STAT 309: MATHEMATICAL COMPUTATIONS I
FALL 2015
LECTURE 2

- norms are of great importance in numerical computations because they allow us to measure the size of errors
- that is, how far are we from the true solution that we are seeking
- this is important because we do everything in numerical computations in the presence of rounding errors but we still want to guarantee that the solution we found is accurate to some degree

1. more vector norms

- example: variation of the p-norm is the *weighted p-norm*, defined by
$$\|\mathbf{x}\|_{p,\mathbf{w}} = \left(\sum_{i=1}^{n} w_i |x_i|^p\right)^{1/p}$$
  - again it can be shown that this is a norm as long as the *weights* w_i, $i = 1, \dots, n$, are strictly positive real numbers
- example: a vast generalization of all of the above is the A-*norm* or *Mahalanobis norm*, defined in terms of a matrix A by
$$\|\mathbf{x}\|_A = (\mathbf{x}^* A\mathbf{x})^{1/2} = \left(\sum_{i,j=1}^{n} a_{ij}\overline{x}_i x_j\right)^{1/2}$$
  - this defines a norm provided that the matrix A is positive definite
  - note that if $W = \operatorname{diag}(\mathbf{w})$, then
$$\|\mathbf{x}\|_W = \|\mathbf{x}\|_{2,\mathbf{w}}$$
- we now highlight some additional, and useful, relationships for a norm
  - first of all, the triangle inequality generalizes directly to sums of more than two vectors:
$$\|\mathbf{x} + \mathbf{y} + \mathbf{z}\| \le \|\mathbf{x} + \mathbf{y}\| + \|\mathbf{z}\| \le \|\mathbf{x}\| + \|\mathbf{y}\| + \|\mathbf{z}\|$$
  - more generally,
$$\left\|\sum_{i=1}^{m} \mathbf{x}_i\right\| \le \sum_{i=1}^{m} \|\mathbf{x}_i\|$$
  - secondly, what can we say about the norm of the difference of two vectors? we know that $\|\mathbf{x} - \mathbf{y}\| \le \|\mathbf{x}\| + \|\mathbf{y}\|$ but we can obtain a more useful relationship as follows:
$$\|\mathbf{x}\| = \|(\mathbf{x} - \mathbf{y}) + \mathbf{y}\| \le \|\mathbf{x} - \mathbf{y}\| + \|\mathbf{y}\|$$
    we obtain
$$\|\mathbf{x} - \mathbf{y}\| \ge \|\mathbf{x}\| - \|\mathbf{y}\|$$

Date: October 8, 2015, version 1.0. Comments, bug reports: `lekheng@galton.uchicago.edu`.

- thirdly, from

$$\|\mathbf{y}\| = \|\mathbf{y} - \mathbf{x} + \mathbf{x}\| \leq \|\mathbf{x} - \mathbf{y}\| + \|\mathbf{x}\|$$

it follows that

$$\|\mathbf{x} - \mathbf{y}\| \geq \|\mathbf{y}\| - \|\mathbf{x}\|$$

and therefore

$$|\|\mathbf{x}\| - \|\mathbf{y}\|| \leq \|\mathbf{x} - \mathbf{y}\| \tag{1.1}$$

- the inequality (1.1) yields a very important property of norms, namely, they are all (uniformly) continuous functions of the entries of their arguments

- there are also interesting relationships for two different norms
- first and foremost, on finite dimensional spaces (which includes $\mathbb{C}^n$ and $\mathbb{C}^{m\times n}$) all norms are *equivalent*
  - that is, given two norms $\|\cdot\|_\alpha$ and $\|\cdot\|_\beta$, there exist constants c_1 and c_2 with $0 < c_1 < c_2$ such that

$$c_1\|\mathbf{x}\|_\alpha \leq \|\mathbf{x}\|_\beta \leq c_2\|\mathbf{x}\|_\alpha \tag{1.2}$$

  for all $\mathbf{x} \in V$
  - example: from the definition of the ∞-norm, we have

$$\|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|_2 \leq \sqrt{n}\|\mathbf{x}\|_\infty$$

  - example: also not hard to show that

$$\frac{1}{n}\|\mathbf{x}\|_1 \leq \|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|_1$$

