

统计计算

2016 年 3 月 3 日

① 引言

② 矩阵的三角 - 三角分解

统计计算经常遇到各种矩阵计算问题.

① 线性方程组的求解

- ② 处理矩阵问题的一个最基本方法: 将一个结构复杂矩阵分解为形式比较简单或者我们对其性质较熟悉的一些矩阵的乘积.

例如:

- 三角矩阵和对角矩阵是非常简单的矩阵, 其计算 (如求逆或求行列式等) 是非常简单的;
- 正交矩阵或酉阵具有很好的性质.

统计计算经常遇到各种矩阵计算问题.

- ① 线性方程组的求解
- ② 处理矩阵问题的一个最基本方法: 将一个结构复杂矩阵分解为形式比较简单或者我们对其性质较熟悉的一些矩阵的乘积.






例如:

- 三角矩阵和对角矩阵是非常简单的矩阵, 其计算 (如求逆或求行列式等) 是非常简单的;
- 正交矩阵或酉阵具有很好的性质.

- 3 几种常用的矩阵分解
 - 三角分解
 - 正交 - 三角分解
 - 特征值分解或谱分解
 - 奇异值分解
 - ...
- 4 广义逆及其计算
- 5 消去 (扫描) 变换

- ③ 几种常用的矩阵分解
 - 三角分解
 - 正交 - 三角分解
 - 特征值分解或谱分解
 - 奇异值分解
 - ...
- ④ 广义逆及其计算
- ⑤ 消去 (扫描) 变换

- ③ 几种常用的矩阵分解
 - 三角分解
 - 正交 - 三角分解
 - 特征值分解或谱分解
 - 奇异值分解
 - ...
- ④ 广义逆及其计算
- ⑤ 消去 (扫描) 变换

-  A. Ben-Israel and T. N. E. Greville, Generalized Inverses: Theory and Applications (Second Edition), John Wiley & Sons, New York, 2002.
-  J. E. Gentle, Matrix Algebra: Theory, Computations, and Applications in Statistics, Springer, 2007.
-  G. H. Golub and C. F. Van Loan, Matrix Computation (Third Edition), Johns Hopkins University Press, Baltimore, 1996.
-  R. A. Horn and C. R. Johnson, Topics in Matrix Analysis, Cambridge University Press, 1991.
-  程元鹏, 矩阵论, 西北工业大学出版社, 2000.

1 引言

2 矩阵的三角 - 三角分解

若一个矩阵分解为两个三角矩阵乘积, 则称之为三角-三角分解. 具体地可以考虑以下四种形式:

① LDR 分解

$$A = LDR,$$

其中, L 是单位下三角矩阵, D 是对角矩阵, R 是单位上三角矩阵.

② LR 分解 (Doolittle 分解)

$$A = LR,$$

其中, L 是单位下三角矩阵, R 是上三角矩阵.

③ Crout 分解

$$\mathbf{A} = \mathbf{LR},$$

其中, \mathbf{L} 是下三角矩阵, \mathbf{R} 是单位上三角矩阵.

④ Cholesky 分解 (平方根分解)

$$\mathbf{A} = \mathbf{TT}^H,$$

其中, \mathbf{A} 是 Hermite 正定阵, \mathbf{T} 是下三角矩阵.

问题

- ① 矩阵的 LR 分解是否存在?
- ② 如何给出分解算法?

设 $A \in \mathbb{C}^{n \times n}$. 若 A 的前 $n-1$ 个顺序主子式 $\Delta_i \neq 0$ ($i = 1, \dots, n-1$), 则 A 存在 LR 分解 (LR 分解存在的充分条件).

下面我们给出用 Gauss 消去法求矩阵 LR 分解的详细过程.

设 $\mathbf{A}^{(0)} = \mathbf{A} = (a_{ij}^{(0)})$. 因为 $a_{11}^{(0)} = \Delta_1 \neq 0$, 所以可令 $l_{i1} = a_{i1}^{(0)} / a_{11}^{(0)}$ ($i = 2, \dots, n$), 并构造

$$\mathbf{L}_1 = \begin{bmatrix} 1 & & & \\ l_{21} & 1 & & \\ \vdots & & \ddots & \\ l_{n1} & & & 1 \end{bmatrix}.$$

记 $\mathbf{A}^{(1)} = \mathbf{L}_1^{-1} \mathbf{A}^{(0)}$, 则有

$$\mathbf{A}^{(1)} = \begin{bmatrix} a_{11}^{(0)} & a_{12}^{(0)} & \cdots & a_{1n}^{(0)} \\ & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} \\ & \vdots & & \vdots \\ & a_{n2}^{(1)} & \cdots & a_{nn}^{(1)} \end{bmatrix}.$$

显然 \mathbf{A} 的二阶顺序主子式 $\Delta_2 = a_{11}^{(0)} a_{22}^{(1)}$.

