Matrix & Vector Norms

Recall that the 2-norm of a vector is the "length":

$$\|\underline{x}\|_2 = (x_1^2 + x_2^2 + \cdots + x_n^2)^{1/2}$$

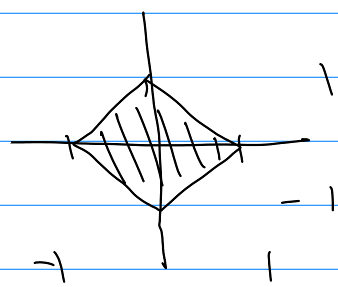
Generalize to p-norms:

$$\|\underline{x}\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p} \qquad 1 \le p < \infty$$

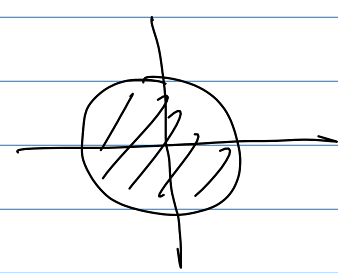
In 2D, the norms can be given by areas

Let $\underline{x} = (x_1, x_2)$ w/ $x_1^2 + x_2^2 \le 1$

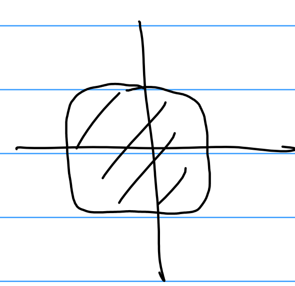
$$\|\underline{x}\|_1 = \sum_{i=1}^{n} |x_i|$$



$$\|\underline{x}\|_2 = \left(\sum^{n} |x_i|^2\right)^{1/2}$$

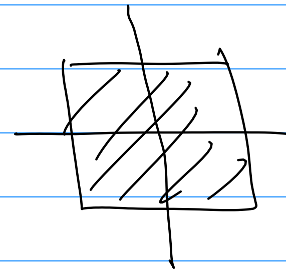


$$\|\underline{x}\|_p = \left(\sum^{n} |x_i|^p\right)^{1/p}$$



Special Norm

$\|\underline{x}\|_\infty = \max |x_i|$



Each of these norms obey:

1) $\|\underline{x}\| \geq 0$ & $\|\underline{x}\| = 0$ iff $\underline{x} = \underline{0}$
2) $\|\underline{x} + \underline{y}\| \leq \|\underline{x}\| + \|\underline{y}\|$
3) $\|\alpha \underline{x}\| = |\alpha| \, \|\underline{x}\|$

Matrix Norms

Vector-induced matrix norms:
those that result from the application of a matrix.

Let $\underline{A} \in M_{mn}$. The induced matrix norm is the number C such that

$$\|\underline{A}\,\underline{x}\|_{(m)} \leq C \, \|\underline{x}\|_{(n)} \quad \text{for all } \underline{x} \in \mathbb{R}^n$$

Note: $\|\cdot\|_{(m)}$ & $\|\cdot\|_{(n)}$ are <u>not</u> the m-norm or n-norm, but the norm in that m or n space.

ex.) 1-norm of a matrix.

Let $\underline{x} \in \mathbb{R}^n$ such that $\|\underline{x}\|_1 \leq 1$

$$\|\underline{A}\,\underline{x}\|_1 = \left\|\sum_{j=1}^{n} x_j \underline{a}_j\right\|_1 \leq \sum_{j=1}^{n} |x_j| \|\underline{a}_j\|_1 \leq \max_j \|\underline{a}_j\|_1$$

$\underline{a}_j$ = Column of $\underline{A}$

Let $\underline{x} = \underline{e}_j$ where $\underline{e}_j$ is the vector that maximizes $\|a_j\|_1$

$$\Rightarrow \|\underline{A}\|_1 = \max_{1 \leq j \leq n} \|\underline{a}_j\|_1 \leftarrow \text{maximum column sum}$$

ex.) $\|\underline{A}\|_\infty = \max_{1 \leq j \leq m} \|a_j^*\|_1 \leftarrow a_j^* = j^{th}$ row of $\underline{A}$.

