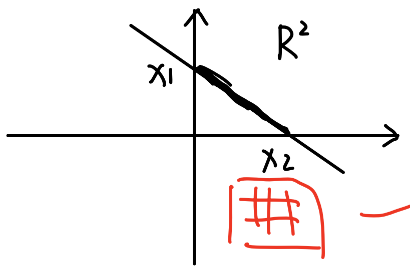
1. 仿射集. Affine Sets （附：教材附录1和附录2仔细看）

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

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



直线是仿射集，线段不是仿射集．R^2是仿射集．

不是仿射集.

(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

泛化

设 $x_1, \cdots, x_k \in C$，$\theta_1, \cdots, \theta_k \in R$，$\theta_1 + \cdots + \theta_k = 1$，仿射组合 $\theta_1 x_1 + \cdots + \theta_k x_k \in C$

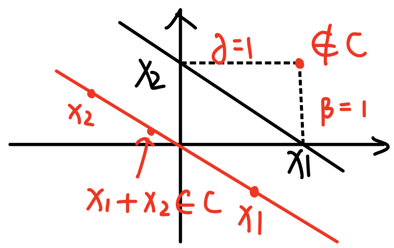
证明：仿射集 C．$x_1, x_2, x_3 \in C$，$\theta_1, \theta_2, \theta_3 \in R$，$\theta_1 + \theta_2 + \theta_3 = 1$

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

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

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

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



$\partial x_1 + \beta x_2 \overset{?}{\in} C$，$\partial, \beta \in R$

$V = C - x_0 = \{x - x_0 \mid x \in C\}$，$\forall x_0 \in C$

与 C 相关的子空间（前提：C 必须是仿射集）

证明：V 是仿射集．（$\forall v_1, v_2 \in V$，$\forall \partial, \beta \in R$，$\partial v_1 + \beta v_2 \in V$）

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

$$\Leftarrow \quad \partial(\underbrace{v_1 + x_0}_{\in C}) + \beta(\underbrace{v_2 + x_0}_{\in C}) + (\underbrace{1-\partial-\beta}_{\in C}) x_0 \in C$$

(5) 线性方程组的解集是仿射集☆☆☆（很重要）

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

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

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

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

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

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

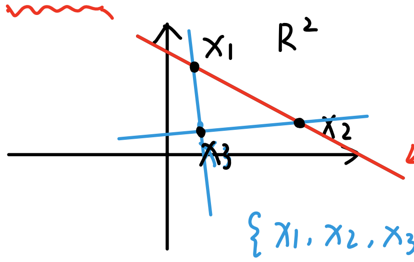
(与C相关的子空间)

$C = \{x \mid Ax = b\} \xrightarrow{平移} = \{y \mid Ay = 0\}$ 化零空间：集合中任意点乘A都为0.

(7) ∀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

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



$\{x_1, x_2\}$的仿射包

$\{x_1, x_2, x_3\}$的仿射包是整个二维平面.

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然在C内.

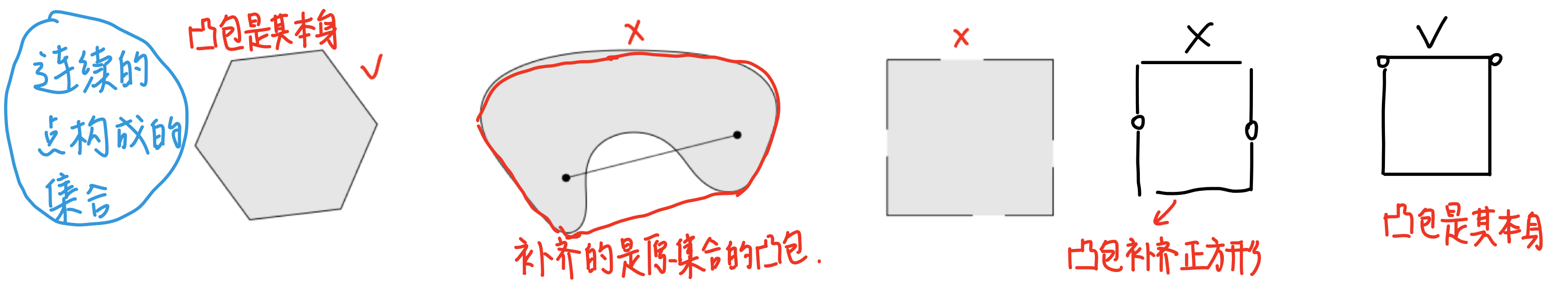
C为凸集 ⟺ $\forall x_1, x_2 \in C$, $\forall \theta \in [0,1]$, $\theta x_1 + (1-\theta) x_2 \in C$.

