1、内积inner product

> < >
> 标量

- 柯西不等式

$$|x^{\top}y| \leq \|x\|_2\|y\|_2, x, y \in \mathbf{R}^n$$

- 夹角

$$\angle(x, y) = \cos^{-1}\left(\frac{x^{\top}y}{\|x\|_2\|y\|_2}\right), x, y \in \mathbf{R}^n$$

- 矩阵内积

> trace：把矩阵对角元素求和
>
> 把对应元素相乘求和，和向量是一回事

$$\langle X, Y\rangle = \mathrm{tr}(X^\top Y) = \sum_{i=1}^{m}\sum_{j=1}^{n} X_{ij} Y_{ij}$$

- 对称矩阵

$$\langle X, Y\rangle = \mathrm{tr}(XY) = \sum_{i=1}^{n}\sum_{j=1}^{n} X_{ij} Y_{ij} = \sum_{i=1}^{n} X_{ii} Y_{ii} + 2\sum_{i<j} X_{ij} Y_{ij}$$

2、函数是范数

- [ ] 条件：

1. 非负
2. 正定，仅f（0）= 0
3. 绝对值线性
4. 满足三角不等式

A function $f: \mathbf{R}^n \to \mathbf{R}$ with $\mathrm{dom}\, f = \mathbf{R}^n$ is called a norm if

- f is nonnegative: $f(x) \geq 0$ for all $x \in \mathbf{R}^n$
- f is definite: $f(x) = 0$ only if $x = 0$
- f is homogeneous: $f(tx) = |t| f(x)$, for all $x \in \mathbf{R}^n$ and $t \in \mathbf{R}$
- f satisfies the triangle inequality:
  $f(x+y) \leq f(x) + f(y)$, for all $x, y \in \mathbf{R}^n$

3、单位球unit ball

> 用范数构建单位球
>
> 用二范数构建的球才是圆的
>
> 单位球反过来也可以定义范数

$$\mathcal{B} = \{x \in \mathbf{R}^n \mid \|x\| \leq 1\}$$

1. 关于原点对称
2. 凸的convex

3. 闭合的，有界的，有非空内部

常见范数

1.绝对值求和范数l1

$$\|x\|_1 = |x_1| + \cdots + |x_n|, x \in \mathbf{R}^n$$

2.l无穷

$$\|x\|_\infty = \max\{|x_1|, \ldots, |x_n|\}$$

3.lp

> p>=1,p可以趋向无穷（l无穷）

$$\|x\|_p = (|x_1|^p + \cdots + |x_n|^p)^{1/p}$$

4.二次范数

> P是对称n*n正定矩阵，P可开根
>
> 从P范数转化为2范数

For $P \in \mathbf{S}^n_{++}$, P-quadratic norm is

$$\|x\|_P = (x^\top P x)^{1/2} = \|P^{1/2} x\|_2$$

5.矩阵绝对值求和

> 对应向量的l1范数
>
> 把矩阵拉直成向量

$$\|X\|_{\text{sav}} = \sum_{i=1}^{m} \sum_{j=1}^{n} |X_{ij}|$$

6.矩阵绝对值最大

> 对应向量的l无穷范数
>
> 把矩阵拉直成向量

$$\|X\|_{\text{mav}} = \max\{|X_{ij}| \mid i = 1, \dots, m, j = 1, \dots, n\}$$

7.矩阵的Frobenius norm

> 对应向量的二范数

$$\|X\|_F = \left(\text{tr}(X^\top X)\right)^{1/2} = \left(\sum_{i=1}^{m}\sum_{j=1}^{n} X_{ij}^2\right)^{1/2}$$

4、范数等价性

> 范数之间差的只是常数倍
>
> 极限分析才考虑等价性
>
> 机器学习有限时间计算时，等价性用处不大，谨慎使用

Equivalence of norms

- Suppose that $\|\cdot\|_a$ and $\|\cdot\|_b$ are norms on $\mathbf{R}^n$, there exist positive constants a and β, for all $x \in \mathbf{R}^n$

$$\alpha\|x\|_a \le \|x\|_b \le \beta\|x\|_a$$

- If $\|\cdot\|$ is any norm on $\mathbf{R}^n$, then there exists a quadratic norm $\|\cdot\|_P$ for which