因为 $\Delta_2 \neq 0$, 所以 $a_{22}^{(1)} \neq 0$. 令 $l_{i2} = a_{i2}^{(1)} / a_{22}^{(1)}$ ($i = 3, \dots, n$), 并构造

$$\mathbf{L}_2 = \begin{bmatrix} 1 & & & & \\ & 1 & & & \\ & l_{32} & 1 & & \\ & \vdots & & \ddots & \\ & l_{n2} & & & 1 \end{bmatrix}.$$

记 $\mathbf{A}^{(2)} = \mathbf{L}_2^{-1} \mathbf{A}^{(1)}$, 则有

$$\mathbf{A}^{(2)} = \begin{bmatrix} a_{11}^{(0)} & a_{12}^{(0)} & a_{13}^{(0)} & \cdots & a_{1n}^{(0)} \\ & a_{22}^{(1)} & a_{23}^{(1)} & \cdots & a_{2n}^{(1)} \\ & & a_{33}^{(2)} & \cdots & a_{3n}^{(2)} \\ & & \vdots & & \vdots \\ & & a_{n3}^{(2)} & \cdots & a_{nn}^{(2)} \end{bmatrix}.$$

此时, \mathbf{A} 的三阶顺序主子式 $\Delta_3 = a_{11}^{(0)} a_{22}^{(1)} a_{33}^{(2)}$.

因为 $\Delta_i \neq 0 (i = 3, \dots, n-1)$, 所以上述步骤可继续下去直到第 $n-1$ 步为止, 从而可得单位下三角矩阵 $\mathbf{L}_1, \dots, \mathbf{L}_{n-1}$.

在此过程中有

$$\Delta_{i+1} = a_{11}^{(0)} a_{22}^{(1)} \cdots a_{ii}^{(i-1)} a_{i+1,i+1}^{(i)}, \quad i = 1, \dots, n-1.$$

令

$$\mathbf{L} = \mathbf{L}_1 \mathbf{L}_2 \cdots \mathbf{L}_{n-1},$$

及

$$\mathbf{R} = \mathbf{L}^{-1} \mathbf{A} = \begin{bmatrix} a_{11}^{(0)} & a_{12}^{(0)} & \cdots & a_{1,n-1}^{(0)} & a_{1n}^{(0)} \\ & a_{22}^{(1)} & \cdots & a_{1,n-1}^{(1)} & a_{2n}^{(1)} \\ & & \ddots & \vdots & \vdots \\ & & & a_{n-1,n-1}^{(n-2)} & a_{n-1,n}^{(n-2)} \\ & & & & a_{nn}^{(n-1)} \end{bmatrix}.$$

则 \mathbf{L} 为单位下三角矩阵, \mathbf{R} 为上三角矩阵, 从而 $\mathbf{A} = \mathbf{LR}$ 成为矩阵 \mathbf{A} 的 LR 分解.

矩阵的 LR 分解的存在唯一性

定理

矩阵 $A \in \mathbb{C}^{n \times n}$ 有唯一 LR 分解的充要条件是 A 的前 $n-1$ 个顺序主子式 $\Delta_i \neq 0 (i = 1, \dots, n-1)$.

推论

设 $A \in \mathbb{C}^{n \times n}$ 非奇异, 则 A 可以分解为 n 阶下三角矩阵与 n 阶上三角矩阵之积的充要条件是 A 的各阶顺序主子式 $\Delta_i \neq 0 (i = 1, \dots, n)$.

上述推论中的三角分解形式不唯一. 事实上, 若 A 有 LR 分解 $A = LR$, 取 D 为任意 n 阶非奇异对角矩阵, 则 $\tilde{L} = LD$ 和 $\tilde{R} = D^{-1}R$ 分别为下三角和上三角矩阵, 且 $A = \tilde{L}\tilde{R}$.

推论

设 $A \in \mathbb{C}^{n \times n}$ 非奇异, 则 A 可以分解为 n 阶下三角矩阵与 n 阶上三角矩阵之积的充要条件是 A 的各阶顺序主子式 $\Delta_i \neq 0 (i = 1, \dots, n)$.

