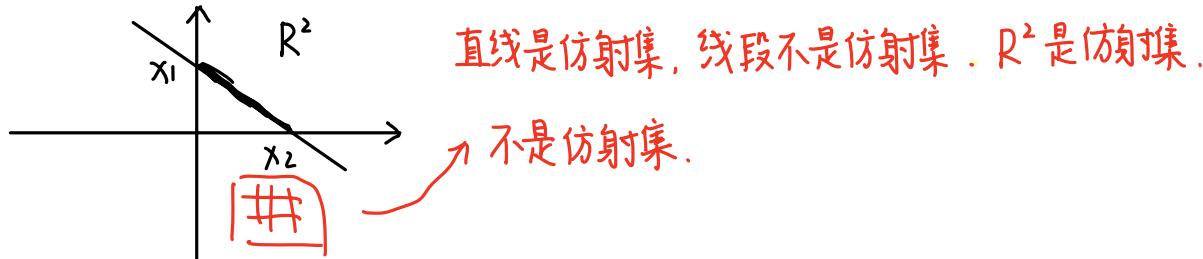


1. 仿射集. Affine Sets (附: 教材附录1和附录2仔细看)

(1) 直线: $x_1, x_2 \in \mathbb{R}^n, \theta \in \mathbb{R}$. $y = \theta x_1 + (1-\theta)x_2 = x_2 + \theta(x_1 - x_2)$

(2) 线段: $y = \theta x_1 + (1-\theta)x_2, \theta \in \mathbb{R}, \theta \in [0,1]$



(3) 仿射集: $\forall x_1, x_2 \in C$, 则连接 x_1 和 x_2 的直线 $\in C$, \Rightarrow 集合 C 为仿射集

\downarrow 泛化
设 $x_1, \dots, x_k \in C, \theta_1, \dots, \theta_k \in \mathbb{R}, \theta_1 + \dots + \theta_k = 1$, 仿射组合 $\theta_1 x_1 + \dots + \theta_k x_k \in C$

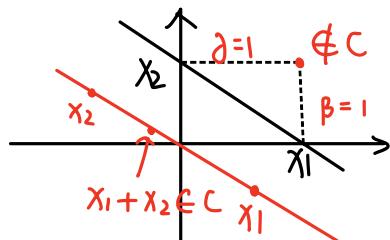
证明: 仿射集 C . $x_1, x_2, x_3 \in C, \theta_1, \theta_2, \theta_3 \in \mathbb{R}, \theta_1 + \theta_2 + \theta_3 = 1$

$$\frac{\theta_1}{\theta_1 + \theta_2} x_1 + \frac{\theta_2}{\theta_1 + \theta_2} x_2 \in C$$

$$(\theta_1 + \theta_2) \left(\frac{\theta_1}{\theta_1 + \theta_2} x_1 + \frac{\theta_2}{\theta_1 + \theta_2} x_2 \right) + (1 - \theta_1 - \theta_2) x_3 \in C$$

$$\theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3 \in C$$

(4) 仿射集的性质. $x_1, x_2 \in C$, C 仿射集, $\theta x_1 + (1-\theta)x_2 \in C$.



$\alpha x_1 + \beta x_2 \notin C, \alpha, \beta \in \mathbb{R}$
 $V = C - x_0 = \{x - x_0 \mid x \in C\}, \forall x_0 \in C$
 与 C 相关的子空间 (前提: C 必须是仿射集)

证明: V 是仿射集. ($\forall v_1, v_2 \in V, \forall \alpha, \beta \in \mathbb{R}, \alpha v_1 + \beta v_2 \in V$)

$$\alpha v_1 + \beta v_2 + x_0 \in C \quad (\forall \text{仿射集都有子空间, 且经过原点})$$

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

(5) 线性方程组的解集是仿射集 (很重要)

$$C = \{x \mid Ax = b\}, A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m, x \in \mathbb{R}^n.$$

$$\forall x_1, x_2 \in C, Ax_1 = b, Ax_2 = b.$$

$$\theta \in \mathbb{R}, \theta x_1 + (1-\theta)x_2 \in C \Rightarrow A(\theta x_1 + (1-\theta)x_2) \in C$$

$\theta \cdot Ax_1 + (1-\theta)Ax_2 = b$. 得证.

