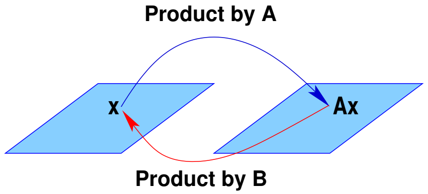
INVERSE OF A MATRIX [2.2]

The inverse of a matrix: Introduction

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

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



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

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

➤ Recall that $I_n x = x$ for all x.

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

$$A^{-1}A = I_n$$

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

$$AA^{-1} = I_n$$

➤ Finding an inverse to A is not always possible. When it is we say that the matrix 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

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

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

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

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

➤ Recall: A is one-to-one iff its columns are linearly independent.

✍ Show: If 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 $(A^T)^{-1} \equiv A^{-T}$

Elementary matrices

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

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

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

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

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

➤ Matrices like E (elementary elimination matrix) and 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:

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

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

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

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

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

➤ The first transformation ($row_2 := row_2 - \frac{1}{2} \times row_1$) 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 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

➤ 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

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

In words: *the i-th row of PA is row number $p(i)$ of A.*

➤ What does this mean?

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

3rd row of PA equals $p(3)-$th row of A

Why is this true?

What can you say of the j-th column of AP?

What is the matrix PA when

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

Back to elementary matrices

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

Prove: If u, v, w (3 columns of A) are independent then the columns E_1u, E_1v, E_1w are independent where 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.

➤ This will help us establish the main results on inverses of matrices

Existence of the inverse and related properties

We are now prepared to prove the following theorem.

Existence Theorem. The 4 following statements are equivalent

(1) A is invertible

(2) The columns of A are linearly independent

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

(4) rref(A) is the identity matrix

1 2 3 4

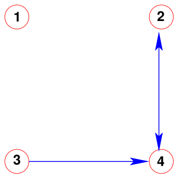
(2) ↔ (4).

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

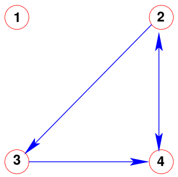
⟹ Only way in which the $rref(A) \neq I$ is by having at least one free variable. Form the augmented system $[A, 0]$. Set this free variable to one (other free var. to zero) and solve for the basic variables. Result: a nontrivial sol. to the systsm $Ax = 0$ → Contradiction

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

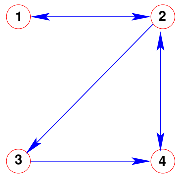
(**) Let A an $n \times n$ matrix with independent columns and $b \in \mathbb{R}^n$ a right-hand side. Apply rref to $[A, b]$. What do A and b become? [Hint: use result of 1st part of proof above]. Consequence?



Proof: *(3) → (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.



Proof: *(2) → (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 → contradiction. ■



(2) ↔ (1)

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

⟹ 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 $E_p E_{p-1} \cdots E_2 E_1 A = I$ (Step 1: left-multuply A by E_1; Step 2: left-multuply result by E_2; etc..)

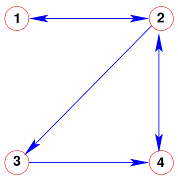
Call C the matrix $E_p E_{p-1} \cdots E_1$. Then $CA = I$. So A has a 'left-inverse'.

➤ It also has a right inverse X (s.t. $AX = I$) because any system $Ax = b$ has a solution (See exercise ✍ (**) seen earlier).

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, $X = C$. Indeed $CA = I \rightarrow C(AX) = X$ (because $AX = I$). So $C = X$.

⟸ 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: Is the Existence Theorem proved?
A: Yes.

➤ Here is what you need to remember:

A invertible ⇔ $rref(A) = I$ ⇔ cols(A) Lin. independ

⇕

cols(A) Span $\mathbb{R}^n$

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 → Too complicated!

- Instead get the reduced echelon form of 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

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

Example of application: Classical Crypto

- Idea of cryptography: A mapping from some space to itself.

Encoding = applying the mapping.

Decoding = applying the inverse mapping.

- Simple example: Hill's cipher [linear]

WIl describe a simplification of the scheme

- Associate a number to every letter [e.g., 0–25]:
  A $\rightarrow$ 0; B $\rightarrow$ 1 ; C $\rightarrow$ 2 ; ....; Z $\rightarrow$ 25
- 1st step: translate message with these numbers.

Example: "BUY GOOGLE TODAY"

Translates to (note: '26' is for space)
1, 20, 24, 26, 6, 14, 14, 6, 11, 4, 26, 13 14, 22, 26

- 2nd step: Put that into a matrix of size $3 \times ??$

$$\text{Message} = X = \begin{bmatrix} 1 & 26 & 14 & 4 & 14 \\ 20 & 6 & 6 & 26 & 22 \\ 24 & 14 & 11 & 13 & 26 \end{bmatrix}$$

- 3rd step: Scramble message with Encoding matrix:

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

➤ This means multiply X by A to get the encoded message:

$$Y = AX = \begin{bmatrix} -159 & -152 & -104 & -142 & -212 \\ 44 & 20 & 17 & 39 & 48 \\ 160 & 178 & 118 & 146 & 226 \end{bmatrix}$$

... which is transmitted.

• 4th step: The receiver must now decode the message by applying the inverse of A which in this case is:

$$A^{-1} = \begin{bmatrix} 1 & 0 & 1 \\ 4 & 4 & 3 \\ -4 & -3 & -3 \end{bmatrix}$$

➤ Decoded message : $X = A^{-1}Y = \begin{bmatrix} 1 & 26 & 14 & 4 & 14 \\ 20 & 6 & 6 & 26 & 22 \\ 24 & 14 & 11 & 13 & 26 \end{bmatrix}$

➤ To break the code all you need is the mapping A

➤ Then compute A^{-1} (easy)

➤ Mapping is linear and so it is easy to find A.

✍ How would you proceed to get A? [Recall Practice exercise sets 8 & 9]

✍ How many messages do you need to intercept to do this? Is the message “Hello” enough? How about “Good morning”?

➤ Nonlinear codes are much harder to break..

➤ Hill’s cipher adds a ‘modulo’ operation by translating Y into letters first. For example, 226 will become Mod(226,25)=1 which gives ‘B’ .... more complicated.