Essence of the lecture (3/4)

仿射集	凸集	凸锥
$\theta_1 + \cdots + \theta_k = 1$ $\theta_1, \ldots, \theta_k \in R$	$\theta_1 + \cdots + \theta_k = 1$ $\theta_1, \ldots, \theta_k \in [0,1]$	$\theta_1, \ldots, \theta_k \geq 0$

直线：

$x_1 \neq x_2 \in R^n,\ \theta \in R$

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

线段：

$\theta \in R,\ \theta \in [0,1]$

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

仿射集（Affine Sets）：若 $\forall x_1, x_2 \in C$，则连接 x_1, x_2 的直线也在集合内

$\Rightarrow$ 因此直线是仿射集，而线段不是

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

$\Rightarrow$ 若 C 为仿射集，则仿射组合也在 C 内，证明如下

$$\frac{\theta_1}{\theta_1+\theta_2} x_1 + \frac{\theta_2}{\theta_1+\theta_2} x_2 \in C$$
$$(\theta_1+\theta_2)\left(\frac{\theta_1}{\theta_1+\theta_2} x_1 + \frac{\theta_2}{\theta_1+\theta_2} x_2\right) + (1-\theta_1-\theta_2) x_3 \in C$$
$$\Leftrightarrow \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3 \in C, \quad \theta_1 + \theta_2 + \theta_3 = 1$$

仿射包：对任意集合 C，仿射包 $(aff\ C)$ 是包含 C 的最小仿射集，

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

$V = C - x_0 = \{x - x_0 | x \in C\}\ \forall x_0 \in C \Rightarrow$ **与 C 相关的子空间**，即将 C 平移经过原点

求证：$v_1,\ v_2 \in V, \forall \alpha, \beta \in R \quad \alpha v_1 + \beta v_2 \in V$

即证：$\alpha v_1 + \beta v_2 + x_0 \in C \Longleftrightarrow \alpha(v_1 + x_0) + \beta(v_2 + x_0) + (1-\alpha-\beta) x_0 \in C$

由于 $v_1 + x_0,\ v_2 + x_0,\ x_0 \in C$，得证

线性方程组的解集是仿射集

证明：$C = \{x | Ax = b\},\ A \in R^{m \times n},\ b \in R^m,\ x \in R^n$

由 $Ax_1 = b,\ Ax_2 = b$ 有 $\theta x_1 + (1-\theta) x_2 \in C$

凸集： 一个集合为凸集，当且仅当任意两点的线段在 C 内

即 $\forall x_1, x_2 \in C,\ \forall\theta \in [0,1],\ \theta x_1 + (1-\theta)x_2 \in C$

凸组合： $x_1, \ldots, x_k \in C,\ \theta_1, \ldots, \theta_k \in [0,1],\ \theta_1 + \cdots + \theta_k = 1,\ \theta_1 x_1 + \cdots + \theta_k x_k$

凸集包含其任意元素的凸组合

凸包： $Conv(C) = \{\theta_1 x_1 + \cdots + \theta_k x_k | x_1, \ldots, x_k \in C,\ \theta_1, \ldots, \theta_k \in [0,1],\ \theta_1 + \cdots + \theta_k = 1\}$

即对任意集合 C，包含 C 的最小凸集

锥： $\forall x \in C,\ \theta \geq 0 \Rightarrow \theta x \in C$

凸锥： $\forall x_1, x_2 \in C,\ \theta_1, \theta_2 \geq 0 \Rightarrow \theta_1 x_1 + \theta_2 x_2 \in C$

凸锥组合： $\theta_1 x_1 + \cdots + \theta_k x_k,\ \theta_1, \ldots, \theta_k \geq 0$

凸锥包： $\{\theta_1 x_1 + \cdots + \theta_k x_k | x_1, \ldots, x_k \in C, \theta_1, \ldots, \theta_k \geq 0\}$

仿射集一定是凸集，凸锥一定是凸集；

空集既是仿射集，又是凸集，又是凸锥；

只有一个元素的集合一定是仿射集和凸集，是否是凸集，取决于该元素是否为原点（因为凸锥一定过原点）

Essence of the lecture (5/6)

几种重要的凸集：

- R^n 空间，R^n 空间的子空间
- 任意直线（若过原点也为凸锥），任意线段，射线 $\{x_0+\theta v|\theta \geq 0,\ x\in R^n,\ \theta\in R,\ v\in R^n\}$
- 超平面与半空间
- 球和椭球
- 多面体（Polyhedron）和单纯形（Simplex）
- 对称矩阵集合，对称半正定矩阵集合，对称正定矩阵集合

超平面：$\{x \mid a^T x = b\}$，凸集，仿射集，是否过原点（凸锥）

半空间：超平面的衍生概念，$\{x \mid a^T \geq b\}$，$\{x \mid a^T x \leq b\}$，凸集，非仿射集，是否过原点（凸锥）

球：$B(x_c, r) = \{x \mid \|x-x_c\|_2 \leq r\} = \{x \mid \sqrt{(x-x_c)^T(x-x_c)} \leq r\}$ 为凸集

