矩阵理论与方法

内容提要 CONTENTS

- □课程信息
- □课程介绍
- □ 矩阵理论与方法

第2章 范数理论及其应用

第2章 范数理论及其应用

在计算数学中,特别是在数值代数中,研究数值方法的收敛性、稳定性及误差分析等问题时,范数理论显得十分重要.本章主要讨论 n 维向量空间 C^n 中的向量范数与矩阵空间 $C^{m\times n}$ 中的矩阵范数的理论及其性质.

$$T^k \Rightarrow A^k$$

$$e^T \Rightarrow e^A$$

极限

导数

积分



$$E_k = \frac{1}{2} \sum_{j=1}^{l} (\hat{y}_j^k - y_j^k)^2$$
.

第5章 神经网络

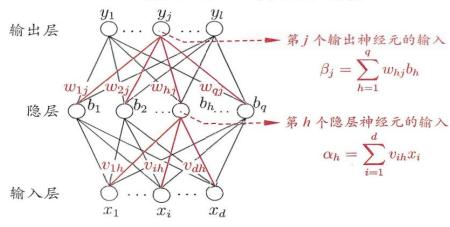


图 5.7 BP 网络及算法中的变量符号

BP 算法基于梯度下降(gradient descent)策略, 以目标的负梯度方向对参数进行调整. 对式(5.4)的误差 E_k , 给定学习率 η , 有

$$\Delta w_{hj} = -\eta \frac{\partial E_k}{\partial w_{hj}} \ . \tag{5.6}$$

我们常会谈到两种"最优": "局部极小"(local minimum)和"全局最小"(global minimum). 对 \boldsymbol{w}^* 和 θ^* , 若存在 $\epsilon > 0$ 使得

$$\forall (\boldsymbol{w}; \theta) \in \{(\boldsymbol{w}; \theta) \mid ||(\boldsymbol{w}; \theta) - (\boldsymbol{w}^*; \theta^*)|| \leq \epsilon\},$$

$$T^k \Rightarrow A^k$$

$$e^T \Rightarrow e^A$$

极限

导数

积分

$$A^{(k)} \to A \quad (k \to +\infty)$$

$$x^{(k)} \rightarrow x \quad (k \rightarrow +\infty), \quad x^{(k)}, \quad x \in V$$

$$x^{(k)} \rightarrow x \quad (k \rightarrow +\infty), \quad x^{(k)}, \quad x \in V$$

$$\Leftrightarrow \lim_{k \to +\infty} x^{(k)} = x$$

$$\forall \varepsilon > 0, \exists N > 0, k > N$$
时,有

$$||x^k - x|| < \varepsilon$$

?

1、向量范数的概念

a) R^n 上的向量范数

b) 线型空间V上的向量范数

c) $R^{m \times n}$ 上的矩阵范数

§ 2.1 向量范数及其性质

一、向量范数的概念及1,范数

设给定了n维向量空间 \mathbb{R}^n 中的向量序列 $\{x^{(k)}\}$,其中 $x^{(k)}=(\xi_1^{(k)},\xi_2^{(k)},\cdots,\xi_n^{(k)})$ $(k=1,2,3,\cdots)$. 如果每一个分量 $\xi_i^{(k)}$,当 $k \to \infty$ 时都有极限 ξ_i ,即

$$\lim_{k\to\infty} \boldsymbol{\xi}_i^{(k)} = \boldsymbol{\xi}_i \quad (i=1, 2, \dots, n)$$

记 $x = (\xi_1, \xi_2, \dots, \xi_n)$,则称向量序列 $\{x^{(k)}\}$ 有极限x,或称 $x^{(k)}$ 收敛于x,简称 $\{x^{(k)}\}$ **收敛**,记为

$$\lim_{k\to\infty} x^{(k)} = x \quad \text{id} \quad x^{(k)} \to x$$

不收敛的向量序列称为是发散的. 例如向量序列

$$\mathbf{x}^{(k)} = \begin{bmatrix} \frac{1}{2^k} \\ \frac{\sin k}{k} \end{bmatrix} \qquad (k = 1, 2, 3, \dots)$$

