第六章 解线性方程组的迭代解法

6.1 Picard 迭代回顾

内容　有$f(x)=0$，将其化为**等价方程**$x=\varphi(x)$，选取初值x_0，令$x_k=\varphi(x_{k-1})$，则有$\lim\limits_{k\to\infty}x_k=x^*$

x^*即为原方程的解。 $|e_k|=|x_k-x^*|\le\frac{L^k}{1-L}|x_1-x_0|$

向量的迭代公式　对于**线性方程组**$Ax=b$，可以写为**等价方程组**$\boldsymbol{x}=\boldsymbol{Bx}+\boldsymbol{f}$

为此构造序列 $x^{(k+1)}=Bx^{(k)}+f$

要考虑**向量列**的**收敛性**、**收敛速度**、**误差**等内容，和B有关。

收敛性判断：迭代过程中前后两次的距离越来越小。

6.2 基本概念

6.2.1 向量与矩阵的范数

考虑问题
1. 能否定义**一个函数（映射）**将**向量或者矩阵**投影成**一个实数**？
2. 如果存在这个映射，这个映射应该满足什么性质？
3. 作为**距离函数**应该满足什么性质？　　正定性、齐次性、三角不等式

6.2.1.1 向量范数

定义 1　设V为线性空间，V上的实值函数$N(x)=\|x\|$满足：

① **正定性**：$\forall x\in V:\|x\|\ge 0$ 且 $\|x\|=0\iff x=0$　**当且仅当零向量时为零**（不存在负距离）

② 齐次性：$\forall k\in R,\forall x\in V:\|kx\|=|k|\|x\|$　空间中任意向量模长伸缩，不存在负距离，加绝对值

③ 三角不等式：$\forall x,y\in V:\|x+y\|\le\|x\|+\|y\|$　**三角形中边长关系**

则称$\boldsymbol{N(x)}=\|\boldsymbol{x}\|$ **为$\boldsymbol{V}$上向量$\boldsymbol{x}$的范数**

性质　① 若$x\ne 0$，$\left\|\frac{x}{\|x\|}\right\|=1$　② $\|-x\|=\|x\|$　③ $\big|\|x\|-\|y\|\big|\le\|x-y\|$

常用范数　基本条件$\forall x=(x_1,x_2,\dots,x_n)\in R^n\ (C^n)$

① **欧式范数（2-范数）**　$\|x\|_2=\sqrt{\sum_{i=1}^n|x_i|^2}=\sqrt{\sum_{i=1}^n\bar{x}_ix_i}$　模长

② **最大模范数（∞-范数）**　$\|x\|_\infty=\max\limits_{1\le i\le n}|x_i|$　绝对值最大的一个

③ **绝对值范数（1-范数）**　$\|x\|_1=\sum_{i=1}^n|x_i|$　绝对值之和（曼哈顿距离）

④ **p范数**　$\boldsymbol{\|x\|_p=\left(\sum_{i=1}^n|x_i|^p\right)^{\frac{1}{p}}}$

注意　① 可以证明上述定义均满足向量的范数定义

② $2-$范数和$1-$范数都是$p-$范数的特例

③ $\infty-$范数也是$p-$范数的特例　（$p\to\infty$，$\|x\|_p\to\|x\|_\infty$）

④ 范数可以认为是函数的函数。

A 的 F 范数　如果认为$m\times n$矩阵是一个$m\times n$维的向量，$\forall A_{m\times n}$：$\|A\|_2=\sqrt{\sum_{i=1}^m\sum_{j=1}^n\left|a_{ij}\right|^2}\triangleq\|A\|_F$

一般**不能简单的**认为矩阵就是向量

6.2.1.2 矩阵范数

条件　简化起见，先考虑方阵情况

定义 2　若$\forall A_{n\times n}$对应一个实数$\|A\|$，满足有：

① $\|A\| \ge 0$, 且 $\|A\| = 0 \Leftrightarrow A = 0$　　② $\|kA\| = |k|\|A\|$

③ $\|A+B\| \le \|A\| + \|B\|$　　④ $\|AB\| \le \|A\|\|B\|$

则称$\|A\|$为方阵的范数：**矩阵范数**　　例如：$\boldsymbol{F}$**范数为矩阵范数**，行列式不是范数

注意　矩阵理化及运算常要考虑**矩阵与向量的乘积**，我们**希望**范数$\|\boldsymbol{Ax}\|_{向量范数} \le \|\boldsymbol{A}\|_{矩阵范数}\|\boldsymbol{x}\|_{向量范数}$

