Chương 2: Hệ Phương Trình Tuyến Tính

Nguyễn Đức Nghĩa, Vũ Văn Thiệu, Trịnh Anh Phúc [1]

[1]Bộ môn Khoa Học Máy Tính, Viện CNTT & TT,
Trường Đại Học Bách Khoa Hà Nội.

Ngày 4 tháng 12 năm 2012

Giới thiệu

1. Thế nào là hệ phương trình tuyến tính ?
2. Ví dụ 3 chiều
3. Ma trận hoán vị và ma trận tam giác
4. Phân tích LU
5. Vai trò của phần tử trụ
6. Hiệu ứng của sai số làm tròn
7. Hệ xác định tồi và số điều kiện của ma trận
8. Giải hệ phương trình tuyến tính bằng phân tích ma trận

Hệ phương trình tuyến tính

Định nghĩa

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1$$

$$a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2$$

$$\dots$$

$$a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m$$

Ký hiệu
$A = (a_{ij})$ với $i = 1, \cdots, m$ và $j = 1, \cdots, n$ là ma trận hệ số A.
$b = (b_1, b_2, \cdots, b_m)^T$ là vectơ vế phải.
$x = (x_1, x_2, \cdots, x_n)^T$ là vectơ biến.
ta có thể viết lại hệ phương trình tuyến tính dưới dạng ma trận

$$Ax = b$$

Hệ phương trình tuyến tính

Ví dụ 1 :

Xét hệ phương trình tuyến tính có

- Ma trận hệ số $A = \begin{pmatrix} 3 & 2 \\ 1 & -1 \end{pmatrix}$
- Vec tơ vế phải là $b = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$

thì hệ có nghiệm duy nhất $x = \begin{pmatrix} 0.2 \\ -0.8 \end{pmatrix}$

Hệ phương trình tuyến tính

Ví dụ 2 :

Xét hệ phương trình tuyến tính có

- Ma trận hệ số $A = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & -5 \end{pmatrix}$
- Vec tơ vế phải là $b = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$

thì hệ có vô số nghiệm $x = \begin{pmatrix} 1-3t \\ 2+5t \\ t \end{pmatrix}$ với mọi $t \in \mathbb{R}$.

Hệ phương trình tuyến tính

Ví dụ 3 :

Xét hệ phương trình tuyến tính có

- Ma trận hệ số $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 3 & 4 \end{pmatrix}$
- Vec tơ vế phải là $b = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$

thì hệ vô nghiệm.

Hệ phương trình tuyến tính

Đối với hệ phương trình tuyến tính có thể xảy ra

- $m = n$: hệ vuông (số phương trình bằng số ẩn, thường có nghiệm duy nhất)
- $m < n$: hệ thiếu (số phương trình ít hơn số ẩn số, hệ thường vô số nghiệm)
- $m > n$: hệ dư (số phương trình nhiều hơn số ẩn số, hệ thường vô nghiệm)

Giải hệ phương trình tuyến tính

Hệ phương trình vuông

$$Ax = b$$

trong đó $A \in \mathbb{R}^{n\times n}$ còn x và b là các vec tơ $\in \mathbb{R}^n$

Giải hệ phương trình vuông

- Nếu ma trận A không suy biến (singular) thì nghiệm duy nhất của phương trình là
$$x = A^{-1}b$$

Matlab

» x=inv(A)*b

Giải hệ phương trình tuyến tính

Ví dụ 4 :

Giải hệ phương trình $A = (7)$ và $b = (21)$ hay phương trình

$$7x = 21$$

- Cách 1 : Giải trực tiếp phép chia $x = 21/7 = 3$
- Cách 2 : Nghịch đảo 7^{-1} rồi nhân với 21 sẽ dẫn đến

$$x = 7^{-1} \times 21 = 0.142857 \times 21 = 2.99997$$

Rõ ràng cách 1 tốt hơn cách 2, thêm nữa cách 2 còn có khối lượng tính toán lớn hơn khi xác định nghịch đảo 7^{-1}.

Giải hệ phương trình tuyến tính

Nhận xét

Ngay cả trong lời giải tổng quát, khi ta xét hệ gồm nhiều phương trình thì việc giải nó thường là tạo ra lời giải trực tiếp mà không qua tính giá trị nghịch đảo A^{-1}.

Chẳng hạn các cách giải :

- Phân tích LU (LU Factorization)
- Phân tích Cholesky (Cholesky Factorization)
- Phân rã QR (QR Decomposition)

Giải hệ phương trình tuyến tính

Toán tử chia ma trận trong Matlab

Nếu A là ma trận bất kỳ và B là ma trận có số hàng giống A thì lời giải của hệ phương trình

$$AX = B$$

thì ta dùng phép *chia trái* $X = A \backslash B$.
Còn lời giải của hệ phương trình

$$XA = B$$

thì ta dùng phép *chia phải* $X = B/A$.