(6) 线性方程组解集的子空间. $V = \{x - x_0 \mid x \in C\}$, $\forall x_0 \in C$

C相关

$$= \{x - x_0 \mid Ax = b\}, Ax_0 = b$$

$$= \{x - x_0 \mid A(x - x_0) = 0\}$$

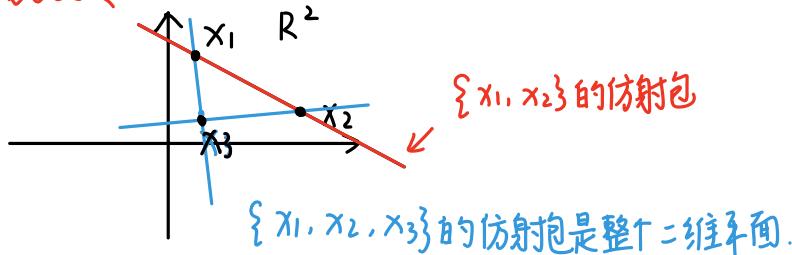
$$C = \{x \mid Ax = b\} \xrightarrow{\text{平移}} = \{y \mid Ay = 0\}$$

(与C相关的子空间)
1维零空间: 集合中任意点乘A都为0.

(7) \forall 集合C, 构造尽可能小的仿射集. (仿射集的仿射包还是原仿射集.)

包含仿射集最
小的仿射集

$$\text{仿射包 } \text{aff } C = \left\{ \theta_1 x_1 + \dots + \theta_k x_k \mid \begin{array}{l} \forall x_1, \dots, x_k \in C \\ \theta_1 + \dots + \theta_k = 1 \end{array} \right\}$$



2. 凸集 Convex Sets

(仿射集是凸集的特例)

(1) 定义: 一个集合C是凸集, 当任意两点之间的线段仍在C内.

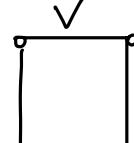
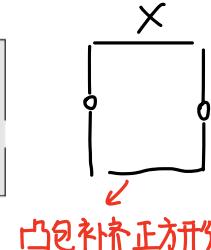
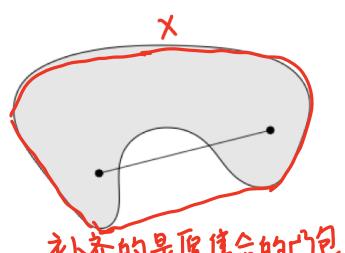
C 为凸集 $\Leftrightarrow \forall x_1, x_2 \in C, \forall \theta \in [0, 1], \theta x_1 + (1-\theta)x_2 \in C$.

(2) 凸组合: x_1, \dots, x_k 的凸组合, $\theta_1, \dots, \theta_k \in \mathbb{R}$, $\theta_1 + \dots + \theta_k = 1$, $\theta_1, \dots, \theta_k \in [0, 1]$

则凸组合为 $\theta_1 x_1 + \theta_2 x_2 + \dots + \theta_k x_k$.

C 为凸集 \Leftrightarrow 任意元素的凸组合 $\in C$.

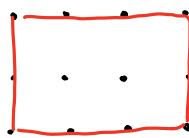
(3) 凸包: $C \in \mathbb{R}^n$, $\text{Conv } C = \left\{ \theta_1 x_1 + \dots + \theta_k x_k \mid \begin{array}{l} \forall x_1, \dots, x_k \in C \\ \forall \theta_1, \dots, \theta_k \in [0, 1] \\ \theta_1 + \dots + \theta_k = 1 \end{array} \right\}$



凸包是其本身

连续的
点构成的
集合

离散的集合：
教室里全部
的同学.



① 不是凸集，因为不能保证任意两个同学中间都有同学.

② 构造凸包：把最外围的同学都用绳子连起来.

3. 锥 (Cone) & 凸锥 (Convex Cone)

(1) 定义： C 是锥， $\Leftrightarrow \forall x \in C, \theta > 0$, 有 $\theta x \in C$.