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

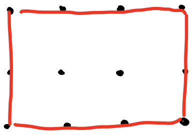
则凸组合为 $\theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_k x_k$.

C为凸集 ⟺ 任意元素的凸组合 ∈ C.

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



离散集合：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.

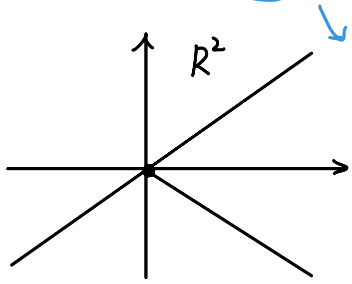
3. 锥(Cone) 及凸锥 (Convex Cone)

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

注意范围

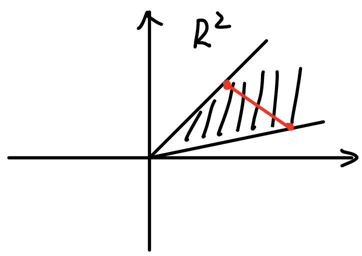
C是凸锥 $\Leftrightarrow$ $\forall x_1, x_2 \in C$，$\theta_1, \theta_2 \geq 0$，有 $\theta_1 x_1 + \theta_2 x_2 \in C$

锥的定义 + 凸集的定义

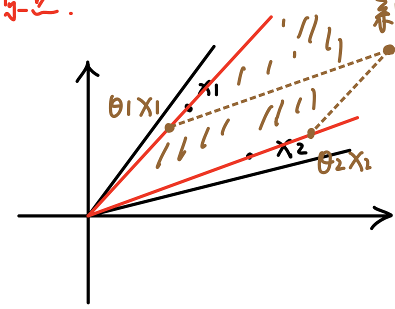


三条射线构成的集合是锥.（不一定连续，分成n支也可以）

必须保证射线的头部在原点.



既是锥，又是凸锥

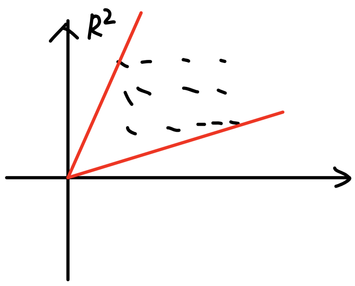


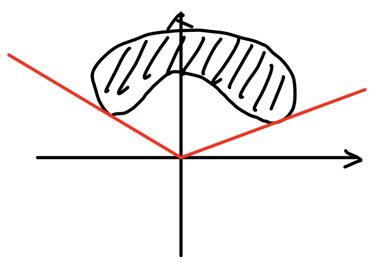
$\theta_1 x_1 + \theta_2 x_2$

$\theta_1, \theta_2 \geq 0$

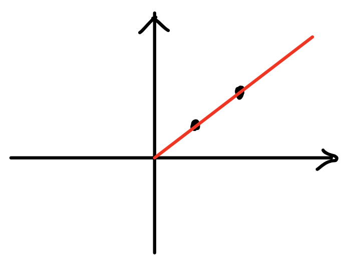
(2) 凸锥组合： $\theta_1 x_1 + \cdots + \theta_k x_k$，$\theta_1, \cdots, \theta_k \geq 0$

(3) 凸锥包： $x_1, \cdots, x_k \in C$，$\left\{ \theta_1 x_1 + \cdots + \theta_k x_k \;\middle|\; \begin{array}{l} x_1, \cdots, x_k \in C \\ \theta_1, \cdots, \theta_k \geq 0 \end{array} \right\}$





直线是凸锥



任意选取k个点，xx组合还在集合内，则称为xx集.

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

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

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

包：即便集合不是xx集，选取k个点构造xx组合.

特例：① $C=\{x\}$　$\theta_1 x+\theta_2 x=x$ 在C内，仍然是仿射集 ⇒ 一定是凸集 ⇒ 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) ① R^n 空间一定是仿射集、凸集、凸锥.

①过原点 ②四则运算仍然在空间内.

② R^n 空间的子空间（非与集合C相关的子空间）一定是仿射集、凸集、凸锥.

