Ill Conditioned Matrices

1. Vector and matrix norms
2. Ill conditioning
3. In-class exercise
4. Condition number

Alan Turing
1948

Ill Conditioned Matrices

Vector and Matrix Norms

Vector Norms

- A scalar measure of vector magnitude
- Notation: $\|\mathbf{x}\|$
- Common norms

$$
\begin{array}{ll}
1-\text{norm} & \|\mathbf{x}\|_1 = |x_1| + |x_2| + \cdots + |x_n| \\
2-\text{norm (Euclidean)} & \|\mathbf{x}\|_2 = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} \\
\infty-\text{norm} & \|\mathbf{x}\|_\infty = \max |x_j| \\
\text{p}-\text{norm} & \|\mathbf{x}\|_p = \left(|x_1|^p + |x_2|^p + \cdots + |x_n|^p\right)^{1/p}
\end{array}
$$

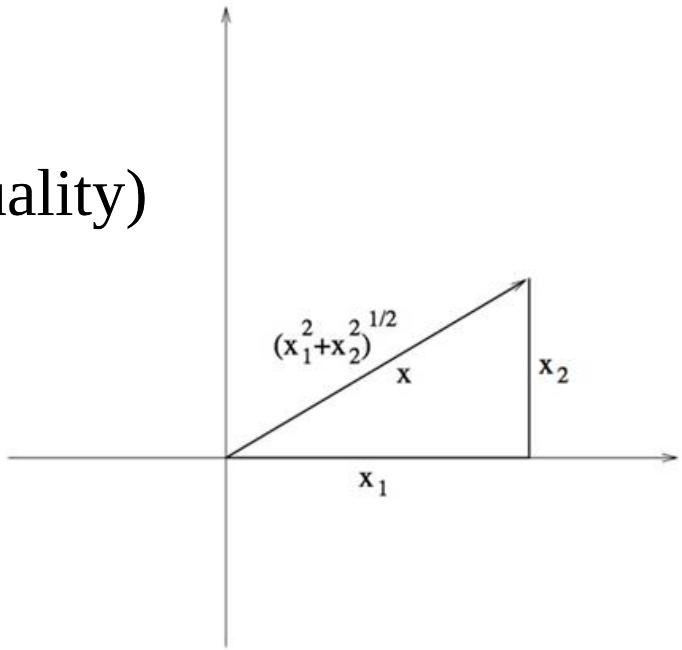
- Example: $\mathbf{x}^T = [2\ \ -3\ \ 0\ \ 1\ \ -4]$

$$
\begin{array}{ll}
1-\text{norm} & \|\mathbf{x}\|_1 = |2| + |-3| + |0| + |1| + |-4| = 10 \\
2-\text{norm (Euclidean)} & \|\mathbf{x}\|_2 = \sqrt{(2)^2 + (-3)^2 + (0)^2 + (1)^2 + (-4)^2} = \sqrt{30} \\
\infty-\text{norm} & \|\mathbf{x}\|_\infty = \max |x_j| = 4
\end{array}
$$

Vector Norms

- Properties
  - » $\|\mathbf{x}\|$ is a non-negative real number
  - » $\|\mathbf{x}\| = 0$ if and only if $\mathbf{x} = 0$
  - » $\|k\mathbf{x}\| = |k|\,\|\mathbf{x}\|$
  - » $\|\mathbf{x} + \mathbf{y}\| <= \|\mathbf{x}\| + \|\mathbf{y}\|$ (triangular inequality)
- Example: Euclidean norm

$$\|\mathbf{x}\|_2 = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$



  - » First three properties hold trivially
  - » Triangular inequality can be proved (see text)

Matrix Norms

- A scalar measure of square matrix magnitude
- Notation: $\|\mathbf{A}\|$
- For any matrix **A** there exists a constant *c* such that:

$$\|\mathbf{Ax}\| \leq c\|\mathbf{x}\|$$

- Equivalent definitions of matrix norm:

$$\|\mathbf{A}\| = \max_{\|\mathbf{x}\|} \frac{\|\mathbf{Ax}\|}{\|\mathbf{x}\|} \quad \|\mathbf{A}\| = \max_{\|\mathbf{x}\|=1} \|\mathbf{Ax}\|$$

Matrix Norms

- Common norms

$$1-\text{norm (column sum)} \quad \|\mathbf{A}\|_1 = \max_k \sum_{j=1}^{n} |a_{jk}|$$

$$\infty-\text{norm (row sum)} \quad \|\mathbf{A}\|_\infty = \max_j \sum_{k=1}^{n} |a_{jk}|$$

$$\text{Frobenius norm} \quad \|\mathbf{A}\|_F = \sqrt{\sum_{j=1}^{n}\sum_{k=1}^{n} a_{jk}^2}$$

