计算机问题求解 – 论题3–14 - 矩阵计算

2014年12月15日

自学问题:
什么是forward substitution?

矩阵的逆与线性方程组的解

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n \end{aligned}$$

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

or, equivalently, letting $A = (a_{ij})$, $x = (x_i)$, and $b = (b_i)$, as

$$\boxed{Ax = b\,.}$$

If A is nonsingular, it possesses an inverse A^{-1}, and

$$\boxed{x = A^{-1}b}$$

逆矩阵存在的条件

A square matrix has full rank if and only if it is nonsingular.

A matrix A has full column rank if and only if it does not have a null vector.

A square matrix A is singular if and only if it has a null vector.

An $n \times n$ matrix A is singular if and only if $\det(A) = 0$.

这是什么意思？

问题2:
如何计算非奇异矩阵的逆?

1：矩阵A的逆=A的伴随矩阵/行列式A的值

2：矩阵A的逆：对(A|E)进行行初等变换得到$(E|A^{-1})$

例：求3阶方阵$A=\begin{pmatrix}2&2&1\\3&1&5\\3&2&3\end{pmatrix}$的逆矩阵.

解：$|A|=1$，$M_{11}=-7,\ M_{12}=-6,\ M_{13}=3,$

$M_{21}=4,\ M_{22}=3,\ M_{23}=-2,$

$M_{31}=9,\ M_{32}=7,\ M_{33}=-4,$

则

$$A^{-1}=\frac{1}{|A|}A^{*}=A^{*}=\begin{pmatrix}A_{11}&A_{21}&A_{31}\\A_{12}&A_{22}&A_{32}\\A_{13}&A_{23}&A_{33}\end{pmatrix}$$

$$=\begin{pmatrix}M_{11}&-M_{21}&M_{31}\\-M_{12}&M_{22}&-M_{32}\\M_{13}&-M_{23}&M_{33}\end{pmatrix}=\begin{pmatrix}-7&-4&9\\6&3&-7\\3&2&-4\end{pmatrix}$$

例 1 设 $A=\begin{pmatrix}1&2&3\\2&2&1\\3&4&3\end{pmatrix}$, 求 A^{-1}.

解 $$(A\,|\,E)=\left(\begin{array}{ccc|ccc}1&2&3&1&0&0\\2&2&1&0&1&0\\3&4&3&0&0&1\end{array}\right)$$

$$\underset{r_3-3r_1}{\overset{r_2-2r_1}{\sim}}\left(\begin{array}{ccc|ccc}1&2&3&1&0&0\\0&-2&-5&-2&1&0\\0&-2&-6&-3&0&1\end{array}\right)\underset{r_3-r_2}{\overset{r_1+r_2}{\sim}}$$

问题3：

为什么通常不直接用求逆矩阵的办法来解线性方程组？

高斯消元法过程中可能出现的现象！

$$\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 4 & 1 \end{bmatrix} X = \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 1 & -2 \\ 0 & 1 & 4 \\ 0 & 0 & 3 \end{bmatrix} X = \begin{bmatrix} 1 \\ 3 \\ 6 \end{bmatrix}$$

问题4：

三角阵会给解线性方程组带来什么便利？

问题5：

你能否借助右边的图解释一下用LUP分解方法解线性方程组的基本思想？

$$PA = LU \Rightarrow$$

$$Ax = b$$

$$\Downarrow$$

$$PAx = Pb$$

$$\Downarrow$$

$$LUx = Pb$$

$$\Downarrow$$

$$Ly = Pb$$

$$\Downarrow$$

$$Ux = y$$

$$Ly = Pb$$

$$\begin{array}{lllllllllll}
y_1 & & & & & & & & & = & b_{\pi[1]}\,, \\
l_{21}y_1 & + & y_2 & & & & & & & = & b_{\pi[2]}\,, \\
l_{31}y_1 & + & l_{32}y_2 & + & y_3 & & & & & = & b_{\pi[3]}\,, \\
 & & & & & & & & & \vdots & \\
l_{n1}y_1 & + & l_{n2}y_2 & + & l_{n3}y_3 & + & \cdots & + & y_n & = & b_{\pi[n]}\,.
\end{array}$$

