Matrices, sub-espacios vectoriales y transformaciones lineales

Credit: This notes are 100% from chapter 3 of the book entitled *Linear Algebra. A Modern Introduction* by David Poole. Thomson. Australia. 2006.

Introduction

We will use matrices to solve system of linear equations (SLE). Let us transform the following SLE

$$x - y - z = 2 \tag{1}$$
$$3x - 3y + 2z = 16 \tag{2}$$
$$2x - y + z = 9 \tag{3}$$

in matrix notation,

$$\begin{bmatrix} 1 & -1 & -1 \\ 3 & -3 & 2 \\ 2 & -1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 \\ 16 \\ 9 \end{bmatrix} \tag{4}$$

This equation can be think as a certain type of function that act on a vector, by transforming it into another vector. This matrix transformation will begin to play a key role in our study of linear algebra. We will review their algebraic properties.

Matrix Operations

About matrix and its size: a matrix is a rectangular array of numbers called the entries, of elements, of the matrix. The size of a matrix is a description of the number of rows and columns it has. A matrix is called $m \times n$ if it has m rows and n columns. A $1 \times m$ matrix is called a row matrix, and an $n \times 1$ matrix is called a column matrix.

Some special matrices: If $m = n$ A is called a square matrix. A square matrix whose non diagonal entries are all zero is called a diagonal matrix. A diagonal matrix, all of whose diagonal entries are the same is called a scalar matrix. If the scalar on the diagonal is 1, the scalar matrix is called an identity matrix. The $n \times n$ identity matrix is denoted by I_n.

Equal: tow matrices A and B are equal if they have the same size and if their corresponding entries are equal, i.e. if $A = [a_{ij}]_{m\times n}$ and $B = [b_{ij}]_{r\times s}$, then $A = B$ if and only if $m = r$ and $n = s$ and $a_{ij} = b_{ij}$ for all i and j.

Sum: $A = [a_{ij}]$ and $B = [b_{ij}]$ are $m \times n$ matrices, their sum $A + B$ is the $m \times n$ matrix $A + B = [a_{ij}] + [b_{ij}]$. If A and B are not the same size, then $A + B$ is not defined.

Scalar multiple: If A is an $m \times n$ matrix and c is a scalar, then the scalar multiple cA is the $m \times n$ matrix $cA = c[a_{ij}] = [ca_{ij}]$.

Difference: The matrix $(-1)A$ is written as $-A$ and called the negative of A. If A and B are the same size, then $A - B = A + (-B)$.

Zero matrix: A matrix all of whose entries are zero is called zero matrix and denoted by O, and: (i) $A + O = A = O + A$, (ii) $A - A = O = -A + A$.

Product: If A is an $m \times n$ matrix and B is an $n \times r$ matrix, then the product $C = AB$ is an $m \times r$ matrix. The (i, j) entry of the product is computed as

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} = \sum_{k=1}^{n} a_{ik}b_{kj} \tag{5}$$

SLE: Every linear system with m equations and n variables can be written in the form

$$A\bar{x} = \bar{b} \tag{6}$$

with A a matrix $m \times n$, $\bar{x}$ a vector of $n \times 1$ and $\bar{b}$ a vector of $m \times 1$.

Pick out a single row or column: The pre or post multiplication of the matrix A by an appropriate versor, gives a row or column of A:

$$\bar{e}_2 A = [0\ 1] \begin{bmatrix} 4 & 2 & 1 \\ 0 & 5 & -1 \end{bmatrix} = [0\ 5\ -1] \tag{7}$$

$$A\bar{e}_3 = \begin{bmatrix} 4 & 2 & 1 \\ 0 & 5 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \tag{8}$$

Theorem 3.1: Let A be a matrix $m \times n$, $\bar{e}_i$ a standard unit vector of $1 \times m$, and $\bar{e}_j$ a standard unit vector of $n \times 1$, then

a) $\bar{e}_i A$ is the ith row of A and

b) $A\bar{e}_j$ is the jth column of A.

Proof: Proof of (b) (the proof of (a) is left as and exercise). If the vectors $\bar{a}_1, \cdots, \bar{a}_n$ are the columns of A, then

$$A\bar{e}_j = 0\bar{a}_1 + 0\bar{a}_2 + \cdots + 0\bar{a}_{j-1} + 1\bar{a}_j + 0\bar{a}_{j+1} + 0\bar{a}_n = \bar{a}_j \tag{9}$$

or

$$A\bar{e}_j = \begin{bmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \cdots & a_{mj} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{bmatrix} \tag{10}$$

Blocks: It will often be convenient to regard a matrix as being composed of a number of smaller sub matrices. By introducing vertical and horizontal lines into a matrix, we can partition it into blocks. Then the original matrix has as entries smaller matrices. Example

$$A = \begin{bmatrix} 1 & 0 & 0 & 2 & -1 \\ 0 & 1 & 0 & 1 & 3 \\ 0 & 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 1 & 7 \\ 0 & 0 & 0 & 7 & 2 \end{bmatrix} \tag{11}$$

$$= \left[\begin{array}{ccc|cc} 1 & 0 & 0 & 2 & -1 \\ 0 & 1 & 0 & 1 & 3 \\ 0 & 0 & 1 & 4 & 0 \\ \hline 0 & 0 & 0 & 1 & 7 \\ 0 & 0 & 0 & 7 & 2 \end{array}\right] \tag{12}$$

$$= \begin{bmatrix} I & B \\ O & C \end{bmatrix} \tag{13}$$

When matrices are being multiplied, there is often an advantage to be gained by viewing them as partitioned matrices, since some of the sub matrices may be the identity of the nil matrix.

