Inverse Matrices, cont'd

Def: An $n \times n$ (square!) matrix A is **invertible** if there exists another $n \times n$ matrix B such that $AB = I_n = BA$.

Remark: Since $AB \neq BA$ in general, you have to require $AB = I_n = BA$. But:

Fact: If A and B are $n \times n$ matrices and $AB = I_n$ or $BA = I_n$, then $B = A^{-1}$.

So the definition above is a bit pedantic...

Remark: A non-square matrix does not admit both a left- and right-inverse, so not invertible. (Can't solve $AB = I_m$ and $CA = I_n$ unless A is square.)

This is why we only treat invertibility of square matrices.

Fact: $(A^{-1})^{-1} = A$

because $AB = I_n$ means $B = A^{-1}$ and $A = B^{-1}$

Fact: If A & B are invertible, then so is AB, and $(AB)^{-1} = \cancel{A^{-1}B^{-1}}\ B^{-1}A^{-1}$.

Check: $(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = A(I_n)A^{-1} = AA^{-1} = I_n$ ✓

Thm: Let A be an $n \times n$ matrix. The Following Are Equivalent: (TFAE)

either all are true or all are false

(1) A is invertible

(2) The RREF of A is I_n

(3) A has a pivot in every row / every column.

We'll see why a bit later.

Eg: $A = \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix} \xrightarrow{\text{RREF}} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2$ invertible

= pivots

Eg: $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ is in RREF, $\neq I_2$ singular

How do you compute the inverse?

Algorithm (Matrix Inversion):

Input: A square matrix.

Output: The inverse matrix, or "singular"

Procedure:

(a) Form the augmented matrix $[A \mid I_n]$

(b) Run Gauss-Jordan on $[A \mid I_n]$.

(c) If the output is $[I_n \mid B]$ then $B = A^{-1}$.

Otherwise A is singular.

why does this work? See below.

Eg: Compute $\begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix}^{-1}$.

$$\left[\begin{array}{cc|cc} 2 & 3 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{array}\right] \xrightarrow{R_2 -= \frac{1}{2}R_1} \left[\begin{array}{cc|cc} 2 & 3 & 1 & 0 \\ 0 & 1/2 & -1/2 & 1 \end{array}\right]$$

$$\xrightarrow{R_2 \times= 2} \left[\begin{array}{cc|cc} 2 & 3 & 1 & 0 \\ 0 & 1 & -1 & 2 \end{array}\right]$$

$$\xrightarrow{R_1 -= 3R_2} \left[\begin{array}{cc|cc} 2 & 0 & 4 & -6 \\ 0 & 1 & -1 & 2 \end{array}\right]$$

$$\xrightarrow{R_1 \div= 2} \left[\begin{array}{cc|cc} 1 & 0 & 2 & -3 \\ 0 & 1 & -1 & 2 \end{array}\right]$$

This has the form $\left[\begin{array}{c|cc} I_2 & 2 & -3 \\ & -1 & 2 \end{array}\right]$

So $\begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix}^{-1} = \begin{bmatrix} 2 & -3 \\ -1 & 2 \end{bmatrix}$.

Eg: Compute $\begin{bmatrix} 1 & 3 \\ -1 & -3 \end{bmatrix}^{-1}$.

$$\left[\begin{array}{cc|cc} 1 & 3 & 1 & 0 \\ -1 & -3 & 0 & 1 \end{array}\right] \xrightarrow{R_2 += R_1} \left[\begin{array}{cc|cc} 1 & 3 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{array}\right]$$

$$\xrightarrow{R_1 -= R_2} \left[\begin{array}{cc|cc} 1 & 3 & 0 & -1 \\ 0 & 0 & 1 & 1 \end{array}\right]$$

But this is in RREF and it does not have the form $[I_2 \mid B] \Longrightarrow$ singular

NB: We knew this after the first step: no pivot in the second column.

Actually there's a shortcut for 2x2 matrices:

Fact: $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is invertible $\iff$ $ad-bc \neq 0$,

in which case

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Eg: $$\begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix}^{-1} = \frac{1}{2\cdot 2 - 3\cdot 1}\begin{bmatrix} 2 & -3 \\ -1 & 2 \end{bmatrix}$$

Check: $$\frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

$$= \frac{1}{ad-bc}\begin{bmatrix} ad-bc & bd-bd \\ ac-ac & ad-bc \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \checkmark$$

What is this good for?

Suppose A is invertible. Let's solve $Ax=b$.

$$\begin{aligned} Ax=b &\iff A^{-1}(Ax) = A^{-1}b \\ &\iff (A^{-1}A)x = A^{-1}b \\ &\iff x = A^{-1}b \end{aligned}$$

For invertible A:

$$Ax=b \iff x=A^{-1}b$$

In particular, $Ax=b$ has exactly one solution for any b, and we have an expression for b in terms of x.

