计算方法 B 复习课
数值代数部分

吴越

中国科学技术大学

2022 年 5 月 23 日

目录

目录

向量范数

定义

在线性空间 $V(\mathbb{R})$ 上，称映射 $\|\cdot\| : V \mapsto [0, +\infty)$ 是一个**范数**，若其满足：

1. 正定性：$\forall \boldsymbol{x} \in V,\ \|\boldsymbol{x}\| \geqslant 0$，且 $\|\boldsymbol{x}\| = 0 \iff \boldsymbol{x} = \mathbf{0}$
2. 齐次性：$\forall a \in \mathbb{R}, \forall \boldsymbol{x} \in V,\ \|a\boldsymbol{x}\| = |a|\,\|\boldsymbol{x}\|$
3. 三角不等式：$\forall \boldsymbol{x}, \boldsymbol{y} \in V,\ \|\boldsymbol{x}\| + \|\boldsymbol{y}\| \geqslant \|\boldsymbol{x} + \boldsymbol{y}\|$

例

实线性空间 $\mathbb{R}^n(n \in \mathbb{N}_+)$ 上的向量 $\boldsymbol{x} = (x_1, \cdots, x_n)^T$ 的 p-范数 ($p \in [1, +\infty]$) 为：

$$\|\boldsymbol{x}\|_p = \begin{cases} \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}}, & p \in [1, +\infty) \\ \max\limits_{i=1,\cdots,n} |x_i|, & p = \infty \end{cases} \tag{1}$$

- 当 $p \in [1, +\infty]$ 时，容易验证 $\|\cdot\|_p$ 是范数。
- 当 $p \in (0, 1)$ 时，三角不等式不成立。

例

把 $\mathbb{R}^{m\times n}$ 视作 $\mathbb{R}^{mn}$，有：

- Frobenuis 范数：$\boldsymbol{A}\in\mathbb{R}^{m\times n}$，$\|A\|_F=\operatorname{tr}\left(\boldsymbol{A}^T\boldsymbol{A}\right)$

定义

$\boldsymbol{A}\in\mathbb{R}^{n\times n}$，$\|\cdot\|$ 是 $\mathbb{R}^n$ 的范数，那么 $\boldsymbol{A}$ 由 $\|\cdot\|$ 诱导的算子范数为：

$$\|\boldsymbol{A}\|=\sup_{\boldsymbol{x}\neq\boldsymbol{0}}\frac{\|\boldsymbol{A}\boldsymbol{x}\|}{\|\boldsymbol{x}\|}=\sup_{\|\boldsymbol{x}\|=1}\|\boldsymbol{A}\boldsymbol{x}\| \tag{2}$$

注记

算子范数一定是范数，反之不成立。

矩阵范数和谱半径 II

例

- 由 $\|\cdot\|_1$ 诱导的 1-范数：$\|\boldsymbol{A}\|_1 = \max\limits_{1\leqslant j\leqslant n}\left\{\sum\limits_{i=1}^{n}|a_{ij}|\right\}$
- 由 $\|\cdot\|_2$ 诱导的 2-范数：$\|\boldsymbol{A}\|_2 = \sqrt{\rho\left(\boldsymbol{A}^T\boldsymbol{A}\right)} = \sigma_1$
- 由 $\|\cdot\|_\infty$ 诱导的 ∞-范数：$\|\boldsymbol{A}\|_\infty = \max\limits_{1\leqslant i\leqslant n}\left\{\sum\limits_{j=1}^{n}|a_{ij}|\right\}$

定义

$\boldsymbol{A} \in \mathbb{R}^{n\times n}$ 的谱半径 $\rho(\boldsymbol{A}) = \max\limits_{1\leqslant i\leqslant n}|\lambda_i|$，其中 λ_i, $i = 1, \cdots, n$ 是全部特征值。