Matrix-column representation: Suppose A is $m \times n$ and B is $n \times r$, so the product AB exist. Let us partition B in term of its columns vectors, as $B = [\bar{b}_1|\cdots|\bar{b}_r]$ then

$$AB = A[\bar{b}_1|\cdots|\bar{b}_r] \tag{14}$$

$$= [A\bar{b}_1|\cdots|A\bar{b}_r] \tag{15}$$

Example,

$$A = \begin{bmatrix} 1 & 3 & 2 \\ 0 & -1 & 1 \end{bmatrix} \tag{16}$$

$$B = \begin{bmatrix} 4 & -1 \\ 1 & 2 \\ 3 & 0 \end{bmatrix} \tag{17}$$

then

$$AB = [A\bar{b}_1|A\bar{b}_2] \tag{18}$$

$$= \left[\begin{array}{c|c} 13 & 5 \\ 2 & -2 \end{array}\right] \tag{19}$$

Linear combination: The matrix-column representation of AB allows us to write each column of AB as a linear combination of the column of A with entries from B as the coefficients. Example

$$A\bar{b}_1 = \begin{bmatrix} 1 & 3 & 2 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix} \tag{20}$$

$$= 4\begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 3 \\ -1 \end{bmatrix} + 3\begin{bmatrix} 2 \\ 1 \end{bmatrix} \tag{21}$$

Row-matrix representation: similarly, if A is $m \times n$ and B is $n \times r$, then the product AB exist. If we partition A in terms of its row vectors, as $\left[\begin{array}{c} A_1 \\ \hline A_2 \\ \hline \vdots \\ \hline A_m \end{array}\right]$ then,

$$AB = \left[\begin{array}{c} A_1 \\ \hline A_2 \\ \hline \vdots \\ \hline A_m \end{array}\right] B \tag{22}$$

$$= \left[\begin{array}{c} A_1B \\ \hline A_2B \\ \hline \vdots \\ \hline A_mB \end{array}\right] \tag{23}$$

Outer product(column-row representation): If we partition the matrix A, $m \times n$ in columns and B, $n \times r$ in row, then the product AB exists and gives

$$A = [\bar{a}_1|\bar{a}_2|\cdots|\bar{a}_n] \tag{24}$$

$$B = \left[\begin{array}{c} \bar{b}_1 \\ \hline \bar{b}_2 \\ \hline \vdots \\ \hline \bar{b}_n \end{array}\right] \tag{25}$$

then

$$AB = [\bar{a}_1|\bar{a}_2|\cdots|\bar{a}_n] \left[\begin{array}{c} \bar{b}_1 \\ \hline \bar{b}_2 \\ \hline \vdots \\ \hline \bar{b}_n \end{array}\right] \tag{26}$$

$$= \bar{a}_1\bar{b}_1 + \bar{a}_2\bar{b}_2 + \cdots + \bar{a}_n\bar{b}_n \tag{27}$$

Each individual term in the above sum is a matrix, i.e. $\bar{a}_i\bar{b}_i$ is the product of $m \times 1$ and a $1 \times r$ matrix, thus $\bar{a}_i\bar{b}_i$ is a $m \times r$ matrix. The product $\bar{a}_i\bar{b}_i$ are called outer product, and Eq. (26) is called the outer product expansion of AB.

Matrix powers: Let A be a matrix of dimension $n \times n$, we define

- $A^k = AA\cdots A$ (k-times)
- $A^0 = I$

with the following properties

- $A^rA^s = A^{r+s}$
- $(A^r)^s = A^{rs}$

Transpose: the transpose of a $m \times n$ matrix A is the $n \times m$ matrix A^T obtained by interchanging the rows and columns of A. That is, the ith column of A^T is the ith row of A for all i or $(A^T)_{ij} = A_{ji}$ for all i and j.

Scalar product: The dot product of two vectors $\bar{u} = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}$ and $\bar{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$ can be written in the form

$$\begin{aligned} \bar{u} \cdot \bar{v} &= u_1 v_1 + \cdots + u_n v_n && (28) \\ &= [u_1 \ \cdots \ u_n] \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} && (29) \\ &= \bar{u}^T \bar{v} && (30) \end{aligned}$$

Symmetric matrix: A square matrix A is symmetric if $A^T = A$, that is, if A is equal to its own transpose or $A_{ij} = A_{ji}$.

Matrix Algebra

Theorem 3.2: Algebraic properties of matrix addition and scalar multiplication. Let A, B and C be matrices of the same size and let c and d be scalars. Then

a) $A + B = B + A$ Commutativity

b) $(A + B) + C = A + (B + C)$ Associativity

c) $A + O = A$

d) $A + (-A) = O$

e) $c(A + B) = cA + cB$ Distributivity

f) $(c + d)A = cA + dA$ Distributivity

g) $c(dA) = (cd)A$

h) $1A = A$

Linear combination: If $A_1, A_2, \cdots, A_k$ are matrices of the same size and $c_1, \cdots, c_k$ are scalars (called coefficients), we may form the linear combination

$$c_1 A_1 + c_2 A_2 + \cdots + c_k A_k \tag{31}$$

Example: given the matrices A_1, A_2, and A_3,

$$A_1 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad A_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad A_3 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \tag{32}$$

find if B and C linear combinations of A_1, A_2 and A_3, if B and C are give by,

$$B \ = \ \begin{bmatrix} 1 & 4 \\ 2 & 1 \end{bmatrix} \quad C = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \tag{33}$$

We write the linear combination $c_1 A_1 + c_2 A_2 + c_3 A_3$, equal it to B and C.