So:

$$y_i = b_{\pi[i]} - \sum_{j=1}^{i-1} l_{ij}\, y_j\,.$$

$Ux = y$

$$
\begin{aligned}
u_{11}x_1 + u_{12}x_2 + \cdots + \quad u_{1,n-2}x_{n-2} + \quad u_{1,n-1}x_{n-1} + \quad u_{1n}x_n &= y_1 \, , \\
u_{22}x_2 + \cdots + \quad u_{2,n-2}x_{n-2} + \quad u_{2,n-1}x_{n-1} + \quad u_{2n}x_n &= y_2 \, , \\
&\vdots \\
u_{n-2,n-2}x_{n-2} + u_{n-2,n-1}x_{n-1} + u_{n-2,n}x_n &= y_{n-2} \, , \\
u_{n-1,n-1}x_{n-1} + u_{n-1,n}x_n &= y_{n-1} \, , \\
u_{n,n}x_n &= y_n \, .
\end{aligned}
$$

$$
x_i = \left(y_i - \sum_{j=i+1}^{n} u_{ij} x_j \right) / u_{ii} \, .
$$

If we have LUP, we can solve the equations in $\Theta(n^2)$

LUP-SOLVE(L, U, π, b)

$n = L.rows$
let x be a new vector of length n
for $i = 1$ **to** n
    $y_i = b_{\pi[i]} - \sum_{j=1}^{i-1} l_{ij} y_j$
for $i = n$ **downto** 1
    $x_i = \left(y_i - \sum_{j=i+1}^{n} u_{ij} x_j\right) / u_{ii}$
return x

But, how can we get LUP?

问题6：从以下的例子中，我们能观察到什么结论?

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ c/a & 1 \end{bmatrix} \times \begin{bmatrix} a & b \\ 0 & d - cb/a \end{bmatrix}$$

$$\begin{bmatrix} a & B \\ C & D \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ C/a & I \end{bmatrix} \times \begin{bmatrix} a & B \\ 0 & D - CB/a \end{bmatrix}$$

假如可以不考虑P

$$A = \left(\begin{array}{c|cccc} a_{11} & a_{12} & \cdots & a_{1n} \\ \hline a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{array}\right)$$

$$= \begin{pmatrix} a_{11} & w^{\mathrm{T}} \\ v & A' \end{pmatrix}$$

问题7：

为什么说这是采用了“高斯消去法”？

$$A = \begin{pmatrix} a_{11} & w^{\mathrm{T}} \\ v & A' \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 \\ v/a_{11} & I_{n-1} \end{pmatrix} \begin{pmatrix} a_{11} & w^{\mathrm{T}} \\ 0 & A' - vw^{\mathrm{T}}/a_{11} \end{pmatrix}$$

问题8：

为什么可以对 $A' - vw^{\mathrm{T}}/a_{11}$ 递归？

We claim that if A is nonsingular, then the Schur complement is nonsingular, too. Why? Suppose that the Schur complement, which is $(n-1) \times (n-1)$, is singular. Then by Theorem D.1, it has row rank strictly less than $n-1$. Because the bottom $n-1$ entries in the first column of the matrix

$$\begin{pmatrix} a_{11} & w^{\mathrm{T}} \\ 0 & A' - vw^{\mathrm{T}}/a_{11} \end{pmatrix}$$

are all 0, the bottom $n-1$ rows of this matrix must have row rank strictly less than $n-1$. The row rank of the entire matrix, therefore, is strictly less than n. Applying Exercise D.2-8 to equation (28.8), A has rank strictly less than n, and from Theorem D.1 we derive the contradiction that A is singular.

$$
\begin{aligned}
A &= \begin{pmatrix} 1 & 0 \\ v/a_{11} & I_{n-1} \end{pmatrix} \begin{pmatrix} a_{11} & w^{\mathrm{T}} \\ 0 & A' - vw^{\mathrm{T}}/a_{11} \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 \\ v/a_{11} & I_{n-1} \end{pmatrix} \begin{pmatrix} a_{11} & w^{\mathrm{T}} \\ 0 & L'U' \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 \\ v/a_{11} & L' \end{pmatrix} \begin{pmatrix} a_{11} & w^{\mathrm{T}} \\ 0 & U' \end{pmatrix} \\
&= LU ,
\end{aligned}
$$

$$\underset{A}{\begin{pmatrix} 2 & 3 & 1 & 5 \\ 6 & 13 & 5 & 19 \\ 2 & 19 & 10 & 23 \\ 4 & 10 & 11 & 31 \end{pmatrix}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 3 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 2 & 0 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} 2 & 3 & 1 & 5 \\ 0 & 4 & 2 & 4 \\ 0 & 16 & 9 & 18 \\ 0 & 4 & 9 & 21 \end{pmatrix}$$

$$\begin{pmatrix} 4 & 2 & 4 \\ 16 & 9 & 18 \\ 4 & 9 & 21 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ 1 & 7 & 1 \end{pmatrix} \times \begin{pmatrix} 4 & 2 & 4 \\ 0 & 1 & 2 \\ 0 & 0 & 3 \end{pmatrix}$$