证明：$\forall\theta \in [0,1]$，取 $f(x) = \|x-x_c\|_2$

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

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

这里主要利用了三角不等式（范数的条件 2），$\|a\|+\|b\| \geq \|a+b\|$

椭球：$\varepsilon(x_c, P) = \{x \mid (x-x_c)^T P^{-1}(x-x_c) \leq 1\}$，$x_c \in R^n$，$P \in S_{++}^n$

其中 P 为对角矩阵，对角线上为矩阵的奇异值的平方，矩阵的奇异值对应了椭球的半轴长

矩阵 A 的奇异值为 $\sqrt{eig(A^TA)}$，需要注意方针的特征值和奇异值也可能不等

多面体（polyhedron）：$\{x \mid a_j^T x \leq b_j,\ j=1,\ldots,m,\ c_j^T = d_j,\ j=1,\ldots,p\}$

可以看作半空间和超平面的交集，所以是凸集

单纯形：R^n 中选择 $v_0,\ldots,v_k$ 共 $k+1$ 个点，使 $v_1-v_0,\ldots,v_k-v_0$ 线性无关，则与上述点相关的单纯形为 $C = Conv\{v_0,\ldots,v_k\} = \{\theta_0 v_0+\cdots+\theta_k v_k \mid \theta \geq 0, 1^T\theta = 1\}$，即找到这 k 个点的凸包

注意 $\{x \mid x \leq 0\}$ 是凸集/多面体/单纯形，即一维空间下取 $x_1 = 0,\ x_2 = -\infty$ 的凸包

求证：单纯形是多面体的一种

证明：定义 $y=[\theta_1,\dots,\theta_k],\ y\ge 0,\ 1^Ty\le 1$(注意舍弃了 θ_0)，$B=[v_1-v_0,\dots,v_k-v_0]\in R^{n\times k}$

$x\in C\Leftarrow x=\theta_0v_0+\cdots+\theta_kv_k=v_0+\theta_1(v_1-v_0)+\cdots+\theta_k(v_k-v_0)=v_0+By$

B 满秩，$rank(B)=k\Rightarrow$ 通过非奇异 (可逆) 矩阵 $A=\begin{bmatrix}A_1\\A_2\end{bmatrix}\in R^{n\times n}$，$\begin{bmatrix}A_1\\A_2\end{bmatrix}B=\begin{bmatrix}I_k\\0\end{bmatrix}$

$Ax=Av_0+ABy\Rightarrow\begin{bmatrix}A_1\\A_2\end{bmatrix}x=\begin{bmatrix}A_1\\A_2\end{bmatrix}v_0+\begin{bmatrix}I_k\\0\end{bmatrix}y$

利用 $y\ge 0,\ 1^Ty\le 1$，有 $A_1x\ge A_1v_0,\ 1^TA_1x\le 1^TAv_0+1$

则单纯形中的 x 可以表示为 $\{x\mid A_1x\ge A_1v_0,\ 1^TA_1x\le 1^TAv_0+1,A_2x=A_2v_0\}$

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：$S^n_+=\{x\in R^{n\times n}\mid X=X^T,\ X\succeq 0\}$

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

求证：S^n_+ 是凸锥证明：$\forall\theta_0,\theta_1\ge 0,\ \forall A,B\in S^n_+$，对于 $\theta_1A+\theta_2B$，首先对称性显然成立

$\forall x\in R^n,\ X^TAX\ge 0,\ X^TBX\ge 0,\ \theta_1X^TAX+\theta_2X^TBX\ge 0$，$\theta_1A+\theta_2B\succ 0$，得证

求证：S^n_{++} 不是凸锥

证明：首先考虑 n=1 的情况，不过原点

高维情况下 $\theta_1X^TAX+\theta_2X^TBX\ge 0\notin S^n_{++}$，因为 θ_1,θ_2 可同时为 0

Essence of the lecture (7/8)

几种重要的保凸运算：

- 凸集的交操作
- 仿射变换（缩放与位移，线性矩阵不等式）
- 透视函数
- 线性分数函数（仿射和透视的融合）

交集：若 S_0 为凸集，$\forall a \in A$，则 $\bigcap\limits_{a\in A} S_a$ 为凸集

仿射函数：$f: R^n \to R^m$ 是仿射的 (线性映射)，当 $f = Ax + b$，$A \in R^{m\times n}$，$b \in R^m$

若 $S \in R^n$ 为凸，则 $f(S) = \{f(x) \mid x \in S\}$ 为凸

缩放与位移：$\alpha S = \{\alpha x \mid x \in S\}$ $\quad S + a = \{x + a \mid x \in S\}$

