_1 的直线上。如下图所示：



仿射集合

如果通过集合 $C \subseteq R^N$ 中任意两个不同点的直线仍然在集合 C 中，那么称集合 C 是仿射的。也就是说， $C \subseteq R^N$ 是仿射等价于：对于任意 $x_1, x_2 \in C, \theta \in R$ 有 $\theta x_1 + (1 - \theta)x_2 \in C$ 。换言之， C 包含了 C 中任意两点的系数之和为 1 的线性组合。

ATTENTION! 这里并没有限制 θ 的大小，也就是说经过两点的直线仍在 C 中。

扩展到多个点。如果 $\theta_1 + \cdots + \theta_k = 1$, 称具有 $\theta_1 x_1 + \cdots + \theta_k x_k$ 形式的点为 x_1, \dots, x_k 的仿射组合。

结论: 一个仿射集合包含其中任意点的仿射组合, 如果 C 是一个仿射集合, $x_1, \dots, x_k \in C, \theta_1 + \dots + \theta_k = 1$, 那么 $\theta_1 x_1 + \cdots + \theta_k x_k$ 仍然在 C 中。

如果 C 是一个仿射集合且 $x_0 \in C$, 则集合

$$V = C - x_0 = \{x - x_0 | x \in C\}$$

是一个子空间, 即关于加法和数乘是封闭的。

设 $v_1, v_2 \in V, \alpha, \beta \in R$ 则有 $v_1 + x_0 \in C, v_2 + x_0 \in C$, 因为 C 是仿射的, 且 $\alpha + \beta + (1 - \alpha - \beta) = 1$, 所以

$$\alpha v_1 + \beta v_2 + x_0 = \alpha(v_1 + x_0) + \beta(v_2 + x_0) + (1 - \alpha - \beta)x_0 \in C$$

由 $\alpha v_1 + \beta v_2 + x_0 \in C$, 可知 $\alpha v_1 + \beta v_2 \in V$

因此仿射集合 C 可以表示为:

$$C = V + x_0 = \{v + x_0 | v \in V\}$$

即一个子空间加上一个偏移。与仿射集合 C 相关联的子空间 V 与 x_0 的选取无关, 所以 x_0 可以是 C 中的任意一点。

我们定义仿射集合 C 的维数为子空间 $V = C - x_0$ 的维数, 其中 x_0 是 C 中的任意元素

我们称由集合 $C \subseteq R^n$ 中的点所有仿射组合组成的集合为 C 的仿射包, 记为 $\text{aff } C$:

$$\text{aff } C = \{\theta_1 x_1 + \cdots + \theta_k x_k | x_1, \dots, x_k \in C, \theta_1 + \cdots + \theta_k = 1\}$$

仿射包是包含 C 的最小的仿射集合, 也就是说: 如果 S 是满足 $C \subseteq S$ 的仿射集合, 那么 $\text{aff } C \subseteq S$

仿射维数与相对内部

我们定义集合 C 的仿射维数为其仿射包的维数。

如果集合 $C \subseteq R^n$ 的仿射维数小于 n , 那么这个集合在仿射集合 $aff C \neq R^n$ 中。我们定义集合 C 的相对内部为 $aff C$ 的内部, 记为 $relint C$, 即

$$relint C = \{x \in C | B(x, r) \cap aff C \subseteq C \text{ for some } r > 0\}$$

其中 $B(x, r) = \{y | \|y - x\| \leq r\}$ 即半径为 r , 中心为 x 并由范数 $\|\cdot\|$ 定义的球。

我们定义集合 C 的相对边界为 $cl C$
 $relint C$, 此处 $cl C$ 表示 C 的闭包

凸集

集合 C 被称为凸集, 如果 C 中任意两点间的线段仍然在 C 中, 即对任意 $x_1, x_2 \in C$ 和满足 $0 \leq \theta \leq 1$ 的 θ 都有

$$\theta x_1 + (1 - \theta)x_2 \in C$$

如果集合中的没一点都可以被其他点沿着他们之间一条无阻碍的路径看见, 那么这个集合就是凸集。所谓无阻碍, 是指整条路径都在集合中。由于仿射集合包含穿过集合中任意不同两点的整条直线, 任意不同两点间的线段自然也在集合中。因而仿射集是凸集。