For B we get

$$c_1 \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} + c_2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} c_3 \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 4 \\ 2 & 1 \end{bmatrix} \tag{34}$$

then

$$\begin{bmatrix} c_2 + c_3 & c_1 + c_3 \\ -c1 + c_3 & c_2 + c_3 \end{bmatrix} = \begin{bmatrix} 1 & 4 \\ 2 & 1 \end{bmatrix} \tag{35}$$

For which we built the system

$$\begin{aligned} c_2 + c_3 &= 1 && (36) \\ c_1 + c_3 &= 4 && (37) \\ -c1 + c_3 &= 2 && (38) \\ c_2 + c_3 &= 1 && (39) \end{aligned}$$

Next, we built the extended matrix and applied Gauss or Gauss-Jordan reduction.

The solutions are

- For B: $c_1 = 1, c_2 = -2, c_3 = 3$, i.e. B is the following linear combination of A_i: $A_1 - 2A_2 + 3A_3 = B$
- For C: C is not a LC since the system is inconsistent

Span: the span of the above matrices A_1, A_2 and A_3 is given by all matrices of the form $\begin{bmatrix} w & x \\ y & z \end{bmatrix}$ such that,

$$\begin{aligned} c_1A_1 + c_2A_2 + c_3A_3 &= c_1 \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} + c_2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} c_3 \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} && (40) \\ &= \begin{bmatrix} c_2 + c_3 & c_1 + c_3 \\ -c1 + c_3 & c_2 + c_3 \end{bmatrix} && (41) \\ &= \begin{bmatrix} w & x \\ y & z \end{bmatrix} && (42) \end{aligned}$$

which defined the following SE

$$\begin{aligned} c_2 + c_3 &= w && (43) \\ c_1 + c_3 &= x && (44) \\ -c1 + c_3 &= y && (45) \\ c_2 + c_3 &= z && (46) \end{aligned}$$

with the following extended matrix

$$\left[\begin{array}{ccc|c} 0 & 1 & 1 & w \\ 1 & 0 & 1 & x \\ -1 & 0 & 1 & y \\ 0 & 1 & 1 & z \end{array}\right] \tag{47}$$

The row reduction process gives

$$\left[\begin{array}{ccc|c} 0 & 0 & 0 & \frac{x-y}{2} \\ 0 & 1 & 0 & -\frac{x+y}{2} + w \\ 0 & 0 & 1 & \frac{x+y}{2} \\ 0 & 0 & 0 & w - z \end{array}\right] \tag{48}$$

the only restriction comes from the last column $w = z$, A_1, A_2, A_3 expand all matrices of the form $\begin{bmatrix} w & x \\ y & w \end{bmatrix}$ with x, y, w arbitrary.
Using this information we could have been concluded, that the above matrix B is a linear combination of A_1, A_2, A_3, while C is not.

Linearly independent (LI): We say that matrices $A_1, A_2, \cdots, A_k$ of the same size are LI if the only solution of the equation

$$c_1A_1 + c_2A_2 + \cdots + c_kA_k = O \tag{49}$$

is the trivial one: $c_1 = \cdots = c_k = 0$.

Linearly dependent (LD): We say that matrices $A_1, A_2, \cdots, A_k$ of the same size are LD if there are nontrivial coefficients $c_1, \cdots, c_k$ that satisfy

$$c_1A_1 + c_2A_2 + \cdots + c_kA_k = O \tag{50}$$

Example: determine whether the above matrices A_1, A_2, A_3 are LI.

$$\begin{aligned} c_1A_1 + c_2A_2 + \cdots + c_kA_k &= O & (51) \\ c_1 \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} + c_2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} c_3 \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} &= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} & (52) \end{aligned}$$

we have to solve the homogeneous LS. From the previous calculations we have

$$\left[\begin{array}{ccc|c} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \tag{53}$$

then, $c_1 = c_2 = c_3 = 0$, i.e. they are LI.

Multiplication: The multiplication is not commutative $AB \neq BA$ (if $B \neq A$). Then $(A + B)^2 \neq A^2 + 2AB + B^2$ if A does not commute with B, i.e. $AB \neq BA$.
If $A^2 = O$ does not implies $A = O$.

Theorem 3.3: Properties of Matrix multiplication. Let A, B and C be matrices (whose sizes are such that the indicated operations can be performed) and let k be a scalar. Then

a) $A(BC) = (AB)C$ Associativity

b) $A(B + C) = AB + AC$ Left distributivity

c) $(A + B)C = AC + BC$ Right distributivity

e) $k(AB) = (kA)B = A(kB)$

f) $I_mA = A = AI_n$ if A is $m \times n$ Multiplicative identity

Theorem 3.4: Properties of the Transpose. Let A and B be matrices (whose sizes are such that the indicated operations can be performed) and let k be a scalar. Then

a) $(A^T)^T = A$

b) $(A+B)^T = A^T + B^T$

c) $(kA)^T = k(A^T)$

d) $(AB)^T = B^T A^T$

e) $(A^r)^T = (A^T)^r$ for all non negative integers r

Properties (b) and (d) can be generalized (assuming that the size of the matrices are such that all of the operations can be performed),

$$\begin{aligned} (A_1 + \cdots + A_k)^T &= A_1^T + \cdots + A_k^T && (54) \\ (A_1 \cdots A_k)^T &= A_k^T \cdots A_1^T && (55) \end{aligned}$$

Theorem 3.5:

a) if A is a square matrix, then $A + A^T$ is a symmetric matrix

b) For any matrix A, AA^T and $A^T A$ are symmetric matrices

Proof: We prove (a) and leave proving (b) as Exercise 34. Let us consider

$$\begin{aligned} (A + A^T)^T &= A^T + (A^T)^T = A^T + A && (56) \\ &= A + A^T && (57) \end{aligned}$$

we have used the properties of the transpose and the commutativity of matrix addition. Thus $A + A^T$ is equal to its own transpose and so, by definition, is symmetric.