Giải hệ phương trình tuyến tính

Ví dụ 5 : Chia trái

» **A=[3 2;1 -1];b=[-1;1];**
» **x = A\b;**
x = 0.2000
-0.8000

Ví dụ 6 : Chia phải

» **AA=A';bb=b';**
» **xx = bb/AA;**
xx = 0.2000 -0.8000

Giải hệ phương trình tuyến tính

Các định lượng cơ bản khi giải hệ phương trình vuông

- **Định thức (Determinant)** để xác định hình thức nghiệm (vô nghiệm hoặc vô số nghiệm, hay nghiệm duy nhất).
- **Vết (Trace)** là tổng các phần tử đường chéo chính.
- **Hạng (Rank)** là số dòng hay cột độc lập tuyến tính lớn nhất của ma trận.

Matlab

» D=det(A)
» T=trace(A)
» R=rank(A)

Giải hệ phương trình tuyến tính

Định lý Kronecker-Capelli

Hệ phương trình tuyến tính $Ax = b$ có nghiệm khi và chỉ khi

$$rank(A) = rank(Ab)$$

Ví dụ 7 : cùng hạng

```
» A=[1 2 3; 4 5 6; 8 10 12];
» b=[5;6;12];
» rA=rank(A);
» rAb=rank([Ab])
rA = 2
rAb = 2
```

Giải hệ phương trình tuyến tính vuông

Các tình huống xảy ra khi có giá trị định thức khi giải $Ax = b$

- Hệ phương trình có nghiệm duy nhất nếu $det(A) \neq 0$.
- Khi $det(A) = 0$ hệ phương trình có thể có vô số nghiệm hoặc vô nghiệm (Ta có thể áp dụng định lý Kronecker-Capelli để xác định rõ nó vô nghiệm hay vô số nghiệm).
- Khi $det(A) \neq 0$ thì tồn tại ma trận nghịch đảo của A và A được gọi là ma trận không suy biến.
- Khi $det(A) = 0$ thì nghịch đảo A^{-1} không tồn tại và A được gọi là ma trận suy biến.

Giải hệ phương trình tuyến tính vuông

Ví dụ 8 :

» A1=[-1 1; -2 2];b1=[1 ; 0];
% hệ vô nghiệm
» x1 = A1\b1, D1 = det(A1)
Warning : Matrix is singular to working precision.
x1 = Inf
Inf
D1=0

Giải hệ phương trình tuyến tính vuông

Ví dụ 9 :

```
» A2=[-1 1; -2 2];b2=[1 ; 2];
% hệ vô số nghiệm
» x2 = A2\b2
Warning : Matrix is singular to working precision.
x2 = -1
      0
```

Giải hệ phương trình tuyến tính vuông

Giải hệ phương trình tuyến tính vuông

Ví dụ 3 chiều

Cho hệ phương trình cấp 3 sau :

$$\begin{pmatrix} 10 & -7 & 0 \\ -3 & 2 & 6 \\ 5 & -1 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 7 \\ 4 \\ 6 \end{pmatrix}$$

ta viết lại dưới dạng hệ phương trình tuyến tính

$$\begin{aligned} 10x_1 - 7x_2 \qquad &= 7 \\ -3x_1 + 2x_2 + 6x_3 &= 4 \\ 5x_1 - x_2 + 5x_3 &= 6 \end{aligned}$$

Giải hệ phương trình tuyến tính vuông

Ví dụ 3 chiều (tiếp)

Hệ phương trình tuyến tính

$$
\begin{aligned}
10x_1 - 7x_2 \qquad &= 7 \quad (1)\\
-3x_1 + 2x_2 + 6x_3 &= 4 \quad (2)\\
5x_1 - x_2 + 5x_3 &= 6 \quad (3)
\end{aligned}
$$

ta tiến hành giải

- Khử $x_1 \Rightarrow (1) \times -0.3 + (2)$ và $(1) \times 0.5 - (3)$

Hệ số 10 của ẩn x_1 trong (1) đc gọi là **phần tử trụ**, các hệ số -0.3 và 0.5 đc gọi là **nhân tử**.

Giải hệ phương trình tuyến tính vuông

Ví dụ 3 chiều (tiếp)

Hệ phương trình tuyến tính sau khi khử x_1

$$
\begin{aligned}
10x_1 - 7x_2 &= 7 && (4)\\
-0.1x_2 + 6x_3 &= 6.1 && (5)\\
2.5x_2 + 5x_3 &= 2.5 && (6)
\end{aligned}
$$

ta tiếp tục giải

- Khử $x_2 \Rightarrow$ do phần tử trụ của x_2 trong (5) là -0.1 có trị tuyệt đối nhỏ, ta tiến hành đổi chỗ hai phương trình (5) và (6) rồi mới tiến hành khử x_2.