是收敛的. 因为当 $k \to \infty$ 时, $\frac{1}{2^k} \to 0$, $\frac{\sin k}{k} \to 0$, 所以

$$\lim_{k \to \infty} \mathbf{x}^{(k)} = \begin{bmatrix} \lim_{k \to \infty} \frac{1}{2^k} \\ \lim_{k \to \infty} \frac{\sin k}{k} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

而向量序列

$$\mathbf{x}^{(k)} = \begin{bmatrix} \sum_{i=1}^{k} \frac{1}{2^{i}} \\ \sum_{i=1}^{k} \frac{1}{i} \end{bmatrix} \qquad (k = 1, 2, 3, \dots)$$

是发散的. 因为
$$\sum_{i=1}^{k} \frac{1}{2^i} = \frac{1}{2} \frac{1 - (1/2)^{k+1}}{1 - 1/2} \to 1$$
,而 $\sum_{i=1}^{k} \frac{1}{i} \to \infty$.

显然,如果向量序列 $\{x^{(k)}\}$ 收敛到向量 x,则向量序列 $\{x^{(k)}-x\} = \{(\xi_1^{(k)}-\xi_1,\xi_2^{(k)}-\xi_2,\cdots,\xi_n^{(k)}-\xi_n)\}$ 一定收敛到零向量 $\{0,0,\cdots,0\}$;反之亦然. 当 $\lim_{k\to\infty} x^{(k)} = x$ 时,向量 $x^{(k)}-x$ 的欧氏长度

 $||x^{(k)} - x|| = \sqrt{(\xi_1^{(k)} - \xi_1)^2 + (\xi_2^{(k)} - \xi_2)^2 + \dots + (\xi_n^{(k)} - \xi_n)^2}$ 收敛于零;

显然,如果向量序列 $\{x^{(k)}\}$ 收敛到向量x,则向量序列

$$\{x^{(k)}-x\}=\{(\xi_1^{(k)}-\xi_1,\,\xi_2^{(k)}-\xi_2,\,\cdots,\,\xi_n^{(k)}-\xi_n)\}$$

一定收敛到零向量(0,0,…,0);反之亦然.

$$||x^{(k)}-x|| = \sqrt{(\xi_1^{(k)}-\xi_1)^2+(\xi_2^{(k)}-\xi_2)^2+\cdots+(\xi_n^{(k)}-\xi_n)^2}$$

收敛于零;反之,若有一向量序列的欧氏长度收敛于零,则它的每一个分量一定收敛于零,从而该向量序列收敛于零向量.

由上述可见,向量的长度可用来刻画收敛的性质.但是对于一般的线性空间,如何定义向量的长度呢?这就是所谓范数的概念.范数是比长度更为广泛的概念,现定义于下.

定义 2.1 如果 V 是数域 K 上的线性空间,且对于 V 的任一向量 x,对应一个实值函数 $\|x\|$,它满足以下三个条件:

- (1) 非负性: 当 $x \neq 0$ 时, ||x|| > 0; 当x = 0时, ||x|| = 0;
- (2) 齐次性: $||ax|| = |a| ||x|| (a \in K, x \in V)$;
- (3) 三角不等式: $||x+y|| \le ||x|| + ||y||$ ($x, y \in V$). 则称 ||x|| 为 V 上向量x 的范数,简称向量范数.

向量的范数具有下列简单性质:

(1)
$$||x|| \neq 0$$
 $||x|| \neq 0$ $||x|| = 1$ $||x|| = \frac{1}{||x||} ||x|| = \frac{1}{||x||} ||x|| = 1$

(2)
$$\forall x \in V$$
 , $||-x|| = ||x||$

$$\left\| \frac{1}{\|x\|} x \right\| = \frac{1}{\|x\|} \|x\| = 1$$

(2)
$$\forall x \in V$$
 , $||-x|| = ||x||$ $||-x|| = |-1|||x|| = ||x||$

向量的范数具有下列简单性质:

(1)
$$\stackrel{\text{if}}{=} \|x\| \neq 0$$
 $\stackrel{\text{if}}{=} \frac{1}{\|x\|} \|x\| = 1$ $\therefore \left\| \frac{1}{\|x\|} x \right\| = \frac{1}{\|x\|} \|x\| = 1$