The inverse of a Matrix

Let us return to the matrix description of the SLE

$$A\bar{x} = \bar{b} \qquad (58)$$

and let us look for ways to use matrix algebra to solve the system.

If there would exist a matrix A' such that $A'A = I$, then we could formally get

$$\begin{aligned} A\bar{x} &= \bar{b} && (59) \\ A'(A\bar{x}) &= A'\bar{b} && (60) \\ \bar{x} &= A'\bar{b} && (61) \end{aligned}$$

This A' could be the inverse of A. In this section we will answer the following two questions:

1. How can we know when a matrix has an inverse?

2. If a matrix does have an inverse, how can we find it?

Inverse: If A is an $n \times n$ matrix, an inverse of A is an $n \times n$ matrix A' with the property that $AA' = I$ and $A'A = I$ where $I = I_n$ is the $n \times n$ identity matrix. If such an A' exists, then A is called invertible. Even though we have seen that matrix multiplication is not, in general, commutative, A' (if it exists) must satisfy $A'A = AA'$.

Theorem 3.6: Unique inverse. If A is an invertible matrix, then its inverse is unique and it is denoted as A^{-1}.
Proof: Suppose that A has two inverses, say, A' and A''. Then

$$\begin{aligned} AA' &= I = A'A && (62) \\ AA'' &= I = A''A && (63) \end{aligned}$$

thus,

$$A' = A'I = A'(AA'') = (A'A)A'' = IA'' = A'' \quad (64)$$

hence, $A' = A''$, and the inverse is unique.

Theorem 3.7: SLE solution. If A is an invertible $n \times n$ matrix, then the SLE given $A\bar{x} = \bar{b}$ has the unique solution $\bar{x} = A^{-1}\bar{b}$ for any $\bar{b}$ in $\mathbb{R}^n$.
Proof: We are asked to prove two things: that $A\bar{x} = \bar{b}$ has a solution and that it has only one solution.
Let us first verify that $\bar{x} = A^{-1}\bar{b}$ is a solution of $A\bar{x} = \bar{b}$:

$$\begin{aligned} A\bar{x} &= A(A^{-1}\bar{b}) = (AA^{-1})\bar{b} = I\bar{b} && (65) \\ &= \bar{b} && (66) \end{aligned}$$

In order to show that it is unique, let us suppose that there is another solution $\bar{y}$, then $A\bar{y} = \bar{b}$. Working we have

$$\begin{aligned} A^{-1}(A\bar{y}) &= A^{-1}\bar{b} && (67) \\ (A^{-1}A)\bar{y} &= A^{-1}\bar{b} && (68) \\ \bar{y} &= A^{-1}\bar{b} && (69) \\ &= \bar{x} && (70) \end{aligned}$$

thus, $\bar{y}$ is the same solution as before, and therefore the solution is unique.

Theorem 3.8: Inverse of a matrix of dimension 2. If $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, then A is invertible if $ad - bc \neq 0$, in which case

$$\begin{aligned} A^{-1} &= \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} && (71) \\ &= \frac{1}{detA}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} && (72) \end{aligned}$$

Where the magnitude $ad - bc = detA$ is called determinant of A. Then, a 2×2 matrix A is invertible if and only if $detA \neq 0$. If $ad - bc = 0$, then A is not invertible.
Proof: See demonstration in book, pag. 163, if needed.

Example: use the inverse of the coefficients matrix to solve the linear system

$$x + 2y = 3 \tag{73}$$
$$3x + 4y = -2 \tag{74}$$

The coefficient matrix is $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ whose inverse is $\begin{bmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix}$ then $\bar{x} = A^{-1}\bar{b}$ gives

$$\bar{x} = \begin{bmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 3 \\ -2 \end{bmatrix} \tag{75}$$
$$= \begin{bmatrix} -8 \\ \frac{11}{2} \end{bmatrix} \tag{76}$$
$$\tag{77}$$

Theorem 3.9: Properties of invertible matrices

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

$$(A^{-1})^{-1} = A \tag{78}$$

b) If A is an invertible matrix and c is a nonzero scalar, then cA is an invertible matrix and

$$(cA)^{-1} = \frac{1}{c}A^{-1} \tag{79}$$

c) If A and B are invertible matrices of the same size, the AB is invertible and

$$(AB)^{-1} = B^{-1}A^{-1} \tag{80}$$

d) If A is an invertible matrix, the A^T is invertible and

$$(A^T)^{-1} = (A^{-1})^T \tag{81}$$

e) If A is an invertible matrix, then A^n is invertible for all non negative integers n and

$$(A^n)^{-1} = (A^{-1})^n \tag{82}$$

Proof: See book if needed, pag. 165 and forward.

Property (c) generalizes to

$$(A_1 \cdots A_n)^{-1} = A_n^{-1} \cdots A_1^{-1} \tag{83}$$

Property (e) allows us to define negative integer powers of an invertible matrix,

$$A^{-n} = (A^{-1})^n = (A^n)^{-1} \tag{84}$$

Example: solve the following matrix equation for X assuming that the matrices involved are such that all of the indicated operations are defined:

$$A^{-1}(BX)^{-1} = (A^{-1}B^3)^2 \tag{85}$$

The solution is

$$X = B^{-4}AB^{-3} \tag{86}$$

Elementary matrix: An elementary matrix is any matrix that can be obtained by performing an elementary row operation (review row operations) on an identity matrix. Since there are three types of elementary row operations, there are three corresponding types of elementary matrices. The property of an elementary matrix is the following: let us assume that the elementary matrix E was obtained from the identity through the transformation R. The product of E times any matrix A produces the same transformation R in the matrix A.