③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）

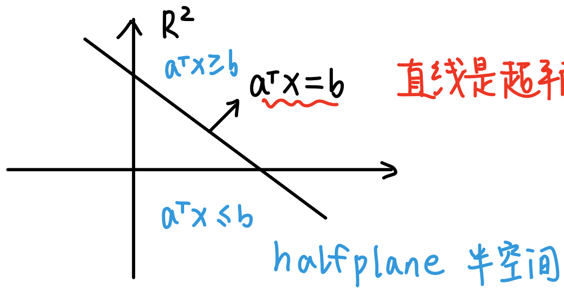
④任意线段都是凸集，如果线段缩为点则为仿射集，缩为点且为原点则为凸锥.

⑤ $\{x_0+\theta v \mid \theta\ge 0\}$，$x_0\in R^n$，$\theta\in R$，$v\in R^n$．（$x_0$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.（本质上是个集合）

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

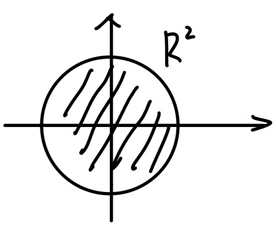


直线是超平面（因为看到的维度，超平面可以是任何维度）

- 超平面是仿射集、凸集，不一定是凸锥.
- 半空间是凸集，但不一定是仿射集和凸锥.

(3) 球和椭球.（欧氏空间里）

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



证明：（球是凸集）

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

$\Leftarrow \|\theta x_1+(1-\theta)x_2-x_c\|_2\le r$

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

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

☆三角不等式很重要，一定要记住.

$\leq \theta r + (1-\theta) r = r$ 得证.

通过P加权的二范数.

n×n维的对称的

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

正定: 全部奇异值大于0

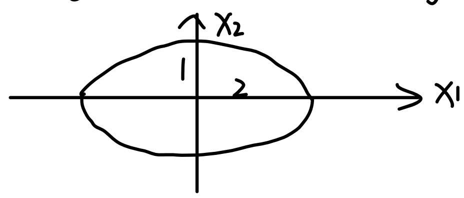
补充: 方阵才有特征值(eig), 但是所有矩阵都有奇异值.

奇异值 $= \sqrt{eig(A^TA)}$ (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)

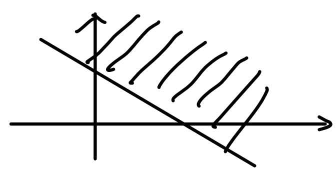
例: $\mathcal{E} = \left\{x \mid x^T \begin{pmatrix}4&0\\0&1\end{pmatrix}^{-1} x \leq 1\right\} = \left\{(x_1,x_2) \mid \frac{1}{4} x_1^2 + x_2^2 \leq 1\right\}$



(还有"有界多面体"的概念)

(4) 多面体: $P=\left\{x \;\middle|\; \begin{matrix} a_j^T x \leq b_j, & j=1,2,\dots,m \\ c_j^T x = d_j & j=1,2,\dots,n \end{matrix}\right\}$ 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

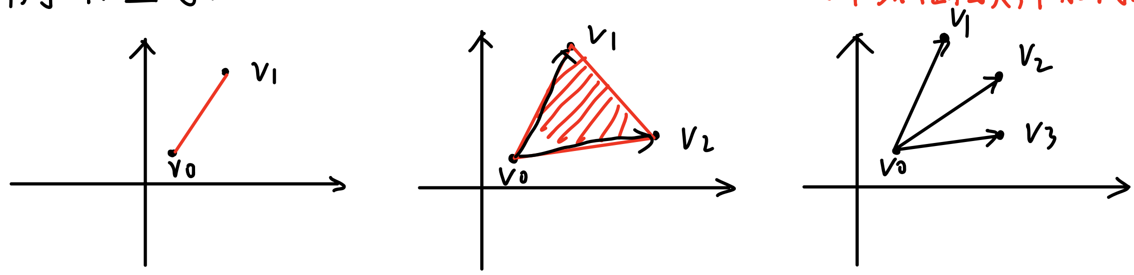
过这些点的凸包就是单纯形.

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.

则与上述点相关的单纯形为: $C=conv\{v_0,\dots,v_k\} = \{\theta_0 v_0+\dots+\theta_k v_k,\ \theta\succeq 0,\ 1^T\theta=1\}$

例: 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.

$x\in C\in R^n$, C为Simplex $\Leftrightarrow x = \theta_0 v_0+\dots+\theta_k v_k$, $1^T\theta=1$, $\theta\succeq 0$

$v_1-v_0,\dots,v_k-v_0$ 线性无关.

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

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

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

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

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

$\Leftrightarrow Ax = Av_0 + ABy$

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

$\Leftrightarrow \begin{cases} A_1 x = A_1 v_0 + y \\ A_2 x = A_2 v_0 \end{cases}$ y的作用是遍历整个集合.