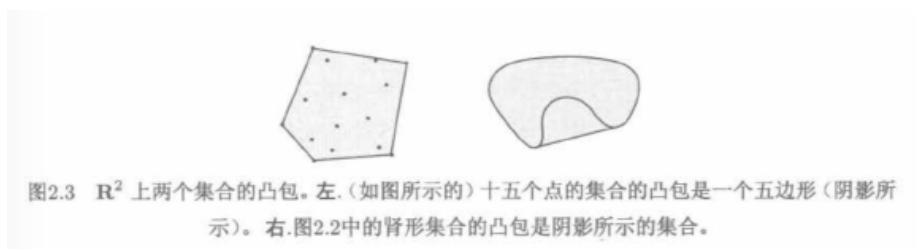
我们称点 $\theta_1 x_1 + \dots + \theta_k x_k$ 为点 x_1, \dots, x_k 的一个凸组合, 其中 $\theta_1 + \dots + \theta_k = 1, \theta_i \geq 0, i = 1, \dots, k$, 一个集合是凸集等价于集合包含其中所有点的凸组合。



我们称集合 C 中所有点的凸组合的集合为其凸包, 记为 $conv C$:

$$conv C = \{\theta_1 x_1 + \dots + \theta_k x_k | x_i \in C, \theta_i \geq 0, i = 1, \dots, k, \theta_1 + \dots + \theta_k = 1\}$$

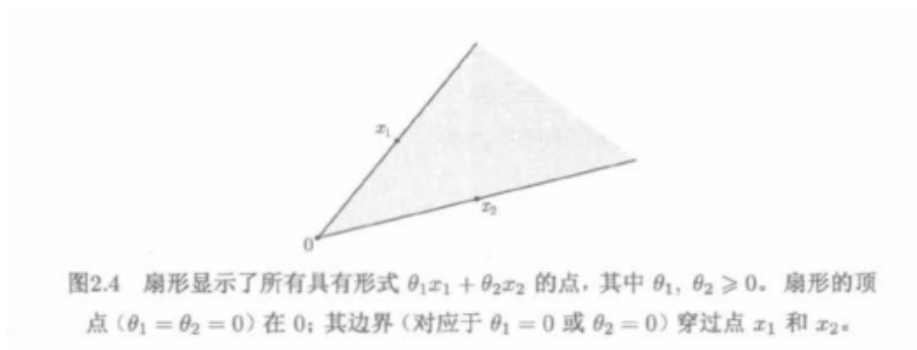
凸包 $conv C$ 总是凸的。它是包含 C 的最小的凸集。也就是说, 如果 B 是包含 C 的凸集, 那么 $conv C \subseteq B$



锥

如果对于任意 $x \in C, \theta \geq 0$ 都有 $\theta x \in C$, 我们称集合 C 是锥或者是非负齐次。如果集合 C 是锥, 并且是凸的, 则称 C 为凸锥, 对于任意 $x_1, x_2 \in C, \theta_1, \theta_2 \geq 0$ 都有

$$\theta_1 x_1 + \theta_2 x_2 \in C$$



具有 $\theta_1 x_1 + \cdots + \theta_k x_k, \theta_1, \dots, \theta_k \geq 0$ 形式的点成为 x_1, \dots, x_k 的锥组合 (非负线性组合)

集合 C 的锥包是 C 中元素的所有锥组合的集合, 即:

$$\{\theta_1 x_1 + \cdots + \theta_k x_k \mid x_i \in C, \theta_i \geq 0, i = 1, \dots, k\}$$