- Example

$$\mathbf{A} = \begin{bmatrix} -2 & 1 \\ -1 & 1 \end{bmatrix} \quad \mathbf{A}^{-1} = \begin{bmatrix} -1 & 1 \\ -1 & 2 \end{bmatrix}$$

$$1-\text{norm (column sum)} \quad \|\mathbf{A}\|_1 = \max_k \sum_{j=1}^{2} |a_{jk}| = 3 \quad \|\mathbf{A}^{-1}\|_1 = 3$$

$$\infty-\text{norm (row sum)} \quad \|\mathbf{A}\|_\infty = \max_j \sum_{k=1}^{2} |a_{jk}| = 3 \quad \|\mathbf{A}^{-1}\|_\infty = 3$$

Matrix Norms

- Must satisfy same properties as vector norm
  - » $\|\mathbf{A}\|$ is a non-negative real number
  - » $\|\mathbf{A}\| = 0$ if and only if $\mathbf{A} = \mathbf{0}$
  - » $\|k\mathbf{A}\| = |k|\,\|\mathbf{A}\|$
  - » $\|\mathbf{A} + \mathbf{B}\| <= \|\mathbf{A}\| + \|\mathbf{B}\|$ (triangular inequality)
- Additional properties

$$\|\mathbf{Ax}\| \le \|\mathbf{A}\|\|\mathbf{x}\|$$

$$\|\mathbf{AB}\| \le \|\mathbf{A}\|\|\mathbf{B}\|$$

$$\|\mathbf{A}^n\| \le \|\mathbf{A}\|^n$$

Ill Conditioned Matrices

Ill Conditioning

Ill-Conditioned Matrices

- Matrix inversion: $\mathbf{Ax} = \mathbf{b} \rightarrow \mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$
  - » Assume **A** is a perfectly known matrix
  - » Consider **b** to be obtained from measurement with some error
- Terminology
  - » Well-conditioned problem: “small” changes in the data **b** produce “small” changes in the solution **x**
  - » Ill-conditioned problem: “small” changes in the data **b** produce “large” changes in the solution **x**

Ill-Conditioned Matrices

- Caused by nearly linearly dependent equations → nearly linearly dependent rows and columns
- Characterized by a nearly singular **A** matrix
- Solution is not reliable
- Common problem for large linear algebraic systems
- Ill-conditioning quantified by the condition number

Ill-Conditioned Matrix Example

- Example

$$\begin{bmatrix} 0.9999 & -1.0001 \\ 1 & -1 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 1 \\ 1+\varepsilon \end{bmatrix}$$

  » ε represents measurement error in b_2

  » Two rows (columns) are nearly linearly dependent

- Perform Gauss-Jordan elimination to find $\mathbf{A}^{-1}$

$$[\mathbf{A} \quad \mathbf{I}] = \begin{bmatrix} 0.9999 & -1.0001 & 1 & 0 \\ 1.0000 & -1.0000 & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 0.9999 & -1.0001 & 1 & 0 \\ 0 & 2.0002x10^{-4} & -1.0001 & 1 \end{bmatrix}$$

Ill-Conditioned Matrix Example

- Gauss-Jordan elimination cont.

$$\begin{bmatrix} 1 & -1.0002 & 1.0001 & 0 \\ 0 & 1 & -5000 & 4999.5 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & -5000 & 5000.5 \\ 0 & 1 & -5000 & 4999.5 \end{bmatrix}$$

- Solution

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b} = \begin{bmatrix} -5000 & 5000.5 \\ -5000 & 4999.5 \end{bmatrix} \begin{bmatrix} 1 \\ 1+\varepsilon \end{bmatrix} = \begin{bmatrix} 0.5 + 5000.5\varepsilon \\ -0.5 + 4999.5\varepsilon \end{bmatrix}$$

UMASS.

Ill-Conditioned Matrix Example

- Actual solution

$$\mathbf{x} = \begin{bmatrix} 0.5 + 5000.5\varepsilon \\ -0.5 + 4999.5\varepsilon \end{bmatrix} \xrightarrow{\varepsilon=0} \mathbf{x} = \begin{bmatrix} 0.5 \\ -0.5 \end{bmatrix}$$

- 10% error ($\varepsilon = 0.1$)

$$\mathbf{x} = \begin{bmatrix} 0.5 + (5000.5)(0.1) \\ -0.5 + (4999.5)(0.1) \end{bmatrix} = \begin{bmatrix} 500.55 \\ 499.45 \end{bmatrix}$$

- Small error in the data produced a large change in the solution
- Need a measure of matrix ill-conditioning to identify this problem

