Condition Number

Sensitivity of linear system

Outline

- Vector norms
- Matrix norms
- Condition numbers
- Usage of condition number
- Estimation of condition number

Norms

- Norm: A metric for measuring entities
  - Vectors, matrices, …
  - Norm ≈ length
- Assuming ***x*** is the entity, then a norm ||***x***|| must obey the following rules
  1. $\|x\| \geq 0$
  2. $\|\alpha * x\| = |\alpha| * \|x\|$, where α is a scalar.
  3. $\|x + y\| \leq \|x\| + \|y\|$,
  4. $\|x * y\| \leq \|x\| * \|y\|$, where '*' denotes a product operator.

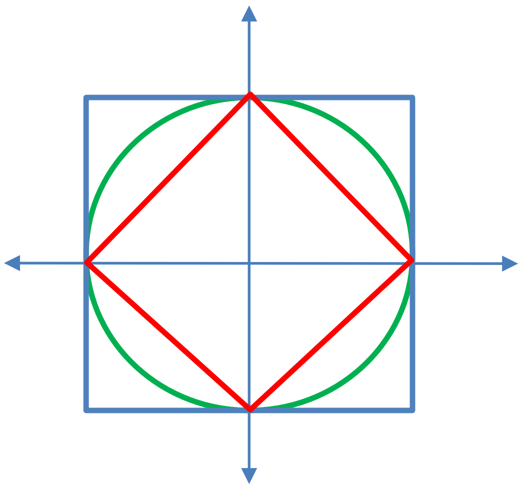
Vector Norms

- 1-norm
  $\vec{x} = [x_0 \quad \dots \quad x_{n-1}]^T$,
  $\|\vec{x}\|_1 = \sum_{i=0}^{n-1} |x_i|$.
- 2-norm
  $\|\vec{x}\|_2 = (\sum_{i=0}^{n-1} {x_i}^2)^{\frac{1}{2}}$
  $= \sqrt{x_0^2 + \cdots + x_{n-1}^2}$
- ∞-norm
  $\|\vec{x}\|_\infty = \max_i |x_i|$,
- p-norm
  $\|\vec{x}\|_p = (\sum_{i=0}^{n-1} x_i^p)^{\frac{1}{p}}$.

Assuming $\|\vec{x}\| \leq 1$ ($\vec{x}$: *positional vector*), based on different norms, the regions containing $\vec{x}$ are different:

- Red diamond: 1-norm
- Green circle: 2-norm
- Blue square: ∞-norm

Regions confined by other norms are between the blue square and green circle.



Questions

- $\vec{x} = [x_0 \quad \cdots \quad x_{n-1}]^T$, which norm of ***x*** is the max?

Answer: the 1-norm

- Given a vector ***x***, which norm of ***x*** is the min?

Answer: ∞-norm

- Why use different norms?

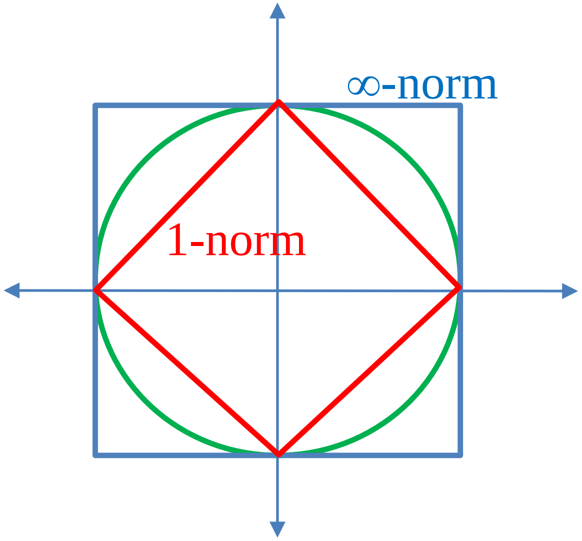
Answer: depending on applications

- Given an error vector e, which norm makes the following condition be the most difficult to achieve $\|e\| \leq \varepsilon$?

Answer:

  - 1-norm is the hardest
  - ∞-norm is the easiest.

- 1-norm, $\|\vec{x}\|_1 = \sum_{i=0}^{n-1} |x_i|$.
- 2-norm, $\|\vec{x}\|_2 = \sqrt{x_0^2 + \cdots + x_{n-1}^2}$
- ∞-norm, $\|\vec{x}\|_\infty = \max_i |x_i|$,



Matrix Norms

- Using *vector-induced* matrix norms
- Assuming that ***A*** is an *n* by *n* matrix, then a vector-induced norm of ***A*** is defined as

  $\|A\| = \sup(\|Ax\|), x \in R^n \; and \; \|x\| = 1.$
- Shortcuts

  $\|A\|_1 = \max_j \sum_{i=0}^{n-1} |a_{ij}|$, max column-wise sum of entry magnitudes.

  $\|A\|_\infty = \max_i \sum_{j=0}^{n-1} |a_{ij}|$, max row-wise sum of entry magnitudes.