它是包含 C 的最小的凸锥

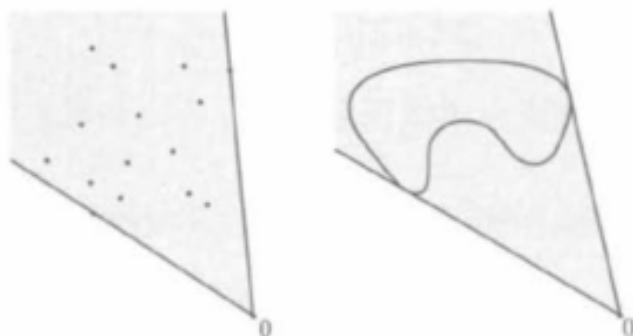


图2.5 图2.3中两个集合的锥包(阴影所示)。

1.2 重要的例子

超平面与半空间

超平面是具有以下形式的集合

$$\{x | a^T x = b\}$$

其中 $a \in R^n, a \neq 0, b \in R$. 超平面是关于 x 的非平凡线性方程的解空间(因此是一个仿射集合)。几何上, 超平面可以解释为与给定向量 a 的内积为常数的点的集合; 也可以看成法线方向为 a 的超平面, 而常数 $b \in R$ 决定了这个平面从原点的偏移。

超平面还可以表示成如下形式:

$$\{x | a^T (x - x_0) = 0\}$$

其中 x_0 是超平面上的任意一点。

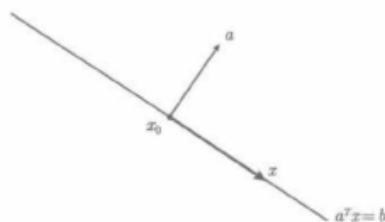
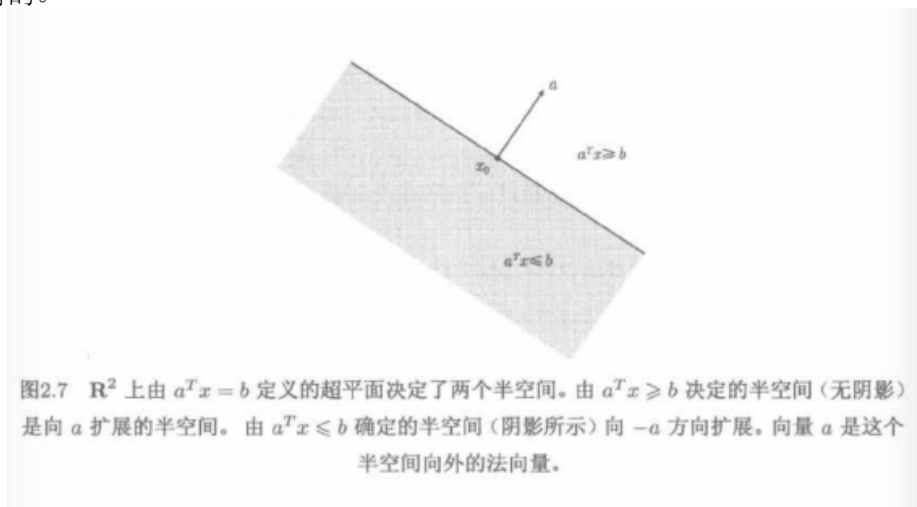


