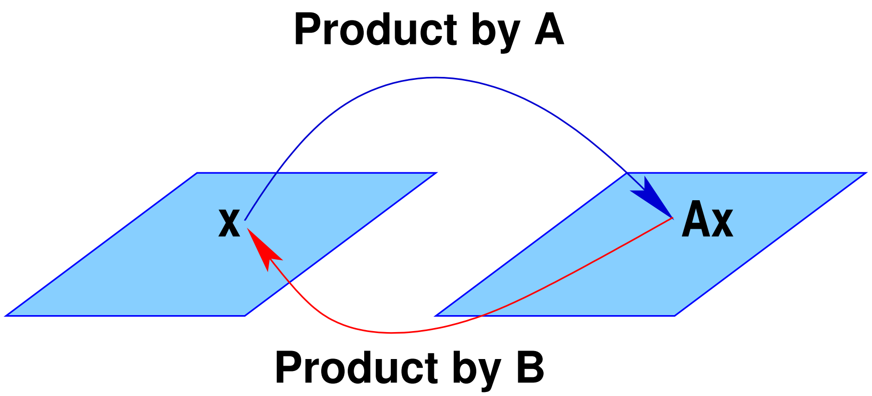
INVERSE OF A MATRIX [2.2]

The inverse of a matrix: Introduction

➤ We have a mapping from $\mathbb{R}^n$ to $\mathbb{R}^n$ represented by a matrix $\boldsymbol{A}$.

➤ Can we invert this mapping? i.e. can we find a matrix (call it $\boldsymbol{B}$ for now) such that when $\boldsymbol{B}$ is applied to $\boldsymbol{Ax}$ the result is $\boldsymbol{x}$?



➤ Example: blurring operation. We want to 'revert' blurring, i.e., to deblur. So: Blurring: $\boldsymbol{A}$; Deblurring: $\boldsymbol{B}$.

➤ $\boldsymbol{B}$ is the inverse of $\boldsymbol{A}$ and is denoted by $\boldsymbol{A}^{-1}$.

- Recall that $\boldsymbol{I_n x = x}$ for all $\boldsymbol{x}$.

- Since we want $\boldsymbol{A^{-1}(Ax) = x}$ for all $\boldsymbol{x}$ this means, we need to have

$$\boxed{\boldsymbol{A^{-1}A = I_n}}$$

- Naturally the inverse of $\boldsymbol{A^{-1}}$ should be $\boldsymbol{A}$ so we also want

$$\boxed{\boldsymbol{AA^{-1} = I_n}}$$

- Finding an inverse to $\boldsymbol{A}$ is not always possible. When it is we say that the matrix $\boldsymbol{A}$ is invertible

- Next: details.

The inverse of a matrix

- An $n \times n$ matrix A is said to be invertible if there is an $n \times n$ matrix B such that $BA = I$ and $AB = I$ where $I = I_n$, the $n \times n$ identity matrix.

- In this case, B is an inverse of A. In fact, B is uniquely determined by A: If C were another inverse of A, then

$$C = CI = C(AB) = (CA)B = IB = B$$

- This unique inverse is denoted by A^{-1} -so that

$$\boxed{AA^{-1} = A^{-1}A = I}$$

Matrix inverse - the 2×2 case

➤ Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. If $ad - bc \neq 0$ then A is invertible and

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

✍ Verify the result

➤ If $ad - bc = 0$ then A is not invertible (does not have an inverse)

➤ The quantity $ad - bc$ is called the determinant of A ($\det(A)$)

➤ The above says that a 2×2 matrix is invertible if and only if $\det(A) \neq 0$.

Matrix inverse - Properties

Theorem If $\boldsymbol{A}$ is invertible, then for each $\boldsymbol{b}$ in $\mathbb{R}^n$, the equation $\boldsymbol{Ax} = \boldsymbol{b}$ has the unique solution $\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b}$.

$\boxed{\text{Proof:}}$ Take any b in $\mathbb{R}^n$. A solution exists because if $\boldsymbol{A}^{-1}\boldsymbol{b}$ is substituted for x, then $\boldsymbol{Ax} = \boldsymbol{A}(\boldsymbol{A}^{-1}\boldsymbol{b}) = (\boldsymbol{A}^{-1}\boldsymbol{A})\boldsymbol{b} = \boldsymbol{Ib} = \boldsymbol{b}.$ So $\boldsymbol{A}^{-1}\boldsymbol{b}$ is a solution.