Việc làm này gọi là **phép xoay**

Ví dụ 3 chiều (tiếp)

Hệ phương trình tuyến tính sau khi thực hiện phép xoay

$$10x_1 - 7x_2 = 7 \quad (7)$$

$$2.5x_2 + 5x_3 = 2.5 \quad (8)$$

$$-0.1x_2 + 6x_3 = 6.1 \quad (9)$$

ta tiếp tục giải

- Khử $x_2 \Rightarrow (8) \times 0.04 + (9)$

Giải hệ phương trình tuyến tính vuông

Ví dụ 3 chiều (tiếp)

Hệ phương trình tuyến tính sau khi thực hiện phép khử x_2

$$
\begin{aligned}
10x_1 - 7x_2 \qquad &= 7 && (10)\\
2.5x_2 + 5x_3 &= 2.5 && (11)\\
6.2x_3 &= 6.2 && (12)
\end{aligned}
$$

ta tiếp tục giải

- Từ phương trình (12) $\Rightarrow x_3 = 1$.
- thay x_3 vào (11) thì $2.5x_2 + 5 \times (1) = 2.5 \Rightarrow x_2 = -1$.
- thay x_2 vào (10) thì $10x_1 - 7 \times (-1) = 7 \Rightarrow x_1 = 0$.

Giải hệ phương trình tuyến tính vuông

Các ma trận L,U,P

Toàn bộ cách giải vừa trình bày có thể được gói gọn trong các ma trận sau

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 0.5 & 1 & 0 \\ -0.3 & 0.04 & 1 \end{pmatrix}, U = \begin{pmatrix} 10 & -7 & 0 \\ 0 & 2.5 & 5 \\ 0 & 0 & 6.2 \end{pmatrix}, P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

Với

- L là ma trận chứa các nhân tử
- U là ma trận hệ số cuối cùng
- P là ma trận hoán vị mô tả phép xoay

thì chúng ta có

$$LU = PA$$

Giải hệ phương trình tuyến tính vuông

Ma trận hoán vị

Ma trận hoán vị là thu được từ ma trận đơn vị I bằng cách hoán vị các hàng của nó.

- Với ma trận hoán vị : $P^{-1} = P^T$
- Phép nhân PX dùng để hoán vị hàng ma trận X
- Phép nhân XP dùng để hoán vị cột ma trận X

Ví dụ 10

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

Giải hệ phương trình tuyến tính vuông

Ví dụ 10 (tiếp)

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

Dạng rút gọn trong Matlab

$$p = [2 \;\; 4 \;\; 3 \;\; 1]$$

Giải hệ phương trình tuyến tính vuông

Ma trận tam giác

Ma trận $X \in \mathbb{R}^{n \times n}$ là **ma trận tam giác trên** nếu $x_{ij} = 0 \in X \quad \forall i < j$ nghĩa là có dạng

$$X = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ 0 & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & x_{nn} \end{pmatrix}$$

- Khi ma trận này có các phần tử đường chéo chính $x_{ii} = 1 \;\; \forall i = 1, \cdots, n$ thì đc gọi là **ma trận tam giác trên đơn vị**.
- Định thức ma trận tam giác trên khác không khi và chỉ khi tất cả các phần tử nằm trên đường chéo chính là khác không.

$\Rightarrow$ Định nghĩa tương tự ta có **ma trận tam giác dưới** và **ma trận tam giác dưới đơn vị**.

Giải hệ phương trình tuyến tính vuông

Ma trận tam giác

Việc giải hệ phương trình ma trận tam giác đều dễ dàng. Ta giải hệ phương trình hàng cuối cùng để có ẩn số cuối, sau đó lần lượt thay vào các phương trình trên để tìm ra các ẩn số còn lại.

Ví dụ 11 :

Giải hệ phương trình tam giác trên $Ux = b$

```
»x = zeros(n,1);
for k = n:-1:1
    x(k) = b(k)/U(k,k);
    i=(1:k-1)';
    b(i) = b(i) - x(k) * U(i,k);
end
```

Giải hệ phương trình tuyến tính vuông

Ví dụ 12 :

Giải hệ phương trình tuyến tính

$$\begin{aligned} 3x_1 + 4x_2 + 5x_3 &= 7 \\ 2x_2 - 3x_3 &= 8 \\ 5x_3 &= 11 \end{aligned}$$

% Chương trình Matlab

```
» U=[3,4,5;0,2,-3;0,0,5]; b = [7;8;11];n=3;x=zeros(n,1);
» for k=n:-1:1
»     x(k) = b(k)/U(k,k);
»     i=(1:k-1)';
»     b(i) = b(i) - x(k) * U(i,k);
» end;
```

Giải hệ phương trình tuyến tính vuông

Phân tích LU

Thuật toán phổ biến dùng để giải hệ phương trình tuyến tính vuông có hai giai đoạn

- Khử xuôi (Forward elimination) chính là phép chuyển ma trận vuông về dạng tam giác trên dùng để khử từng ẩn số, với nhân tử, phần tử trụ tương thích kết hợp phép xoay.
  - gồm $n-1$ bước
  - tại bước $k = 1, \cdots, n-1$ nhân phương trình thứ k với nhân tử rồi trừ các phương trình còn lại để khử ẩn số x_k.
- Thế ngược (Backward subtitution) giải phương trình hàng cuối cùng để tìm ẩn cuối cùng, sau đó thế ngược lần lượt lên các hàng trên để tìm ra các ẩn còn lại. (xem lại ví dụ 12)

Phân tích LU (tiếp)

- Gọi P_k là các ma trận hoán vị tại các bước $k = 1, \cdots, n-1$
- Gọi M_k là các ma trận tam giác dưới đơn vị thu được bằng cách chèn các nhân tử được sử dụng ở bước k xuống dưới vị trí đường chéo của cột k của ma trận đơn vị.
- Gọi U là ma trận tam giác trên thu được cuối cùng khi kêt thúc giai đoạn khử xuôi.