$\|\underline{A}\|_\infty$ = maximum row sum

Other Common Matrix-Norm

Frobenius Norm:

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

Matrix norms also follow:

1) $\|\underline{A}\| \geq 0$, $\|\underline{A}\| = 0$ iff $\underline{A} = \underline{0}$

2) $\|\underline{A} + \underline{B}\| \leq \|\underline{A}\| + \|\underline{B}\|$

3) $\|\alpha \underline{A}\| = |\alpha| \, \|\underline{A}\|$

Condition Number

Consider the numeric solution to $\underline{A}\,\underline{x} = \underline{b}$

At a minimum, $\underline{b}$ has some error in it,

we actually are solving is

$$\underline{A}(\underline{x} + \underline{\Delta x}) = \underline{b} + \underline{\Delta b}$$

$\underline{\Delta b}$ = error

$\underline{\Delta x}$ = error

$$\underline{A}(\underline{x} + \underline{\Delta x}) = \underline{b} + \underline{\Delta b}$$

$\underline{A}\,\underline{x} + \underline{A}\,\underline{\Delta x} = \underline{b} + \underline{\Delta b}$ but, $\underline{A}\,\underline{x} = \underline{b}$

$\Rightarrow \underline{A}\,\underline{\Delta x} = \underline{\Delta b}$

Assume $\underline{A}^{-1}$ exists $\Rightarrow$ $\underline{\Delta x} = \underline{A}^{-1}\underline{\Delta b}$

$$\|\underline{\Delta x}\| = \|\underline{A}^{-1}\underline{\Delta b}\| \leq \|\underline{A}^{-1}\|\,\|\underline{\Delta b}\|$$

Now look at $\underline{A}\,\underline{x} = \underline{b}$ in exact math

$$\|\underline{b}\| = \|\underline{A}\,\underline{x}\| \leq \|\underline{A}\|\,\|\underline{x}\|$$

$$\Rightarrow \frac{1}{\|\underline{x}\|} \leq \|\underline{A}\|\,\frac{1}{\|\underline{b}\|}$$

Normalized error:

$$\frac{\|\underline{\Delta x}\|}{\|\underline{x}\|} \leq \|\underline{A}\|\,\|\underline{A}^{-1}\|\,\frac{\|\underline{\Delta b}\|}{\|\underline{b}\|}$$

If the normalized error of $\underline{b}$ is $\frac{\|\underline{\Delta b}\|}{\|\underline{b}\|}$, the error of the solution

will scale by $\|\underline{A}\|\,\|\underline{A}^{-1}\|$

$\Rightarrow$ this is the condition number:

$$K(\underline{A}) = \|\underline{A}\|\,\|\underline{A}^{-1}\|$$

ex1) let $\|\underline{\Delta b}\|/\|\underline{b}\| \sim 10^{-16}$, but $K(\underline{A}) \sim 10^{6}$

$\Rightarrow \|\underline{\Delta x}\|/\|\underline{x}\| \sim 10^{-10}$ (6 <u>orders</u> higher)

In general, $\underline{A}$ will also have errors:

$$\frac{\|\Delta \underline{x}\|}{\|\underline{x}\|} \le \kappa(\underline{A})\left(\frac{\|\Delta \underline{A}\|}{\|\underline{A}\|} + \frac{\|\Delta \underline{b}\|}{\|\underline{b}\|}\right)$$

LU Decomposition.

Motivate by looking at row echelon form

Take a generic $\underline{A}$ & convert to an upper diagonal matrix $\underline{U}$.

ex1) $\underline{A} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 6 & 10 \\ 3 & 14 & 28 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 4 \\ 3 & 14 & 28 \end{bmatrix}$

Each step of this process (called Gaussian elimination) can be written as a matrix-matrix product.

$$\underline{E}_1 = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\underline{E}_1\underline{A} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 4 \\ 3 & 14 & 28 \end{bmatrix}$$

All elimination matrices, $\underline{E}$, are lower triangular.

(1) The multiplication of lower triangular matrices, results in a lower triangular matrix:

$$\underline{L}_1 \underline{L}_2 = \underline{L}_3$$

(2) The inverse of a lower triangular matrix is also lower triangular:

$$\underline{L}_1^{-1} = \underline{L}_2$$

Gaussian Elimination is nothing but repeated multipliation by elimination matrices:

$$\underbrace{\underline{E}_n \underline{E}_{n-1} \cdots \underline{E}_1}_{\text{A lower triangular matrix}} \underline{A} = \underline{U} \leftarrow \text{row echlon form upper triangular,}$$

$$\underline{L}^{-1} \underline{A} = \underline{U}$$

$$\underline{L}\,\underline{L}^{-1} \underline{A} = \underline{L}\,\underline{U}$$

$$\underline{A} = \underline{L}\,\underline{U} \leftarrow \text{LU Decomposition.}$$

(1) The diagonal of $\underline{L}$ $\underline{\text{must}}$ have all $\underline{1}$'s.