Ill Conditioned Matrices

In-class Exercise

Ill Conditioned Matrices

Condition Number

Condition Number

- Definition: $\kappa(\mathbf{A}) = \|\mathbf{A}\| \, \|\mathbf{A}^{-1}\|$
- A "large" condition number (> 1000) indicates an ill-conditioned matrix
- Well conditioned matrix example

$$\mathbf{A} = \begin{bmatrix} 5 & 1 & 1 \\ 1 & 4 & 2 \\ 1 & 2 & 4 \end{bmatrix} \qquad \mathbf{A}^{-1} = \frac{1}{56}\begin{bmatrix} 12 & -2 & -2 \\ -2 & 19 & -9 \\ -2 & -9 & 19 \end{bmatrix}$$

$$\kappa(\mathbf{A}) = \|\mathbf{A}\|_1 \|\mathbf{A}^{-1}\|_1 = (7)\frac{1}{56}(30) = 3.75$$

Condition Number Example 1

- Ill-conditioned matrix example

$$\mathbf{A} = \begin{bmatrix} 1.0001 & 1 \\ 1 & 1.0001 \end{bmatrix} \qquad \mathbf{A}^{-1} = \frac{1}{0.0002}\begin{bmatrix} 1.0001 & -1 \\ -1 & 1.0001 \end{bmatrix}$$

$$\kappa(\mathbf{A}) = \|\mathbf{A}\|_1 \|\mathbf{A}^{-1}\|_1 = (2.0001)\frac{1}{0.0002}(2.0001) = 20002$$

- Effect of ill-conditioning on solution

$$\begin{bmatrix} 1.0001 & 1 \\ 1 & 1.0001 \end{bmatrix}\mathbf{x} = \begin{bmatrix} 1 \\ 1+\varepsilon \end{bmatrix}$$

ε	x_1	x_2
0	0.5	0.5
0.01	-49.5	50.5
0.001	-4.5	5.5
0.0001	0	1.0

Impact of the Condition Number

- Effect of data errors (see text for derivation)

$$\mathbf{A}\mathbf{x} = \mathbf{b} + \delta\mathbf{b} \quad \Rightarrow \quad \frac{\|\delta\mathbf{x}\|}{\|\mathbf{x}\|} \leq \kappa(\mathbf{A}) \frac{\|\delta\mathbf{b}\|}{\|\mathbf{b}\|}$$

  - » δb is the error in the data and δx is the resulting error in the solution
  - » Scalar measures of the errors are obtained using vector norms
  - » The errors are scaled by norms of the data and of the solution
  - » The scaled error in the solution is bounded above by the scaled error in the data

- Implications
  - » The condition number determines how much the data error affects the solution
  - » The upper bound can be very conservative

Condition Number Example 2

$$\begin{bmatrix} 5 & 1 & 1 \\ 1 & 4 & 2 \\ 1 & 2 & 4 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 14 \\ 0 \\ 28 \end{bmatrix} \quad \Rightarrow \quad \mathbf{x} = \begin{bmatrix} 2 \\ -5 \\ 9 \end{bmatrix}$$

$$\begin{bmatrix} 5 & 1 & 1 \\ 1 & 4 & 2 \\ 1 & 2 & 4 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 14 \\ 0 \\ 28 \end{bmatrix} + \begin{bmatrix} 0.1 \\ 0.1 \\ 0.1 \end{bmatrix} \quad \Rightarrow \quad \mathbf{x} = \begin{bmatrix} 2.0143 \\ -4.9857 \\ 9.0143 \end{bmatrix} \quad \Rightarrow \quad \delta\mathbf{x} = \begin{bmatrix} 0.0143 \\ 0.0143 \\ 0.0143 \end{bmatrix}$$

$$\text{Estimate}: \quad \left\|\delta\mathbf{x}\right\|_1 \le \kappa(\mathbf{A}) \frac{\left\|\delta\mathbf{b}\right\|_1}{\left\|\mathbf{b}\right\|_1} \left\|\mathbf{x}\right\|_1 = (3.75)\frac{0.3}{42}(16) = 0.429$$

$$\text{Actual}: \quad \left\|\delta\mathbf{x}\right\|_1 = 0.0429 \ll 0.429$$

Condition Number Example 3

- Hilbert matrix

$$H=\begin{bmatrix} 1 & \frac{1}{2} & \cdots & \frac{1}{n-1} & \frac{1}{n} \\ \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{n} & \frac{1}{n+1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{1}{n-1} & \frac{1}{n} & \cdots & \frac{1}{2n-3} & \frac{1}{2n-2} \\ \frac{1}{n} & \frac{1}{n+1} & \cdots & \frac{1}{2n-2} & \frac{1}{2n-1} \end{bmatrix} \xrightarrow{n=3} H=\begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{bmatrix}$$

- Effect of *n* on condition number

n	κ(H)
2	19
4	1.6×10^{4}
6	1.5×10^{7}
8	1.5×10^{10}

Condition Number Example 4

- Effect of matrix size on condition number
  - » Generated 100 random matrices for each *n* value with *n* = 1, 2, 5, 10, 20, 50, 100, 20, 500, 1000
  - » Averaged the 100 condition numbers for each *n*
- Large matrices are very likely to be ill-conditioned