To prove that the solution is unique, show that if $\boldsymbol{u}$ is any solution, then $\boldsymbol{u}$ must be $\boldsymbol{A}^{-1}\boldsymbol{b}$. If $\boldsymbol{Au} = \boldsymbol{b}$, we can multiply both sides by $\boldsymbol{A}^{-1}$ and obtain $\boldsymbol{A}^{-1}\boldsymbol{Au} = \boldsymbol{Ab}$, $\boldsymbol{Iu} = \boldsymbol{A}^{-1}\boldsymbol{b}$, and $\boldsymbol{u} = \boldsymbol{A}^{-1}\boldsymbol{b}$ ■

✍ Show: If $\boldsymbol{A}$ is invertible then it is one to one, i.e., its columns are linearly independent.

Matrix inverse - Properties

a. If A is an invertible matrix, then A^{-1} is invertible and

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

b. If A and B are $n \times n$ invertible matrices, then so is AB, and we have

$$(AB)^{-1} = B^{-1}A^{-1}$$

c. If A is an invertible matrix, then so is A^T, and the inverse of A^T is the transpose of A^{-1} :

$$(A^T)^{-1} = (A^{-1})^T$$

- Common notation $(A^T)^{-1} \equiv A^{-T}$

Existence of the inverse and related properties

Our next goal is to prove the following theorem.

Existence Theorem. The 4 following statements are equivalent

(1) $\boldsymbol{A}$ is invertible

(2) The columns of $\boldsymbol{A}$ are linearly independent

(3) The Span of the columns of $\boldsymbol{A}$ is $\mathbb{R}^n$

(4) rref($\boldsymbol{A}$) is the identity matrix

Elementary matrices

✍ Consider the matrix on the right and call it $\boldsymbol{E}$. What is the result of the product $\boldsymbol{EA}$ for some matrix $\boldsymbol{A}$?

$$\begin{bmatrix} \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{0} \\ -\boldsymbol{r} & \mathbf{0} & \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} \end{bmatrix}$$

✍ Can this operation result in a change of the linear independence of the columns of $\boldsymbol{A}$? [prove or disprove]

✍ Consider now the matrix on the right [obtained by swapping rows 2 and 4 of $\boldsymbol{I}$]. Call it $\boldsymbol{P}$. Same questions as above.

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

➤ Matrices like $\boldsymbol{E}$ (elementary elimination matrix) and $\boldsymbol{P}$ (permutation matrix) are called 'elementary matrices'

Elimination algorithms and elementary matrices

➤ We will show this:

The following algorithms: Gaussian elimination, Gauss-Jordan, reduction to echelon form, and to reduced row echelon form, are all based on multiplying the original matrix by a sequence of elementary matrices to the left. Each of these transformations preserves linear independence of the columns of the original matrix.

➤ An elementary matrix is one that is obtained by performing a single elementary row operation on an identity matrix.

➤ Let us revisit Gaussian Elimination - Recommended : compare with lecture note example on section 1.1..

Recall: Gaussian Elimination

➤ Consider example seen in section 1.1 – Step 1 must transform:

$$\left[\begin{array}{ccc|c} 2 & 4 & 4 & 2 \\ 1 & 3 & 1 & 1 \\ 1 & 5 & 6 & -6 \end{array}\right] \text{ into: } \left[\begin{array}{ccc|c} x & x & x & x \\ 0 & x & x & x \\ 0 & x & x & x \end{array}\right]$$

$row_2 := row_2 - \frac{1}{2} \times row_1$: $\quad$ $row_3 := row_3 - \frac{1}{2} \times row_1$:

$$\left[\begin{array}{ccc|c} 2 & 4 & 4 & 2 \\ 0 & 1 & -1 & 0 \\ 1 & 5 & 6 & -6 \end{array}\right] \qquad \left[\begin{array}{ccc|c} 2 & 4 & 4 & 2 \\ 0 & 1 & -1 & 0 \\ 0 & 3 & 4 & -7 \end{array}\right]$$

- The first transformation ($\boldsymbol{row_2 := row_2 - \frac{1}{2} \times row_1}$) is equivalent to performing this product:

$$\begin{bmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \times \left[\begin{array}{ccc|c} 2 & 4 & 4 & 2 \\ 1 & 3 & 1 & 1 \\ 1 & 5 & 6 & -6 \end{array}\right] = \left[\begin{array}{ccc|c} 2 & 4 & 4 & 2 \\ 0 & 1 & -1 & 0 \\ 1 & 5 & 6 & -6 \end{array}\right]$$