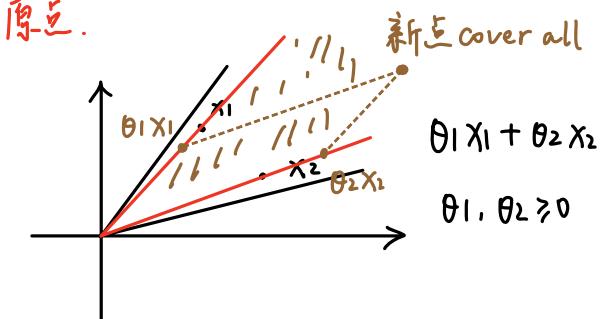
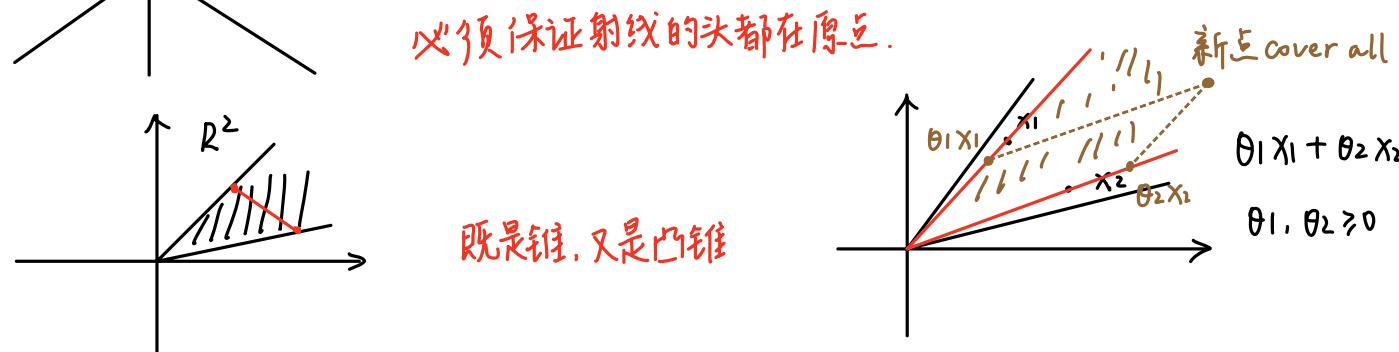
C 是凸锥 $\Leftrightarrow \forall x_1, x_2 \in C, \underline{\theta_1, \theta_2 \geq 0}$, 有 $\theta_1 x_1 + \theta_2 x_2 \in C$

注意范围

锥的定义 + 凸集的定义

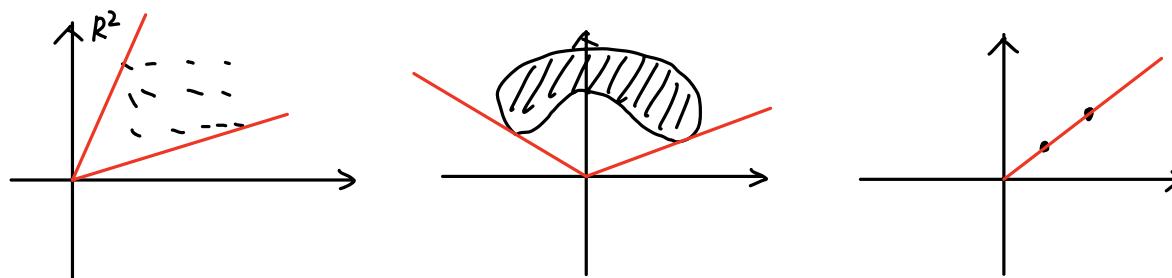
三条射线构成的集合是锥. (不一定连续，分成几段也可以)

必须保证射线的头都在原点.



(2) 凸锥组合： $\theta_1 x_1 + \dots + \theta_k x_k, \theta_1, \dots, \theta_k \geq 0$

(3) 凸锥包： $x_1, \dots, x_k \in C, \left\{ \theta_1 x_1 + \dots + \theta_k x_k \mid \begin{array}{l} x_1, \dots, x_k \in C \\ \theta_1, \dots, \theta_k \geq 0 \end{array} \right\}$ 直线是凸锥



任意选取 k 个点，所有组合还在集合内，则称为凸集.

仿射组合： $\forall \theta_1, \dots, \theta_k, \theta_1 + \dots + \theta_k = 1$ 任意仿射组合都是凸组合.

凸组合： $\forall \theta_1, \dots, \theta_k, \theta_1 + \dots + \theta_k = 1, \theta_1, \dots, \theta_k \in [0, 1]$

凸锥组合： $\forall \theta_1, \dots, \theta_k, \theta_1, \dots, \theta_k \geq 0$ 任意凸锥组合一定是凸的.

