Mathematics Methods for Computer Science

Instructor: Xubo Yang

SJTU-SE DALAB

Lecture

Norms, Sensitivity and Conditioning

Questions

Gaussian elimination works in theory, but what about floating point precision?

How much can we trust $\vec{x}_0$ if

$$0 < \|A\vec{x}_0 - \vec{b}\| \ll 1$$

Recall: Backward Error

反向误差要用于正向误差无法直接计算的时候，实际上可以理解为通过结果之外的其他已知数据来模拟计算误差

Backward Error

The amount a problem statement would have to change to realize a given approximation of its solution

对于根号而言，若取x=2，得到实际结果为1.4，则正向误差为1.41.-1.4
反向误差为2-1.4*1.4=0.04

Example 1: $\sqrt{x}$

Example 2: $A\vec{x} = \vec{b}$

Perturbation Analysis

(扰动)

How does $\vec{x}$ change if we solve

$$(A + \delta A)\vec{x} = \vec{b} + \delta\vec{b}?$$

德塔A和德塔b描述的是A和b受到扰动时产生的误差。

Two viewpoints:

- Thanks to floating point precision, A and $\vec{b}$ are approximate
- If $\vec{x}_0$ isn't the exact solution, what is the backward error?

What is "Small"?

What does it mean for a statement to hold for small $\delta\vec{x}$?

Vector Norm

A function $\|\cdot\| : \mathbb{R}^n \to [0, \infty)$ satisfying:

1. $\|\vec{x}\| = 0 \quad$ iff $\vec{x} = \vec{0}$
2. $\|c\vec{x}\| = |c|\|\vec{x}\| \quad \forall c \in R, \vec{x} \in \mathbb{R}^n$
3. $\|\vec{x} + \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\| \quad \forall \vec{x}, \vec{y} \in \mathbb{R}^n$

三角形法则

p-norm, ∞-norm

p-norm $||\vec{x}||_p \equiv (||x_1||^p + ||x_2||^p + \cdots + ||x_n||^p)^{1/p}, \quad p \geq 1$

2-norm $||\vec{x}||_2 \equiv \sqrt{||x_1||^2 + ||x_2||^2 + \cdots + ||x_n||^2}$ （正常意义上的模）

1-norm $||\vec{x}||_1 \equiv \sum_{k=1}^{n} ||x_k||$ (aka. Manhattan/taxicab norm)

∞-norm $||\vec{x}||_\infty \equiv max(||x_1||, ||x_2||, \cdots, ||x_n||)$

How are Norms Different?

1. 此图中每一个与坐标轴的交点都为(0,1)或(1,0)
2. 对于同一个向量，随着p的增大，其对应的p范数变小

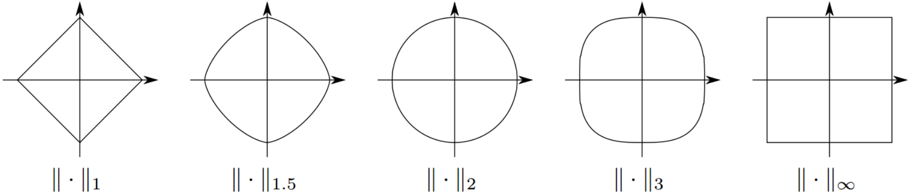

Figure 4.7 The set $\{\vec{x} \in \mathbb{R}^2 : \|\vec{x}\| = 1\}$ for different vector norms $\|\cdot\|$.

How are Norms the Same?

(范数等价性原理)

Equivalent norms

Two norms $\|\cdot\|$ and $\|\cdot\|'$ are equivalent if there exist constants c_{low} and c_{high} such that $c_{low}\|\vec{x}\| \leq \|x\|' \leq c_{high}\|\vec{x}\|$ for all $\vec{x} \in \mathbb{R}^n$.

How are Norms the Same?

Equivalent norms

Two norms $\|\cdot\|$ and $\|\cdot\|'$ are equivalent if there exist constants c_{low} and c_{high} such that $c_{low}\|\vec{x}\| \leq \|x\|' \leq c_{high}\|\vec{x}\|$ for all $\vec{x} \in \mathbb{R}^n$.

Theorem

All norms on $\mathbb{R}^n$ are equivalent.

(实数域上面的所有范数都是相互等价的)

How are Norms the Same?