(2)
$$\forall x \in V$$
 , $||-x|| = ||x||$ $||-x|| = |-1|||x|| = ||x||$

(3)
$$\forall x, y \in V$$
 , $||x|| - ||y|| \le ||x - y||$

$$||x|| = ||(x - y) + y|| \le ||x - y|| + ||y|| \Rightarrow ||x|| - ||y|| \le ||x - y||$$

(4)
$$\forall x, y \in V$$
, $||x|| - ||y|| \le ||x - y||$

$$||x|| = ||(x-y) + y|| \le ||x-y|| + ||y|| \implies ||x|| - ||y|| \le ||x-y||$$

同样
$$||y|| - ||x|| \le ||y - x|| = ||x - y||$$

例 2.1 在 n 维酉空间 \mathbb{C}^n 上,复向量 $x = (\xi_1, \xi_2, \dots, \xi_n)$ 的 长度

$$\|x\| = \sqrt{|\xi_1|^2 + |\xi_2|^2 + \cdots + |\xi_n|^2}$$
 (2.1.1) 就是一种范数.

定义 2.1 如果 V 是数域 K 上的线性空间,且对于 V 的任一向量 x,对应一个实值函数 $\|x\|$,它满足以下三个条件:

- (1) 非负性: 当 $x \neq 0$ 时, ||x|| > 0; 当x = 0时, ||x|| = 0;
- (2) 齐次性: $||ax|| = |a| ||x|| (a \in K, x \in V);$
- (3) 三角不等式: $||x+y|| \le ||x|| + ||y||$ ($x, y \in V$). 则称 ||x|| 为 V 上向量x 的范数,简称向量范数.

例 2.1 在 n 维酉空间 \mathbb{C}^n 上,复向量 $x = (\xi_1, \xi_2, \dots, \xi_n)$ 的长度

$$\|x\| = \sqrt{|\xi_1|^2 + |\xi_2|^2 + \dots + |\xi_n|^2}$$
 (2.1.1) 就是一种范数.

为了说明这里的 $\|x\|$ 是范数,只需验证它满足范数的三个条件就行了.事实上有:

(1) 对于 $\|x\| = \sqrt{|\xi_1|^2 + |\xi_2|^2 + \dots + |\xi_n|^2}$, 当 $x \neq 0$ 时, 显然 $\|x\| > 0$; 当 x = 0 时, 则 $\|x\| = \sqrt{0^2 + \dots + 0^2} = 0$.

例 2.1 在 n 维酉空间 \mathbb{C}^n 上,复向量 $x = (\xi_1, \xi_2, \dots, \xi_n)$ 的长度

$$\|x\| = \sqrt{|\xi_1|^2 + |\xi_2|^2 + \dots + |\xi_n|^2}$$
 (2.1.1)
 就是一种范数.

为了说明这里的 $\|x\|$ 是范数,只需验证它满足范数的三个条件就行了.事实上有:

- (1) 对于 $\|x\| = \sqrt{|\xi_1|^2 + |\xi_2|^2 + \cdots + |\xi_n|^2}$, 当 $x \neq 0$ 时, 显然 $\|x\| > 0$; 当 x = 0 时, 则 $\|x\| = \sqrt{0^2 + \cdots + 0^2} = 0$.
 - (2) 对任意的复数 a, 因为

$$a\mathbf{x}=(a\xi_1, a\xi_2, \dots, a\xi_n)$$

所以

$$||ax|| = \sqrt{|a\xi_1|^2 + |a\xi_2|^2 + \dots + |a\xi_n|^2} = |a| \sqrt{|\xi_1|^2 + |\xi_2|^2 + \dots + |\xi_n|^2} = |a| ||x||$$