图2.6 R^2 中由法向量 a 和超平面上一点 x_0 确定的超平面。对于超平面上任意一点 x , $x - x_0$ (如深色箭头所示) 都垂直于 a 。

一个超平面将 R^n 划分为两个半空间。半空间有如下形式：

$$\{x | a^T x \leq b\}$$

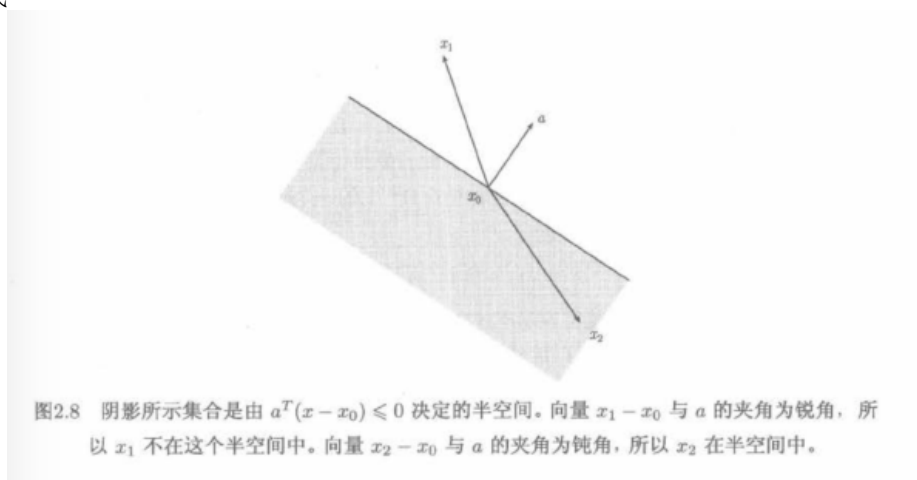
即非平凡的线性不等式的解空间。其中 $a \neq 0$ 。半空间是凸的，但不是仿射的。



半空间也可以表示为

$$\{x | a^T (x - x_0) \leq 0\}$$

几何解释：半空间由 x_0 加上任意与（向外的法）向量 a 呈钝角的向量组成



Euclid 球和椭球

R^n 中的空间 Euclid 球, 具有以下形式:

$$B(x_c, r) = \{x | \|x - x_c\|_2 \leq r\} = \{x | (x - x_c)^T (x - x_c) \leq r^2\}$$

其中 $r > 0$, $\|\cdot\|_2$ 表示 Euclid 范数, 即 $\|u\|_2 = (u^T u)^{1/2}$, 向量 x_c 是球心, 标量 r 为半径。 $B(x_c, r)$ 由距离球心 x_c 距离不超过 r 的所有点组成。另一个常见的表达式为,

$$B(x_c, r) = \{x_c + ru | \|u\|_2 \leq 1\}$$

一类相关的凸集是椭球, 具有如下形式:

$$\varepsilon = \{x | (x - x_c)^T P^{-1} (x - x_c) \leq 1\}$$

其中 $P = P^T \succ 0$, 即 P 为对称正定矩阵。向量 $x_c \in R^n$ 为椭球的中心。矩阵 P 决定了椭球从 x_c 向各个方向扩展的幅度。 ε 的半轴长度由 $\sqrt{\lambda_i}$ 给出, λ_i 为 P 的特征值。球可以看成是 $P = r^2 I$ 的椭球。

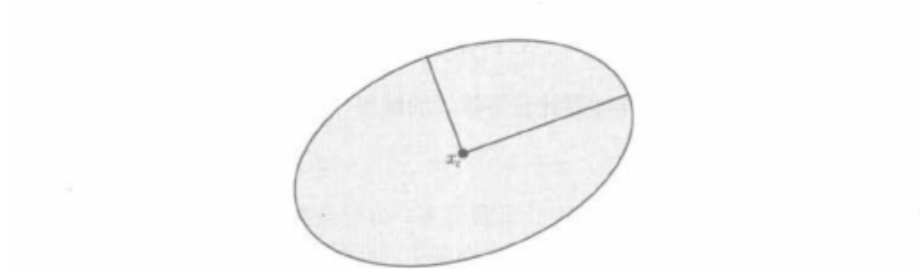


图2.9 阴影所示为 R^2 中的一个椭球。中心 x_c 由点显示, 两个半轴由线段显示

椭球另一个常用的表示形式

$$\varepsilon = \{x_c + Au | \|u\|_2 \leq 1\}$$

其中 A 是非奇异的方阵。

设 $\|\cdot\|$ 是 R^n 中的范数。由范数的一般性质知, 以 r 为半径, x_c 为球心的范数球 $\{x | \|x - x_c\| \leq r\}$ 是凸的。