  - in fact, no matter what crazy choices of norms that we make, say

$$\|x\|_\alpha = \left(\sum_{i=1}^{n} i|x_i|^n\right)^{1/n}, \qquad \|x\|_\beta = \mathbf{x}^\mathsf{T}\begin{bmatrix} 3 & -1 & & \\ -1 & 3 & \ddots & \\ & \ddots & \ddots & -1 \\ & & -1 & 3 \end{bmatrix}\mathbf{x},$$

  we know that there are c_1 and c_2 so that (1.2) holds

- by definition, a sequence of vectors $\mathbf{x}_0, \mathbf{x}_1, \ldots$ converges to a vector $\mathbf{x}$ if and only if

$$\lim_{k\to\infty} \|\mathbf{x}_k - \mathbf{x}\| = 0$$

  for any norm (you may also write down a formal version in terms of ε and N)
- the equivalence of norms on finite dimensional vector spaces tells us that

$$\lim_{k\to\infty} \|\mathbf{x}_k - \mathbf{x}\|_\alpha = 0 \quad \text{if and only if} \quad \lim_{k\to\infty} \|\mathbf{x}_k - \mathbf{x}\|_\beta = 0$$

  for any choice of norms $\|\cdot\|_\alpha$ and $\|\cdot\|_\beta$ (why?)
- if we can establish convergence of an algorithm in a specific norm convergence in every other norm follows automatically
- for this reason, norms are very useful to measure the error in an approximation
- secondly we have a relationship that applies to products of norms, the *Hölder inequality*

$$|\mathbf{x}^*\mathbf{y}| \leq \|\mathbf{x}\|_p\|\mathbf{y}\|_q, \quad \frac{1}{p} + \frac{1}{q} = 1$$

  - a well-known corollary arises when $p = q = 2$, the *Cauchy–Schwarz inequality*

$$|\mathbf{x}^*\mathbf{y}| \leq \|\mathbf{x}\|_2\|\mathbf{y}\|_2$$

- you will see a generalization of Cauchy–Schwarz inequality in the last problem of Homework 0, this is called the *Bessel inequality*
- by setting $\mathbf{x} = [1, 1, \dots, 1]^{\mathsf{T}}$, the Hölder inequality yields the relationships

$$\left|\sum_{i=1}^{n} y_i\right| \le \sum_{i=1}^{n} |y_i|$$

and

$$\left|\sum_{i=1}^{n} y_i\right| \le n \max_{i=1,\dots,n} |y_i|$$

and

$$\left|\sum_{i=1}^{n} y_i\right| \le \sqrt{n} \left(\sum_{i=1}^{n} |y_i|^2\right)^{1/2}$$

2. matrix norms

- note that the space of complex $m \times n$ matrices $\mathbb{C}^{m\times n}$ is a vector space over $\mathbb{C}$ (ditto for real matrices over $\mathbb{R}$) of dimension mn
- we write O for the $m \times n$ zero matrix, i.e., all entries are 0
- a norm on either $\mathbb{C}^{m\times n}$ or $\mathbb{R}^{m\times n}$ is called a *matrix norm*
- recall that these means $\|\cdot\| : \mathbb{C}^{m\times n} \to \mathbb{R}$ satisfies
  (1) $\|A\| \ge 0$ for all $A \in \mathbb{C}^{m\times n}$
  (2) $\|A\| = 0$ if and only if $A = O$
  (3) $\|\alpha A\| = |\alpha| \|A\|$
  (4) $\|A + B\| \le \|A\| + \|B\|$
- often we add a fifth condition that $\|\cdot\|$ satisfies the *submultiplicative property*

$$\|AB\| \le \|A\|\|B\|$$

- submultiplicative is some also called *consistent*

3. Hölder norms

- example: *Frobenius norm*

$$\|A\|_F = \left(\sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}|^2\right)^{1/2}$$

which is submultiplicative since

$$\|AB\|_F^2 = \sum_{i=1}^{m}\sum_{k=1}^{p} \left|\sum_{j=1}^{n} a_{ij} b_{jk}\right|^2 \le \sum_{i=1}^{m}\sum_{k=1}^{p} \left[\left(\sum_{j=1}^{n} |a_{ij}|^2\right)\left(\sum_{j=1}^{n} |b_{jk}|^2\right)\right]$$

by the Cauchy–Schwarz inequality and the last expression is equal to

$$\left(\sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}|^2\right)\left(\sum_{k=1}^{p}\sum_{j=1}^{n} |b_{jk}|^2\right) = \|A\|_F^2 \|B\|_F^2$$