包：即使集合不是凸集，选取 k 个点构造凸组合.

特例①: $C = \{x\}$ $\theta_1 x + \theta_2 x = x$ 在 C 内, 仍然是仿射集 \Rightarrow 一定是凸集 \Rightarrow 如果是原点, 则为凸锥.

② $C = \emptyset$. 空集是仿射集、凸集、凸锥. (不存在 存在于 不存在当中)

4. 几种重要的凸集.

(1) \mathbb{R}^n 空间一定是仿射集、凸集、凸锥.

① 过原点 ② 四则运算仍在空间内.

③ \mathbb{R}^n 空间的子空间 (与集合 C 相关的子空间) 一定是仿射集、凸集、凸锥.

④ 任意直线是仿射集、凸集, 不一定是凸锥. (高维空间的一维子空间)

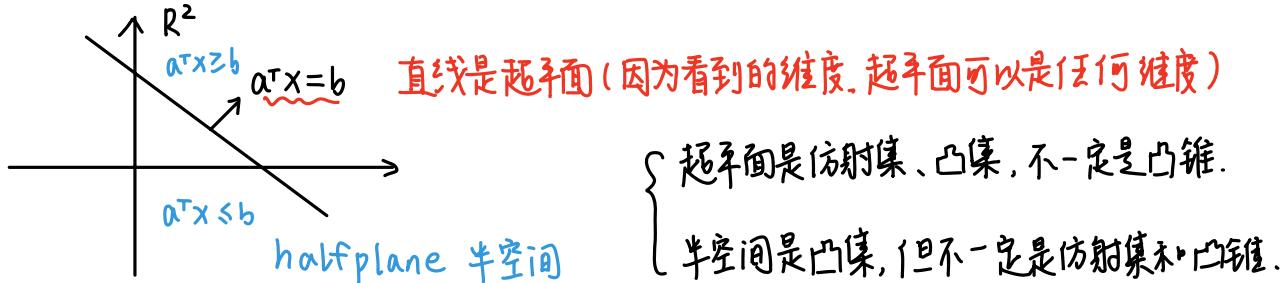
⑤ 任意线段都是凸集, 如果线段缩为点则为仿射集, 缩为点且为原点则为凸锥.

⑥ $\{x_0 + \theta v \mid \theta \geq 0\}, x_0 \in \mathbb{R}^n, \theta \in \mathbb{R}, v \in \mathbb{R}^n$. (x_0 为原点, v 为方向的射线)

不一定是仿射集 (除 $\theta=0$), 不一定是凸锥 (除 $x_0=0$)

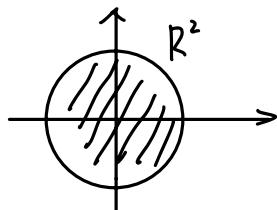
(2) 超平面与半空间. (本质上是集合)

Hyperplane: $\{x \mid a^T x = b\}, x, a \in \mathbb{R}^n, b \in \mathbb{R}, a \neq 0$



(3) 球和椭球. (欧氏空间里)

球: $B(x_c, r) = \{x \mid \|x - x_c\|_2 \leq r\} = \{x \mid \sqrt{(x - x_c)^T (x - x_c)} \leq r\}, x_c \in \mathbb{R}^n, r \in \mathbb{R}$



证明: (球是凸集)

$$\forall x_1, x_2 \in B, \|x_1 - x_c\|_2 \leq r, \|x_2 - x_c\|_2 \leq r, \forall \theta \in [0, 1]$$

$$\Leftrightarrow \|\theta x_1 + (1-\theta)x_2 - x_c\|_2 \leq r$$

☆ 三角不等式很重要,
一定要记住.

$$= \|\theta(x_1 - x_c) + (1-\theta)(x_2 - x_c)\|_2$$

$$\leq \theta \|x_1 - x_c\|_2 + (1-\theta) \|x_2 - x_c\|_2$$

$$\leq \theta r + (1-\theta)r = r \quad \text{得证.}$$

椭球: $\Sigma(x_c, P) = \{x \mid (x-x_c)^T P^{-1} (x-x_c) \leq 1\}$, $x \in \mathbb{R}^n$, $P \in \mathbb{S}_{++}^n$ 正定矩阵

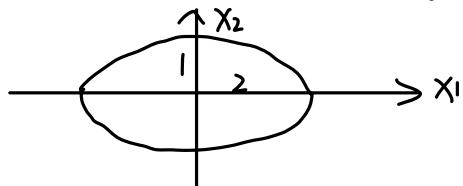
补充: 方阵才有特征值(erg), 但是所有矩阵都有奇异值. 正定=全部奇异值大于0.

$$\text{奇异值} = \sqrt{\text{erg}(A^T A)} \quad (\text{方阵的奇异值和特征值可能不同})$$

eg: $P = r^2 I_n$. 带入椭球后则成为球.