关于范数 $\|\cdot\|$ 的范数锥是集合

$$C = \{(x, t) | \|x\| \leq t\} \subseteq R^{n+1}$$

是一个凸锥

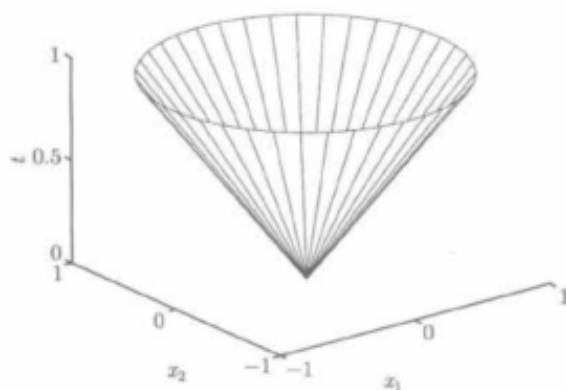


图2.10 \mathbf{R}^3 中二阶锥 $\{(x_1, x_2, t) \mid (x_1^2 + x_2^2)^{1/2} \leq t\}$ 的边界

多面体

多面体被定义为有限个线性等式和不等式的解集

$$P = \{a_j^T x \leq b_j, j = 1, \dots, m, c_j^T x = d_j, j = 1, \dots, p\}$$

因此多面体是有限个半空间和超平面的交集。

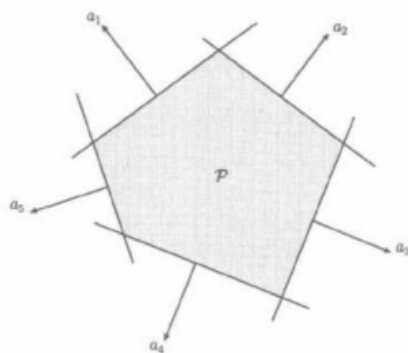


图2.11 多面体 P (阴影所示) 是外法向量为 a_1, \dots, a_5 的五个半空间的交集。

单纯形是一类重要的多面体。设 $k+1$ 个点 $v_0, \dots, v_k \in \mathbf{R}^n$ 仿射独立，即 $v_1 - v_0, \dots, v_k - v_0$ 线性独立，那么这些点决定了一个单纯形，如下

$$C = \text{conv}\{v_0, \dots, v_k\} = \{\theta_0 v_0 + \dots + \theta_k v_k \mid \theta \succeq 0, \mathbf{1}^T \theta = 1\}$$

其中 $\mathbf{1}$ 表示所有分量均为 1 的向量。这个单纯形的仿射维数为 k ，因此也成为 R^n 空间的 k 维单纯形。

思考： $\{\theta_0 v_0 + \dots + \theta_k v_k | \theta \succeq 0, \mathbf{1}^T \theta = 1\}$ 这一串其实就是凸集的定义吧， $\theta \succeq 0$ 其实就是向量 θ 的 $\theta_i \geq 0$ ， $\mathbf{1}^T \theta = 1$ 就是 $\theta_0 + \dots + \theta_k = 1$ 所以 $\theta \succeq 0, \mathbf{1}^T \theta = 1$ ，有 $x = \theta_0 v_0 + \theta_1 v_1 + \dots + \theta_k v_k$ 。