Look at: $\begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix} = \begin{bmatrix} l_{11} & 0 \\ l_{21} & l_{22} \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} \\ 0 & u_{22} \end{bmatrix}$

$$\Rightarrow \left.\begin{aligned} l_{11} u_{11} &= 1 \\ l_{11} u_{12} &= 2 \\ l_{21} u_{11} &= 3 \\ l_{21} u_{12} + l_{22} u_{22} &= 5 \end{aligned}\right\} \begin{array}{l} 6 \text{ unknowns} \\ 4 \text{ equations} \end{array}$$

Set $l_{11} = 1$ & $l_{22} = 1$

(2) Why do we care?

Solve $\underline{A}\,\underline{x} = \underline{b}$ by finding $\underline{A}^{-1}$:

$\underline{x} = \underline{A}^{-1}\underline{b}$ ← $\underline{\underline{\text{Very}}}$ expensive.

Look at $\underline{A}\,\underline{x} = \underline{b}$

$\underline{A}\,\underline{x} = \underline{L}\,\underline{U}\,\underline{x} = \underline{b}$

$\underline{x} = \underline{U}^{-1}\underline{L}^{-1}\underline{b}$

$\underline{U}^{-1}$ & $\underline{L}^{-1}$ are cheap! because they are triangular!

$$\underline{L}\underline{x} = \begin{bmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

1) Compute $\underline{A} = \underline{L}\,\underline{U}$

2) Solve $\underline{L}\,\underline{y} = \underline{b} \Rightarrow \underline{y} = \underline{L}^{-1}\underline{b}$

3) Solve $\underline{U}\,\underline{x} = \underline{y} \Rightarrow \underline{x} = \underline{U}^{-1}\underline{y} = \underline{U}^{-1}\underline{L}^{-1}\underline{b}$

ex1) Compute LU of $\underline{A} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 6 & 10 \\ 3 & 14 & 28 \end{bmatrix}$

1) Eliminate a_{21}

Multiply $\underline{A}$ by $\underline{E}_1 = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

$$\underline{E}_1 \underline{A} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 4 \\ 3 & 14 & 28 \end{bmatrix}$$

2) eliminate location (3, 1)

$$\underline{E}_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix} \Rightarrow \underline{E}_2 \underline{E}_1 \underline{A} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 4 \\ 0 & 8 & 19 \end{bmatrix}$$

3) Eliminate $(3,2)$

$$\underline{E}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -4 & 1 \end{bmatrix}$$

$$\underline{E}_3 \underline{E}_2 \underline{E}_1 \underline{A} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 4 \\ 0 & 0 & 3 \end{bmatrix} = \underline{U}$$

$$\underline{A} = (\underline{E}_3 \underline{E}_2 \underline{E}_1)^{-1} \underline{U}$$

$$\underline{A} = \underline{E}_1^{-1} \underline{E}_2^{-1} \underline{E}_3^{-1} \underline{U}$$

$$\underline{E}_1^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ +2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \underline{E}_2^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ +3 & 0 & 1 \end{bmatrix}$$

$$\underline{E}_3^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & +4 & 1 \end{bmatrix}$$

$$\underline{E}_1^{-1} \underline{E}_2^{-1} \underline{E}_3^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 4 & 1 \end{bmatrix}$$

The negative of the operation, on factor of the value divided by the pivot.

$$\underline{A} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 4 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 4 \\ 0 & 0 & 3 \end{bmatrix}$$

$$\underline{L} \qquad\qquad \underline{U}$$

Generic Algorithm.

Look at eliminating all rows below a pivot:

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 5 & 6 & 7 \\ 0 & 8 & \boxed{9} & \boxed{10} \\ 0 & 11 & \boxed{12} & \boxed{13} \end{bmatrix}$$

1 op ① For each row below a pivot, compute a factor

2 op ② for every column not under a pivot, multiply by that factor & do a subtraction

③ Repeat for all rows except the last one.