描述的是椭球的半轴长(两个奇异值)

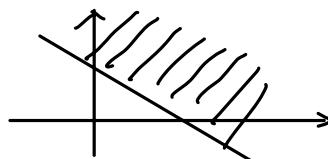
$$\text{例: } \Sigma = \left\{ x \mid x^T \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix}^{-1} x \leq 1 \right\} = \left\{ (x_1, x_2) \mid \frac{1}{4}x_1^2 + x_2^2 \leq 1 \right\}$$



(还有“有界多面体”的概念)

(4) 多面体: $P = \{x \mid \begin{array}{l} a_j^T x \leq b_j, j=1, 2, \dots, m \\ c_j^T x = d_j, j=1, 2, \dots, n \end{array}\}$ 一些超平面和半空间的交集 (不一定有界)

多面体一定是凸集.



(多面体不一定规则有界)

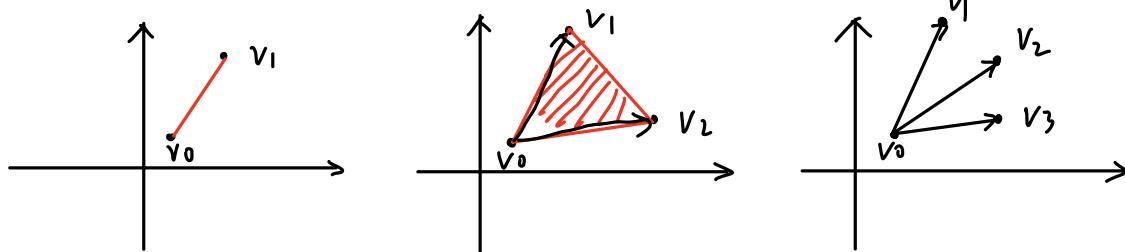
(5) 单纯形. Simplex

过这些点的凸包就是单纯形.

\mathbb{R}^n 空间中选择 v_0, \dots, v_k 共 $k+1$ 个点, v_1-v_0, \dots, v_k-v_0 线性无关.

则马上过点相关的单纯形为: $C = \text{conv}\{v_0, \dots, v_k\} = \{\theta_0 v_0 + \dots + \theta_k v_k, \theta \geq 0, 1^T \theta = 1\}$

例: \mathbb{R}^2 空间.



(不线性相关, 不能构造单纯形)

证明: 单纯形一定是多面体.

$$x \in C \in \mathbb{R}^n, C \text{ 为 Simplex} \Leftrightarrow x = \theta_0 v_0 + \dots + \theta_k v_k, \theta \geq 0, 1^T \theta = 1$$

v_1-v_0, \dots, v_k-v_0 线性无关.

定义: $[\theta_1, \dots, \theta_k]^T = y$, $y \geq 0$, $I^T y \leq 1$

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

$$x \in C \Leftrightarrow \theta_0 v_0 + \dots + \theta_k v_k = v_0 + \theta_1(v_1 - v_0) + \dots + \theta_k(v_k - v_0) = v_0 + By$$

$\text{rank}(B)=k$ (引满秩的矩阵), $k \leq n$. $\begin{bmatrix} I_k \\ 0 \end{bmatrix} \Rightarrow k \times k$
 $\begin{bmatrix} 0 \\ I_{n-k} \end{bmatrix} \Rightarrow (n-k) \times k$

$$\text{非奇异矩阵 } A = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} \in \mathbb{R}^{n \times n}, AB = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} B = \begin{bmatrix} I_k \\ 0 \end{bmatrix}$$

$$\Leftrightarrow Ax = Av_0 + ABy$$

$$\Leftrightarrow \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} x = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} v_0 + \begin{bmatrix} A_1 B \\ A_2 B \end{bmatrix} y$$

$$\Leftrightarrow \begin{cases} A_1 x = A_1 v_0 + y \\ A_2 x = A_2 v_0 \end{cases} \quad \begin{array}{l} \text{作用是遍历整个集合.} \\ \text{把 } y \text{ 的两个性质带进去} \end{array}$$

$$\Leftrightarrow \begin{cases} A_1 x \geq A_1 v_0 \\ |^T A_1 x \leq 1 + |^T A_1 v_0 \\ A_2 x = A_2 v_0 \end{cases}$$

eg: $\{x \mid x \leq 0\}$ 是单纯形

$x_0=0, x_1=-1$ 两个点.