- example: more generally we have Hölder p-norm for any $p \in [1, \infty]$,

$$\|A\|_{H,p} = \left(\sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}|^p\right)^{1/p}$$

and

$$\|A\|_{H,\infty} = \max_{i,j}|a_{ij}|$$

- Hölder norms are obtained by viewing an $m \times n$ matrix $A = [a_{ij}]_{i,j=1}^{m,n} \in \mathbb{C}^{m\times n}$ as a vector $\boldsymbol{\alpha} = [a_{11}, a_{12}, \dots, a_{mn}]^{\mathsf{T}} \in \mathbb{C}^{mn}$ with mn entries, this is often written as $\boldsymbol{\alpha} = \operatorname{vec}(A)$
- we have $\|A\|_{H,p} = \|\operatorname{vec}(A)\|_p$
- clearly $\|A\|_{H,2} = \|A\|_F = \|\operatorname{vec}(A)\|_2$

- in general Hölder p-norms are not submultiplicative for $p \neq 2$
  - example: take $A = \left[\begin{smallmatrix}1&1\\0&0\end{smallmatrix}\right], B = \left[\begin{smallmatrix}1&0\\1&0\end{smallmatrix}\right]$, then $AB = \left[\begin{smallmatrix}2&0\\0&0\end{smallmatrix}\right]$ but
  $$\|AB\|_{H,\infty} = 2 > 1 = \|A\|_{H,\infty}\|B\|_{H,\infty}$$
- note that the submultiplicative property implies that
$$\|A^n\| \le \|A\|^n$$
- so if $\|A\| < 1$, then, as $n \to \infty$, $\|A^n\| \to 0$
- if $\|A^n\| \to 0$, then $A^n \to O$, i.e. each entry of A^n converges to 0, by the continuity of norms
- the condition $\|A\| < 1$ is not necessary for $A^n \to 0$
  - example:
  $$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$
  has $\|A\|_1 = 1$ but $A^2 = A^3 = \cdots = O$
  - the above example is a *nilpotent* matrix, i.e., has the property that $A^n = O$ for some finite $n \in \mathbb{N}$
  - another example:
  $$A = \begin{bmatrix} 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 \end{bmatrix}$$
  has $\|A\|_\infty = 1$ but $A^n \to O$

4. Induced norms

- a very important class of matrix norms are the so called *operator* or *induced* or *natural norms* defined as
$$\|A\|_{\alpha,\beta} := \max_{\mathbf{x}\neq\mathbf{0}} \frac{\|A\mathbf{x}\|_\alpha}{\|\mathbf{x}\|_\beta} \tag{4.1}$$
for any $A \in \mathbb{C}^{m\times n}$ and any vector norms $\|\cdot\|_\alpha : \mathbb{C}^m \to \mathbb{R}$ and $\|\cdot\|_\beta : \mathbb{C}^n \to \mathbb{R}$ defined on the domain and codomain of A respectively
- the induced norm may also be written as
$$\|A\|_{\alpha,\beta} = \max\{\|A\mathbf{x}\|_\alpha : \|\mathbf{x}\|_\beta \le 1\} \tag{4.2}$$
or as
$$\|A\|_{\alpha,\beta} = \max\{\|A\mathbf{x}\|_\alpha : \|\mathbf{x}\|_\beta = 1\} \tag{4.3}$$
- in other words, the induced norm measures how far the operator A sends points in the unit disc (or the unit circle)
- proof is simple, for example, here's how you would prove (4.3):
$$\max_{\mathbf{x}\neq\mathbf{0}} \frac{\|A\mathbf{x}\|_\alpha}{\|\mathbf{x}\|_\beta} = \max_{\mathbf{x}\neq\mathbf{0}} \left\| \frac{1}{\|\mathbf{x}\|_\beta} A\mathbf{x} \right\|_\alpha = \max_{\mathbf{x}\neq\mathbf{0}} \left\| A\left(\frac{\mathbf{x}}{\|\mathbf{x}\|_\beta}\right)\right\|_\alpha = \max_{\|\mathbf{v}\|_\beta=1} \|A\mathbf{v}\|_\alpha,$$
the first equality uses the property that $|\alpha|\|\mathbf{v}\|_\alpha = \|\alpha\mathbf{v}\|_\alpha$, the second equality uses $\alpha A\mathbf{x} = A(\alpha\mathbf{x})$, and the last equality uses the observation that $\mathbf{v} = \mathbf{x}/\|\mathbf{x}\|_\beta$ always has unit β-norm