上述推论中的三角分解形式不唯一. 事实上, 若 A 有 LR 分解 $A = LR$, 取 D 为任意 n 阶非奇异对角矩阵, 则 $\tilde{L} = LD$ 和 $\tilde{R} = D^{-1}R$ 分别为下三角和上三角矩阵, 且 $A = \tilde{L}\tilde{R}$.

例
设

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 2 & 3 \end{bmatrix}$$

则不存在单位下三角矩阵 \mathbf{L} 和上三角矩阵 \mathbf{R} 使得 $\mathbf{A} = \mathbf{LR}$.

非奇异矩阵的 LR 分解 (续)

引理

设 A 是 n 阶非奇异矩阵, 则存在置换矩阵 P 使得 PA 的 n 个顺序主子式非零.

定理

设 A 是 n 阶非奇异矩阵, 则存在置换矩阵 P 使得 PA 有唯一的 LR 分解.

矩阵的 LR 分解算法

设 $\mathbf{A} = (a_{ij}) \in \mathbb{C}^{n \times n}$ 有 LR 分解:

$$\begin{aligned}\mathbf{A} &= \begin{bmatrix} 1 & & & \\ l_{21} & 1 & & \\ \vdots & \vdots & \ddots & \\ l_{n1} & l_{n2} & \cdots & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ & r_{22} & \cdots & r_{2n} \\ & & \ddots & \vdots \\ & & & r_{nn} \end{bmatrix} \\ &= \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ l_{21}r_{11} & l_{21}r_{12} + r_{22} & \cdots & l_{21}r_{1n} + r_{2n} \\ \vdots & \vdots & & \vdots \\ l_{n1}r_{11} & l_{n1}r_{12} + l_{n2}r_{22} & \cdots & l_{n1}r_{1n} + l_{n2}r_{2n} + \cdots + r_{nn} \end{bmatrix}. \end{aligned} \quad (1)$$

矩阵的 LR 分解算法 (续)

由式 (1) 两边第一行和第一列元素分别相等, 可分别求得 \mathbf{R} 的第一行和 \mathbf{L} 的第一列元素:

$$\begin{aligned} r_{1j} &= a_{1j}, & j &= 1, \cdots, n, \\ l_{j1} &= \frac{a_{j1}}{r_{11}}, & j &= 2, \cdots, n. \end{aligned}$$

再由式 (1) 两边第二行和第二列元素分别相等, 可分别求得 \mathbf{R} 的第二行和 \mathbf{L} 的第二列元素:

$$\begin{aligned} r_{2j} &= a_{2j} - l_{21}r_{1j}, & j &= 2, \cdots, n, \\ l_{j2} &= \frac{a_{j2} - l_{j1}r_{12}}{r_{22}}, & j &= 3, \cdots, n. \end{aligned}$$

依此类推, 即可求出矩阵 \mathbf{R} 和 \mathbf{L} .

上述递推过程的计算顺序为:

第 1 行 + 第 1 列 \rightarrow 第 2 行 + 第 2 列 \rightarrow

...

LR 分解算法

对 $i = 1, \dots, n$, 计算

$$\begin{cases} r_{ij} = a_{ij} - \sum_{k=1}^{i-1} l_{ik} r_{kj}, & j = i, \dots, n, \\ l_{ji} = \frac{1}{r_{ii}} \left(a_{ji} - \sum_{k=1}^{i-1} l_{jk} r_{ki} \right), & j = i+1, \dots, n. \end{cases}$$

矩阵的 LR 分解算法 (续)

若对矩阵 A 进行 LR 分解之后不再需要存储 A , 则将因子 L 和 R 按

$$\begin{bmatrix} r_{11} & r_{12} & r_{13} & \cdots & r_{1n} \\ l_{21} & r_{22} & r_{23} & \cdots & r_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ l_{n1} & l_{n2} & l_{n3} & \cdots & r_{nn} \end{bmatrix}$$

的方式直接存放于矩阵 A 的存储单元, 就可得到以下节省存储空间的算法.

节省存储空间的 LR 分解算法

对 $i = 1, \dots, n$ 计算

$$\begin{cases} a_{ij} = a_{ij} - \sum_{k=1}^{i-1} a_{ik} a_{kj}, & j = i, \dots, n, \\ a_{ji} = \frac{1}{a_{ii}} \left(a_{ji} - \sum_{k=1}^{i-1} a_{jk} a_{ki} \right), & j = i+1, \dots, n. \end{cases}$$

问 题?