Examples:

$$E_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{87}$$

$$E_2 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{88}$$

and be the arbitrary matrix A

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{bmatrix} \tag{89}$$

Then

$$E_1 A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 3a_{21} & 3a_{22} & 3a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{bmatrix} \tag{90}$$

and

$$E_2 A = \begin{bmatrix} a_{31} & a_{32} & a_{33} \\ a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \\ a_{41} & a_{42} & a_{43} \end{bmatrix} \tag{91}$$

Theorem 3.10: Let E be the elementary matrix obtained by performing an elementary row operation on I_n. If the same elementary row operation is performed on an $n \times n$ matrix A, the result is the same as the matrix EA.
Comment 1: elementary matrices can provides some valuable insights into invertible matrices and the solution of systems of equations.
Comment 2: as every elementary row operation can be reversed, every elementary matrices is invertible.
Comment 3: the inverse of and elementary matrix is another elementary matrix.
Comment 4: the inverse of and elementary matrix is the elementary matrix which results from reversing the mathematical operation which gives origin to the elementary matrix.

Example 1: The elementary matrix E_1

$$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \tag{92}$$

corresponds to $R_2 \leftrightarrow R_3$, which ins undone by doing $R_2 \leftrightarrow R_3$ again. It also can be seen as the elementary matrix which reverse the original process, i.e. $R_3 \leftrightarrow R_2$ from the identity matrix. Thus, $E_1^{-1} = E_1$ (check this by showing that $E_1^2 = E_1 E_1 = I$).

Example 2: The elementary matrix E_2

$$E_2 = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{93}$$

corresponds to $R_2 \to R_2 - 2R_1$, which ins undone by doing the transformation $R_2 \to R_2 + 2R_1$ to the identity matrix

$$E_2^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{94}$$

Thus, $E_2 E_2^{-1} = I$

$$\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{95}$$

Example 3: The elementary matrix E_3

$$E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{96}$$

corresponds to $R_2 \to -2R_2$, which ins undone by doing (from the identity matrix) $R_2 \to -\frac{1}{2}R_2$, then

$$E_3^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{97}$$

then

$$E_3 E_3^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{98}$$

Theorem 3.11: Each elementary matrix is invertible, and its inverse is an elementary matrix of the same type.

Theorem 3.12: The Fundamental theorem (FT) of invertible matrices. Version 1 of 5 This theorem gives a set of equivalent characterizations of what it means for a matrix to be invertible. Let A be an $n \times n$ matrix. The following statements are equivalent:

a) A is invertible.

b) $A\bar{x} = \bar{b}$ has a unique solution for every $\bar{b}$ in $\mathbb{R}^n$.

c) $A\bar{x} = 0$ has only the trivial solution.

d) The reduced row echelon form of A is I_n.

e) A is a product of elementary matrices.

Proof: See book, pag. 171.

Example 3.19: Express A as product of elementary matrices, where

$$A = \begin{bmatrix} 2 & 3 \\ 1 & 3 \end{bmatrix} \tag{99}$$

Then,

$$\begin{aligned} A &= \begin{bmatrix} 2 & 3 \\ 1 & 3 \end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_2} \begin{bmatrix} 1 & 3 \\ 2 & 3 \end{bmatrix} \xrightarrow{R_2 \to R_2 - 2R_1} \begin{bmatrix} 1 & 3 \\ 0 & -3 \end{bmatrix} \\ &\xrightarrow{R_1 \to R_1 + R_2} \begin{bmatrix} 1 & 0 \\ 0 & -3 \end{bmatrix} \xrightarrow{R_2 \to -\frac{1}{3}R_2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2 \end{aligned} \tag{100}$$

These transformation corresponds to the following elementary matrices

$$E_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad E_2 = \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix} \quad E_3 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \quad E_4 = \begin{bmatrix} 1 & 0 \\ 0 & -\frac{1}{3} \end{bmatrix} \tag{101}$$

such that

$$E_4 E_3 E_2 E_1 A = I \tag{102}$$

then

$$A = (E_4 E_3 E_2 E_1)^{-1} I \tag{103}$$
$$= E_1^{-1} E_2^{-1} E_3^{-1} E_4^{-1} I \tag{104}$$
$$\tag{105}$$

Ejercicio: Demonstrate that the right hand side of the previous expression give A. This also shows that the inverse of A is

$$A^{-1} = E_4 E_3 E_2 E_1 \tag{106}$$

since

$$E_4 E_3 E_2 E_1 A = I \tag{107}$$

Ejercicio: Calcular la matriz inversa de A.

Ejercicio: Comparar con la solución obtenida por otro compañero. Responeder las siguientes dos preguntas: (i) puede ser la descomposición diferente? (ii) puede ser la inversa diferente?

Comment 1: since the sequence of elementary row operations that transform A into I is not unique, neither is the representation of A as a product of elementary matrices.

Comment 2: despite the previous comment A^{-1} must be the same. Check this statement comparing with another elementary transformation.

Theorem 3.13: Let A be a square matrix. If B is a square matrix such that either $AB = I$ or $BA = I$, then A is invertible and $B = A^{-1}$.

