latex 体会体会-代码也有注解

zhangyifeng

2018 年 10 月 29 日

目录

1 Hello China

1.1 Hello Beijing

1.1.1 Hello Dongcheng District

Tian'anmen Square is in the center of Beijing

Chairman Mao is in the center of Tian'men Square

Hello Guangzhou

Sun Yat-sen University is the best university in Guangzhou.
This is my first year in UCL[1]. $\sum_{n=0}^{m}$

[1] University College London

第五章　矩阵函数及其微积分

引　言　怎样讨论矩阵的微积分?

在图像处理, 模式识别或移动通信等领域, 常需要利用特定的线性变换将高维向量压缩成低维向量或者将低维向量还原为高维向量, 并且使误差尽可能小. 描述此类问题的一个数学模型如下例:

例 5.0.1 设 $\alpha \in \mathbb{R}^n, \beta \in \mathbb{R}^m$, 求半正交矩阵 $U \in \mathbb{R}^{m \times n}$, 使得

$$\|U\alpha - \beta\| \tag{5.0.1}$$

最小.

例 5.0.1 相当于求以矩阵 U 为自变量的函数 $J(U) = \|U\alpha - \beta\|$ 在约束条件 $U^T U = I$ 或 $UU^T = I$ 下的最小值点 (矩阵), 这样的优化问题具有普遍意义 (属于运筹学的研究领域). 比较一元或多元微分学可知, 解决此问题的一个可行办法是求函数 $J(U) = \|U\alpha - \beta\|$ 关于未知矩阵 U 的导数, 这就需要研究矩阵函数的微积分 (如果将 U 的元素都作为未知数列出, 则目标函数 $\|U\alpha - \beta\|$ 就是一个 mn 元函数, 可以使用多元微分学来研究此问题, 但那将是什么样的场景!). 我们在大学的许多课程和实践中对微积分的强大作用已经深有体会, 如果矩阵能与微积分相结合, 无疑将会产生更为巨大的作用. 那么如何才能将微积分引入到矩阵的研究中来呢? 比较数学分析或高等数学课程, 我们首先需要研究矩阵序列的收敛性, 这就需要计算两个矩阵之间的距离. 一旦有了距离概念, 就能够和数学分析或高等数学几乎完全平行地讨论矩阵序列的极限和矩阵函数的连续性, 进而讨论矩阵函数的微分学与积分学等理论. 我们在第一章内积空间中已经看到, 距离概念可以由长度或范数导出, 而长度或范数可以由内积导出, 因此研究矩阵函数的微积分实际上可以通过在矩阵空间中引入适当的内积后顺利进行. 但一般的无限维线性空间可能没有内积概念 (所有 $m \times n$ 阶函数矩阵构成的线性空间当然是无限维的), 因此我们将在本章第一节研究比内积导出的范数更为广泛的概念, 以使范数能够应用在更大的范围. 那么, 什么是矩阵的范数呢? 我们先看下面简单的例子:

例 5.0.2 设 x 是复数, 则当 $|x| < 1$ 时有

$$(1-x)^{-1} = 1 + x + x^2 + \cdots + x^m + \cdots \tag{5.0.2}$$

问题: 何时上式对于矩阵也成立, 即设 A 是 n 阶矩阵, 公式

$$(I-A)^{-1} = I + A + A^2 + \cdots + A^m + \cdots \tag{5.0.3}$$

何时成立? 即相当于 $|x| < 1$ 的条件是什么? 容易知道一个充分条件是 $\rho(A) < 1$. 但是矩阵的特征值及其谱半径都是不容易计算的, 我们能否改进这个条件? 答案是肯定的, 只需将 $\rho(A) < 1$ 换成 $\|A\| < 1$, 即 A 的范数小于1, 任何一种范数即可! 因此, 矩阵的范数可以看作是实数的绝对值或者复数的模的推广, 是一种衡量矩阵 (包括向量) 大小的尺度.

另外, 我们对矩阵函数的微积分也不陌生, 比如三元函数 $f(x, y, z)$ 的梯度向量

$$\nabla f(x) = (\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}) \tag{5.0.4}$$

$$
\begin{aligned}
&\min_{x \in \boldsymbol{R}^n} \ f(x) \\
&s.t. \ c_i(x) \leq 0, \ i = 1, 2, \cdots, k \\
&\qquad h_j(x) = 0, \ j = 1, 2, \cdots, l
\end{aligned}
\tag{1}
$$