由这两个点构成的凸包是单纯形.

∴任意单纯形都能由两个不等式的一个表达式表达. 所以单纯形都是多面体.

(6) 对称矩阵集合 $S^n = \{x \in \mathbb{R}^{n \times n} \mid x = x^T\}$ $x \geq 0$ 表示每个元素 ≥ 0 , $x \geq 0$ 表示

对称半正定矩阵集合 $S_+^n = \{x \in \mathbb{R}^{n \times n} \mid x = x^T, x \geq 0\}$ 每个奇异值 ≥ 0 .

对称正定矩阵集合 $S_{++}^n = \{x \in \mathbb{R}^{n \times n} \mid x = x^T, x > 0\}$

证明: S_+^n 是 Convex Cone.

$\forall \theta_1, \theta_2 \geq 0, \forall A, B \in S_+^n$, 证明 $\theta_1 A + \theta_2 B \in S_+^n$

$\forall X \in \mathbb{R}^n, x^T A x \geq 0, x^T B x \geq 0$ (半正定的定义)

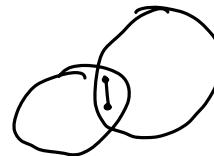
$x^T (\theta_1 A + \theta_2 B) x = x^T \theta_1 A + x^T \theta_2 B x \geq 0$, 证毕.

$n=1$ 时, $S_+^n = \mathbb{R}_+$, $S_{++}^n = \mathbb{R}_{++}$ (不是凸锥), $S^n = \mathbb{R}$

对称的半正定矩阵.

$n=2$ 时, (-一个矩阵空间变化为若干实数空间), $S_+^n = \left\{ \begin{bmatrix} x & y \\ y & z \end{bmatrix} \mid x \geq 0, z \geq 0, x \geq y^2 \right\}$

5. 凸集的交集 & 保凸运算



(1) 交集. 若 S_1, S_2 为凸, 则 $S_1 \cap S_2$ 为凸.

若 S_α 为凸集, $\forall \alpha \in A$, 则 $\bigcap_{\alpha \in A} S_\alpha$ 为凸集. (该结论应用于并集不成立)

(2) 仿射函数 (线性的映射)

$f: R^n \rightarrow R^m$ 是仿射的, 当 $f(x) = Ax + b$, $A \in R^{m \times n}$, $b \in R^m$.

凸集经过仿射函数的变换仍然为凸集. (i.e. $f(S) = \{f(x) | x \in S\}$ 为凸)

n 维空间中的凸集

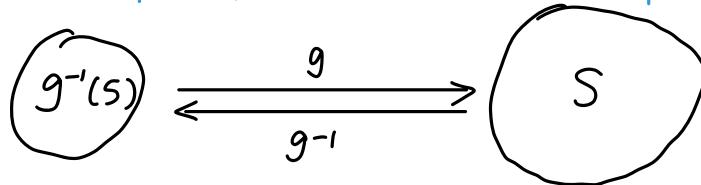
m 维空间中的凸集.



$g: R^k \rightarrow R^n$ 为仿射, 当 $g^{-1}(S) = \{x | f(x) \in S\}$ (从 n 维空间映射为 k 维空间)

k 维空间中的凸集

n 维空间中的凸集



{ 缩放: $\lambda S = \{\lambda x | x \in S\}$

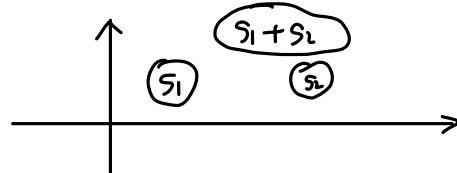
(3) 性质: ① 缩放、移位是保持凸性的. { 移位: $S + a = \{x + a | x \in S\}$

② 两个凸集的和仍然为凸集. $S_1 + S_2 = \{x + y | x \in S_1, y \in S_2\}$

↓ 定义 $S_1 \times S_2 = \{(x, y) | x \in S_1, y \in S_2\}$ 为凸集.