$$\|x\|_P \le \|x\| \le \sqrt{n}\,\|x\|_P$$

holds for all x

5、算子范数operator norms

> sup=max
>
> 矩阵被a和b两个范数定义了算子范数
>
> b范数用来选择控制u，Xu用a范数算

Operator norms

- Suppose $\|\cdot\|_a$ and $\|\cdot\|_b$ are norms on $\mathbf{R}^m$ and $\mathbf{R}^n$, respectively. Operator norm of $X \in \mathbf{R}^{m\times n}$ induced by $\|\cdot\|_a$ and $\|\cdot\|_b$ is

$$\|X\|_{a,b} = \sup\{\|Xu\|_a \mid \|u\|_b \leq 1\}$$

- When $\|\cdot\|_a$ and $\|\cdot\|_b$ are Euclidean norms, the operator norm of X is its maximum singular value, and is denoted $\|X\|_2$

$$\|X\|_2 = \sigma_{\max}(X) = \left(\lambda_{\max}(X^\top X)\right)^{1/2}$$

  ✓ Spectral norm or ℓ_2-norm

> σ奇异值，2范数就是求最大奇异值

- 特殊：把a，b设为1或无穷范数

6、对偶范数dual norms

> ||·||表示Rn上任意范数

$$\|z\|_* = \sup\{z^\top x \mid \|x\| \leq 1\}$$

性质1

- 范数的共轭函数是对偶范数单位球的示性函数

例 3.26　范数。 令 $\|\cdot\|$ 表示 $\mathbf{R}^n$ 上的范数，其对偶范数为 $\|\cdot\|_*$。我们说明 $f(x) = \|x\|$ 的共轭函数为

$$f^*(y) = \begin{cases} 0 & \|y\|_* \leqslant 1 \\ \infty & \text{其他情况}, \end{cases}$$

即范数的共轭函数是对偶范数单位球的示性函数。

如果 $\|y\|_* > 1$，根据对偶范数的定义，存在 $z \in \mathbf{R}^n$，$\|z\| \leqslant 1$ 使得 $y^Tz > 1$。取 $x = tz$，令 $t \to \infty$ 可得

$$y^Tx - \|x\| = t(y^Tz - \|z\|) \to \infty,$$

即 $f^*(y) = \infty$，没有上界。反之，若 $\|y\|_* \leqslant 1$，对任意 x，有 $y^Tx \leqslant \|x\|\|y\|_*$，即对任意 x，$y^Tx - \|x\| \leqslant 0$。因此，在 $x = 0$ 处，$y^Tx - \|x\|$ 达到最大值 0。

性质2

- inequality

> 任意两向量内积
>
> 范数相乘可以是内积上界

$$z^\top x \leq \|x\|\|z\|_*$$

证明：

> x/||x||的范数为1，满足对偶范数定义（换元）
> sup上确界

$$z^\top x = z^\top \frac{x}{\|x\|} \cdot \|x\| \leq \|z\|_* \|x\|$$

$$z^\top \frac{x}{\|x\|} \leq \sup\{z^\top x \mid \|x\| \leq 1\} = \|z\|_*$$

性质3

- 2范数的对偶是自己

$$\sup\{z^\top x \mid \|x\|_2 \leq 1\} = \|z\|_2$$

- 无穷的对偶是1

$$\sup\{z^\top x \mid \|x\|_\infty \leq 1\} = \|z\|_1$$

- 对偶的对偶等于自己
- 计算对偶范数公式

The dual of ℓ_p-norm is the ℓ_q-norm such that

$$\frac{1}{p} + \frac{1}{q} = 1$$

矩阵的对偶范数

> 只需要加一个trace变成内积
> 2范数的对偶范数，核范数，是最大的一个内积，等于奇异值之和（结论）

The dual of the ℓ_2-norm on $\mathbf{R}^{m\times n}$ is the nuclear norm

$$\|Z\|_{2*} = \sup\{\mathrm{tr}(Z^\top X) \mid \|X\|_2 \leq 1\}$$
$$= \sigma_1(Z) + \cdots + \sigma_r(Z) = \mathrm{tr}\left[(Z^\top Z)^{1/2}\right]$$

> 秩r，有r个非零奇异值
> σ1也表示最大奇异值
> 最后一行是两个结论