Proof: Suppose $BA = I$. Consider the equation $A\bar{x} = \bar{0}$.
First, let us left multiply by B, then, $BA\bar{x} = B\bar{0}$ then $I\bar{x} = \bar{0}$, i.e. the unique solution is $\bar{x} = \bar{0}$. From the equivalence of (c) and (a) in the Fundamental Theorem (FT), A is invertible, which means that A^{-1} exist and satisfies $AA^{-1} = I = A^{-1}A$.
Second, let us right multiply $BA = I$ by A^{-1}, then $BAA^{-1} = IA^{-1} \Rightarrow B = A^{-1}$.
The proof in the case $AB = I$ is propose as the Exercise 41.

Theorem 3.14: efficient method of computing inverses. Let A be a square matrix. If a sequence of elementary row operations reduces A to I, then the same sequence of elementary row operations transform I into A^{-1}.
Proof: If A is row equivalent to I, then we can achieve the reduction by left-multiplying by a sequence $E_1, \cdots, E_k$ of elementary matrices. Therefore, $E_k \cdots E_1 A = I$. Setting $B = E_k \cdots E_1$ gives $BA = I$. By Theorem 3.13, A is invertible and $A^{-1} = B$. Now applying the same sequence of elementary row operations to I is equivalent to left-multiplying I by $E_k \cdots E_1 = B$. Then $E_k \cdots E_1 I = BI = B = A^{-1}$.
Thus, I is transformed into A^{-1} by the same sequence of elementary row operations, i.e. one can simultaneously transform A into I and I into A^{-1}. Is not this great?

The Gauss-Jordan method for computing the inverse. We construct the super-augmented matrix $[A|I]$. Theorem 3.14 shows that if A is row equivalent to I (which, by the FT (d)$\Leftrightarrow$(a), means that A is invertible), then elementary row operations will yield $[A|I] \to [I|A^{-1}]$. If A cannot be reduced to I, then the FT guarantees us that A is not invertible.

Example 3.30: Find the inverse of $A = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 2 & 4 \\ 1 & 3 & -3 \end{bmatrix}$

Starting from

$$[A|I] = \left[\begin{array}{ccc|ccc} 1 & 2 & -1 & 1 & 0 & 0 \\ 2 & 2 & 4 & 0 & 1 & 0 \\ 1 & 3 & -3 & 0 & 0 & 1 \end{array}\right] \tag{108}$$

we should get

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 9 & -\frac{3}{2} & -5 \\ 0 & 1 & 0 & -5 & 1 & 3 \\ 0 & 0 & 1 & -2 & \frac{1}{2} & 1 \end{array}\right] = [I|A^{-1}] \tag{109}$$

Example 3.31: Verify that the following matrix is not invertible $A = \begin{bmatrix} 2 & 1 & -4 \\ -4 & -1 & 6 \\ -2 & 2 & -2 \end{bmatrix}$

We should get

$$\left[\begin{array}{ccc|ccc} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & 1 & -3 & 2 & 1 & 0 \\ 0 & 0 & 0 & -5 & -3 & 1 \end{array}\right] \tag{110}$$

The LU factorization

Matrix factorization is any representation of a matrix as a product of two or more other matrices.

Example 3.33 Let $A = \begin{bmatrix} 2 & 1 & 3 \\ 4 & -1 & 3 \\ -2 & 5 & 5 \end{bmatrix}$

Applies the following rows reductions,

$$\begin{aligned} A &= \begin{bmatrix} 2 & 1 & 3 \\ 4 & -1 & 3 \\ -2 & 5 & 5 \end{bmatrix} \xrightarrow{R_2-2R_1;R_3+R_1} \begin{bmatrix} 2 & 1 & 3 \\ 0 & -3 & -3 \\ 0 & 6 & 8 \end{bmatrix} \\ &\xrightarrow{R_3+2R_2} \begin{bmatrix} 2 & 1 & 3 \\ 0 & -3 & -3 \\ 0 & 0 & 2 \end{bmatrix} = U \end{aligned} \tag{111}$$

The three elementary matrices are

$$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \quad E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix} \tag{112}$$

Hence

$$E_3E_2E_1A = U \Rightarrow A = E_1^{-1}E_2^{-1}E_3^{-1}U \tag{113}$$

with

$$E_1^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad E_2^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \quad E_3^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix} \tag{114}$$

then

$$\begin{aligned} A &= E_1^{-1}E_2^{-1}E_3^{-1}U \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -2 & 1 \end{bmatrix} U \\ &= LU \end{aligned} \tag{115}$$

Then the matrix A can be factored as

$$A = LU \tag{116}$$

with L the unit lower triangular matrix and U an upper triangular matrix.

LU factorization:

Let A be a square matrix. A factorization as $A = LU$, where L is unit lower triangular and U is upper triangular, is called an LU factorization of A.

Comment 1: In our previous example no row interchanges were needed in the row reduction of A.

Comment 2: If a zero had appeared in a pivot position at any step, we would have had to swap rows to get a nonzero pivot and the L would be no longer unit lower triangular.

Comment 3: Inverses and products of unit lower triangular matrices are also unit lower triangular (See Exercises 29 and 30).

Comment 4: The notion of an LU factorization can be generalized to non square matrices by requiring U to be a matrix in row echelon form. (See Exercises 13 and 14.)

Comment 5: Some books define and LU factorization of a square matrix A to be any factorization $A = LU$, were L is lower triangular and U is upper triangular. Notice that the 'unit' condition over L was relaxed.

Theorem 3.15: If A is a square matrix that can be reduced to row echelon form without using any row interchanges, the A has an LU factorization.

Application: To see why the LU factorization is useful, consider a linear system $A\bar{x} = \bar{b}$, where the coefficient matrix has an LU factorization $A = LU$. Then

$$\begin{aligned} A\bar{x} &= \bar{b} && (117)\\ LU\bar{x} &= L(U\bar{x}) = \bar{b} && (118)\\ & && (119) \end{aligned}$$

Let us define $\bar{y} = U\bar{x}$ then

- First we find $\bar{y}$ by solving $L\bar{y} = \bar{b}$ using forward substitution.
- Then we solve $U\bar{x} = \bar{y}$ for $\bar{x}$ using back substitution.