看成其中的一个元素而非两个元素.

线性映射 $f(x, y) = x + y$



再推广: 一个凸集和一个非凸集的和仍然有可能为凸集.

(4) 线性矩阵不等式. LM7.

一旦被矩阵搞
晕，就把矩阵想
象成林量($m \times n$)

$$A(x) = \boxed{x_1 A_1 + \dots + x_n A_n \leq B}, \quad B, A_i, x_i \in S^m$$

$$(A(x) - B) \leq 0$$

线性矩阵不等式的解集 $\{x \mid A(x) \leq B\}$ 为凸集.

证明：(借助仿射函数) 每个点都是矩阵

高维(一堆矩阵)

低维(对称矩阵)

定义仿射空间 $f(x) \triangleq B - A(x)$ 每个点由一个矩阵构成

又因为 S_+^n 为凸集.

低维到高维.

$f^{-1}(S_+^n) = \{x \mid B - A(x) \geq 0\}$ 半正是 仍然是凸集.

因此，线性矩阵不等式的解集是凸集.

(5) 杆有球是球的仿射映射.

$$\Sigma = \{x \mid (x - x_c)^T P^{-1} (x - x_c) \leq 1\}, \quad P \in S_+^{++}$$

单位球 $\{u \mid \|u\|_2 \leq 1\}$

$f(u) = P^{\frac{1}{2}}u + \underline{x_c}$ 移位. , 定义 $x \triangleq P^{\frac{1}{2}}u + x_c$, 则 $u = P^{-\frac{1}{2}}(x - x_c)$

$$\{f(u) \mid \|u\|_2 \leq 1\} = \{P^{\frac{1}{2}}u + x_c \mid \|u\|_2 \leq 1\}$$

$$= \{x \mid \|P^{-\frac{1}{2}}(x - x_c)\|_2 \leq 1\}$$

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

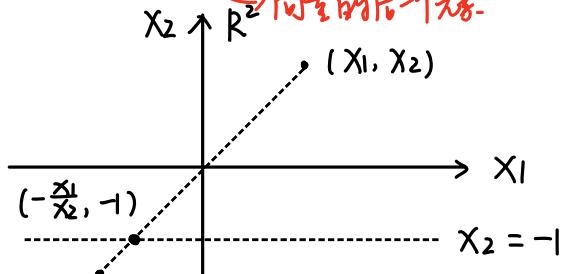
降维

(6) 透视函数.

$P: \mathbb{R}^{n+1} \rightarrow \mathbb{R}^n$, $\text{dom } P = \mathbb{R}^n \times \mathbb{R}^{++}$ (定义域, 前 n 个元素随意取值, 最后元素为正)

$P(z, t) = \frac{z}{t}$, $z \in \mathbb{R}^n$, $t \in \mathbb{R}^{++}$

$x_2 \uparrow \mathbb{R}^2$ 向量的后一个元素.