$$\underset{A}{\begin{pmatrix} 2 & 3 & 1 & 5 \\ 6 & 13 & 5 & 19 \\ 2 & 19 & 10 & 23 \\ 4 & 10 & 11 & 31 \end{pmatrix}} = \underset{L}{\begin{pmatrix} 1 & 0 & 0 & 0 \\ 3 & 1 & 0 & 0 \\ 1 & 4 & 1 & 0 \\ 2 & 1 & 7 & 1 \end{pmatrix}} \underset{U}{\begin{pmatrix} 2 & 3 & 1 & 5 \\ 0 & 4 & 2 & 4 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 3 \end{pmatrix}}$$

LU-DECOMPOSITION(A)

```
n = A.rows
let L and U be new n × n matrices
initialize U with 0s below the diagonal
initialize L with 1s on the diagonal and 0s above the diagonal
for k = 1 to n
    u_kk = a_kk
    for i = k + 1 to n
        l_ik = a_ik / u_kk          // l_ik holds v_i
        u_ki = a_ki                 // u_ki holds w_i^T
    for i = k + 1 to n
        for j = k + 1 to n
            a_ij = a_ij − l_ik u_kj
return L and U
```

问题9:

为什么算法中并没有用递归?

$$\begin{array}{cccc} 2 & 3 & 1 & 5 \\ 6 & 13 & 5 & 19 \\ 2 & 19 & 10 & 23 \\ 4 & 10 & 11 & 31 \end{array}$$

(a)

$$\begin{array}{c|ccc} 2 & 3 & 1 & 5 \\ \hline 3 & 4 & 2 & 4 \\ 1 & 16 & 9 & 18 \\ 2 & 4 & 9 & 21 \end{array}$$

(b)

$$\begin{array}{cccc} 2 & 3 & 1 & 5 \\ 3 & 4 & 2 & 4 \\ 1 & 4 & 1 & 2 \\ 2 & 1 & 7 & 17 \end{array}$$

(c)

$$\begin{array}{cccc} 2 & 3 & 1 & 5 \\ 3 & 4 & 2 & 4 \\ 1 & 4 & 1 & 2 \\ 2 & 1 & 7 & 3 \end{array}$$

(d)

$$\underset{A}{\begin{pmatrix} 2 & 3 & 1 & 5 \\ 6 & 13 & 5 & 19 \\ 2 & 19 & 10 & 23 \\ 4 & 10 & 11 & 31 \end{pmatrix}} = \underset{L}{\begin{pmatrix} 1 & 0 & 0 & 0 \\ 3 & 1 & 0 & 0 \\ 1 & 4 & 1 & 0 \\ 2 & 1 & 7 & 1 \end{pmatrix}} \underset{U}{\begin{pmatrix} 2 & 3 & 1 & 5 \\ 0 & 4 & 2 & 4 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 3 \end{pmatrix}}$$

(e)

问题9:

为什么范例中没有看到L,U矩阵?

问题10:

为什么需要置换矩阵？为什么一定能够找到可置换的行？

$$QA = \begin{pmatrix} a_{k1} & w^{\mathrm{T}} \\ v & A' \end{pmatrix}$$

初等行变换

$$\begin{aligned} QA &= \begin{pmatrix} a_{k1} & w^{\mathrm{T}} \\ v & A' \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ v/a_{k1} & I_{n-1} \end{pmatrix} \begin{pmatrix} a_{k1} & w^{\mathrm{T}} \\ 0 & A' - vw^{\mathrm{T}}/a_{k1} \end{pmatrix} \end{aligned}$$

递归

$$P'(A' - vw^{\mathrm{T}}/a_{k1}) = L'U'$$

Define

$$P = \begin{pmatrix} 1 & 0 \\ 0 & P' \end{pmatrix} Q$$

$$\begin{aligned} PA &= \begin{pmatrix} 1 & 0 \\ 0 & P' \end{pmatrix} QA \\ &= \begin{pmatrix} 1 & 0 \\ 0 & P' \end{pmatrix} \begin{pmatrix} 1 & 0 \\ v/a_{k1} & I_{n-1} \end{pmatrix} \begin{pmatrix} a_{k1} & w^{\mathrm{T}} \\ 0 & A' - vw^{\mathrm{T}}/a_{k1} \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ P'v/a_{k1} & P' \end{pmatrix} \begin{pmatrix} a_{k1} & w^{\mathrm{T}} \\ 0 & A' - vw^{\mathrm{T}}/a_{k1} \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ P'v/a_{k1} & I_{n-1} \end{pmatrix} \begin{pmatrix} a_{k1} & w^{\mathrm{T}} \\ 0 & P'(A' - vw^{\mathrm{T}}/a_{k1}) \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ P'v/a_{k1} & I_{n-1} \end{pmatrix} \begin{pmatrix} a_{k1} & w^{\mathrm{T}} \\ 0 & L'U' \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ P'v/a_{k1} & L' \end{pmatrix} \begin{pmatrix} a_{k1} & w^{\mathrm{T}} \\ 0 & U' \end{pmatrix} \\ &= LU , \end{aligned}$$