Quá trình khử được viết lại dưới dạng ma trận như sau

$$U = M_{n-1}P_{n-1}\cdots M_1P_1A$$

Phân tích LU (tiếp)

Phương trình có thể viêt lại tương đương như sau

$$L_1 L_2 \cdots L_{n-1} U = P_{n-1} \cdots P_1 A$$

trong đó L_k thu được từ M_k bằng cách hoán vị và đổi dấu các nhân tử ở dưới đường chéo. Vậy nếu ta đặt

$$L = L_1 L_2 \cdots L_{n-1}$$
$$P = P_{n-1} \cdots P_2 P_1$$

thì ta thu được công thức ban đầu

$$LU = PA$$

Giải hệ phương trình tuyến tính vuông

Ví dụ 12 :

Quay lại ví dụ 3 chiều đầu tiên, ta có $A = \begin{pmatrix} 10 & -7 & 0 \\ -3 & 2 & 6 \\ 5 & -1 & 5 \end{pmatrix}$ thì các ma trận xác định trong quá trình khử xuôi là

$$P_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, M_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0.3 & 1 & 0 \\ -0.5 & 0 & 1 \end{pmatrix}$$

$$P_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, M_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0.04 & 1 \end{pmatrix}$$

Giải hệ phương trình tuyến tính vuông

Ví dụ 12 (tiếp):

Các ma trận L_1, L_2 tương ứng là

$$L_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0.5 & 1 & 0 \\ -0.3 & 0 & 1 \end{pmatrix}, L_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -0.04 & 1 \end{pmatrix}$$

Chú ý :

Khi tính toán giai đoạn khử, ta sẽ tính toán trực tiếp trên các hàng của ma trận chứ không thực hiện phép nhân ma trận như trên.

Phân tích LU

- Hệ thức $LU = PA$ vừa trình bày được gọi là *phân tích LU* hay *phân rã tam giác* của ma trận A.

Giải hệ phương trình tuyến tính vuông

Giải hệ phương trình bằng phân tích LU

Với hệ phương trình

$$Ax = b$$

với ma trận A là không suy biến đồng thới $PA = LU$ là phân tích LU của A thì hệ phương trình có thể được giải bới hai bước.

- **Khử xuôi** Giải hệ

  $$Ly = Pb$$

  để tìm y, do L là ma trận đơn vị dưới nên y có thể tìm được nhở phép khử xuôi (từ trên xuống dưới).
- **Thế ngược** Giải hệ

  $$Ux = y$$

  bằng cách thế ngược để tìm được x.

Vai trò của phần tử trụ

Phần tử trụ

- Các phần tử nằm trên đường chéo chính của ma trận U.
- Phần tử trụ thứ k là hệ số của ẩn x_k trong phương trình thứ k tại bước k của giai đoạn khử.
- Trong cả hai bước khử xuôi và thế ngược đều cần chia cho phần tử trụ nên chúng không thể có giá trị không.

Trực giác :

hệ phương trình giải tồi nếu phần tử trụ gần không.

Vai trò của phần tử trụ

Ví dụ 13 :

Thay đổi chút ít tại hàng hai trong các ví dụ trên

$$\begin{pmatrix} 10 & -7 & 0 \\ -3 & 2.099 & 6 \\ 5 & -1 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 7 \\ 3.901 \\ 6 \end{pmatrix}$$

Như vậy, giả sử mọi tính toán chính xác đến 4 chữ số thập phân

- Hệ số ẩn x_2 tại hàng hai thay đổi từ 2.000 thành 2.099
- Đồng thời vế phải tương ứng thay đổi từ 4.000 thành 3.901

mục đích là giữ nguyên nghiệm $(0, -1, 1)^T$ của hệ phương trình.

Vai trò của phần tử trụ

Ví dụ 13 (tiếp) :

Bước đầu tiên của giai đoạn khử

$$\begin{pmatrix} 10 & -7 & 0 \\ 0 & -0.001 & 6 \\ 0 & 2.5 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 7 \\ 6.001 \\ 2.5 \end{pmatrix}$$

tiếp tục thực hiện khử dù phần tử trụ dù -0.001 là nhỏ so với các hệ số khác của ma trận mà không thực hiện phép xoay. Vậy ta

- Nhân phương trình hàng thứ hai với 2.5×10^3 rồi cộng với hàng thứ ba.
- Vế phải của phương trình này, khi nhân 6.001 với 2.5×10^3 thì kết quả 1.50025×10^4 làm tròn thành 1.5002×10^4

Vai trò của phần tử trụ

Ví dụ 13 (tiếp) :

$$\begin{pmatrix} 10 & -7 & 0 \\ 0 & -0.001 & 6 \\ 0 & 0 & 1.5005 \times 10^4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 7 \\ 6.001 \\ 1.5004 \times 10^4 \end{pmatrix}$$

tiếp tục ...