Eg: Solve $\begin{aligned} 2x_1+3x_2&=b_1 \\ x_1+2x_2&=b_2 \end{aligned}$.

$$A=\begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix} \qquad A^{-1}=\begin{bmatrix} 2 & -3 \\ -1 & 2 \end{bmatrix}$$

$$A\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}=\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \iff \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}=A^{-1}\begin{bmatrix} b_1 \\ b_2 \end{bmatrix}=\begin{bmatrix} 2 & -3 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

$$\iff \begin{aligned} x_1 &= 2b_1-3b_2 \\ x_2 &= -b_1+2b_2 \end{aligned}$$

So if you want to solve $\begin{aligned} 2x_1+3x_2&=3 \\ x_1+2x_2&=4 \end{aligned}$

$$\Rightarrow \begin{aligned} x_1 &= 2(3)-3(4)=-6 \\ x_2 &= -(3)+2(4)=5 \end{aligned}$$

Elementary Matrices

These give a way to do row operations by matrix multiplication!

Def: An elementary matrix is a matrix obtained from I_n by doing one row operation.

Eg:

- $R_1 += 2R_2$ $\begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

- $R_2 \times= 2$ $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

- $R_1 \leftrightarrow R_2$ $\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

Fact: If E is an $m\times m$ elementary matrix and A is an $m\times n$ matrix, then

$$E\cdot A = \left(\begin{array}{l}\text{what you get by performing that}\\ \text{row operation on } A.\end{array}\right)$$

$$\boxed{(\text{row operations}) \equiv \left(\begin{array}{l}\text{left-multiplication by}\\ \text{elementary matrices}\end{array}\right)}$$

Eg: $\begin{bmatrix} -2 & -3 & 0 & -7 \\ 1 & 2 & 3 & 4 \\ 0 & 1 & 0 & 1 \end{bmatrix} \xrightarrow{R_1 += 2R_2} \begin{bmatrix} 0 & 1 & 6 & 1 \\ 1 & 2 & 3 & 4 \\ 0 & 1 & 0 & 1 \end{bmatrix}$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{R_1 += 2R_2} \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

same!

$$\begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -2 & -3 & 0 & -7 \\ 1 & 2 & 3 & 4 \\ 0 & 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 6 & 1 \\ 1 & 2 & 3 & 4 \\ 0 & 1 & 0 & 1 \end{bmatrix}$$

Left-multiplication by $\begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ does $R_1 += 2R_2$.

Fact: An elementary matrix is invertible. Its inverse corresponds to the elementary matrix that un-does the row operation.

Why?

$$E_1 = (\text{matrix for } R_1 += 2R_2) = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$E_2 = (\text{matrix for } R_1 -= 2R_2) = \begin{bmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$E_1E_2 = E_1E_2I_n = E_1(\text{do } R_1 -= 2R_2 \text{ to } I_n)$$

$$= \left(\begin{array}{l} \text{first do } R_1 -= 2R_2 \text{ to } I_n, \\ \text{then do } R_1 += 2R_2 \text{ to } I_n \end{array} \right) = I_n.$$

Check: $\begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

What if you do multiple row operations?

Consider these row operations & their elementary matrices:

$E_1: R_1 += 2R_2 \qquad E_2: R_2 \times= 2 \qquad E_3: R_2 \leftrightarrow R_3$

Apply in order to A

$$A \xrightarrow{R_1 += 2R_2} E_1A \xrightarrow{R_2 \times = 2} E_2(E_1A) \xrightarrow{R_2 \leftrightarrow R_3} E_3(E_2(E_1A))$$
$$= (E_3E_2E_1)A$$

The elementary matrices ended up in the opposite order! Why?

$$E_3E_2E_1A = E_3E_2(E_1A): \text{ first multiply by } E_1$$

Application to Invertibility:

Suppose $RREF(A) = I_n$.

So there are some number of row ops to transform $A \rightsquigarrow I_n$.

Let $E_1, \dots, E_r$ be their elementary matrices.

$$I_n = (E_r E_{r-1} \cdots E_1) A$$
$$\Rightarrow \boxed{A^{-1} = E_r E_{r-1} \cdots E_1}$$

In particular, A is invertible: justifies (part of) the Thm above (p.2)

This also justifies the algorithm for computing A^{-1}:

$$[A \mid I_n] \xrightarrow{\text{row ops}} [I_n \mid B]$$

Then
$$[I_n \mid B] = (E_r \cdots E_1)[A \mid I_n]$$
$$= \big[(E_r \cdots E_1)A \mid (E_r \cdots E_1)\big]$$

column ~ 1st matrix multiplication

$\Downarrow$

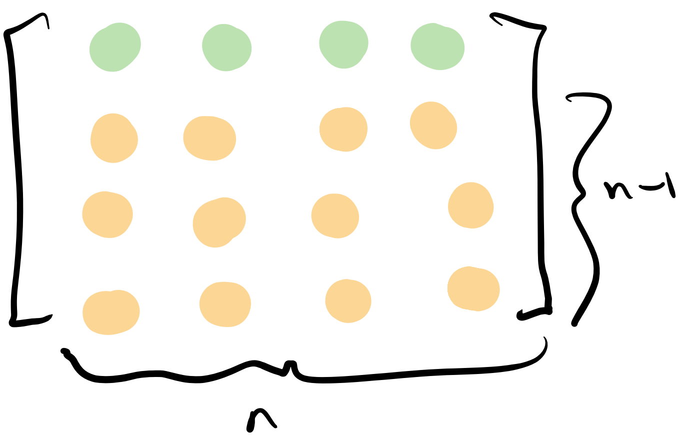
$C[A|B] = [CA|CB]$

$$\Rightarrow B = E_r \cdots E_1 = A^{-1}$$

LU Decomposition

How much computer time does Gauss-Jordan take?
(computational complexity)