- Similarly, operation of $\boldsymbol{row_3}$ is equivalent to product:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\frac{1}{2} & 0 & 1 \end{bmatrix} \times \left[\begin{array}{ccc|c} 2 & 4 & 4 & 2 \\ 0 & 1 & -1 & 0 \\ 1 & 5 & 6 & -6 \end{array}\right] = \left[\begin{array}{ccc|c} 2 & 4 & 4 & 2 \\ 0 & 1 & -1 & 0 \\ 0 & 3 & 4 & -7 \end{array}\right]$$

- Hint: Use the row-wise form of the matrix products
- Matrix on the left is called an Elementaty elimination matrix
- Do the same thing for 2nd (and last) step of GE.

Another type of elementary matrices: Permutations

- We used these in partial pivoting.
- A permutation matrix is a matrix obtained from the identity matrix by permuting its rows
- For example for the permutation $p = \{3, 1, 4, 2\}$ we obtain $\longrightarrow$

$$P = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

- Important observation: the matrix PA is obtained from A by permuting its rows with the permutation p

$$(PA)_{i,:} = A_{p(i),:}$$

➤ What does this mean?

It means that for example the 3rd row of $\boldsymbol{PA}$ is simply row number $\boldsymbol{p(3)}$ which is 4, of the original matrix $\boldsymbol{A}$.

3rd row of $\boldsymbol{PA}$ equals $\boldsymbol{p(3)}$ $-$th row of $\boldsymbol{A}$

✍ Why is this true?

✍ What can you can of the $\boldsymbol{j}$-th column of $\boldsymbol{AP}$?

✍ What is the matrix $\boldsymbol{PA}$ when

$$\boldsymbol{P = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix} \quad A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 0 & -1 & 2 \\ -3 & 4 & -5 & 6 \end{pmatrix}} \ ?$$

Back to elementary matrices

➤ Do the elementary matrices $\boldsymbol{E_1, E_2, ..., E_{n-1}}$ (including permutations) change linear independence of the columns?

✍ Prove: If $\boldsymbol{u, v, w}$ (3 columns of $\boldsymbol{A}$) are independent then the columns $\boldsymbol{E_1u, E_1v, E_1w}$ are independent where $\boldsymbol{E_1}$ is an elementary matrix (elimination matrix or a permutation matrix).

➤ So: (*Very important*) Elimination operations (Gaussian elimination, Gauss-Jordan, reduction to echelon form, and to rref) preserve the linear independence of the columns.

➤ Consequence: Gaussian elimination with partial pivoting cannot fail when the columns of $\boldsymbol{A}$ are linearly independent.

Conclusion: When $\boldsymbol{A}$ has independent columns, the linear system $\boldsymbol{Ax = b}$ always has a unique solution.

Proof: Only way in which GE can fail is that the algorithm reaches a system with a matrix U like the one in left hand side of:

$$\begin{bmatrix} x & x & x & x & x & x \\ & x & x & x & x & x \\ & & x & x & x & x \\ & & & 0 & x & x \\ & & & 0 & x & x \\ & & & 0 & x & x \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

➤ Then the columns of original A must be dependent.

➤ Can find a solution to $Ux = 0$ with x not a zero vector: Take $x_4 = 1$ and $x_i = 0$ for $i > k$. Then back-solve to get $x_{k-1}, x_{k-2}, \cdots, x_2, x_1$ (in example $k = 4$). Done.

➤ Therefore the columns of original A must be dependent too ■

✍ Argument is a little simpler with Gauss-Jordan. Get to the same conclusion using Gauss-Jordan with pivoting (permute below row k!).

Theorem: Let A be an $n \times n$ matrix. Then the columns of A are linearly independent iff its reduced echelon form is the identity matrix

$\boxed{\Longrightarrow}$ Only way in which the $rref(A) \neq I$ is by having at least one free variable. Form the augmented system $[A, 0]$. Set this free variable to one (other free var. to zero) and solve for the basic variables. Result: a nontrivial sol. to the systsm $Ax = 0$ $\rightarrow$ Contradiction

$\boxed{\Longleftarrow}$ If $rref(A) = I$ then columns of A are independent since the elementary operations do not alter linear dependence. ■

Theorem: Let A be an $n \times n$ matrix. Then A has independent columns if and only if A is invertible.