Each of these two linear systems is straightforward to solve because the coefficient matrices L and U are both triangular.

Example: Solve the system $A\bar{x} = \bar{b}$ with $A = \begin{bmatrix} 2 & 1 & 3 \\ 4 & -1 & 3 \\ -2 & 5 & 5 \end{bmatrix}$ and $\bar{b} = \begin{bmatrix} 1 \\ -4 \\ 9 \end{bmatrix}$

From example 115 we have

$$\begin{aligned} A &= E_1^{-1}E_2^{-1}E_3^{-1}U \\ &= LU \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -2 & 1 \end{bmatrix} U && (120) \end{aligned}$$

then

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -2 & 1 \end{bmatrix} \qquad (121)$$

and $L\bar{y} = \bar{b}$ implies

$$\begin{aligned} y_1 &= 1 && (122)\\ 2y_1 + y_2 &= -4 && (123)\\ -y_1 - 2y_2 + y_3 &= 9 && (124) \end{aligned}$$

which give $y_1 = 1$, $y_2 = -6$, and $y_3 = -2$.

Then $U\bar{x} = \bar{y}$ implies

$$\begin{aligned} 2x_1 + x_2 + 3x_3 &= 1 && (125)\\ -3x_2 - 3x_3 &= -6 && (126)\\ 2x_3 &= -2 && (127) \end{aligned}$$

which give $x_3 = -1$, $x_2 = 3$, and $x_1 = \frac{1}{2}$.

An easy way to find LU factorization:

L can be computed directly from the row reduction process. This is valid assuming that A can be reduced to row echelon form without using any row interchanges. Then the entire row reduction process can be done using only elementary row operations of the form $R_i - kR_j$, where k is the multiplier factor.
In the example 115 we used the following elementary row operations:

$$R_2 - 2R_1 \Rightarrow k = 2 \tag{128}$$
$$R_3 + R_1 \Rightarrow k = -1 \tag{129}$$
$$R_3 + 2R_1 \Rightarrow k = -2 \tag{130}$$

The unit triangular matrix L is build from the multiplier using the index i as labeling the row and j the column,

$$R_2 - 2R_1 \Rightarrow k = 2 \Rightarrow L_{21} = 2 \tag{131}$$
$$R_3 + R_1 \Rightarrow k = -1 \Rightarrow L_{31} = -1 \tag{132}$$
$$R_3 + 2R_2 \Rightarrow k = -2 \Rightarrow L_{32} = -2 \tag{133}$$

then

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -2 & 1 \end{bmatrix} \tag{134}$$

Comment: in applying this method, it is important to note that the elementary row operations $R_i - kR_j$ must be performed from top to bottom within each column (using the diagonal entry as the pivot), and column by column from left to right.

Example 3.35: Find (in class) the LU factorization of $A = \begin{bmatrix} 3 & 1 & 3 & -4 \\ 6 & 4 & 8 & -10 \\ 3 & 2 & 5 & -1 \\ -9 & 5 & -2 & -4 \end{bmatrix}$ we should get $L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 1 & \frac{1}{2} & 1 & 0 \\ -3 & 4 & -1 & 1 \end{bmatrix}$ $U = \begin{bmatrix} 3 & 1 & 3 & -4 \\ 0 & 2 & 2 & -2 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & -4 \end{bmatrix}$

Theorem 3.16: If A is an invertible matrix that has an LU factorization, then L and U are unique.
Proof: See book, pag. 184.

The $P^T LU$ factorization

If during the reduction we realize that row interchanges is necessary, we first interchange the rows in the original matrix and then we proceed to the factorization. If again, during the factorization it is found a new row interchange has to be performed, we return to the already row-permuted matrix and perform the new row interchange. Then we restart the factorization. This has to be done as many times is needed. Since the row interchange could be done in different ways, the $P^T LU$ will be not unique; however, one P has been determined, L and U are unique.

Let us assume we have to make k permutation over the original matrix A

$$A' = P_k P_{k-1} \cdots P_2 P_1 A = PA \tag{135}$$

and then we make the LU factorization to the resulting matrix A'

$$A' = LU \tag{136}$$

In order to get the factorization of A we need the inverse of P,

$$PA = A' = LU \Rightarrow A = P^{-1}LU = P^T LU \tag{137}$$

where the following theorem have been used.

Theorem 3.17: If P is a permutation matrix, then $P^{-1} = P^T$.
Proof: See book, pag. 185.

Definition: Let A be a square matrix. A factorization of A as $A = P^T LU$, where P is a permutation matrix, L is unit lower triangular, and U is upper triangular, is called a $P^T LU$ factorization of A.

Example 3.36: Find the $P^T LU$ of $A = \begin{bmatrix} 0 & 0 & 6 \\ 1 & 2 & 3 \\ 2 & 1 & 4 \end{bmatrix}$.
We will need two row interchanges: $R_1 \leftrightarrow R_2$ and $R_2 \leftrightarrow R_3$:

$$\begin{aligned} P &= P_2 P_1 \quad (138) \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (139) \\ &= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \quad (140) \end{aligned}$$

then

$$P^T = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \tag{141}$$

The matrix to be factorized is

$$PA = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 4 \\ 0 & 0 & 6 \end{bmatrix} \tag{142}$$

Making $R_2 - 2R_1$ we get U,

$$PA \to \begin{bmatrix} 1 & 2 & 3 \\ 0 & -3 & -2 \\ 0 & 0 & 6 \end{bmatrix} = U \tag{143}$$

and L comes from the multiplier of the single transformation, $L_{21} = 2$,

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{144}$$

Then, the factorization of A is given by $A = P^T LU$ with P^T, L and U as given above,

$$A = P^T LU = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 0 & -3 & -2 \\ 0 & 0 & 6 \end{bmatrix} \tag{145}$$

Theorem 3.18: Every square matrix has a (non-unique) $P^T LU$ factorization.