如果定义 $y = (\theta_1, \dots, \theta_k)$ 和

$$B = [v_1 - v_0 \ \dots \ v_k - v_0] \in R^{n \times k}$$

$x \in C$ 的充要条件是

$$x = v_0 + By$$

对于 $y \succeq 0, \mathbf{1}^T y \leq 1$ 成立

$$\begin{aligned} x &= v_0 + By = v_0 + (v_1 - v_0)\theta_1 + \dots + (v_k - v_0)\theta_k \\ &= v_0 + v_1\theta_1 + \dots + v_k\theta_k - (\theta_1 + \dots + \theta_k)v_0 \\ &= v_0\theta_0 - v_0\theta_0 + v_0 + v_1\theta_1 + \dots + v_k\theta_k - (\theta_1 + \dots + \theta_k)v_0 \\ &= (v_0\theta_0 + v_1\theta_1 + \dots + v_k\theta_k) + v_0 - (\theta_0 + \dots + \theta_k)v_0 \\ &= (v_0\theta_0 + v_1\theta_1 + \dots + v_k\theta_k) + v_0 - v_0 \\ &= (v_0\theta_0 + v_1\theta_1 + \dots + v_k\theta_k) \end{aligned}$$

v_0, \dots, v_k 仿射独立意味着矩阵 B 的秩为 k 。因此，存在非奇异矩阵 $A = (A_1, A_2) \in R^{n \times n}$ 使得

$$AB = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} B = \begin{bmatrix} I \\ 0 \end{bmatrix}$$

所以

$$Ax = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} x = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} (v_0 + By) = \begin{bmatrix} A_1 v_0 + y \\ A_2 v_0 \end{bmatrix}$$

所以 $x \in C$ 当且仅当 $A_2 x = A_2 v_0$ 并且向量 $y = A_1 x - A_1 v_0$ 满足 $y \succeq 0, \mathbf{1}^T y \leq 1$

换言之，我们得到了 $x \in C$ 的充要条件

$$A_2x = A_2v_0, A_1x \succeq A_1v_0, \mathbf{1}^T A_1x \leq 1 + \mathbf{1}^T A_1v_0$$

这些是 x 的线性等式和不等式，因此描述了一个多面体
多面体的凸包描述：有限集合 $\{v_1, \dots, v_k\}$ 的凸包是

$$\text{conv}\{v_1, \dots, v_k\} = \{\theta_1 v_1 + \dots + \theta_k v_k \mid \theta \succeq 0, \mathbf{1}^T \theta = 1\}$$

它是一个有界的多面体。

1.3 保凸运算

1. 交集

交集运算时保凸的：如果 S_1, S_2 是凸集，那么 $S_1 \cap S_2$ 也是凸集。

可以通过将集合表示为半空间的交集来表示集合的凸性。反过来，每一个闭的凸集 S 是半空间的交集

2. 仿射函数

假设函数 $f : R^n \rightarrow R^m$ 是仿射的，具有 $f(x) = Ax + b$ 的形式，其中 $A \in R^{m \times n}, b \in R^m$ 。假设 $S \subseteq R^n$ 是凸的，并且 $f : R^n \rightarrow R^m$ 是仿射函数。那么 S 在 f 下的象

$$f(S) = \{f(x) \mid x \in S\}$$

是凸的。

3. 线性分式及透视函数

透视函数：定义 $P : R^{n+1} \rightarrow R^n, P(z, t) = z/t$ ，定义域为 $\text{dom} P = R^n \times R_{++}$ 透视函数对向量进行伸缩，或称为规范化，使得最后一维分量为 1 并舍弃。

注释 2.1 我们用小孔成像来解释透视函数。 (\mathbf{R}^3) 中的小孔照相机由一个不透明的水平面 $x_3 = 0$ 和一个在原点的小孔组成, 光线通过这个孔在 $x_3 = -1$ 呈现出一个水平图像。在相机上方 x ($x_3 > 0$) 处的一个物体, 在相平面的点 $-(x_1/x_3, x_2/x_3, 1)$ 处形成一个图像。忽略象点的最后一维分量 (因为它恒等于 -1), x 处的点的象在象平面上呈现于 $y = -(x_1/x_3, x_2/x_3) = -P(x)$ 处。图2.15显示了这个过程。