定理

- $\rho(\boldsymbol{A}) = \inf\limits_{\|\cdot\| \text{ is an operator norm}} \|\boldsymbol{A}\|$
- *(Gelfand)* 当 $\|\cdot\|$ 是算子范数时，数列 $\left\{\left\|\boldsymbol{A}^k\right\|^{\frac{1}{k}}\right\}_{k=1}^{\infty}$ 单调递减，收敛于 $\rho(\boldsymbol{A})$

矩阵条件数

定义

$\boldsymbol{A} \in \mathbb{R}^{n\times n}$，$\|\cdot\|$ 是算子范数，则条件数定义为：

$$\mathrm{Cond}(\boldsymbol{A}) = \|\boldsymbol{A}\| \left\|\boldsymbol{A}^{-1}\right\| \tag{3}$$

命题

假设 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ 且 $\boldsymbol{A}(\boldsymbol{x} + \Delta\boldsymbol{x}) = \boldsymbol{b} + \Delta\boldsymbol{b}$，那么：

$$\frac{\|\Delta\boldsymbol{x}\|}{\|\boldsymbol{x}\|} \leqslant \mathrm{Cond}(\boldsymbol{A})\frac{\|\Delta\boldsymbol{b}\|}{\|\boldsymbol{b}\|} \tag{4}$$

目录

上三角和下三角方程组

下三角方程组

$\boldsymbol{Ax}=\boldsymbol{b}$，$\boldsymbol{A}$ 下三角. forward substituting 算法：依次求出 $x_1,x_2,\cdots,x_n$，求 x_k 时需要利用 $x_1,\cdots,x_{k-1}$ 的值。

下三角方程组

$\boldsymbol{Ax}=\boldsymbol{b}$，$\boldsymbol{A}$ 下三角. backward substituting 算法：依次求出 $x_n,x_{n-1},\cdots,x_1$，求 x_k 时需要利用 $x_{k+1},\cdots,x_n$ 的值。

注记

计算量 $\Theta\left(n^2\right)$.

定理

(Doolittle-LU 分解存在唯一性) $\boldsymbol{A} \in \mathbb{R}^{n\times n}$，以下三个命题等价：

1. 存在唯一的可逆单位下三角阵 $\boldsymbol{L}$ 和可逆上三角阵 $\boldsymbol{U}$，使得 $\boldsymbol{A} = \boldsymbol{LU}$
2. $\boldsymbol{A}$ 的顺序主子式非零
3. 可以通过普通的 *Gauss* 消元法求解 $\boldsymbol{Ax} = \boldsymbol{b}$ *(*只做行变换*)*.

命题

当 $\boldsymbol{A}$ 主对角严格占优时，$\boldsymbol{A}$ 可以进行 *LU* 分解。

证明.

任取 $\boldsymbol{A}$ 的顺序主子阵 $\widetilde{\boldsymbol{A}}$，它总是主对角严格占优的，则其可逆。于是，由上面的定理，$\boldsymbol{A}$ 可以 *LU* 分解。 □

LU 分解算法 II

两种算法

Doolittle-LU 分解算法 (L 单位下三角)，Crout-LU 分解算法 (U 单位上三角)，计算量 $\Theta\left(n^3\right)$.

对称矩阵

- 对称正定阵可以做 Cholesky 分解 $\boldsymbol{A}=\boldsymbol{L}\boldsymbol{L}^T$，其中 $\boldsymbol{L}$ 可逆下三角。
- 对称可逆阵可以做 LDLT 分解 $\boldsymbol{A}=\boldsymbol{L}\boldsymbol{D}\boldsymbol{L}^T$，其中 $\boldsymbol{L}$ 单位下三角，$\boldsymbol{D}$ 可逆对角。

利用 LU 分解求解线性方程组

对于 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$，先计算 $\boldsymbol{A}=\boldsymbol{L}\boldsymbol{U}$，问题等价为 $\boldsymbol{L}\boldsymbol{U}\boldsymbol{x}=\boldsymbol{b}$，也等价为：

$$\begin{cases} \boldsymbol{L}\boldsymbol{y}=\boldsymbol{b} \\ \boldsymbol{U}\boldsymbol{x}=\boldsymbol{y} \end{cases} \tag{5}$$

先用下三角方法求解 $\boldsymbol{L}\boldsymbol{y}=\boldsymbol{b}$，再用上三角方法求解 $\boldsymbol{U}\boldsymbol{x}=\boldsymbol{y}$，计算量 $\Theta\left(n^3\right)+\Theta\left(n^2\right)=\Theta\left(n^3\right)$.

