Numerical Analysis FMN011

Carmen Arévalo

Lund University

carmen@maths.lth.se

Lecture 4

Linear Systems $Ax = b$

A is $n \times n$ matrix, b is given n-vector, x is unknown solution n-vector.

$A^{n\times n}$ is non-singular (invertible) if it has any one of the following properties:

- A has an inverse
- $\det(A) \neq 0$.
- $\mathrm{rank}(A) = n$
- The unique solution of $Ax = 0$ is $x = 0$.
- The system $Ax = b$ has a unique solution.

If A is singular, then the system $Ax = b$ has either infinitely many solutions or no solution at all.

Solving Triangular Linear Systems

Upper triangular matrix: $a_{ij} = 0$ if $i > j$.

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n-1} & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n-1} & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3n-1} & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & a_{n-1n-1} & a_{n-1n} \\ 0 & 0 & 0 & \ldots & 0 & a_{nn} \end{bmatrix}$$

Back substitution

```
function x=back(A,b)

%   x = back(A,b)
%      Performs back substitution on system Ax=b.
%      A is assumed to be in upper triangular form.

n = length(b);

x(n,1) = b(n)/A(n,n);

for i= n-1:-1:1
   x(i,1) = (b(i) - A(i,i+1:n)*x(i+1:n,1))/A(i,i);
end
```

Forward Substitution

Lower triangular matrix: $a_{ij} = 0$ if $i < j$.

```
function x = forward( A, b )
%
%       Performs forward substitution on system Ax=b.
%       A is assumed to be in lower triangular form.

n = length(b);

x(1,1) = b(1)/A(1,1);

for i = 2:n
   x(i,1) = (b(i)-A(i,1:i-1)*x(1:i-1,1))/A(i,i);
end
```

Computational complexity of substitution

$$\boxed{\sum_{k=1}^{N} k = \frac{N(N+1)}{2}}$$

Back and forward substitution have same number of FLOPs (sum, difference, multiplication, division or square root).

Number of FLOPs:

$$1 + \sum_{i=2}^{n} [(i-1) + i] = 1 + \sum_{i=1}^{n-1} (2i+1) = n^2$$

Elementary transformations

Equivalent systems have the same solution

Operations on **equations** that yield an equivalent system:

- Interchanges (order of equations can be changed)
- Scaling (equation can be multiplied by a constant)
- Replacement (add a multiple of another equation)

Operations on **rows** that yield an equivalent system:

- Row interchanges
- Multiplication by a constant
- $\text{row}_r = \text{row}_r - m_{rp} \times \text{row}_p$

Gaussian elimination

Converts a system into one with a $\triangle$ matrix by means of elementary transformations.

$$\begin{aligned} 2x_1 + 4x_2 - 6x_3 &= -4 \\ x_1 + 5x_2 + 3x_3 &= 10 \\ 3x_1 + 9x_2 + 6x_3 &= 15 \end{aligned}$$

Augmented matrix:

$$\begin{array}{r} \text{pivot} \rightarrow \\ m_{21} = 1/2 \\ m_{31} = 3/2 \end{array} \left[\begin{array}{ccc|c} \underline{\mathbf{2}} & 4 & -6 & -4 \\ 1 & 5 & 3 & 10 \\ 3 & 9 & 6 & 15 \end{array}\right]$$

m_{ij} are called multipliers

Replace rows 2 and 3:

$$\begin{array}{r} \\ \textbf{pivot} \rightarrow \\ m_{32} = 1 \end{array} \left[\begin{array}{ccc|c} 2 & 4 & -6 & -4 \\ 0 & \underline{\mathbf{3}} & 6 & 12 \\ 0 & 3 & 15 & 21 \end{array}\right]$$

Replace row 3:

$$\left[\begin{array}{ccc|c} 2 & 4 & -6 & -4 \\ 0 & 3 & 6 & 12 \\ 0 & 0 & 9 & 9 \end{array}\right]$$

The system matrix is now upper $\triangle$

Solving a system of equations

With upper $\triangle$ matrix, solve by **back substitution**.

To solve a system:

1. Perform a Gaussian elimination
2. Perform a back substitution

In previous example:

$$\begin{aligned} x_3 &= 9/9 = 1 \\ x_2 &= (12 - 6 \cdot 1)/3 = 2 \\ x_1 &= (-4 - 4 \cdot 2 + 6 \cdot 1)/2 = -3 \end{aligned}$$

Pivoting: If during the process a pivot is zero, **exchange rows**. If this is not possible because all elements below it are also zero, the system is singular.