- Kết quả vế phải phương trình hai làm tròn 1.5001×10^4 được cộng với 2.5 là vế phải của phương trình thứ ba và lại được làm tròn.

Vậy phương trình ba trở thành $1.5005 \times 10^4 x_3 = 1.5004 \times 10^4$ giải ra ta có

$$x_3 = \frac{1.5004 \times 10^4}{1.5005 \times 10^4} = 0.99993$$

Rõ ràng, với giá trị chính xác của ẩn số $x_3 = 1$ thì giá trị giải được bởi phương trình này chưa đáng ngại.

Vai trò của phần tử trụ

Ví dụ 13 (tiếp) :

$$\begin{pmatrix} 10 & -7 & 0 \\ 0 & -0.001 & 6 \\ 0 & 0 & 1.5005\times 10^4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 7 \\ 6.001 \\ 1.5004\times 10^4 \end{pmatrix}$$

tiếp tục ...

- đối với ẩn số x_2

$$-0.001x_2 + 6\times(0.99993) = 6.001$$

  nên $x_2 = \frac{1.5\times 10^{-3}}{-1.0\times 10^{-3}} = -1.5$

- Cuối cùng thế lên phương trình đầu tìm ẩn x_1

$$10x_1 - 7\times(-1.5) = 7$$

  suy ra $x_1 = -3.5$

Vai trò của phần tử trụ

Ví dụ 13 (tiếp) :

Như vậy khi không thực hiện phép xoay chọn phần tử trụ

$$\begin{pmatrix} 10 & -7 & 0 \\ -3 & -0.001 & 6 \\ 5 & 2.5 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 7 \\ 6.001 \\ 2.5 \end{pmatrix}$$

thay vì được nghiệm $(0, -1, 1)^T$ ta lại được nghiệm $(-0.35, -1.5, 0.99993)^T$.

Vì sao có sự cố này ?

Sai số là do chúng ta chọn **phần tử trụ quá nhỏ**. Vậy ta nên chọn phần tử trụ có trị tuyệt đối lớn nhất tại mỗi bước khử k.

Hiệu ứng của sai số làm tròn

Cách đo sự khác biệt

Thông thường, khi thu được lời giải x^* khác với lời giải đúng x ta thường sử dụng hai cách đo sự khác biệt

- Sai số : $e = x - x^*$
- Độ lệch : $r = b - Ax^*$

Về lý thuyết nếu A không suy biến thì hai đại lượng này cùng bằng không, tuy nhiên khi tính toán trong máy tính hai đại lượng này không đồng điệu.

Hiệu ứng của sai số làm tròn

Ví dụ 14 :

Xét hệ phương trình

$$0.780x_1 + 0.563x_2 = 0.217$$
$$0.913x_1 + 0.659x_2 = 0.254$$

Khử Gauss như ví dụ trước, áp dụng quy tắc chọn phần tử trụ lớn nhất tuy nhiên mọi tính toán chỉ chính xác đến 3 chữ số thập phân.

Hiệu ứng của sai số làm tròn

Ví dụ 14 (tiếp):

tiếp tục ....

- Thực hiện phép xoay, để 0.913 trở thành phần tử trụ.

$$0.913x_1 + 0.659x_2 = 0.254$$
$$0.780x_1 + 0.563x_2 = 0.217$$

  Tính hệ số $0.780/0.913 = 0.854$

- Nhân hệ số 0.854 với pt thứ nhất rồi trừ đi pt thứ hai. Ta được

$$0.913x_1 + 0.659x_2 = 0.254$$
$$0.001x_2 = 0.001$$

Hiệu ứng của sai số làm tròn

Ví dụ 14 (tiếp):

$$\begin{aligned} 0.913x_1 + 0.659x_2 &= 0.254 \\ 0.001x_2 &= 0.001 \end{aligned}$$

tiếp tục ....

- Ẩn $x_2 = 0.001/0.001 = 1.000$ (chính xác)
- Thế lên pt trên, $x_1 = (0.254 - 0.659x_2)/0.913 = -0.443$

Cuối cùng ta thu được nghiệm $x^* = (-0.443, 1.000)^T$

Hiệu ứng của sai số làm tròn

Ví dụ 14 (tiếp):

Đo sự khác biệt, rõ ràng nghiệm đúng của hệ $x = (1, -1)^T$

- Sai số : $e = x - x^* = (1.433, -2)^T$
- Độ lệch :

$$r = b - Ax^* = \begin{pmatrix} 0.217 - (0.780(-0.443) + 0.563(1.000)) \\ 0.254 - (0.913(-0.443) + 0.659(1.000)) \end{pmatrix}$$
$$= \begin{pmatrix} -0.000460 \\ -0.000541 \end{pmatrix}$$

Rõ ràng trong khi độ lệch chấp nhận được khi ta làm tròn sai số 3 số thập phân sau dấu phẩy thì sai số thậm chí còn lớn hơn cả lời giải.

