Computational Linear Algebra: Measuring Lengths and Distances with Vector and Matrix Norms

David Shuman

February 8, 2022

Vector Norms

Given a vector $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ \vec{x}_n \end{bmatrix} \in \mathbb{R}^n$ a vector norm is a rule that assigns a real number $||\vec{x}||$ to each vector $\vec{x} \in \mathbb{R}^n$ satisfying

(a) $||\vec{x}|| \geq 0$ for all $\vec{x}$ and $||\vec{x}|| = 0$ if and only if $\vec{x} = \vec{0}$

(b) $||\alpha\vec{x}|| = |\alpha|||\vec{x}||$

(c) $||\vec{x} + \vec{y}|| \leq ||\vec{x}|| + ||\vec{y}||$ (triangle inequality)

The norm is a measure of the *length of the vector*

Eg. 1. The Euclidean or ℓ_2-norm,

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

Eg. 2. The p-norm,

$$||\vec{x}||_p = (|x_1|^p + |x_2|^p + \cdots + |x_n|^p)^{1/p}$$

note the need for the absolute values when p is odd

Vector Norms (cont.)

The p-norm for $1 \leq p \leq \infty$

$$||\vec{x}||_p = (|x_1|^p + |x_2|^p + \cdots + |x_n|^p)^{1/p}$$

The three most common p-norms are $p = 1, 2, \infty$ since they are the easiest to compute with and in some sense are the most natural

- $p = 1$ (the Manhattan or taxicab norm)

$$||\vec{x}||_1 = |x_1| + |x_2| + \cdots + |x_n|$$

- $p = 2$ (the Euclidean norm)

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

- $p = \infty$

$$||\vec{x}||_\infty = \max(|x_1|, |x_2|, \ldots, |x_n|)$$

- Our R code for the p-norm is **vnorm**. Be sure you understand it

Vector Norms (cont.)

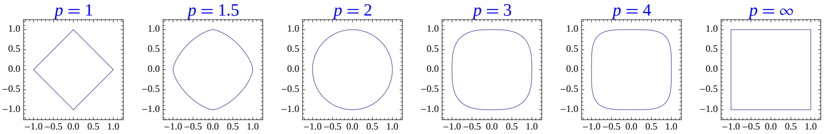
- Note that: $\lim_{p\to\infty} ||\vec{x}||_p = ||\vec{x}||_\infty$

  For example: consider $\vec{v} = (3, -2, 2, 3, 1, 4, 1, 2, 3)$. Then,

p	1	1.5	2	3	4	5	6	7	∞
$\lvert\lvert\vec{v}\rvert\rvert_p$	21	10.51	7.55	5.55	4.84	4.50	4.31	4.2	4

  Notice how it converges (quickly) to the ∞-norm

- Draw "unit circles" in each norm



Matrix Norms

A matrix norm assigns to each $n \times n$ matrix A a real number $||A||$ such that

(a) $||A|| \geq 0$ for all A and $||A|| = 0$ if and only if $A = 0$
(b) $||\alpha A|| = |\alpha| ||A|$
(c) $||A + B|| \leq ||A|| + ||B||$ (triangle inequality)

The goal is to assign a notion of size to the matrix

Def. The Frobenius Norm is a natural matrix norm that is analogous to the 2-norm of a vector and is given by:

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

Def. The Max Norm is another natural matrix norm that is analogous to the ∞-norm of a vector and is given by:

$$||A||_{max} = \max_{i,j}(|a_{ij}|)$$

Neither of these are very useful in describing errors in $Ax = b$

Matrix Norms (cont.)

Def. The Matrix p-norm is given by

$$||A||_p = \max_{\vec{x} \neq \vec{0}} \frac{||A\vec{x}||_p}{||\vec{x}||_p} = \max_{||\vec{x}||_p = 1} ||A\vec{x}||_p$$

Comments

- Notice that this norm is inherited[1] from the vector norm
- $||A||_p$ gives the maximum relative expansion by A
- Apply A to the unit sphere in $\mathbb{R}^n$. Then $||A||_p$ is the length of the vector that is farthest from the origin in the image
- Key property: $||A\vec{x}||_p \leq ||A||_p ||x||_p$
- Computing this is a hard optimization problem, in general!
- Reason we can maximize over the unit sphere:

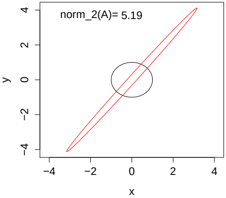
$$\frac{||A(\alpha\vec{x})||_p}{||(\alpha\vec{x})||_p} = \frac{|\alpha|\ ||A\vec{x}||_p}{|\alpha|\ ||\vec{x}||_p} = \frac{||A\vec{x}||_p}{||\vec{x}||_p}$$

[1] Actually, people say that the matrix norm is subordinate to the vector norm

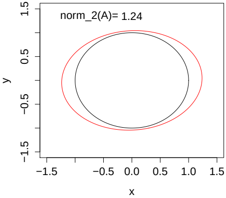
Examples of Matrix 2-Norms for 2x2 Matrices

- Image of the unit sphere in the 2-norm under a linear mapping A is a hyperellipse
- The matrix norm $||A||_2$ gives the length of the longest principal semiaxis

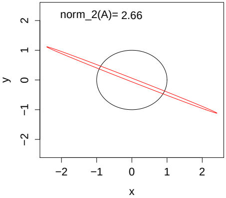
$$\begin{pmatrix} 3 & 1 \\ 4 & 1 \end{pmatrix}$$



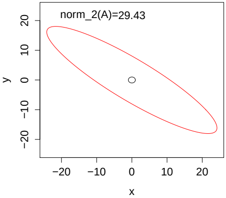
$$\begin{pmatrix} 0.3 & 1.2 \\ -1.0 & 0.3 \end{pmatrix}$$



$$\begin{pmatrix} 2.1 & -1.2 \\ -1.0 & 0.5 \end{pmatrix}$$



$$\begin{pmatrix} 21 & -12 \\ -10 & 15 \end{pmatrix}$$



Computing the Matrix Norms $p = 1, 2, \infty$

The matrix norms $p = 1, 2, \infty$ are "easily computable"

(a) $||A||_1 = \max_{1 \le j \le n} \sum_{i=1}^{n} |a_{ij}|$ = maximum absolute column sum

(b) $||A||_\infty = \max_{1 \le i \le n} \sum_{j=1}^{n} |a_{ij}|$ = maximum absolute row sum

(c) $||A||_2 = \sqrt{\max\{\text{eigenvalue}(A^T A)\}} = \max\{\text{sing}(A)\}$

Eg., Compute $||A||$ in the following examples. Try using norm in R

$$\begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}, \quad \begin{pmatrix} 2 & -1 & 1 \\ 1 & 0 & 1 \\ 3 & -1 & 4 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 5 & 0 & 0 \\ 0 & -10 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad \begin{pmatrix} .1 & .3 \\ .2 & .4 \end{pmatrix}$$