Equivalent norms

Two norms $||\cdot||$ and $||\cdot||'$ are equivalent if there exist constants c_{low} and c_{high} such that $c_{low}||\vec{x}|| \leq ||x||' \leq c_{high}||\vec{x}||$ for all $\vec{x} \in \mathbb{R}^n$.

Theorem

All norms on $\mathbb{R}^n$ are equivalent.

$$(10000, 1000, 1000) \ vs. \ (10000, 0, 0)?$$

Matrix Norms: "Unrolled" Construction

Convert to vector, and use vector p-norm:

思路1：
将m*n的矩阵展开为mn*1的向量

$$A \in \mathbb{R}^{m\times n} \leftrightarrow \mathrm{a}[:] \in \mathbb{R}^{mn}$$

Special Case: Frobenius norm $(p = 2)$:

$$\|A\|_{\mathrm{Fro}} \equiv \sqrt{\sum_{ij} a_{ij}^2}$$

Matrix Norms: "Induced" Construction

Maximum stretching of a unit vector by A:

$$\|A\| \equiv \max\{\|A\vec{x}\| : \|\vec{x}\| = 1\}$$

这个等式的意思是说：在x的模恒为1的条件下，找到一个矩阵A，使得Ax的模达到最大值。
而结合等式||A||=||AX||/||x||可以得到，||AX||最大时，||A||实际上也就达到了最大值。所以这个等式对于p=2就可以理解为：对于一个单位圆，施加一个变换矩阵A，求得变化之后的图形的最大宽度就是矩阵A的模。

Matrix Norms: "Induced" Construction

Maximum stretching of a unit vector by A:

$$\|A\| \equiv \max\{\|A\vec{x}\| : \|\vec{x}\| = 1\}$$

Different matrix norms induced by different vector p-norms.

Case $p = 2$: What is the norm induced by $\| \cdot \|_2$?

Matrix Norms: $\|A\|_2$ Visualization

对标注的式子两边同时乘以x^T,在取模可以得到：
λ ^2=||A||^2(因为||x||=1)

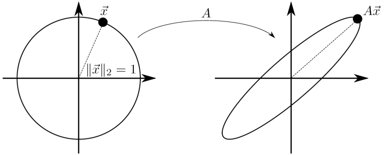

Figure 4.8 The norm $\|\cdot\|_2$ induces a matrix norm measuring the largest distortion of any point on the unit circle after applying A.

Induced two-norm, or spectral norm, of $A \in \mathbb{R}^{n\times n}$ is the square root of the largest eigenvalue of A^TA:

$$\|A\|_2^2 = \max\{\lambda : \text{There exists } \in \mathbb{R}^n \text{ with } A^TA\vec{x} = \lambda\vec{x}\}$$

找到Gram矩阵对应的最大特征值λ

Other Induced Norms

$$\|A\|_1 \equiv \max_j \sum_i |a_{ij}|$$

查找列和最大时对应的列

$$\|A\|_\infty \equiv \max_i \sum_j |a_{ij}|$$

查找行和最大时对应的行

Recall: Condition Number

Condition number

Ratio of forward to backward error

Root-finding example:

$$\frac{1}{f'(x^*)}$$

Model Problem

$$(A + \varepsilon \delta A)\vec{x}(\varepsilon) = \vec{b} + \varepsilon \delta \vec{b}$$

Simplification

取 ε =0时有Ax=b

$$\left.\frac{d\vec{x}}{d\varepsilon}\right|_{\varepsilon=0} = A^{-1}(\delta\vec{b} - \delta A\vec{x}(0))$$

$$\frac{\|\vec{x}(\varepsilon) - \vec{x}(0)\|}{\|\vec{x}(0)\|} \leq |\varepsilon| \, \|A^{-1}\| \, \|A\| \left(\frac{\|\delta\vec{b}\|}{\|\vec{b}\|} + \frac{\|\delta A\|}{\|A\|} \right) + O\left(\varepsilon^2\right)$$

前向相对误差　　　　后项相对误差

大致推导过程为利用泰勒展开并代入Ax=b以及对于 ε 在0处的导数进行化简，并利用||A||*||B|| >= ||A*B||进行放缩(这个不等式可以理解为：p=2时，为柯西-施瓦茨不等式的矩阵形式)