$\boxed{\Longrightarrow}$ From previous theorem, A can be reduced to the identity matrix with the reduced echelon form procedure. There are elementary matrices $E_1, E_2, \ldots., E_p$ such that

$$\underbrace{E_p E_{p-1} \cdots E_2 \underbrace{E_1 A}_{step1}}_{step2} = I$$

Call C the matrix $E_p E_{p-1} \cdots E_1$. Then $CA = I$. So A has a 'left-inverse'.

It also has a right inverse X (s.t. $AX = I$) because any system $Ax = b$ has a solution (Gaussian elimination will not fail since columns are linearly independent)

Therefore we can solve $Ax_i = e_i$, where e_i is the i-th col. if I. For $X = [x_1, x_2, \cdots, x_n]$ this gives $AX = I$.

Finally, $X = C$ because:

$$CA = I \to C\underbrace{(AX)}_{I} = X \to C = X$$

$\boxed{\Longleftarrow}$ Let A be invertible. Its columns are lin. independent if (by definition) $Ax = 0$ implies $x = 0$ - this is trivially true as can be seen by multiplying $Ax = 0$ to the left by A^{-1}. ■

Q: Can we now prove the Existence Theorem?

Existence Theorem. The 4 following statements are equivalent

(1) $\boldsymbol{A}$ is invertible

(2) The columns of $\boldsymbol{A}$ are linearly independent

(3) The Span of the columns of $\boldsymbol{A}$ is $\mathbb{R}^n$

(4) rref($\boldsymbol{A}$) is the identity matrix

- We have proved (1) iff (2) and also (2) iff (4)
- Easy to show (2) $\rightarrow$ (3) and then (3) $\rightarrow$ (4)
- ✍ Is this enough to prove theorem?
- The most important result to remember is:

$\boldsymbol{A}$ invertible $\Leftrightarrow \boldsymbol{rref(A) = I} \Leftrightarrow$ cols($\boldsymbol{A}$) Lin. independ.

Proof:

(3) $\rightarrow$ (4). As was seen before – (3) implies that there is a pivot in every row. Since the matrix is $n \times n$ the only possible rref echelon matrix of this type is I.

(2) $\rightarrow$ (3) Proof by contradiction. Assume A has linearly independent columns. And assume that some system $Ax = b$ does not have a solution. Then A, b will have a reduced row echelon form in which b will become a pivot. So there is a zero row in the A part of the echelon matrix.. This means we have at least a free variable - So systems $Ax = 0$ will have nontrivial solutions $\rightarrow$ contradiction. ■

Computing the inverse

Q: How do I compute the inverse of a matrix A?

A: Two common strategies [not necessarily the best]

- Using the reduced row echelon form
- Solving the n systems $Ax = e_i$ for $i = 1, \cdots, n$

How to use the echelon form?

➤ Could record the product of the E_i's as suggested by one of the previous theorems $\longrightarrow$ Too complicated!

- Instead perform the echelon form on the augmented matrix

$$[A, I]$$

- Assuming A is invertible result is of the form

$$[I, C]$$

- The inverse is C.

✍ Explain why.

✍ What will happen if A is not invertible?

Example: Compute the inverse of $\begin{bmatrix} 0 & \frac{1}{2} & -1 \\ \frac{1}{2} & \frac{3}{4} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{4} & \frac{3}{2} \end{bmatrix}$

Solution. First form the augmented matrix

$$\left[\begin{array}{ccc|ccc} 0 & \frac{1}{2} & -1 & 1 & 0 & 0 \\ \frac{1}{2} & \frac{3}{4} & \frac{1}{2} & 0 & 1 & 0 \\ \frac{1}{2} & -\frac{1}{4} & \frac{3}{2} & 0 & 0 & 1 \end{array}\right]$$

➤ Then get reduced echelon form:

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 5 & -2 & 4 \\ 0 & 1 & 0 & -2 & 2 & -2 \\ 0 & 0 & 1 & -2 & 1 & -1 \end{array}\right]$$

Inverse is

$$C = \begin{bmatrix} 5 & -2 & 4 \\ -2 & 2 & -2 \\ -2 & 1 & -1 \end{bmatrix}$$