  $\|A\|_2 = \max_i |\lambda_i|$, λ_i is the ith eigenvalue, max eigenvalue magnitude.

Example

- Given a matrix, compute its 1-norm, 2-norm, and ∞-norm.

$$A = \begin{bmatrix} 6 & 2 & 3 \\ 2 & 8 & 1 \\ 3 & 1 & 5 \end{bmatrix},$$

- The norms are as follows

Summations of columns: 11, 11, 9

Summations of rows: 11, 11, 9

Eigen values: {10.4753, 6.1242, 2.4005}

1-norm: 11

2-norm: 10.4753

∞-norm: 11

Questions

- For vector norms:
  $\|\vec{x}\|_1 \geq \|\vec{x}\|_2 \geq \|\vec{x}\|_\infty$
- Do these relations hold for matrix norms?
  Answer: No.

- Are there other matrix norms?
  Answer: Yes.
- Entry-wise matrix norm:
  $\|A\|_p = \left(\sum\sum\left|a_{ij}\right|^p\right)^{1/p}$. $p = 1, 2, \ldots$
- Trace norm of matrix: $\|\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\|_2$,
  - Assume $\boldsymbol{A}$ is a real matrix

Perturbed Right Hand Side

- Given a linear system $Ax = b$, if the rhs b is perturbed then
  - The system $Ax = b$ becomes $Ax = (b + \Delta b)$.
- The solution will not be x but $x + \Delta x$ such that
  $A(x + \Delta x) = (b + \Delta b)$.
- What is the relation between the perturbation Δb and the error Δx?
  - Depending on the sensitivity of $\boldsymbol{A}$.

Perturbed Right Hand Side

- What's the effect of the perturbation upon the solution?

  $A(x + \Delta x) = (b + \Delta b).$

  $x + \Delta x = A^{-1}(b + \Delta b).$

  $\Delta x = A^{-1}b + A^{-1}\Delta b - x = A^{-1}\Delta b.$

  $\|\Delta x\| \leq \|A^{-1}\| \cdot \|\Delta b\|.$

  - Divided both sides by $\|x\|$, we have

  $$\frac{\|\Delta x\|}{\|x\|} \leq \|A^{-1}\| \left(\frac{\|\Delta b\|}{\|x\|}\right).$$

  - Multiply the right side by $\frac{\|b\|}{\|b\|}$, we have

  $$\frac{\|\Delta x\|}{\|x\|} \leq \|A^{-1}\| \left(\frac{\|\Delta b\|}{\|b\|}\right)\left(\frac{\|b\|}{\|x\|}\right).$$

  - Since $b = Ax$, we have $\|b\| \leq \|A\| \cdot \|x\|$ and

  $$\frac{\|\Delta x\|}{\|x\|} \leq \|A^{-1}\| \left(\frac{\|\Delta b\|}{\|b\|}\right)\left(\frac{\|A\| \cdot \|x\|}{\|x\|}\right) = \|A^{-1}\| \cdot \|A\| \left(\frac{\|\Delta b\|}{\|b\|}\right).$$

The Condition Number

- Previous result shows:
  $\frac{\|\Delta x\|}{\|x\|} \leq \|A^{-1}\| \cdot \|A\| \left(\frac{\|\Delta b\|}{\|b\|}\right)$. //The relative error of x is bounded by a multiple of the relative error of b .
- The following term is called the **condition number** of this system
  $cond(A) = \|A^{-1}\| \cdot \|A\|.$

- The condition number is used to measure the sensitivity of the linear system:
  - If a small perturbation is added into the right hand side, how much does the solution be affected?

  $\frac{\|\Delta x\|}{\|x\|} \leq cond(A) \left(\frac{\|\Delta b\|}{\|b\|}\right).$
- Large condition number → unstable system → loss of significant digits.
- Small condition number → stable system → x can be solved more accurately. (Or the dimension n can be larger under the same error criterion.)

Perturbation of A

- Assume $\boldsymbol{A}$ is perturbed $\boldsymbol{A} \rightarrow \boldsymbol{A} + \boldsymbol{\Delta A}$. We have
  $(A + \Delta A)(x + \Delta x) = b.$
- How much does the solution be influenced?
  - Let $\tilde{A} = (A + \Delta A)$ and $\tilde{x} = (x + \Delta x).$

  $\tilde{A}\tilde{x} = \mathrm{b},$

  $$\tilde{x} = \tilde{A}^{-1}b = \tilde{A}^{-1}(Ax) = \tilde{A}^{-1}\left(A + \tilde{A} - \tilde{A}\right)x$$
  $$= x + \tilde{A}^{-1}\left(A - \tilde{A}\right)x = x - \tilde{A}^{-1}(\Delta A)x\ .$$
- Move x to the left side
  $$\tilde{x} - x = \Delta x = -\tilde{A}^{-1}(\Delta A)x\ .$$
- Take norm on both sides and divide the norms by $\|x\|$
  $$\frac{\|\Delta x\|}{\|x\|} = \left\|\tilde{A}^{-1}(\Delta A)\right\| \leq \left\|\tilde{A}^{-1}\right\| \cdot \left\|\tilde{A}\right\| \frac{\|\Delta A\|}{\|\tilde{A}\|} = cond(\tilde{A})\left(\frac{\|\Delta A\|}{\|\tilde{A}\|}\right).$$