Matrix Condition Number

Condition Number

The condition number of $A \in \mathbb{R}^{n\times n}$ for a given matrix norm $\|\cdot\|$ is $cond\ A \equiv \kappa \equiv \|A\|\|A^{-1}\|$.
If A is not invertible, $cond\ A \equiv \infty$

若为无穷大，则反向误差与正向误差之间就没有推断关系，从而为poorly-conditioned，这与前面讲过的结论也符合。

Matrix Condition Number

Condition Number

The condition number of $A \in \mathbb{R}^{n \times n}$ for a given matrix norm $\| \cdot \|$ is $cond\ A \equiv \kappa \equiv \|A\|\|A^{-1}\|$.
If A is not invertible, $cond\ A \equiv \infty$

$$\text{Relative change: } D \equiv \frac{\|\delta\vec{b}\|}{\|\vec{b}\|} + \frac{\|\delta A\|}{\|A\|}$$

$$\frac{\|\vec{x}(\varepsilon) - \vec{x}(0)\|}{\|\vec{x}(0)\|} \leq \varepsilon \cdot D \cdot \kappa + O\left(\varepsilon^2\right)$$

Matrix Condition Number

Condition Number

The condition number of $A \in \mathbb{R}^{n \times n}$ for a given matrix norm $\|\cdot\|$ is $cond\ A \equiv \kappa \equiv \|A\|\|A^{-1}\|$.
If A is not invertible, $cond\ A \equiv \infty$

Relative change: $D \equiv \frac{\|\delta\vec{b}\|}{\|\vec{b}\|} + \frac{\|\delta A\|}{\|A\|}$

$$\frac{\|\vec{x}(\varepsilon) - \vec{x}(0)\|}{\|\vec{x}(0)\|} \leq \varepsilon \cdot D \cdot \kappa + O\left(\varepsilon^2\right)$$

Invariant to scaling (unlike determinant!); equals one for the identity.

Condition Number of Induced Norm

$$\operatorname{cond} A = \frac{\max_{\vec{x} \neq 0} \frac{\|A\vec{x}\|}{\|\vec{x}\|}}{\min_{\vec{y} \neq \vec{0}} \frac{\|A\vec{y}\|}{\|\vec{y}\|}} = \frac{\max_{\|\vec{x}\|=1} \|A\vec{x}\|}{\min_{\|\vec{y}\|=1} \|A\vec{y}\|}$$

Condition Number: Visualization

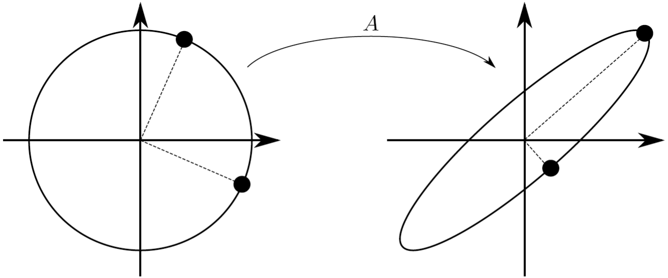

Figure 4.9 The condition number of A measures the ratio of the largest to smallest distortion of any two points on the unit circle mapped under A.

Experiments with an ill-conditioned Vandermonde matrix

Chicken $\Leftrightarrow$ Egg

$$\operatorname{cond} A \equiv \|A\|\,\|A^{-1}\|$$

Computing $\|A^{-1}\|$ typically requires solving $A\vec{x} = \vec{b}$, but how do we know the reliability of $\vec{x}$?

To Avoid...

What is the condition number of computing the condition number of A?

To Avoid...

What is the condition number of computing the condition number of A?

What is the condition number of computing what the condition number is of computing the condition number of A?

因为计算A的时候本身就会引入误差，而在A^-1的时候也会引入新的误差，从而导致了计算条件数的问题变成了一个死循环。。。因为每一步计算若是不精确的，我们都希望用条件数来衡量精确程度。

Instead

Bound the condition number.

- Below: Problem is at least this hard
- Above: Problem is at most this hard

Potential for Approximation

$$\|A^{-1}\vec{x}\| \leq \|A^{-1}\| \, \|\vec{x}\|$$

$$\Downarrow$$

$$\text{cond } A = \|A\| \, \|A^{-1}\| \geq \frac{\|A\| \, \|A^{-1}\vec{x}\|}{\|\vec{x}\|}$$