(3) 对于任意两个复向量 $x = (\xi_1, \xi_2, \dots, \xi_n), y = (\eta_1, \eta_2, \dots, \eta_n),$ 有

$$x + y = (\xi_1 + \eta_1, \xi_2 + \eta_2, \dots, \xi_n + \eta_n)$$

所以

借助于 C"中内积式(1.3.24) 及其性质,可得

$$||x+y||^2 = (x+y, x+y) =$$
 $(x, x) + 2\operatorname{Re}(x, y) + (y, y)$

因为

Re
$$(x, y) \leqslant \mid (x, y) \mid \leqslant \sqrt{(x, x)(y, y)} = \parallel x \parallel \parallel y \parallel$$
 所以

$$\| \mathbf{x} + \mathbf{y} \|^{2} \le \| \mathbf{x} \|^{2} + 2 \| \mathbf{x} \| \| \mathbf{y} \| + \| \mathbf{y} \|^{2} = (\| \mathbf{x} \| + \| \mathbf{y} \|)^{2}$$

$$\mathbb{P} \|x+y\| \leqslant \|x\| + \|y\|.$$

例 2.1 在 n 维酉空间 \mathbb{C}^n 上,复向量 $x = (\xi_1, \xi_2, \dots, \xi_n)$ 的长度

$$\|x\| = \sqrt{|\xi_1|^2 + |\xi_2|^2 + \cdots + |\xi_n|^2}$$
 (2.1.1)
 就是一种范数.

因此,式(2.1.1) 是 C^n 上的一种范数. 通常称这种范数为 2 - 范数,记作 $\|x\|_2$. 式(2.1.1) 也是欧氏空间 R^n 上的一种范数,这只要把复数域 C 改为实数域 R 即可.

 \mathbf{M} 2.2 证明 $\|x\| = \max_{i} |\xi_{i}|$ 是 \mathbf{C}^{n} 上的一种范数,这里 $x = (\xi_{1}, \xi_{2}, \dots, \xi_{n}) \in \mathbf{C}^{n}$.

称例 2. 2 中的范数为
$$\infty$$
 -范数,记为 $\|x\|_{\infty}$,即 $\|x\|_{\infty} = \max |\xi_i|$ (2. 1. 4)

例 2.2 证明 $||x|| = \max_{i} |\xi_{i}|$ 是 \mathbb{C}^{n} 上的一种范数,这里 $x = (\xi_{1}, \xi_{2}, \dots, \xi_{n}) \in \mathbb{C}^{n}$.

证 当 $x \neq 0$ 时,有 $||x|| = \max_{i} |\xi_{i}| > 0$;当x = 0时,显然有 ||x|| = 0.

又对任意的 $a \in C$,有

 $||ax|| = \max_{i} |a\xi_{i}| = |a| \max_{i} |\xi_{i}| = |a| ||x||$

对 \mathbb{C}^n 的任意两个向量 $x = (\xi_1, \xi_2, \dots, \xi_n), y = (\eta_1, \eta_2, \dots, \eta_n),$ 有

$$||x + y|| = \max_{i} |\xi_{i} + \eta_{i}| \le \max_{i} |\xi_{i}| + \max_{i} |\eta_{i}| = ||x|| + ||y||$$

因此, $||x|| = \max |\xi_i|$ 是 \mathbb{C}^n 上的一种范数.

称例 2.2 中的范数为 ∞ -范数,记为 $\|x\|_{\infty}$,即

$$\|x\|_{\infty} = \max_{i} |\xi_{i}| \qquad (2.1.4)$$

例 2. 3 证明 $||x|| = \sum_{i=1}^{n} |\xi_i|$ 也是 \mathbb{C}^n 上的一种范数,其中 $x = (\xi_1, \xi_2, \dots, \xi_n) \in \mathbb{C}^n$.

称例 2.3 中的范数为 $1-范数, 记为 ||x||_1, 即$

$$\| \mathbf{x} \|_{1} = \sum_{i=1}^{n} | \xi_{i} |$$
 (2.1.5)

例 2.3 证明
$$||x|| = \sum_{i=1}^n |\xi_i|$$
 也是 \mathbb{C}^n 上的一种范数,其中 $x = (\xi_1, \xi_2, \dots, \xi_n) \in \mathbb{C}^n$.

证 当 $x \neq 0$ 时,显然 $||x|| = \sum_{i=1}^{n} |\xi_i| > 0$;当 x = 0 时,由于 x 的每一分量都是零,故 ||x|| = 0.