求证：两个凸集的和仍旧是凸集，定义凸集的和为 $(S_1 + S_2 = \{x + y \mid x \in S_1,\ y \in S_2\})$

证明：因为仿射映射是从一个集合的角度应用的，所以要将两个集合融合为一个

定义 $S_1 \times S_2 = \{(x, y) \mid x \in S_1,\ y \in S_2\}$，由凸集定义显然这是一个凸集

则令 $f(x, y) = x + y$，有 $S_1 + S_2 = f(S_1 \times S_2)$，由仿射变换保凸知，凸集和仍为凸集

线性矩阵不等性（LMI）：$A(x) = x_1A_1 + \cdots + x_nA_n \preceq B$，$B$，$A_i$，$x_i \in S^m$

求证：LMI 的解为凸集，即 $\{x \mid A(x) \preceq B\}$ 为凸集

证明：定义仿射变换 $f(x) \triangleq B - A(x)$，注意 $f(x)$ 中的 $x = [A_1, A_2, \ldots, A_n]$，是多个矩阵，

而 $B - A(x)$ 返回的是一个矩阵，f 是从高维向低维的一个映射

由 S_+^n 为凸，所以 $f^{-1}(S_+^n) = \{x \mid B - A(x) \succeq 0\} \Leftrightarrow \{x \mid A(x) \preceq B\}$ 为凸，得证

求证：椭球是球的仿射映射

证明：椭球：$\epsilon = \{x \mid (x - x_c)^T p^{-1}(x - x_c) \le 1\}$，$p \in S_{++}^m$，单位球：$\{u \mid \|u\|_2 \le 1\}$

定义 $f(u) = p^{\frac{1}{2}}u + x_c$，则 $\{f(u) \mid \|u\|_2 \le 1\} = \{p^{\frac{1}{2}}u + x_c \mid \|u\|_2 \le 1\}$

令 $x \triangleq p^{\frac{1}{2}}u + x_c \Rightarrow u = p^{-\frac{1}{2}}(x - x_c)$（由于 $p \in S_{++}^m$，p 可以求逆）

回代得 $\{x \mid \|p^{-\frac{1}{2}}(x - x_c)\|_2 \le 1\} = \{x \mid (x - x_c)^T p^{-1}(x - x_c) \le 1\}$

透视函数（perspective function）：$P:\ R^{n+1} \to R^n$，$dom\ P:\ R^n \times R_{++}$

$P(z, t) = \frac{z}{t}$，$z \in R^n$，$t \in R_{++}$，凸集通过透视变换仍为凸集

考虑 R^n 内的线段，$x=(\tilde{x},x_{n+1})$，$y=(\tilde{y},y_{n+1})$，$\tilde{x},\tilde{y}\in R^n$, $x_{n+1},y_{n+1}\in R_{++}$, $\theta\geq 0$，线段为 $\theta x+(1-\theta)y$

求证：任意线段通过透视函数后仍为凸集

证明：P 是透视函数，证明如下

$$\begin{aligned}P(\theta x+(1-\theta)y)&=\frac{\theta\tilde{x}+(1-\theta)\tilde{y}}{\theta x_{n+1}+(1-\theta)y_{n+1}}\\&=\frac{\theta x_{n+1}}{\theta x_{n+1}+(1-\theta)y_{n+1}}\frac{\tilde{x}}{x_{n+1}}+\frac{(1-\theta)y_{n+1}}{\theta x_{n+1}+(1-\theta)y_{n+1}}\frac{\tilde{y}}{y_{n+1}}\\&=\mu P(x)+(1-\mu)P(y)\end{aligned}$$

求证：任意凸集的反透视映射仍为凸集，即 $P^{-1}(C)=\{(x,t)\in R^{n+1}\mid\frac{x}{t}\in C,\ t>0\}$

证明：考虑该该反透视映射集合中的两点 $(x,t)\in P^{-1}(C)$，$(y,s)\in P^{-1}(C)$，$0\leq\theta\leq 1$

即证：$P(\theta(x,t)+(1-\theta)(y,s))\in C$

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

线性分数函数：仿射映射和透视映射的结合

$\delta:\ R^n\to R^{m+1}$ 为仿射映射，$\delta(x)=\begin{bmatrix}A\\C^T\end{bmatrix}x+\begin{bmatrix}b\\d\end{bmatrix}$，$A\in R^{m\times n},b\in R^m,C\in R^n,d\in R$

$P:\ R^{m+1}\to R^m$，线性分数函数 $f:\ R^n\to R^m\triangleq P\circ\delta$

$f(x)=\dfrac{Ax+b}{cx+d},\ dom f=\{x\mid c^Tx+d>0\}$

经过 2 次保凸运算后，结果仍为凸集，即线性分数函数保凸。

其一定是拟凸函数，但不一定是凸函数，见 α-sublevel set(17/18)

求证：两个随机变量（$u\in\{1,2,\ldots,n\}$，$v\in\{1,2,\ldots,m\}$）的联合概率 $\to$ 条件概率保凸

证明：$P_{ij}=P(u=i,v=j)$，$f_{ij}=P(u=i\mid v=j)\Rightarrow f_{ij}=\dfrac{P_{ij}}{\sum_{k=1}^n P_{kj}}$

这是一个线性分数函数，x 是 $[P_{1j},\ldots,P_{nj}]$，分子由 $[0,\ldots,1,\ldots,0]$ 点乘 x，

分母由 1 向量点乘得来，因此该映射是一个保凸的映射。