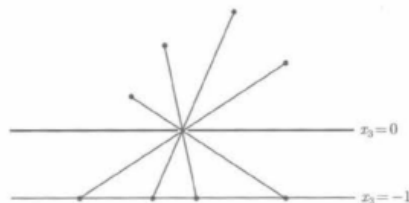


图2.15 透视函数的小孔成像解释。深色的水平直线表示 \mathbf{R}^3 中的平面 $x_3 = 0$, 除了在原点处有一个小孔外, 它是不透光的。平面之上的物体或光源呈现在浅色水平直线所示的象平面 $x_3 = -1$ 上。源位置向其象位置的映射对应于透视函数。

线性分式函数: 由透视函数和仿射函数复合而成。函数 $f: \mathbf{R}^n \rightarrow \mathbf{R}^m$

$$f(x) = (Ax + b)/(c^T x + d), \text{dom} f = \{x | c^T x + d > 0\}$$

1.4 广义不等式

称锥 $K \subseteq \mathbf{R}^n$ 为正常锥, 如果满足以下条件: 1.K 是凸的 2.K 是闭的 3.K 是实的, 具有非空内部 4.K 是尖的, 即不包含直线

正常锥 K 可以用来定义广义不等式, 即 \mathbf{R}^n 上的偏序关系。

$$x \preceq_K y \leftrightarrow y - x \in K$$

严格的偏序关系

$$x \prec_K y \iff y - x \in \text{int}K$$

最小与极小元

如果对于每个 $y \in S$, 均有 $x \prec_K y$, 称 $x \in S$ 是 S 的最小元。如果一个集合有最小, 且是唯一的, 就是极小元。一个集合可以有多个极小元。

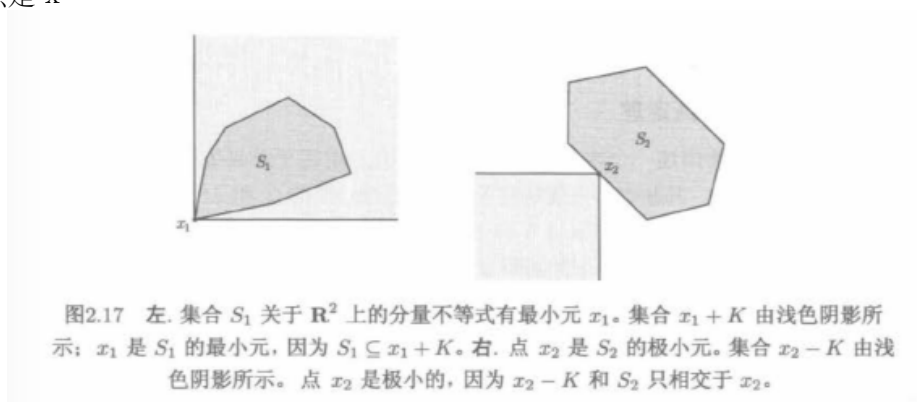
元素 $x \in S$ 是 S 中的一个最小元, 当且仅当

$$S \subseteq x + K$$

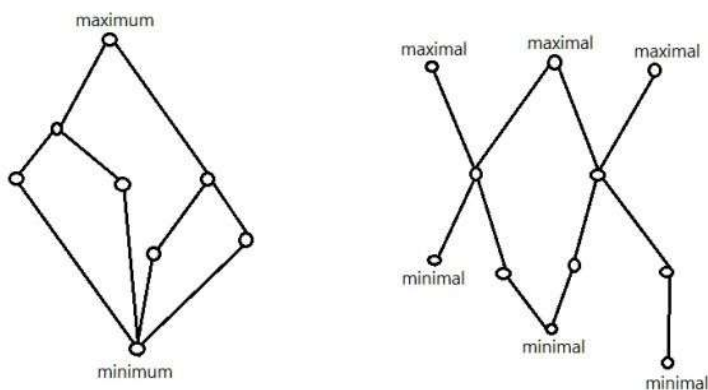
$x+K$ 表示可以与 x 相比并且大于或等于 x 的所有元素
 元素 $x \in S$ 是 S 中的极小元, 当且仅当