Operation count for Gauss elimination

$$\boxed{\sum_{k=1}^{N} k^2 = \frac{N(N+1)(2N+1)}{6}}$$

```
for j=1:n-1
   for i=j+1:n
      m=a(i,j)/a(j,j);
      b(i)=b(i)-m*b(j);
      for k=j+1:n
         a(i,k)=a(i,k)-m*a(j,k);
      end
   end
end
```

Number of FLOPs for the elimination step:

$$\sum_{j=1}^{n-1}(3+2(n-j)(n-j)) = \frac{2}{3}n^3 + \frac{1}{2}n^2 - \frac{7}{6}n$$

Properties of vector norms

1. $||x|| > 0$ if $x \neq 0$
2. $||cx|| = |c| \cdot ||x||$ for any scalar c
3. $||x + y|| \leq ||x|| + ||y||$

Vector norms

The p norm of a vector $\boxed{||x||_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}}$

- 1-norm: $||x||_1 = \sum_{i=1}^n |x_i|$
- 2-norm: $\left(\sum_{i=1}^n |x_i|^2\right)^{1/2}$ (Euclidean)
- ∞-norm: $\max_i |x_i|$

$$\boxed{||x||_\infty \le ||x||_2 \le ||x||_1 \le \sqrt{n}||x||_2 \le n||x||_\infty}$$

Matrix Norms induced by vector norms

$$||A|| = \max_{||x||=1} ||Ax||$$

Properties:

1. $||A|| > 0$ if $A \neq 0$

2. $||cA|| = |c| \cdot ||A||$ for any scalar c

3. $||A + B|| \leq ||A|| + ||B||$

4. $||AB|| \leq ||A|| \cdot ||B||$

Norms of some interesting matrices

$$\|A\|_1 = \max_j \sum_{i=1}^{n} |a_{ij}|, \quad \|A\|_\infty = \max_i \sum_{j=1}^{n} |a_{ij}|$$

$$\begin{aligned} \|0\| &= 0 \\ \|I\| &= 1 \\ \|P\| &= 1 \quad \text{if} \quad P \quad \text{is a permutation of the rows of} \quad I \\ \|D\| &= \max\{|d_i|\} \end{aligned}$$

Ill conditioning

$Ax = b$ is ill conditioned if small perturbations in the coefficients of A or b produce large changes in x

$$\begin{aligned} 34x_1 + 55x_2 &= 21 \\ 55x_1 + 89x_2 &= 34 \end{aligned} \qquad \Rightarrow \qquad x_1 = -1,\ x_2 = 1$$

Add 1% error to entry a_{21}

$$\begin{aligned} 34x_1 + 55x_2 &= 21 \\ 55.55x_1 + 89x_2 &= 34 \end{aligned} \qquad \Rightarrow \qquad \hat{x}_1 = 0.03419,\ \hat{x}_2 = 0.3607.$$

The relative error (in $\max$-norm) is $|-1-0.034188|/|1| = 103.4\%$

$$\begin{aligned} 34 \cdot (0.034188) + 55 \cdot (0.36068) &= 21 \\ 55 \cdot (0.034188) + 89 \cdot (0.36068) &= 33.9812 \end{aligned}$$

Relative error in residual is $\|A\hat{x} - b\|/\|b\| \approx 0.05\%$

The system is ill conditioned (A is ill conditioned)

A small residual does not point to a small error!

Error magnification factor

The error magnification factor for $Ax = b$ is the ratio between the relative error norm and the relative residual norm,

$$\text{emf } = \frac{\|x - \hat{x}\|/|x\|}{\|b - A\hat{x}\|/\|b\|}$$

It tells us how much the error is affected by the size of the residual.

The maximum possible error magnification factor, over all right-hand sides b, is called the condition number

Condition number

Condition number of a matrix relative to a norm $||\cdot||_p$:

$$\boxed{\kappa_p(A) = ||A||_p \cdot ||A^{-1}||_p}$$

If $\kappa(A) \approx 10^k$, about k significant digits will be lost in solving $Ax = b$. In the previous example, $k = 4$, so we need to have an input with at least 5 correct significant digits.

- $\kappa(A) \geq 1$
- $\kappa(I) = 1$
- $\kappa(P) = 1 \quad$ if P is a permutation of the rows of I
- $\kappa(cA) = \kappa(A)$
- $\kappa(D) = \frac{\max |d_i|}{\min |d_i|}$

Swamping

- Exact solution

$$\begin{pmatrix} 10^{-20} & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 4 \end{pmatrix} \Rightarrow y = \frac{1 - 4 \times 10^{-20}}{1 - 2 \times 10^{-20}} \approx 1;\ x \approx 2$$

- With double precision

$$\begin{pmatrix} 10^{-20} & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 4 \end{pmatrix} \Rightarrow y = 1,\ x = 0$$

- After row exchange

$$\begin{pmatrix} 1 & 2 \\ 10^{-20} & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 4 \\ 1 \end{pmatrix} \Rightarrow y = 1,\ x = 2$$

Multipliers should be small (pivots should be large)