把y的两个性质带进去

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

eg: $\{x \mid x \leq 0\}$ 是单纯形

$x_0 = 0,\ x_1 = -\infty$ 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等式表达. 所以单纯形都是多面体.

(6) 对称矩阵集合 $S^n = \{x \in R^{n\times n} \mid x = x^T\}$

对称半正定矩阵集合 $S^n_+ = \{x \in R^{n\times n} \mid x = x^T,\ x \succeq 0\}$

$x \geq 0$ 表示每个元素 ≥ 0，$x \succeq 0$ 表示每个奇异值 ≥ 0.

特征值 > 0.

对称正定矩阵集合 $S^n_{++} = \{x \in R^{n\times n} \mid x = x^T,\ x \succ 0\}$

证明：S^n_+ 是 Convex Cone.

$\forall \theta_1, \theta_2 \geq 0$，$\forall A, B \in S^n_+$，证明 $\theta_1 A + \theta_2 B \in S^n_+$

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

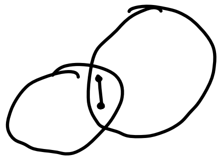
$x^T(\theta_1 A + \theta_2 B)x = x^T \theta_1 A + x^T \theta_2 B x \geq 0$，证毕.

对称的半正定矩阵.

$n=1$ 时，$S^n_+ = R_+$，$S^n_{++} = R_{++}$（不是凸锥），$S^n = R$

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

5. 凸集的交集 & 保凸运算



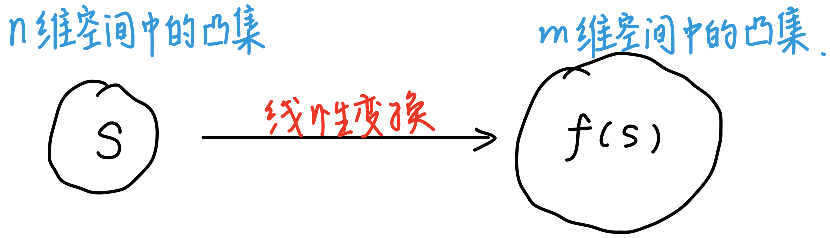
(1) 交集. 若S_1,S_2为凸，则$S_1\cap S_2$为凸.

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）

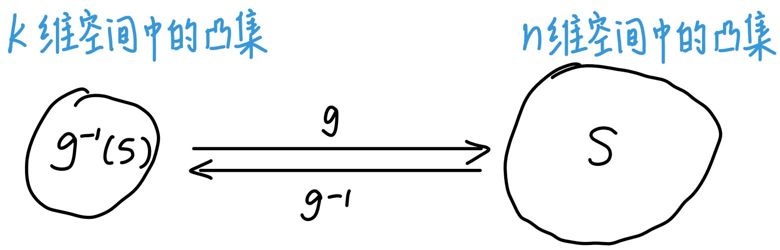
(2) 仿射函数（线性的映射）

$f: R^n\to R^m$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. $f(S)=\{f(x)\mid x\in S\}$为凸.）



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



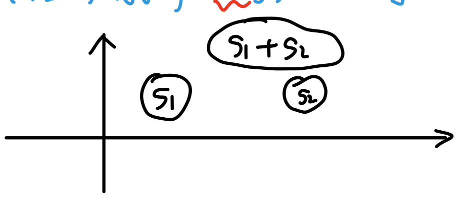
(3) 性质：① 缩放、移位是保持凸性的.
- 缩放：$\alpha S=\{\alpha x\mid x\in S\}$
- 移位：$S+a=\{x+a\mid x\in S\}$

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

定义 $\overset{\in R^2}{S_1\times S_2}=\{(\overset{\in R}{x},\overset{\in R}{y})\mid x\in S_1,\ y\in S_2\}$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. LMI.

一旦被矩阵搞晕,就把矩阵想象成标量(1×1)

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

$$A(x) = \boxed{x_1A_1 + \cdots + x_nA_n \preceq B}, \quad B, A_i, x_i \in S^m$$

线性矩阵不等式的解集 $\{x \mid A(x) \preceq B\}$ 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. 半正定

$$f^{-1}(S^n_+) = \{x \mid B - A(x) \succeq 0\}$$ 仍然是凸集.

因此,线性矩阵不等式的解集是凸集.

(5)椭球是球的仿射映射.

$$\varepsilon = \{x \mid (x-x_c)^T P^{-1}(x-x_c) \le 1\}, \quad P \in S^n_{++}$$

单位球 $\{u \mid \|u\|_2 \le 1\}$

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

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

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

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

(6)透视函数.