$$(x - K) \cap S = \{x\}$$

$x-K$ 可以表示与 x 相比并且小于等于 x 的所有元素, 它与 S 的唯一共同点是 x



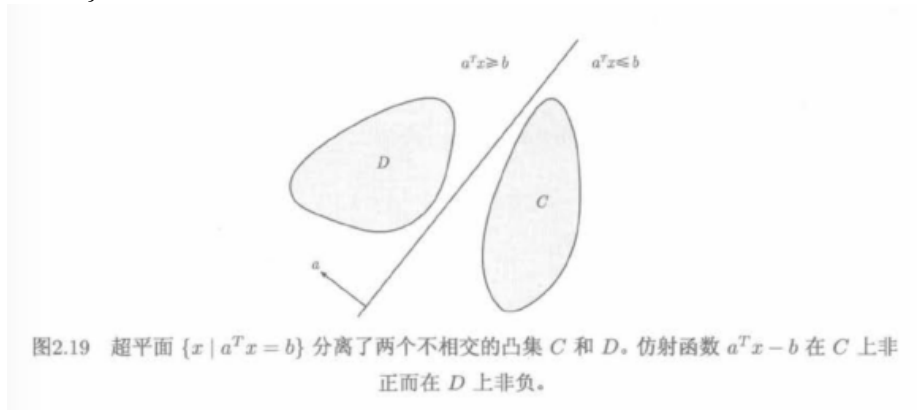
补充: 最小 (minimum): 表示偏序集中其他的点都比它大; 极小 (minimal), 表示该偏序集中没有比它更小的点, 即该偏序集中的其他点要么比它大, 要么和它不能够比较大小



1.5 分离与支撑超平面

超平面分离定理

假设 C 和 D 是两个不相交的凸集, 即 $C \cap D = \phi$, 那么存在 $a \neq 0, b$ 使得所有 $x \in C$ 有 $a^T x \leq b$, 对于所有的 $x \in D$ 有 $a^T x \geq b$. 超平面 $\{x | a^T x = b\}$ 成为集合 C 和 D 的分离超平面。

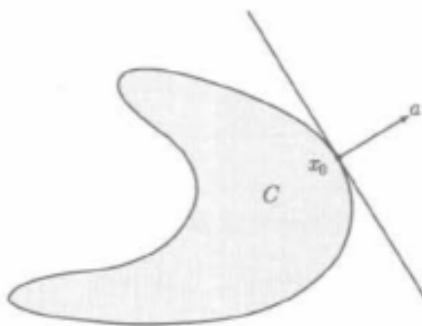


支撑超平面

设 $C \subseteq R^n$ 而 x_0 是其边界 bdC 上的一点, 即

$$x_0 \in bdC = clC \cap intC$$

如果 $a \neq 0$, 并且对任意 $x \in C$ 满足 $a^T x \leq a^T x_0$, 那么称超平面 $\{x | a^T x = a^T x_0\}$ 为集合 C 在点 x_0 处的支撑超平面。几何解释是超平面 $\{x | a^T x = a^T x_0\}$ 与 C 相切于点 x_0 , 而且半空间 $\{x | a^T x \leq a^T x_0\}$ 包含 C 。



1.6 对偶锥与广义不等式

令 K 为一个锥。集合

$$K^* = \{y | x^T y \geq 0, \forall x \in K\}$$

成为 K 的对偶锥。 K^* 是一个锥，并且总是凸的，即使 K 不是凸锥。

从几何上看， $y \in K^*$ 当且仅当 $-y$ 是 K 在原点的一个之称超平面的法线，如图



图2.22 左. 以 y 为内法向量的半空间包含锥 K ，因此， $y \in K^*$ 。
右. 以 z 为内法向量的半空间不包含 K ，因此， $z \notin K^*$ 。