- exercise: prove (4.3) and (4.2) are equal
- another exercise: prove that

$$\|A\mathbf{x}\|_\alpha \le \|A\|_{\alpha,\beta}\|\mathbf{x}\|_\beta \tag{4.4}$$

  for any $\mathbf{x} \in \mathbb{C}^n$
- a note on the use of *supremum* and *maximum*: for $S \subseteq \mathbb{C}^n$ and a real-valued function f whose domain includes S,
  - we write $\sup_{\mathbf{x}\in S} f(\mathbf{x})$ for the smallest $\mu \in \mathbb{R}$ such that $f(\mathbf{x}) \le \mu$ for every $\mathbf{x} \in S$ (and we set $\mu = +\infty$ if f is unbounded on S)
  - we write $\max_{\mathbf{x}\in S} f(\mathbf{x})$ if the supremum is attained by some element in S, i.e., there is an $\mathbf{x}_{\max} \in S$ such that $f(\mathbf{x}_{\max}) = \sup_{\mathbf{x}\in S} f(\mathbf{x})$
  - $\mathbf{x}_{\max}$ is called a *maximizer* of f on S
  - likewise for infimum and minimum (and minimizer)
  - by the extreme value theorem, if f is continuous and S is compact, then supremum and infimum are always attained
- in the above $S = \{\mathbf{x} \in \mathbb{C} : \|\mathbf{x}\|_\beta \le 1\}$ and $S = \{\mathbf{x} \in \mathbb{C} : \|\mathbf{x}\|_\beta = 1\}$ are compact and the function $f = \|\cdot\|_\alpha : \mathbb{C}^m \to \mathbb{R}$ is continuous
- in other words, we can always find an $\mathbf{x}_{\max}$ with $\|\mathbf{x}_{\max}\|_\beta = 1$ such that

$$\|A\mathbf{x}_{\max}\|_\alpha = \|A\|_{\alpha,\beta}$$

- that's why we may always write max in (4.3) and (4.2), and therefore in (4.1); although strictly speaking we should have written (4.1)

$$\|A\|_{\alpha,\beta} := \sup_{\mathbf{x}\neq\mathbf{0}} \frac{\|A\mathbf{x}\|_\alpha}{\|\mathbf{x}\|_\beta}$$

- the induced norm is *not* submultiplicative in general: take

$$A = \begin{bmatrix} 2 & 2 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$$

  since every $\mathbf{x} \in \mathbb{R}^2$ with $\|\mathbf{x}\| = 1$ has the form $\mathbf{x} = (\cos\theta, \sin\theta)^{\mathsf{T}}$, we see that

$$\begin{aligned} \|A\|_{\infty,2} &= \max_{\|\mathbf{x}\|_2=1} \|A\mathbf{x}\|_\infty = \max_\theta |2\cos\theta + 2\sin\theta| = 2\sqrt{2} \\ \|B\|_{\infty,2} &= \max_{\|\mathbf{x}\|_2=1} \|B\mathbf{x}\|_\infty = \max_\theta |\cos\theta| = 1 \\ \|AB\|_{\infty,2} &= \max_{\|\mathbf{x}\|_2=1} \|AB\mathbf{x}\|_\infty = \max_\theta |4\cos\theta| = 4 \end{aligned}$$

  but

$$\|AB\|_{\infty,2} = 4 > 2\sqrt{2} = \|A\|_{\infty,2}\|B\|_{\infty,2}$$

  (thanks to Lijun Ding for this example)
- however given $A \in \mathbb{C}^{m\times n}$ and $B \in \mathbb{C}^{n\times p}$ it is always true that

$$\|AB\|_{\alpha,\gamma} \le \|A\|_{\alpha,\beta}\|B\|_{\beta,\gamma}$$

  for any norms $\|\cdot\|_\gamma$ on $\mathbb{C}^p$, $\|\cdot\|_\beta$ on $\mathbb{C}^n$, $\|\cdot\|_\alpha$ on $\mathbb{C}^m$
- the most interesting induced norms are the ones obtained when $\|\cdot\|_\alpha$ and $\|\cdot\|_\beta$ are vector ℓ^p-norms, we write