例

利用 Doolittle-LU 分解求解 $\boldsymbol{Ax}=\boldsymbol{b}$，其中 $\boldsymbol{A}=\begin{pmatrix}1&2&3\\4&5&6\\7&8&10\end{pmatrix}$，$\boldsymbol{b}=\begin{pmatrix}12\\24\\39\end{pmatrix}$.

解

先求 $\boldsymbol{A}$ 的 LU 分解:

$$\left(\ \boldsymbol{L}^{(0)}\ \middle|\ \boldsymbol{U}^{(0)}\ \right)=\left(\begin{array}{ccc|ccc}1&0&0&1&2&3\\0&1&0&4&5&6\\0&0&1&7&8&10\end{array}\right) \tag{6}$$

$$\left(\ \boldsymbol{L}^{(1)}\ \middle|\ \boldsymbol{U}^{(1)}\ \right)=\left(\begin{array}{ccc|ccc}1&0&0&1&2&3\\4&1&0&0&-3&-6\\7&0&1&0&-6&-11\end{array}\right) \tag{7}$$

$$\left(\ \boldsymbol{L}^{(2)}\ \middle|\ \boldsymbol{U}^{(2)}\ \right)=\left(\begin{array}{ccc|ccc}1&0&0&1&2&3\\4&1&0&0&-3&-6\\7&2&1&0&0&1\end{array}\right) \tag{8}$$

然后求 $\boldsymbol{Ly}=\boldsymbol{b}$，得 $\boldsymbol{y}=(12,-24,3)^T$. 最后求 $\boldsymbol{Ux}=\boldsymbol{y}$，得 $\boldsymbol{x}=(-1,2,3)^T$.

目录

基于矩阵分裂的迭代法

线性方程组迭代求解算法很多，例如 Jacobi, Gauss-Seidel, SOR, GMRES, MINRES, (P)CG, Gradent Descent，本课程介绍了基于矩阵分裂的迭代法，其迭代格式固定，收敛性取决且仅取决于迭代矩阵的谱半径。
对于 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$，设 $\boldsymbol{A}=\boldsymbol{N}-\boldsymbol{P}$ 且 $\boldsymbol{N}$ 可逆，那么方程组等价为 $\boldsymbol{x}=\boldsymbol{N}^{-1}\boldsymbol{P}\boldsymbol{x}+\boldsymbol{N}^{-1}\boldsymbol{b}$.
令 $\boldsymbol{M}=\boldsymbol{N}^{-1}\boldsymbol{P}$，$\boldsymbol{g}=\boldsymbol{N}^{-1}\boldsymbol{b}$，那么等价为 $\boldsymbol{x}=\boldsymbol{M}\boldsymbol{x}+\boldsymbol{g}$.

定理

不动点迭代 $x^{(k+1)}=\boldsymbol{M}\boldsymbol{x}^{(k)}+\boldsymbol{g}$ 收敛当且仅当 $\rho(\boldsymbol{M})<1$，也即 $\rho\left(\boldsymbol{I}-\boldsymbol{N}^{-1}\boldsymbol{A}\right)<1$.