又对于任意 $a \in C$,有

$$||ax|| = \sum_{i=1}^{n} |a\xi_{i}| = |a| \sum_{i=1}^{n} |\xi_{i}| = |a| ||x||$$

对任意两个向量 $x, y \in \mathbb{C}^n$,有

$$||x + y|| = \sum_{i=1}^{n} |\xi_{i} + \eta_{i}| \leq \sum_{i=1}^{n} (|\xi_{i}| + |\eta_{i}|) =$$

$$\sum_{i=1}^{n} |\xi_{i}| + \sum_{i=1}^{n} |\eta_{i}| = ||x|| + ||y||$$

于是由定义 2.1 知 $\|x\| = \sum_{i=1}^{n} |\xi_i|$ 是 \mathbb{C}^n 上的一种范数.

线性空间
$$C^n$$
,设 $x = (x_1, x_2, \dots, x_n)^T \in C^n$

1:
$$\|x\|_{1} = \sum |\xi_{i}|$$
 是一种向量范数,记为**1**-范数

2:
$$||x|| = \sqrt{(x,x)}$$
 是一种向量范数,记2-范数

3:
$$\|x\| = \max_{i} |x_{i}|$$
 是一种向量范数,记为 ∞ -范数

线性空间
$$C^n$$
,设 $x = (x_1, x_2, \dots, x_n)^T \in C^n$

1:
$$\|x\|_{1} = \sum |\xi_{i}|$$
 是一种向量范数,记为**1-**范数

2:
$$||x|| = \sqrt{(x,x)}$$
 是一种向量范数,记2-范数

3:
$$\|x\| = \max_{i} |x_{i}|$$
 是一种向量范数,记为 ∞ -范数

例:求向量
$$e = (1,2,-3)$$
的 l_1,l_2 和 l_∞ 范数

线性空间
$$C^n$$
,设 $x = (x_1, x_2, \dots, x_n)^T \in C^n$

1:
$$\|x\|_{1} = \sum |\xi_{i}|$$
 是一种向量范数,记为**1**-范数

2:
$$||x|| = \sqrt{(x,x)}$$
 是一种向量范数,记2-范数

3:
$$\|x\| = \max_{i} |x_{i}|$$
 是一种向量范数,记为 ∞ -范数

例:求向量
$$e = (1,2,-3)$$
的 l_1, l_2 和 l_∞ 范数

解:
$$||e||_1 = 6$$
, $||e||_2 = \sqrt{14}$, $||e||_{\infty} = 3$,

lp范数

线性空间
$$C^n$$
,设 $x = (x_1, x_2, \dots, x_n)^T \in C^n$

1:
$$\|x\|_{1} = \sum |\xi_{i}|$$
 是一种向量范数,记为**1-**范数

2:
$$||x|| = \sqrt{(x,x)}$$
 是一种向量范数,记2-范数

3:
$$\|x\| = \max_{i} |x_{i}|$$
 是一种向量范数,记为 ∞ -范数

4:
$$\|x\|_p = \left(\sum_{i=1}^n |x_i|^p\right)^{\frac{1}{p}} \quad \left(1 \le p < \infty\right)$$

是一种向量范数,记为p-范数或 l_p 范数

lp范数

称
$$\left(\sum_{i=1}^{n} |\xi_{i}|^{p}\right)^{1/p}$$
为向量 x 的 p -范数或 l_{p} 范数,记为 $||x||_{p}$, $(1 \leq p < +\infty)$ $||x||_{p} = \left(\sum_{i=1}^{n} |\xi_{i}|^{p}\right)^{1/p}$ (2.1.7)

在式(2.1.7) 中,令
$$p=1$$
,便得 $\|x\|_1$;令 $p=2$,便得 $\|x\|_2$;并且还有 $\|x\|_{\infty}=\lim_{p\to\infty}\|x\|_p$

向量空间中向量范数

前面讲的是С"的向量范数

问题: V"的向量范数怎么构造?

向量空间中向量范数

例2.6 给定线性空间 V^n 的基 x_1, x_2, \dots, x_n ,设 $x \in V^n$ 在该基下的坐标向量为 $\tilde{x} = (\xi_1, \xi_2, \dots, \xi_n)^T$,那么

$$\|x\|_{p} = \|\tilde{x}\|_{p} \quad (1 \leqslant p < +\infty)$$

满足范数定义的三个条件.因此,它是 V^* 上的范数,也称为x的p-范数.

