LECTURE 5

Xuxing cover   Olson Hall   2:10 – 3:00 pm

(no review problem today – no "in-class work")

Matrix Norm

Given any vector norm $\|\cdot\|$, we can define an associated <u>operator norm</u>

<u>note</u> "Operator": does an operation. Matrix operates on vectors, matrices

eg. "Row operation"

$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & b \\ a+c & b+d \end{bmatrix}$$

Covered on Monday

<u>defn.</u> $$\|A\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|}$$

"What's the most A stretches a vector?"

? Supremum — may need to review

<u>ex.</u> Check that matrix norm is really a norm (vector norm)

- $\|A\| \geq 0 \ \forall A$, $\|A\| = 0$ iff $A = 0$
- $\|\alpha A\| = |\alpha| \, \|A\|$, $\alpha \in \mathbb{R}$
- $\|A + B\| \leq \|A\| + \|B\|$ (triangle inequality)

<u>prop 2.1</u> (a) $\|Ax\| \leq \|A\| \, \|x\|$ and (b) $\|AB\| \leq \|A\| \, \|B\|$

Pf. (a) $$\|A\| := \sup_{y \neq 0} \frac{\|Ay\|}{\|y\|} \geq \frac{\|Ax\|}{\|x\|}$$

(b) $$\|ABx\| \leq \|A\| \, \|Bx\| \leq \|A\| \, \|B\| \, \|x\|$$

$$\Rightarrow \frac{\|ABx\|}{\|x\|} \leq \|A\| \, \|B\| \quad \text{for } \underline{\text{all}} \ x.$$

$$\Rightarrow \|AB\| := \sup_{x \neq 0} \frac{\|ABx\|}{\|x\|} \leq \|A\| \, \|B\|$$

$\blacksquare$

Notable Matrix Norms:

- $\|A\|_2 = \left(\max_{i \in [n]} \lambda_i(A^TA)\right)^{1/2}$

  = square root of the largest eigenvalue of A^TA

  = biggest "stretch" in any direction

  * why are all eigenvalues of A^TA nonnegative? $(\lambda_i \geq 0)$

- If $A \in \mathbb{R}^{m \times n}$ (m rows, n columns)
  - $\|A\|_1 = \max_{1 \leq j \leq n} \underbrace{\sum_{i=1}^{m} |a_{ij}|}_{\text{sum of abs. values of entries in column } j}$ * why?
  - $\|A\|_\infty = \max_{1 \leq i \leq m} \underbrace{\sum_{j=1}^{n} |a_{ij}|}_{\text{sum of abs values of entries in row } i}$ * why?

Another matrix norm: the Frobenius Norm ($A \in \mathbb{R}^{m \times n}$ again)

$$\|A\|_F = \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2}$$

ie, just the $\|\cdot\|_2$-norm if we view $A \in \mathbb{R}^{mn}$.

In fact, this is equivalent to

$$= \|A\|_F^2 = tr(A^TA).$$

* why?

Recall that the trace $tr(B)$ of a square matrix $B \in \mathbb{R}^{n \times n}$ is the sum of the diagonal entries:

$$tr(B) = \sum_{i=1}^{n} b_{ii}$$

* This doesn't come from a vector norm. (So we don't consider it an operator norm.)

Quick Review: Linear independence, Bases

Let $\{v_i\}_{i=1}^{k}$ be a set of vectors $v_i \in \mathbb{R}^m$

- $\text{span}(v_1, v_2, \ldots, v_k) = \{ y \mid y = \sum_{i=1}^{k} \alpha_i v_i \}$
  $= \{ \sum_{i=1}^{k} \alpha_i v_i \mid \alpha_i \in \mathbb{R} \text{ for all } i \}$

- The set of vectors $\{v_i\}_{i=1}^{k}$ is <u>linearly independent</u> if
  $\sum_{i=1}^{k} \alpha_i v_i = 0$ iff $\alpha_i = 0 \ \forall i$.

  In other words, there is no way to write v_j as a linear combo of the $\{v_i \mid i \neq j\}$.

- If $n > m$ (# vectors in the set > dim of vector space), then the set of vectors is necessarily linearly dependent.

  (by the pigeonhole principle)

- A basis for $\mathbb{R}^m$ is a set of m linearly independent vectors. $\{b_i\}_{i=1}^{m}$
  $\Rightarrow \mathbb{R}^m = \text{span}(\{b_i\}_{i=1}^{m})$,

  ie all vectors $y \in \mathbb{R}^n$ can be written as linear combos of the vectors $b_1, \ldots, b_m$.

eg. If you are given vectors as signals
(eg. samples of handwriting of the digit "5")

- The space of possible signals is very high-dimensional
- In your samples, the size of the largest set of "almost linearly independent vectors" tells you how much variety there is in the way people write "5".
  - "almost" because in practice, we usually are ok with not having "perfect" or "exact" linear dependence.