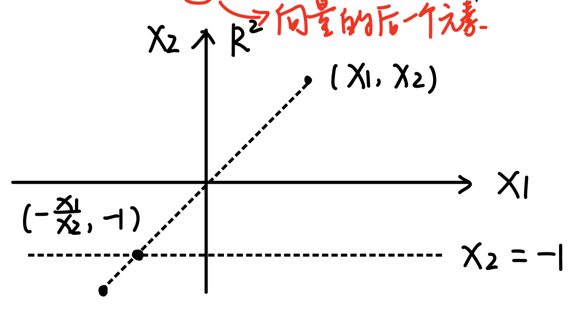
降维

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

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

z → 向量的前n个元素

t → 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然为凸集.

丢掉了一维.

(7) 考虑 R^{n+1} 内线段. $x=(\tilde{x}, x_{n+1})$, $y=(\tilde{y}, y_{n+1})$　$\tilde{x} \in R^n$　$x_{n+1} \in R_{++}$　$\tilde{y} \in R^n$　$y_{n+1} \in R_{++}$

$\theta \geq 0$, 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然为凸集和线段.

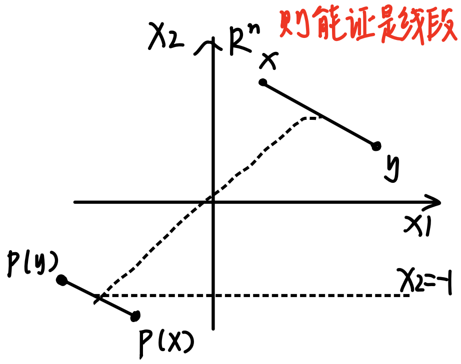
$$x \xrightarrow{P} P(x),\quad y \xrightarrow{P} P(y),\quad \theta x+(1-\theta)y \xrightarrow{P} P(\theta x+(1-\theta)y)$$

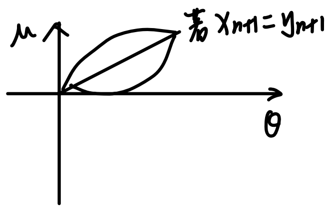
$$P(\theta x+(1-\theta)y) = \frac{\theta\tilde{x}+(1-\theta)\tilde{y}}{\theta x_{n+1}+(1-\theta)y_{n+1}}$$

只要证明 μ 和 θ 是一一映射. 则能证是线段.

$$= \underbrace{\frac{\theta x_{n+1}}{\theta x_{n+1}+(1-\theta)y_{n+1}}}_{\mu}\ \underbrace{\frac{\tilde{x}}{x_{n+1}}}_{P(x)} + \underbrace{\frac{(1-\theta)y_{n+1}}{\theta x_{n+1}+(1-\theta)y_{n+1}}}_{1-\mu}\ \underbrace{\frac{\tilde{y}}{y_{n+1}}}_{P(y)}$$

$0 \leq \mu \leq 1$





变化 μ 得到线段.

线段由 μ 决定, 间接由 θ 决定.

(8) 任意凸集的反透视映射仍然为凸集. 升维.

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

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

$$\frac{\theta x+(1-\theta)y}{\theta t+(1-\theta)s} \overset{?}{\in} C$$

$$\underbrace{\frac{\theta t}{\theta t+(1-\theta)s}}_{\mu}\ \underbrace{\frac{x}{t}}_{\in C} + \underbrace{\left(1-\frac{\theta t}{\theta t+(1-\theta)s}\right)}_{1-\mu}\ \underbrace{\frac{y}{s}}_{\in C} \in C$$

所以, 任意凸集的反透视映射仍然为凸集.

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

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

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

$$f(x) = \frac{Ax+b}{c^Tx+d}$$ ☆. $\mathrm{dom} f = \{x \mid c^Tx+d>0\}$

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 v $\{1,\cdots,n\}$ $\{1\cdots m\}$

$p_{ij}=p(u=i, v=j)$, $f_{ij}=p(u=i \mid v=j)$, then $\overset{标量}{f_{ij}}=\frac{p_{ij}}{\sum_{k=1}^{n} p_{kj}}$

需要证明联合概率是凸集、 $\sum_{k=1}^{n} p_{kj}$ 实质上是对向量 $(p_{1j},\cdots,p_{nj})$ 求和.

p_{ij} 实质上是将其点乘 $(0\cdots\overset{第i个}{1}\cdots 0)$

f_{ij} 则是从n维向量到标量的线性函数变换.