Perturbation

The other relation:

$$\frac{\|\Delta x\|}{\|x\|} \leq \frac{cond(A)}{1-cond(A)\frac{\|\Delta A\|}{\|A\|}}\left(\frac{\|\Delta A\|}{\|A\|}\right).$$

Consider the *scaling factor*:

$$S = \frac{cond(A)}{1-cond(A)\frac{\|\Delta A\|}{\|A\|}}:$$

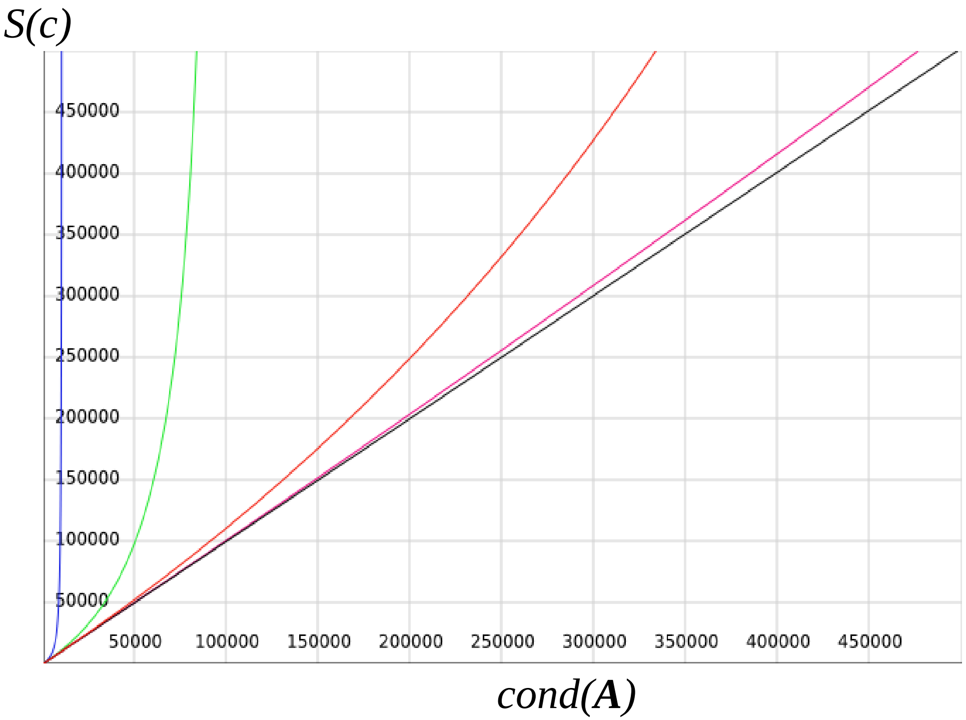
Assume the perturbation is small,

$$S(c) = \frac{c}{1-\epsilon c}.$$

Or

$$\frac{\|\Delta x\|}{\|x+\Delta x\|} \leq cond(A)\left(\frac{\|\Delta A\|}{\|A\|}\right).$$

black: $\epsilon = 10^{-8}$, purple: $\epsilon = 10^{-7}$, red: $\epsilon = 10^{-6}$, green: $\epsilon = 10^{-5}$, blue: $\epsilon = 10^{-4}$.



S(c) and the condition number cond(A)

Condition Number, Hilbert Matrix

- Hilbert matrix

$$a_{ij} = \frac{1}{i+j+1}, 0 \leq i, j \leq n-1.$$

- Condition numbers of Hilbert matrices

n = 2, cond($\boldsymbol{A}$) = 19.282

n = 4, cond($\boldsymbol{A}$) = 1551.4

n = 8, cond($\boldsymbol{A}$) = 1.5258e+10

n = 16, cond($\boldsymbol{A}$) = 6.5341e+17

$$A = \begin{bmatrix} 1 & 1/2 & 1/3 \\ 1/2 & 1/3 & 1/4 \\ 1/3 & 1/4 & 1/5 \end{bmatrix}$$

Estimation of cond($\boldsymbol{A}$)

- The condition number is defined as
  $cond(A) = \|A^{-1}\| \cdot \|A\|$
- The norm of $\boldsymbol{A}$ can be computed. (Using the shortcut methods).
- How to compute $\|A^{-1}\|$?
  $Ax = b, x = A^{-1}b$
  $\|x\| \leq \|A^{-1}\| \cdot \|b\|$. Thus $\|A^{-1}\| \geq \frac{\|x\|}{\|b\|}$.
  - Random generate several rhs b_i and solve the system
    $Ax_i = b_i$, i=1,2,…,k.
    - LU-decomposition is helpful.
  - Compute $\max_i \left(\frac{\|x_i\|}{\|b_i\|}\right)$;

  $cond(A) \approx \|A\| \cdot [\max_i \left(\frac{\|x_i\|}{\|b_i\|}\right)]$;