Gaussian Elimination on an $n \times n$ matrix takes:

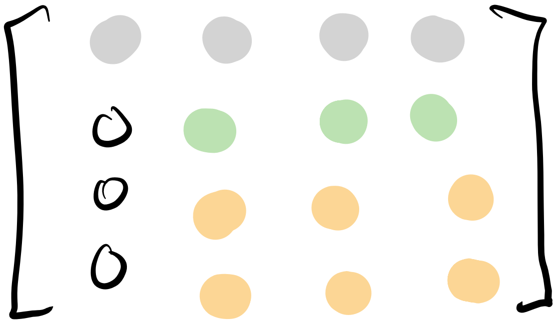


Step 1: Each row replacement requires $n-1$ multiplications & $n-1$ additions. (no computation in 1st col: just write "0")

Do for $n-1$ lower rows:

$(n-1)(n-1)$ mult
$(n-1)(n-1)$ add

$2(n-1)^2$ flops ← (floating point operations)



Step 2: Each row replacement requires $n-2$ multiplications & $n-2$ additions. Must do this for $n-2$ remaining rows

$$\begin{array}{rl} & (n-2)(n-2) \text{ mult} \\ + & (n-2)(n-2) \text{ add} \\ \hline & 2(n-2)^2 \text{ flops} \end{array}$$

etc.

Total: $2[\underbrace{(n-1)^2+(n-2)^2+\cdots+1^2}_{\text{pyramidal number}}]$

$$= 2\cdot\frac{n(n-1)(2n-1)}{6} \approx \frac{2}{3}n^3 \text{ flops}$$

Back-Substitution

$x_n =$ — 1 mult $= 1$ flop

$x_{n-1} + x_n =$ — 2 mult, 1 add $= 3$ flops

(substitute x_n, $\times$, subtract, $\div$)

$x_{n-2} + x_{n-1} + x_n =$ — 3 mult, 2 add $= 5$ flops

(substitute x_n & x_{n-1}, $\times$, $\times$, subtract, $\div$)

$x_1 + \cdots + x_n =$ — n mult, $(n-1)$ add $= 2n-1$ flops

Total: $1+3+5+\cdots+(2n-1) = n^2$ flops

NB: $\frac{2}{3}n^3$ is a lot more than n^2!

For a 1000×1000 matrix, $\frac{2}{3}n^3 \approx \frac{2}{3}$ gigaflops but $n^2 = 1$ megaflop. If we want to solve $Ax=b$ for 1000 values of b, doing elimination each time takes $\frac{2}{3}$ teraflops!

Upshot: We want to do elimination only once!

This is what LU decomposition does.

Def: A matrix is upper/lower triangular if all entries below/above the diagonal are zero.

upper triangular

$$\begin{bmatrix} \bullet & \bullet & \bullet & \bullet \\ 0 & \bullet & \bullet & \bullet \\ 0 & 0 & \bullet & \bullet \end{bmatrix}$$

lower triangular

$$\begin{bmatrix} \bullet & 0 & 0 & 0 \\ \bullet & \bullet & 0 & 0 \\ \bullet & \bullet & \bullet & 0 \end{bmatrix}$$

$\bullet$ = anything

A matrix is unitriangular if it is triangular and all diagonal entries = 1.

upper unitriangular

$$\begin{bmatrix} 1 & \bullet & \bullet & \bullet \\ 0 & 1 & \bullet & \bullet \\ 0 & 0 & 1 & \bullet \end{bmatrix}$$

lower unitriangular

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ \bullet & 1 & 0 & 0 \\ \bullet & \bullet & 1 & 0 \end{bmatrix}$$

Eg: A matrix in REF is upper-Δular

Eg: If E is the elementary matrix for $R_i += c \cdot R_j$ for $i > j$ (add a higher row to a lower row) then E is lower-uniΔular.

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{R_3 += 2R_1} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix} = E$$

Fact: Products & inverses of square upper/lower (uni) triangular matrices are again upper/lower (uni) triangular.

Suppose A can be reduced to REF (using the Gaussian elimination algorithm) without doing any row swaps. Then the only row operations necessary are $R_i += c R_j$ for $i > j$.

U = output of Gaussian elimination (in REF)

$$U = \underbrace{E_r \cdots E_1}_{\text{lower-uni}\Delta\text{ular}} A \Rightarrow A = \underbrace{(E_r \cdots E_1)^{-1}}_{\text{lower-uni}\Delta\text{ular}} U = LU$$

So we've shown:

Fact: If Gaussian elimination on A requires no row swaps, then

$$A = LU$$

for L lower-uniΔular and U in REF.