Vector and Matrix Norms
LARP/2018

Norms

To meter the lengths of vectors in a vector space we need the idea of a norm.

Norm is a function that maps x to a nonnegative real number

$$\| \, . \, \| : \quad F \rightarrow R^{+}$$

A Norm must satisfy following properties:

$$1 - \text{Positivity} \quad \|x\| > 0 \;\; , \forall \, x \neq 0$$

$$2 - \text{Homogeneity} \quad \|\alpha x\| = |\alpha| \|x\| \; , \forall \, x \in F \text{ and } \forall \, \alpha \in C$$

$$3 - \text{Triangle inequality} \quad \|x + y\| \leq \|x\| + \|y\| \; , \forall \, x, y \in F$$

Norm of vectors

p-norm is: $\|x\|_p = \left(\sum_i |a_i|^p \right)^{\frac{1}{p}}$ $\quad p \geq 1$

For p=1 we have 1-norm or sum norm $\|x\|_1 = \left(\sum_i |a_i| \right)$

For p=2 we have 2-norm or euclidian norm $\|x\|_2 = \left(\sum_i |a_i|^2 \right)^{1/2}$

For p=∞ we have ∞-norm or max norm $\|x\|_\infty = \max_i \{ |a_i| \}$

Norm of vectors

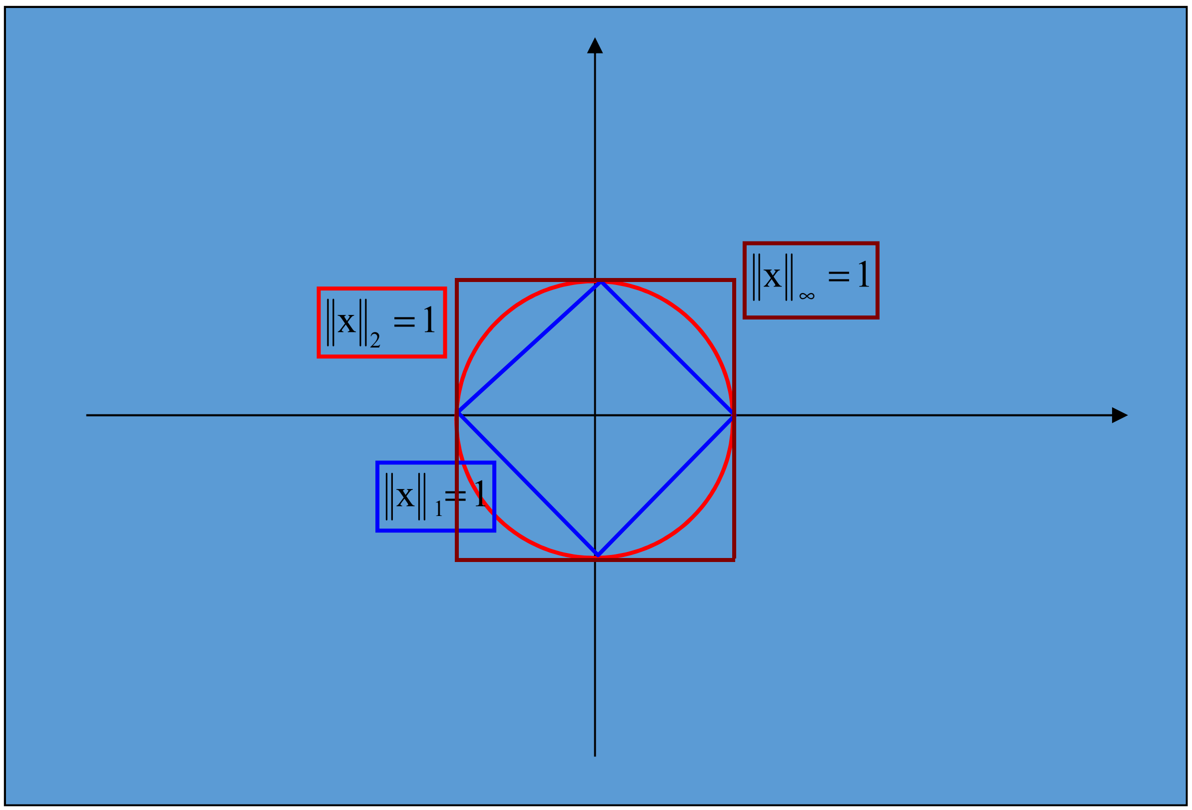
Let $x = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$

Then

$$\|x\|_1 = (1+1+2) = 4$$

$$\|x\|_2 = \sqrt{1^2 + 1^2 + 2^2} = \sqrt{6}$$

$$\|x\|_\infty = \max(1,1,2) = 2$$



- For any vector norm:

$$\|x\| > 0 \text{ if } x \neq 0$$

$$\|\gamma x\| = |\gamma| \cdot \|x\| \text{ for any scalar } \gamma$$

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

- These properties define a vector norm

Norm of real functions

Consider continuous functions on the interval [0,1] over the field of real numbers (R)

1-norm is defined as $$\|f(t)\|_\infty = \sup_{t\in[0,1]} |f(t)|$$

2-norm is defined as $$\|f(t)\|_2 = \left(\int_0^1 |f(t)|^2 dt \right)^{\frac{1}{2}}$$

The l_p-Norm

The l_p- Norm for a vector x is defined as ($p\geq1$):

$$\|x\|_{l_p} = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p}$$

Examples:

- for p=2 we have the ordinary euclidian norm: $$\|x\|_{l_2} = \sqrt{x^T x}$$

- for p= ∞ the definition is $$\|x\|_{l_\infty} = \max_{1\leq i\leq n} |x_i|$$

- a norm for matrices is induced via $$\|A\| = \max_{x\neq0} \frac{\|Ax\|}{\|x\|}$$

- for l_2 this means :
  $||A||_2$=maximum eigenvalue of A^TA

Norm of matrices

We can extend norm of vectors to matrices

Sum matrix norm (extension of 1-norm of vectors) is: $\|A\|_{sum} = \sum_{i,j} |a_{ij}|$

Frobenius norm (extension of 2-norm of vectors) is: $\|A\|_F = \sqrt{\sum_{i,j} |a_{ij}|^2}$

$$\|A\|_F = \sqrt{Tr(AA^T)}$$

Max element norm (extension of max norm of vectors) is: $\|A\|_{\max} = \max_{i,j} |a_{ij}|$

Matrix norm

A norm of a matrix is called matrix norm if it satisfy

$$\|AB\| \leq \|A\|.\|B\|$$

The induced-norm of a matrix A is defined as follows:

$$\|A\|_{ip} = \max_{\|x\|_p = 1} \|Ax\|_p$$

Any induced-norm of a matrix A is a matrix norm

Matrix norm for matrices

If we put p=1 so we have

$$\|A\|_{i1} = \max_{\|x\|_1=1} \|Ax\|_1 = \max_j \sum_i |a_{ij}|$$

Maximum column sum

Maximum absolute ***column*** sum of the matrix

If we put p= ∞ , we have

$$\|A\|_{i\infty} = \max_{\|x\|_\infty=1} \|Ax\|_\infty = \max_i \sum_j |a_{ij}|$$

Maximum row sum

Maximum absolute ***row*** sum of the matrix

Matrix norm for matrices

$$\|A\|_{ip} = \max_{\|x\|_p = 1} \|Ax\|_p$$

If we put p=1 so we have

$$\|A\|_{i1} = \max_{\|x\|_1 = 1} \|Ax\|_1 = \max_j \sum_i |a_{ij}|$$

Maximum column sum

If we put p=inf so we have

$$\|A\|_{i\infty} = \max_{\|x\|_\infty = 1} \|Ax\|_\infty = \max_i \sum_j |a_{ij}|$$

Maximum row sum

If we put p=2 so we have

$$\|A\|_{i2} = \max_{\|x\|_2 = 1} \|Ax\|_2 = \max_{\|x\|_2 \neq 1} \frac{\|Ax\|_2}{\|x\|_2} = \sigma_1(A) = \sigma_{\max}(A) = \overline{\sigma}(A)$$

- We will only use matrix norms "induced" by vector norms:

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

- 1-norm:

$$\|A\|_1 = \max_j \sum_{i=1}^{n} |A_{ij}| \ \text{(max absolute column sum)}$$

- ∞-norm:

$$\|A\|_\infty = \max_i \sum_{j=1}^{n} |A_{ij}| \ \text{(max absolute row sum)}$$

Properties of Matrix Norms

- These induced matrix norms satisfy:

$$\|A\| > 0 \text{ if A} \neq 0$$

$$\|\gamma A\| = |\gamma| \cdot \|A\| \text{ for any scalar } \gamma$$

$$\|A + B\| \leq \|A\| + \|B\| \text{ (triangle inequality)}$$

$$\|AB\| \leq \|A\| \cdot \|B\|$$

$$\|Ax\| \leq \|A\| \cdot \|x\| \text{ for any vector } x$$

Condition Number

- If A is square and nonsingular, then

$$\text{cond}(A) = \|A\| \cdot \|A^{-1}\|$$

- If A is singular, then cond(A) = ∞
- If A is nearly singular, then cond(A) is large.
- The condition number measures the ratio of maximum stretch to maximum shrinkage:

$$\|A\| \cdot \|A^{-1}\| = \left(\max_{x \neq 0} \frac{\|Ax\|}{\|x\|} \right) \cdot \left(\min_{x \neq 0} \frac{\|Ax\|}{\|x\|} \right)^{-1}$$

Properties of Condition Number

- For any matrix A, $\text{cond}(A) \geq 1$
- For the identity matrix, $\text{cond}(I) = 1$
- For any permutation matrix, $\text{cond}(P) = 1$
- For any scalar α, $\text{cond}(\alpha A) = \text{cond}(A)$
- For any diagonal matrix D,

$$\text{cond}(D) = \left(\max |D_{ii}|\right) / \left(\min |D_{ii}|\right)$$

Errors and Residuals

- Residual for an approximate solution y to $Ax = b$ is defined as $r = b - Ay$
- If A is nonsingular, then $||x - y|| = 0$ if and only if $||r|| = 0$.
  - Does not imply that if $||r||<\varepsilon$, then $||x-y||$ is small.

Estimating Accuracy

- Let *x* be the solution to *Ax* = *b*
- Let *y* be the solution to *Ay* = *c*
- Then a simple analysis shows that

$$\frac{\|x-y\|}{\|x\|} \leq \mathrm{cond}(A)\frac{\|b-c\|}{\|c\|}$$

- Errors in the data (b) are magnified by *cond(A*)
- Likewise for errors in *A*