$$\|A\|_{p,q} := \max_{\mathbf{x}\neq\mathbf{0}} \frac{\|A\mathbf{x}\|_p}{\|\mathbf{x}\|_q} \quad \text{and} \quad \|A\|_p := \max_{\mathbf{x}\neq\mathbf{0}} \frac{\|A\mathbf{x}\|_p}{\|\mathbf{x}\|_p}$$

  for any $A \in \mathbb{C}^{m\times n}$ and $p, q \in [1, \infty]$
- we call $\|\cdot\|_{p,q}$ the matrix (p, q)-norm and $\|\cdot\|_p$ the matrix p-norm

- the matrix 2-norm

$$\|A\|_2 = \max_{\mathbf{x}\neq\mathbf{0}} \frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2}$$

is very widely used and has its own special name, *spectral norm*, because of its relation to the spectrum of a matrix (i.e., the eigenvalues); we will discuss it in the next two lectures
- the matrix 1-norm and ∞-norm are also very widely used, largely because, they can be easily computed
- let $A = [a_{ij}]_{i,j=1}^{m,n} \in \mathbb{C}^{m\times n}$, then

$$\|A\|_1 = \max_{j=1,\dots,n} \left[\sum_{i=1}^{m} |a_{ij}|\right] \tag{4.5}$$

and

$$\|A\|_\infty = \max_{i=1,\dots,m} \left[\sum_{j=1}^{n} |a_{ij}|\right] \tag{4.6}$$

- an easy way to remember these is that $\|A\|_1$ is the maximum column sum and $\|A\|_\infty$ is the maximum row sum of A
- let us prove (4.6) and leave (4.5) as an exercise:
  - we use (4.3), so

$$\begin{aligned}\|A\|_\infty &= \max\{\|A\mathbf{x}\|_\infty : \|\mathbf{x}\|_\infty = 1\}\\ &= \max_{\|\mathbf{x}\|_\infty=1} \left\{ \max_{i=1,\dots,m} \left| \sum_{j=1}^{n} a_{ij}x_j \right| \right\}\\ &\le \max_{\|\mathbf{x}\|_\infty=1} \left\{ \max_{i=1,\dots,m} \left[\sum_{j=1}^{n} |a_{ij}||x_j| \right] \right\}\\ &\le \max_{i=1,\dots,m} \left[\sum_{j=1}^{n} |a_{ij}|\right] \qquad (4.7)\end{aligned}$$

    where the last inequality follows because $\|\mathbf{x}\|_\infty = 1$ and so we must have $|x_j| \le 1$
  - to show equality, we just need to exhibit one single $\mathbf{x}^*$ with $\|\mathbf{x}^*\|_\infty = 1$ so that

$$\|A\mathbf{x}^*\|_\infty \ge \max_{i=1,\dots,m} \left[\sum_{j=1}^{n} |a_{ij}|\right]$$

  - we know that the maximum in (4.7) is attained by some row $i = k \in \{1,\dots,m\}$, so

$$\max_{i=1,\dots,m} \left[\sum_{j=1}^{n} |a_{ij}|\right] = \sum_{j=1}^{n} |a_{kj}|$$

  - now we define $\mathbf{x}^* = [x_1^*,\dots,x_n^*] \in \mathbb{C}^n$ as the vector whose coordinates are given by

$$x_j^* = \begin{cases} |a_{kj}|/a_{kj} & \text{if } a_{kj} \neq 0,\\ 0 & \text{if } a_{kj} = 0,\end{cases}$$

    for $j = 1,\dots,n$

– observe that $\mathbf{x}^*$ has $\|\mathbf{x}^*\|_\infty = 1$ as well as the effect of attaining the requisite bound

$$\|A\mathbf{x}^*\|_\infty = \max_{i=1,\dots,m} \left| \sum_{j=1}^{n} a_{ij} x_j^* \right| \geq \sum_{j=1}^{n} |a_{kj}| = \max_{i=1,\dots,m} \left[\sum_{j=1}^{n} |a_{ij}| \right]$$