Hiệu ứng của sai số làm tròn

Các câu hỏi đặt ra khi có hiện tượng sai số khi làm tròn

- Vì sao độ lệch lại có giá trị nhỏ ?
- Vì sao sai số lại có giá trị quá lớn ?
- Định thức của hệ $0.780 \times 0.659 - 0.913 \times 0.563 = 10^{-6}$ gần không có phải là nguyên nhân gây hiện tượng này ?

Hiệu ứng của sai số làm tròn

Ví dụ 14 (tiếp):

Thay giả thiết làm tròn với 3 số thập phân thành làm tròn với 6 số thập phân sau dấu phẩy, ta thu được hệ phương trình sau khi khử Gauss

$$\begin{pmatrix} 0.913000 & 0.659000 \\ 0.000000 & 0.000001 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0.254000 \\ -0.000001 \end{pmatrix}$$

Chú ý, giá trị vế phải của pt thứ hai đã thay đổi. Thực tế **nghiệm gần đúng cũng là nghiệm chính xác** của hệ

$$x_1 = \frac{-0.000001}{0.000001} = -1.000000$$

$$x_2 = \frac{0.254 - 0.659x_2}{0.913} = 1.000000$$

Hiệu ứng của sai số làm tròn

Giải thích tại sao độ lệch nhỏ

- Độ lệch của phương trình hai nhỏ do ma trận gần suy biến $det(A) = 10^{-6}$ dẫn đến hai phương trình gần như phụ thuộc tuyến tính.
- Vì thế cặp ẩn (x_1, x_2) thỏa mãn phương trình thứ nhất cũng thỏa mãn phương trình thứ hai

$\Rightarrow$ Nếu như ta biết chắc định thức bằng không thì ta không cần phải quan tâm đến phương trình thứ hai vì mọi nghiệm của hệ phương trình thứ nhất đều thỏa mãn hệ phương trình thứ hai.

Kết luận quan trọng :

Khi ta tiếng hành khử Gauss với phần tử trụ tối đại trên cột **đảm bảo** cho độ lệch $r = b - Ax^*$ là nhỏ.

Hiệu ứng của sai số làm tròn

Hệ xác định tồi và số điều kiện của ma trận

Các hệ số ít khi được biết tới một cách chính xác

bởi vì

- Đối với các hệ phương trình xuất hiện trong ứng dụng, các hệ số thường được quy cho giá trị thực nghiệm nên gắn với sai số quan sát.
- Nhiều hệ phương trình khác có các hệ số được tính bởi các công thức và vì thế chúng được biết chính xác tới sai số làm tròn khi tính toán theo công thức đã cho.
- Ngay cả đối với các hệ phương trình được cất giữ chính xác trong máy tính cũng không thể tránh khỏi sai số (Xem lại các cách biểu diễn số nguyên, số nguyên có dấu, số thực dấu phẩy động trong giáo trình Tin học đại cương)

Vậy câu hỏi được đặt ra là :

Nếu có sai số trong biểu diễn các hệ số của hệ phương trình tuyến tính thì điều đó ảnh hưởng đến lời giải như thế nào ? Hay khi giải $Ax = b$ làm sao có thể đo được độ nhạy của x khi có thay đổi trong A, b ?

Hệ xác định tồi và số điều kiện của ma trận

Có một vài nhận xét sau

- Nếu A suy biến thì đối với b nào đó x hoặc vô nghiệm hoặc vô số nghiệm. Trong trường hợp A có đinh thức nhỏ thì một sự thay đổi nhỏ trong A và b có thể dẫn sự thay đổi lớn trong lời giải.
- Hãy nghĩ đến kích thước các phần tử trụ và khái niệm gần suy biến. Bởi vì nếu các phép toán số học được thực hiện chính xác thì tất cả các phần tử trụ khác không khi và chỉ khi ma trận không suy biến. Từ đó rút ra khẳng định sau :'Nếu các phần tử trụ là nhỏ thì ma trận gần suy biến' điều ngược lại không đúng, hay nói cách khác có ma trận gần suy biến mà các phần tử trụ đều không nhỏ.

Chuẩn vec tơ

Định nghĩa : Hàm $v : \mathbb{R}^n \mapsto R$ được gọi là chuẩn vec tơ (vector norm) trên $\mathbb{R}^n$ khi và chỉ khi

1. $v(x) \geq 0 \ \ \forall x \in \mathbb{R}^n$ và $v(x) = 0$ khi và chỉ khi $x = 0$
2. $v(\alpha x) = |\alpha| v(x) \ \ \forall x \in \mathbb{R}^n, \forall \alpha \in \mathbb{R}$
3. $v(x + y) \leq v(x) + v(y) \ \ \forall x, y \in \mathbb{R}^n$ đây là bất đẳng thức tam giác.