Let $\underline{A} \in M_{mn}$

$L = I_{mn}$

```
for i = 1 : m-1
    for j = i+1 : n
1 op        L(j,i) = A(j,i) / A(i,i)
            A(j,i) = 0
            for k = i+1 : n
2 op            A(j,k) = A(j,k) - L(j,i) * A(i,k)
            end
    end
end
U = A(1:n, 1:n)
```

Now look at the Cost of this, in floating point operations FLOPS,

Let $m = n$ (Square matrix)

T = Op Count

$$T = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \left(1 + \sum_{k=i+1}^{n} 2\right)$$

Need: $$\sum_{i=1}^{n-1} i = \frac{n(n-1)}{2}$$

$$\sum_{i=1}^{n-1} i^2 = \frac{n(n-1)(2n-1)}{6}$$

Now:

$$T = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \left(1 + \sum_{k=i+1}^{n} 2\right) = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} 1 + \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \sum_{k=i+1}^{n} 2$$

$$= \sum_{i=1}^{n-1} (n-i) + \sum_{i=1}^{n} \sum_{j=i+1}^{n} 2(n-i)$$

$$= \sum_{i=1}^{n} (n-i) + \sum_{i=1}^{n} 2(n-i)(n-i)$$

$$= \sum_{i=1}^{n} \left[(n-i) + 2(n-i)(n-i) \right]$$

$$= \sum_{i=1}^{n} \left(2n^2 - 4ni + n + 2i^2 - i \right)$$

$$= 2n^2(n-1) - \frac{4n(n-1)n}{2} + n(n-1)$$

$$+ \frac{2n(n-1)(2n-1)}{6} - \frac{(n-1)n}{2}$$

$$T = \frac{2n^3}{6} - \frac{n^2}{2} - \frac{n}{6}$$

"Big O" notation: we say that a function $f(x)$ is $O(g(x))$ if as $x \to a$ there exists a δ & M such that

$$|f(x)| \leq M |g(x)| \text{ for } |x-a| < \delta$$

In our case, all this means is that as n becomes large, the op count T becomes dominated by the n^3 term.

$\Rightarrow T = \frac{2}{3} n^3 + O(n^2)$

$\Rightarrow$ To Do LU Decomposition, the cost scales as n^3!

Once you have $\underline{A} = \underline{L}\,\underline{U}$, solving

$\underline{A}\,\underline{x} = \underline{b} \Rightarrow \underline{L}\,\underline{U}\,\underline{x} = \underline{b}$ is

only $O(n^2)$

$\Rightarrow$ Factorization is the expensive part of LU.

Failure of Gaussian Elimination.

Look at $\underline{A} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$

Full rank $\rightarrow \underline{A}^{-1}$ exists

$\kappa(\underline{A}) \sim 2.618$

Problem 1: how do I eliminate the (1,1) location?

Problem 2: Look at $\underline{A}$ w/ slight perturbation:

$$\underline{A} = \begin{bmatrix} 10^{-20} & 1 \\ 1 & 1 \end{bmatrix}$$

Exact LU is

$$\underline{L} = \begin{bmatrix} 1 & 0 \\ 10^{20} & 1 \end{bmatrix} \qquad \underline{U} = \begin{bmatrix} 10^{-20} & 1 \\ 0 & 1-10^{20} \end{bmatrix}$$

$1-10^{20}$ can not be represented in finite precision.

$$\Rightarrow 1-10^{20} \sim -10^{20}$$

$$\Rightarrow \underline{\hat{L}} = \underline{L} \qquad \underline{\hat{U}} = \begin{bmatrix} 10^{-20} & 1 \\ 0 & -10^{20} \end{bmatrix}$$

$$\underline{\hat{L}}\,\underline{\hat{U}} = \begin{bmatrix} 10^{-20} & 1 \\ 1 & 0 \end{bmatrix} \overset{?}{\neq} \underline{A}$$

Pivots to the rescue!

A pivot matrix is one that simply swaps rows.

(Technically partial pivoting)

$$\underbrace{\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}}_{\underline{P}} \underbrace{\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}}_{\underline{A}} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \underline{U}$$

When partial pivoting is used w/ LU, you actually get

$$\underline{P}\,\underline{A} = \underline{L}\,\underline{U}$$

Solving $\underline{A}\,\underline{x} = \underline{b}$

$$\underline{P}\,\underline{A}\,\underline{x} = \underline{P}\,\underline{b}$$

$$\underline{L}\,\underline{U}\,\underline{x} = \underline{P}\,\underline{b}$$

$$\Rightarrow \underline{x} = \underline{U}^{-1}\,\underline{L}^{-1}\,(\underline{P}\,\underline{b})$$