行置换的处理

for $i = 1$ to n
　　$\pi[i] = i$
for $k = 1$ to n
　　$p = 0$
　　for $i = k$ to n
　　　　if $|a_{ik}| > p$
　　　　　　$p = |a_{ik}|$
　　　　　　$k' = i$
　　if $p == 0$
　　　　error "singular matrix"
　　exchange $\pi[k]$ with $\pi[k']$

问题11:

如何理解数组pi?

问题12：

算法中并没有出现两个三角矩阵和置换矩阵*P*，这些矩阵的值是如何体现的？

```
for k = 1 to n
    ……
    exchange π[k] with π[k′]
    for i = 1 to n
        exchange a_ki with a_k′i
    for i = k + 1 to n
        a_ik = a_ik / a_kk
        for j = k + 1 to n
            a_ij = a_ij − a_ik a_kj
```

1	2	0	2	0.6
2	3	3	4	–2
3	5	5	4	2
4	–1	–2	3.4	–1

(a)

3	5	5	4	2
2	3	3	4	–2
1	2	0	2	0.6
4	–1	–2	3.4	–1

(b)

3	5	5	4	2
2	0.6	0	1.6	–3.2
1	0.4	–2	0.4	–.2
4	–0.2	–1	4.2	–0.6

(c)

3	5	5	4	2
2	0.6	0	1.6	–3.2
1	0.4	–2	0.4	–0.2
4	–0.2	–1	4.2	–0.6

(d)

3	5	5	4	2
1	0.4	–2	0.4	–0.2
2	0.6	0	1.6	–3.2
4	–0.2	–1	4.2	–0.6

(e)

3	5	5	4	2
1	0.4	–2	0.4	–0.2
2	0.6	0	1.6	–3.2
4	–0.2	0.5	4	–0.5

(f)

3	5	5	4	2
1	0.4	–2	0.4	–0.2
2	0.6	0	1.6	–3.2
4	–0.2	0.5	4	–0.5

(g)

3	5	5	4	2
1	0.4	–2	0.4	–0.2
4	–0.2	0.5	4	–0.5
2	0.6	0	1.6	–3.2

(h)

3	5	5	4	2
1	0.4	–2	0.4	–0.2
4	–0.2	0.5	4	–0.5
2	0.6	0	0.4	–3

(i)

$$
\underbrace{\begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix}}_{P}
\underbrace{\begin{pmatrix} 2 & 0 & 2 & 0.6 \\ 3 & 3 & 4 & -2 \\ 5 & 5 & 4 & 2 \\ -1 & -2 & 3.4 & -1 \end{pmatrix}}_{A}
=
\underbrace{\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0.4 & 1 & 0 & 0 \\ -0.2 & 0.5 & 1 & 0 \\ 0.6 & 0 & 0.4 & 1 \end{pmatrix}}_{L}
\underbrace{\begin{pmatrix} 5 & 5 & 4 & 2 \\ 0 & -2 & 0.4 & -0.2 \\ 0 & 0 & 4 & -0.5 \\ 0 & 0 & 0 & -3 \end{pmatrix}}_{U}
$$

问题13：

你能否解释一下，为什么可以利用LUP分解来计算逆矩阵？

In general, once we have the LUP decomposition of A, we can solve, in time $\Theta(kn^2)$, k versions of the equation $Ax = b$ that differ only in b.

We can think of the equation

$$AX = I_n , \tag{28.10}$$

which defines the matrix X, the inverse of A, as a set of n distinct equations of the form $Ax = b$. To be precise, let X_i denote the ith column of X, and recall that the unit vector e_i is the ith column of I_n. We can then solve equation (28.10) for X by using the LUP decomposition for A to solve each equation

$$AX_i = e_i$$

separately for X_i.

问题14：

这有多复杂？

课外作业

- TC Ex.28.1: 2, 3, 6, 7
- TC Ex.28.2: 1, 2, 3
- TC Ex.28.3: 1, 3
- TC Prob.28.1