Thông thường $v(x)$ được ký hiệu là $||x||$

Hệ xác định tồi và số điều kiện của ma trận

Chuẩn vec tơ (tiếp)

Một số chuẩn thường dùng

- $||x||_2 = \sqrt{\sum_{i=1}^{n} x_i^2}$ là (l_2) hay chuẩn Euclid
- $||x||_1 = \sum_{i=1}^{n} |x_i|$ là (l_1)
- $||x||_\infty = \max_{1 \le i \le n} |x_i|$ là (l_∞)
- $||x||_p = (\sum_{i=1}^{n} x_i^2)^{1/p}$ là (l_p)

Matlab

norm(x,p) dùng cho l_p
còn với $p = 2$ thì hàm đơn giản hơn **norm(x)**

Chuẩn ma trận

Định nghĩa : Hàm $||.|| : \mathbb{R}^{n\times n} \mapsto \mathbb{R}$ được gọi là chuẩn ma trận nếu

$$||A|| = \max_{||x||=1, x\in\mathbb{R}^n} ||Ax|| = \max_{||x||=1, x\in\mathbb{R}^n} \frac{||Ax||}{||x||}$$

trong đó $||Ax||$ là chuẩn của vec tơ Ax. Tất nhiên, ta có bất đẳng thức $||Ax|| \leq ||A|| \, ||x||$

Hệ xác định tồi và số điều kiện của ma trận

Chuẩn ma trận (tiếp)

Các tính chất của chuẩn ma trận

1. $||A|| \geq 0; ||A|| = 0$ khi và chỉ khi $A = 0$.
2. $||\alpha A|| = ||\alpha|| ||A||, \alpha \in \mathbb{R}$
3. $||A + B|| \leq ||A|| + ||B||$
4. $||AB|| \leq ||A|| \times ||B||$
5. $||Ax|| \leq ||A|| ||x||$

Hệ xác định tồi và số điều kiện của ma trận

Chuẩn ma trận (tiếp)

Các chuẩn vec tơ sinh ra các chuẩn ma trận tương ứng

- Chuẩn Euclid : $||A||_2 = \max_{||x||_2=1} ||Ax||_2$
- Chuẩn max "tổng dòng" : $||A||_\infty = \max_{||x||_\infty=1} ||Ax||_\infty = \max_{1\leq i\leq n} \sum_{j=1}^{n} |a_{ij}|$
- Chuẩn max "tổng cột" : $||A||_1 = \max_{||x||_1=1} ||Ax||_1 = \max_{1\leq j\leq n} \sum_{i=1}^{n} |a_{ij}|$
- Chuẩn Frobenius : $||A||_F = \left(Tr(A^TA)^{1/2}\right) = \left(\sum_{i,j=1}^{n} a_{ij}^2\right)^{1/2}$

Trong Matlab

norm(A,p) trong đó $p = 1, 2, inf$

Hệ xác định tồi và số điều kiện của ma trận

Số điều kiện của ma trận

Định nghĩa : Số điều kiện (condition number) **cond(A)**, thường được ký hiệu là $\kappa_p(A)$, của ma trận vuông A tính đối với một chuẩn ma trận p cho trước là số

$$cond(A) = ||A|| \cdot ||A^{-1}||$$

trong đó, ta quy ước $cond(A) = \infty$ khi A là suy biến. Bởi vì,

$$||A|| \cdot ||A^{-1}|| = \frac{\max_{x \neq 0} \frac{||Ax||}{||x||}}{\min_{x \neq 0} \frac{||Ax||}{||x||}}$$

nên số điều kiện đo tỷ số giữa độ giãn nở lớn nhất và độ co hẹp lớn nhất mà ma trận có thể tác động đối với một vec tơ khác không.

Hệ xác định tồi và số điều kiện của ma trận

Số điều kiện của ma trận (tiếp)

Số điều kiện cho biết ma trận gần suy biến đến mức độ nào : ma trận càng lớn càng **gần suy biến** (hệ phương trình tương ứng xác định tồi), trái lại ma trận với số điều kiện càng gần 1 càng xa với **gần suy biến**.

Chú ý :

Định thức ma trận không là đặc trưng tốt cho tính gần suy biến. Mặc dù khi $det(A) = 0$ thì ma trận suy biến nhưng độ lớn hay nhỏ của định thức không chứa thông tin về việc ma trận có gần suy biến hay không.
Ví dụ, cho ma trận $det(\alpha \mathbb{I}_n) = \alpha^n$ có thể là số rất nhỏ khi $|\alpha| < 1$ nhưng ma trận $\alpha \mathbb{I}_n$ lại có điều kiện tốt với $cond(\alpha \mathbb{I}_n) = 1$. Trong đó $\mathbb{I}_n$ là ma trận đơn vị n chiều.

Hệ xác định tồi và số điều kiện của ma trận

Một số tính chất số điều kiện của ma trận

1. Với mọi ma trận A : $cond(A) \geq 1$
2. Với mọi ma trận đơn vị $\mathbb{I}$: $cond(\mathbb{I}) = 1$
3. Với mọi ma trận hoán vị P : $cond(P) = 1$
4. Với mọi ma trận A và số thực khác không α : $cond(\alpha A) = cond(A)$
5. Với mọi ma trận đường chéo $D = diag(d_i)$: $cond(D) = \frac{\max\{d_i\}}{\min\{d_i\}}$
6. Số điều kiện có ý nghĩa quan trọng trong việc đánh giá tính chính xác của lời giải của hệ phương trình tuyến tính.