由向量范数出发，寻找矩阵范数，使得满足条件。则称矩阵范数为由向量范数诱导出的范数。

定义 3　设$\|\cdot\|$为向量范数，$\|\cdot\|_M$为矩阵范数。 若$\forall A \in R^{n\times n}, \forall x \in R^n \Rightarrow$ $\|\boldsymbol{Ax}\| \le \|\boldsymbol{A}\|_{\boldsymbol{M}}\|\boldsymbol{x}\|$

则称$\|A\|_M$为与向量范数$\|\cdot\|$**相容的矩阵范数**

例如，$\|A\|_F$与$\|\cdot\|_1$不相容，如果$A = \begin{pmatrix} 1 & 4 \\ 1 & 4 \end{pmatrix}$　$x = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$

定义 4　设$\forall A \in R^{n\times m}$，$\|\cdot\|$为向量范数，称$\|\boldsymbol{A}\| = \max\limits_{\substack{x\in R^n \\ x\neq 0}} \frac{\|\boldsymbol{Ax}\|}{\|\boldsymbol{x}\|}$ **为矩阵$\boldsymbol{A}$的算子范数（诱导范数）**

注意　① $\|A\| = \max\limits_{\|x\|=1} \|Ax\|$

② 可以证明算子范数满足矩阵范数的 4 个条件.故为矩阵范数.

③ 矩阵范数不一定都是算子范数，如 F-范数.

④ 算子范数与向量范数相容。($\|A\| = \max \frac{\|Ax\|}{\|x\|}$，$\|Ax\| = \frac{\|Ax\|}{\|x\|}\|x\| \le \|A\|\|x\|$)

常用范数

① 行范数　$\|A\|_\infty = \max\limits_{\|x\|=1} \|Ax\|_\infty = \max\limits_{1\le i\le n} \sum_{j=1}^n |a_{ij}|$　　行范数(每行相加，取最大)

② 列范数　$\|A\|_1 = \max\limits_{\|x\|=1} \|Ax\|_1 = \max\limits_{1\le j\le n} \sum_{i=1}^n |a_{ij}|$　　列范数(每列相加，取最大)

③ 2 范数　$\|A\|_2 = \max\limits_{\|x\|=1} \|Ax\|_2 = \sqrt{\lambda_{max}(A^TA)}$　　难以计算，但很有用

其中$\lambda_{max}(A^TA)$为矩阵A^TA的绝对值最大的特征值

例题　设$A = \begin{pmatrix} 1 & 2 & 0 \\ -1 & 2 & -1 \\ 0 & 1 & 1 \end{pmatrix}$，求它的三个范数

$\|A\|_1 = \max\{2,5,2\} = 5$　　$\|A\|_\infty = \max\{3,4,2\} = 4$

$A^TA = \begin{pmatrix} 2 & 0 & 1 \\ 0 & 9 & -1 \\ 1 & -1 & 2 \end{pmatrix} \Rightarrow |\lambda I - A^TA| = \begin{vmatrix} \lambda-2 & 0 & -1 \\ 0 & \lambda-9 & 1 \\ -1 & 1 & \lambda-2 \end{vmatrix} = 0$　　则$\lambda^3 - 13\lambda^2 + 38\lambda - 25 = 0$

可得：$\lambda_1 = 9.14, \lambda_2 = 2.92, \lambda_3 = 0.9331$　即$\|A\|_2 = \sqrt{9.14} = 3.023$

定义 5　设$\lambda_i(i = 1,2,\dots,n)$为$A \in R^{n\times n}$的$n$个特征值，称$\boldsymbol{\rho(A)} = \max\limits_{1\le i\le n}|\boldsymbol{\lambda_i}|$为$A$的谱半径

$\rho(A) \le \|A\|_p$ **谱半径小于任意的范数**，可用于**估计特征值的上界**

定理　若$\|A\| < 1$，则$I + A$可逆，且$\|(I+A)^{-1}\| \le \frac{1}{1-\|A\|}$

① 因为$\|A\| < 1$，所以$\rho(A) < 1 \Rightarrow \pm 1$不是 A 的特征值，所以$I + A$可逆

②令$D = (I+A)^{-1}$，则$1 = \|I\| = \|(I+A)D\| = \|D + AD\| \ge \|D\| - \|AD\| \ge \|D\| - \|A\|\|D\|$

$= \|D\|(1 - \|A\|)$ 因为