按照例 2.6 的方式,可以在线性空间 V^n 上定义多种不同的向量范数.这样的向量范数不仅依赖于 $C^n(\mathbb{R}^n)$ 上的向量范数,而且与 V^n 中基的选取密切相关.

$$\lim_{k \to +\infty} x^{(k)} = x \Leftrightarrow$$

$$\forall \varepsilon > 0, \exists N > 0, k > N$$

$$\|x^k - x\| < \varepsilon$$

1、向量范数

2、向量范数的等价性

向量范数的等价性

定义:有限维线性空间 V'' 中任意两个向量范数 $\|x\|_{\alpha}$ 和 $\|x\|_{\beta}$,如果存在着正常数 c_1 和 c_2 ,使得 c_1 $\|x\|_{\beta} \le \|x\|_{\alpha} \le c_2$ $\|x\|_{\beta}$ ($\forall x \in V''$)则称范数 $\|x\|_{\alpha}$ 与 $\|x\|_{\beta}$ 等价

向量范数的等价性

定义:有限维线性空间V"中任意两个向量范数 $\|\mathbf{x}\|_{\alpha}$ 和 $\|\mathbf{x}\|_{\beta}$,如果存在着正常数 c_1 和 c_2 , 使得 $c_1 \|x\|_{\beta} \leq \|x\|_{\alpha} \leq c_2 \|x\|_{\beta}$ $(\forall x \in V^n)$ 则称范数 $\|\mathbf{x}\|_{\alpha}$ 与 $\|\mathbf{x}\|_{\beta}$ 等价

(1) 自反性:
$$1 \cdot \|x\|_{\alpha} \leq \|x\|_{\alpha} \leq 1 \cdot \|x\|_{\alpha}, \forall x \in V^n$$

(2) 对称性:
$$\frac{1}{c_2} \|x\|_{\alpha} \le \|x\|_{\beta} \le \frac{1}{c_1} \|x\|_{\alpha}, \forall x \in V^n$$

(2) 对称性:
$$\frac{1}{c_{2}}\|x\|_{\alpha} \leq \|x\|_{\beta} \leq \frac{1}{c_{1}}\|x\|_{\alpha}, \forall x \in V^{n}$$
(3) 传递性: $c_{1}\|x\|_{\beta} \leq \|x\|_{\alpha} \leq c_{2}\|x\|_{\beta}$
 $c_{3}\|x\|_{\gamma} \leq \|x\|_{\beta} \leq c_{4}\|x\|_{\gamma}$
 $\Rightarrow c_{5}\|x\|_{\gamma} \leq \|x\|_{\alpha} \leq c_{6}\|x\|_{\gamma}$

向量范数的等价性

二、线性空间 V'' 上的向量范数的等价性

前面已经指出,在数域 K 上的线性空间 V,特别是在 \mathbb{C}^n 上可以定义各种各样的向量范数,其数值大小一般不同. 但是,在各种向量范数之间存在下述重要关系.

定理 2.1 设 $\|x\|_a$ 和 $\|x\|_\beta$ 为有限维线性空间 V 的任意 两种向量范数(它们不限于 p -范数),则存在两个与向量 x 无关的 正常数 c_1 和 c_2 ,使得不等式

 $c_1 \| \mathbf{x} \|_{\beta} \leq \| \mathbf{x} \|_{\alpha} \leq c_2 \| \mathbf{x} \|_{\beta} \quad (\forall \mathbf{x} \in V) \quad (2.1.9)$ 成立.

$$\lim_{k \to +\infty} x^{(k)} = x \iff$$

$$\forall \varepsilon > 0, \exists N > 0, k > N$$
时,有

$$\forall \|\cdot\|: \quad \left\|x^k - x\right\| < \varepsilon$$

作业 (第五版)

- 1、定义: 2.1、2.2
- 2、定理: 2.1、2.2
- 3、例题: 2.1-2.3, 2.6
- 4、习题2.1: 1

作业 (第三版)

- 1、定义: 2.1、2.2
- 2、定理: 2.1、2.2
- 3、例题: 2.1-2.3, 2.6
- 4、习题2.1: 1

下课, 谢谢大家!