Matlab với số điều kiện

cond(A,p) để tính $\kappa_p(A)$ với $p = 1, 2, inf$
condest(A) để đánh giá $\kappa_1(A)$
rcond(A) để đánh giá $1/\kappa_1(A)$

Đánh giá sai số khi biết số điều kiện của ma trận

Gọi x là lời giải chính xác của $Ax = b$, còn x^* là lời giải của hệ $Ax^* = b + \Delta b$ (chú ý ta chỉ coi b bị nhiễu cộng). Đặt $\Delta x = x^* - x$, ta có $b + \Delta b = Ax^* = A(x + \Delta x) = Ax + A\Delta x$ do $Ax = b$ thế vào suy ra $\Delta x = A^{-1}\Delta b$.

$$b = Ax \Rightarrow ||b|| \leq ||A||||x|| \tag{13}$$

$$\Delta x = A^{-1}\Delta b \Rightarrow ||\Delta x|| \leq ||A^{-1}||||\Delta b|| \tag{14}$$

Nhân hai bất đẳng thức (13) (14) và sử dụng định nghĩa $cond(A) = ||A||||A^{-1}||$ ta có đánh giá

$$\frac{||\Delta x||}{||x||} \leq cond(A)\frac{||\Delta b||}{||b||}$$

Hệ xác định tồi và số điều kiện của ma trận

Đánh giá sai số khi biết số điều kiện của ma trận (tiếp)

tiếp tục...

$$\frac{||\Delta x||}{||x||} \leq cond(A)\frac{||\Delta b||}{||b||}$$

Vậy số điều kiện cho phép ta xác định khả năng biến đổi sai số tương đối trong lời giải $\frac{||\Delta x||}{||x||}$ khi biết sự thay đổi tương đối trong vế phải $\frac{||\Delta b||}{||b||}$

- Khi $cond(A)$ lớn hay hệ gần suy biến thì sự biến đổi tương đối của vế phải sẽ 'ép' sự thay đổi sai số tương ứng trong lời giải.
- Ngược lại, khi $cond(A)$ tiến về 1 hay hệ có điều kiện tốt thì sự biến đổi tương đương của vế phải và lời giải là như nhau.

Hệ xác định tồi và số điều kiện của ma trận

Đánh giá sai số khi biết số điều kiện của ma trận (Kết luận)

Nếu dữ liệu vào được biểu diễn gần đúng với độ chính xác máy tính thì đánh giá sai số tương đối của lời giải tính được sẽ cho bởi công thức:

$$\frac{||x^* - x||}{||x||} \approx cond(A)\epsilon_M$$

lời giải tính được sẽ mất đi một quãng $\log_{10}(cond(A))$ chữ số thập phân trong sai số tương đối so với độ chính xác của dữ liệu.

Kết luận

Hệ phương trình tuyến tính $Ax = b$ là có điều kiện tồi nếu $cond(A)$ là lớn, khi đó sự thay đổi không lớn của dữ liệu có thể dẫn đến sự thay đổi lớn của lời giải.

Hệ xác định tồi và số điều kiện của ma trận

Ví dụ 15 :

Xét hệ phương trình

$$0.789x_1 + 0.563x_2 = 0.127$$
$$0.913x_1 + 0.659x_2 = 0.254$$

Kết quả khi dùng Matlab

```
» A=[0.789 0.563;0.913 0.659];
» fprintf('cond(A)=%d ; det(A)=%d ',cond(A),det(A))
» cond(A) = 2.193219e+006 ; det(A)=1.000000e-006
```

Hệ xác định tồi và số điều kiện của ma trận

Ví dụ 16 :

Xét hệ phương trình

$$\begin{pmatrix} 4.1 & 2.8 \\ 9.7 & 6.6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 4.1 \\ 9.7 \end{pmatrix}$$

Đây là hệ có điều kiện tồi do $cond(A,1) = 2494.4$ đồng thời nghiệm chính xác của hệ là $x = (1,0)^T$. Nếu ta thay vế phải $b + \Delta b = (4.11, 9.70)^T$ thì nghiệm của hệ sẽ là $x^* = (0.34, 0.97)^T$.

Trong Matlab ta có

» A = [4.1 2.8; 9.7 6.6]; b = [4.1 ; 9.7]; b1=[4.11 ; 9.7];
» x = (A \ b)', x1 = (A \ b1)'
x = 1 0
x1 = 0.3400 0.9700

Hệ xác định tồi và số điều kiện của ma trận

Bài đọc thêm về nhà

Chúng ta có thể sử dụng nhiều phương pháp ngoài phân tích LU để giải hệ phương trình

- Phân tích Cholesky
  - Khái niệm ma trận bán xác định dương
  - Nếu A là ma trận xác định dương thì tồn tại ma trận tam giác dưới dương L sao cho $A = LL^T$
  - Khử xuôi $Ly = b$, thế ngược $L^T x = y$
- Phân rã QR
  - Khái niệm ma trận trực giao
  - Phân rã QR : nếu A có hạng n thì tồn tại $A = QR$
  - Giải bài toán bình phương tối thiểu

$$\min\{||Ax - b||^2 | x \in \mathbb{R}\}$$

vậy nghiệm x là điểm dừng khi cực tiểu hóa bài toán trên.