误差估计

假设 $\boldsymbol{x}^*$ 是真解，那么误差 $\boldsymbol{e}^{(k)}=\boldsymbol{x}^*-\boldsymbol{x}^{(k)}$ 满足 $\boldsymbol{e}^{(k+1)}=\boldsymbol{M}\boldsymbol{e}^{(k)}$，因此：

$$\left\|\boldsymbol{e}^{(k)}\right\| \leqslant \left\|\boldsymbol{M}^k\right\|\left\|\boldsymbol{e}^{(0)}\right\| \leqslant \|\boldsymbol{M}\|^k\left\|\boldsymbol{e}^{(0)}\right\| \tag{9}$$

也即，迭代是线性收敛的。

三种常见的迭代格式 I

设 $\boldsymbol{A}=\boldsymbol{L}+\boldsymbol{D}+\boldsymbol{U}$，其中 $\boldsymbol{L}$ 是严格下三角阵，$\boldsymbol{D}$ 是对角阵，$\boldsymbol{U}$ 是严格上三角阵。

Jacobi 迭代

取 $\boldsymbol{N}=\boldsymbol{D}$，则为 Jacobi 迭代：

$$\boldsymbol{D}\boldsymbol{x}^{(k+1)}=-\left(\boldsymbol{L}+\boldsymbol{U}\right)\boldsymbol{x}^{(k)}+\boldsymbol{b} \tag{10}$$

命题

当 $\boldsymbol{A}$ 严格主对角占优时，*Jacobi* 迭代收敛 *(*迭代矩阵的 ∞-范数小于 1*)*。

Gauss-Seidel 迭代

取 $\boldsymbol{N}=\boldsymbol{D}+\boldsymbol{L}$，则为 Gauss-Seidel 迭代：

$$\left(\boldsymbol{D}+\boldsymbol{L}\right)\boldsymbol{x}^{(k+1)}=-\boldsymbol{U}\boldsymbol{x}^{(k)}+\boldsymbol{b} \tag{11}$$

三种常见的迭代格式 II

注记

- 当 $\boldsymbol{A}$ 严格主对角占优时，*Gauss-Seidel* 迭代收敛。
- *Gauss-Seidel* 迭代矩阵的第一列为 $\mathbf{0}$.
- *Gauss-Seidel* 迭代实质上是在 *Jacobi* 迭代中，每次都使用最新的结果。

SOR 迭代

取 $\boldsymbol{N} = \boldsymbol{D} + \omega \boldsymbol{L}$，则为 SOR 迭代：

$$(\boldsymbol{D} + \omega \boldsymbol{L})\, \boldsymbol{x}^{(k+1)} = -\left((1-\omega)\boldsymbol{L} + \boldsymbol{U}\right) \boldsymbol{x}^{(k)} + \boldsymbol{b} \tag{12}$$

注记

SOR 迭代格式实质上是 *Gauss-Seidel* 迭代结果和旧值的凸组合。当 $\omega = 1$ 时为 *Gauss-Seidel* 迭代。

目录

设 $\boldsymbol{A} \in \mathbb{R}^{n\times n}$，对于初值 $\boldsymbol{x}^{(0)} \in \mathbb{R}^n$，构造序列 $\boldsymbol{x}^{(k+1)} = \boldsymbol{A}\boldsymbol{x}^{(k)}$.

定理

假设 $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ 的 n 个特征值 λ_i 满足 $|\lambda_1| > |\lambda_2| \geqslant |\lambda_i|, i = 3, \cdots, n$，那么任取非零线性泛函 φ，都有：

$$\lim_{k\to\infty} \frac{\varphi\left(\boldsymbol{x}^{(k+1)}\right)}{\varphi\left(\boldsymbol{x}^{(k)}\right)} = \lambda_1 \tag{13}$$

该序列线性收敛，收敛速率为：

$$\left| \frac{\varphi\left(\boldsymbol{x}^{(k+1)}\right)}{\varphi\left(\boldsymbol{x}^{(k)}\right)} - \lambda_1 \right| = \Theta\left(\left| \frac{\lambda_2}{\lambda_1} \right|^k \right) \tag{14}$$

注记

- 实际计算中，为了防止上溢和下溢，需要每一步都进行规范化 *(令某个范数归一)*。
- 当按模最大特征值是重特征值或者是虚数时，需要用 *Jordan* 标准型单独讨论。