$$\left(-\frac{x_1}{x_2}, -1\right) = (-P(x_1, x_2), -1)$$

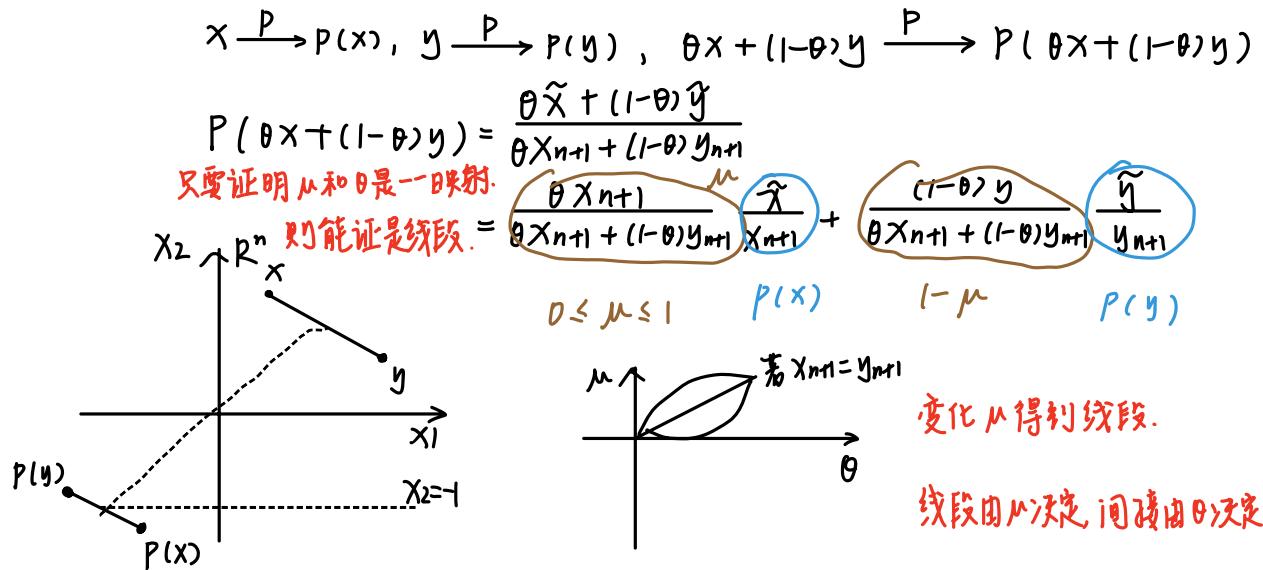
☆任意凸集经过透视函数仍为凸集.

丢掉了一维.

(7) 考虑 \mathbb{R}^{n+1} 内线段 $\overset{\text{ER}^n}{\overrightarrow{x}}, \overset{\text{ER}^n}{\overrightarrow{y}}$, $x = (\overset{\text{ER}^n}{x}, x_{n+1}), y = (\overset{\text{ER}^n}{y}, y_{n+1})$

$\theta \geq 0$, 线段为 $\theta x + (1-\theta)y$
从 $n+1$ 维的线段到 n 维线段.

证明: 该线段经过透视函数仍为凸集和线段.



(8) 任意凸集的反透视映射仍然为凸集. 开维.

$$P^{-1}(C) = \left\{ (x, t) \in \mathbb{R}^{n+1} \mid \frac{x}{t} \in C, t > 0 \right\}$$

证明: $(x, t) \in P^{-1}(C)$, $(y, s) \in P^{-1}(C)$, $0 \leq \theta \leq 1$

$$\begin{aligned} \frac{\theta x + (1-\theta)y}{\theta t + (1-\theta)s} &\in C \\ \frac{\theta t}{\theta t + (1-\theta)s} \frac{x}{t} + (1 - \frac{\theta t}{\theta t + (1-\theta)s}) \frac{y}{s} &\in C \end{aligned}$$

所以, 任意凸集的反透视映射仍然为凸集.

(9) 线性分数函数 $g: \mathbb{R}^n \rightarrow \mathbb{R}^{m+1}$ 为仿射映射.

$$g(x) = \begin{bmatrix} A \\ C^\top \end{bmatrix} x + \begin{bmatrix} b \\ d \end{bmatrix}, \quad A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m, C \in \mathbb{R}^n, d \in \mathbb{R}$$

定义透视变换 $P: \mathbb{R}^{m+1} \rightarrow \mathbb{R}^m$, $f: \mathbb{R}^n \rightarrow \mathbb{R}^m \triangleq P \circ g$ (先经 g 变为 $m+1$ 维, 再经 P 变为 m 维)

$$f(x) = \frac{Ax+b}{C^\top x+d} \star, \quad \text{dom } f = \{x \mid C^\top x + d > 0\}$$

先仿射, 后透视

证明: 任意凸集经(线性分数变换)仍为凸集.

(10) 两个随机变量的联合概率 → 条件概率. $u \sim \{1, \dots, n\}$ $v \sim \{1, \dots, m\}$

$$P_{ij} = p(u=i, v=j), \quad f_{ij} = p(u=i | v=j), \text{ then } f_{ij} = \frac{P_{ij}}{\sum_{k=1}^n P_{kj}}$$

需要证明联合概率是凸集. $\sum_{k=1}^n P_{kj}$ 实际上是对向量 (P_{1j}, \dots, P_{nj}) 求和.

P_{ij} 实际上是将真点乘 $(0 \dots \overset{\text{取在 } i}{1} \dots 0)$

f_{ij} 则是从 n 维向量到标量的线性函数变换.