Subspaces, Basis, Dimension, and Rank

Subspace of $\mathbb{R}^n$: A subspace of $\mathbb{R}^n$ is any collection S of vectors in $\mathbb{R}^n$ such that

1. The zero vector $\bar{0}$ is in S
2. S is closed under addition, i.e., if $\bar{u}$ and $\bar{v}$ are in S, then $\bar{u}+\bar{v}$ is in S
3. S is closed under scalar multiplication, i.e. if $\bar{u}$ is in S and c is a scalar, the $c\bar{u}$ is in S.

Examples: every line and plane through the origin in $\mathbb{R}^3$ is a subspace of $\mathbb{R}^3$: Let $\mathcal{P}$ be a plane through the origin with direction vectors $\bar{v}_1$ and $\bar{v}_2$. Hence $\mathcal{P} = span(\bar{v}_1, \bar{v}_2)$. The zero vector $\bar{0}$ is in $\mathcal{P}$, since $\bar{0} = 0\bar{v}_1 + 0\bar{v}_2$. Now let

$$\begin{aligned} \bar{u} &= a_1\bar{v}_1 + a_2\bar{v}_2 && (146)\\ \bar{w} &= b_1\bar{v}_1 + b_2\bar{v}_2 && (147)\\ & && (148) \end{aligned}$$

be vectors in $\mathcal{P}$. Then,

$$\bar{u}+\bar{w} = (a_1+b_1)\bar{v}_1 + (a_2+b_2)\bar{v}_2 \tag{149}$$

thus, $\bar{u}+\bar{w}$ is in $\mathcal{P}$.

Now let c be a scalar, then

$$c\bar{u} = (ca_1)\bar{v}_1 + (ca_2)\bar{v}_2 \tag{150}$$

thus, $c\bar{u}$ is in $\mathcal{P}$.

Then, $\mathcal{P}$ is a subspace of $\mathbb{R}^3$.

Comment: the fact that $\bar{v}_1$ and $\bar{v}_2$ are vectors in $\mathbb{R}^3$ played no role at all in the verification of the properties.

Theorem 3.19: Let $\bar{v}_1, \cdots, \bar{v}_k$ be vectors in $\mathbb{R}^n$. Then span$(\bar{v}_1, \cdots, \bar{v}_k)$ is a subspace of $\mathbb{R}^n$. We will refer to span$(\bar{v}_1, \cdots, \bar{v}_k)$ as the subspace spanned by $(\bar{v}_1, \cdots, \bar{v}_k)$.
Proof: See book, pag. 191.

Example 3.38: Let us see that the set of all vectors $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ with $x = 3y$ and $z = -2y$ forms a subspace of $\mathbb{R}^3$.

We have $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3y \\ y \\ -2y \end{bmatrix} = y\begin{bmatrix} 3 \\ 1 \\ -2 \end{bmatrix}$ with y arbitrary and the vector $\bar{v} = \begin{bmatrix} 3 \\ 1 \\ -2 \end{bmatrix}$ in $\mathbb{R}^3$.
Then, by the theorem 3.19 the span$(\bar{v})$ is a subspace of $\mathbb{R}^3$. This subspace is formed by lines through the origin with direction $\bar{v}$.

Definition: Let A be an $m \times n$ matrix

1. The row space of A is the subspace row(A) of $\mathbb{R}^n$ spanned by the rows of A
2. The column space of A is the subspace col(A) of $\mathbb{R}^m$ spanned by the columns of A

Theorem 3.20: Let B be any matrix that is row equivalent to a matrix A. Then row(B)=row(A).
Proof: The matrix A can be transformed into B by a sequence of row operations. Then the row of B are linear combinations of the rows of A. Hence, linear combinations of the rows of B can be express as linear combinations of the rows of A, which implies that row(B) $\subseteq$ row(A). By reversing these row operations transform B into A. Therefore, the above argument implies row(A) $\subseteq$ row(B).
Combining these two results, we have row(B) = row(A).

Theorem 3.21: Let A be an $m\times n$ matrix and let N be the set of solutions of the homogeneous linear system $A\bar{x} = \bar{0}$ (with $\bar{x}$ $n \times 1$ and $\bar{0}$ $m \times 1$). Then N is a subspace of $\mathbb{R}^n$.
Proof:
(i) Since $A\bar{0}_n = \bar{0}_m$, $\bar{0}_n$ is in N.
(ii) Let $\bar{u}$ and $\bar{v}$ be in N, i.e. $A\bar{u} = \bar{0}$ and $A\bar{v} = \bar{0}$, then $A(\bar{u} + \bar{v}) = \bar{0}$, i.e. $\bar{u} + \bar{v}$ is in N
(iii) Finally, for any scalar c, we have $A(c\bar{u}) = \bar{0}$, then $c\bar{u}$ is in N.
From (i), (ii) and (iii) it follows that N is a subspace of $\mathbb{R}^n$.

Null space: Let A be an $m \times n$ matrix. The null space of A is the subspace of $\mathbb{R}^n$ consisting of solutions of the homogeneous linear system $A\bar{x} = \bar{0}$. It is denoted by null(A).

Theorem 3.22: Let A be a matrix whose entries are real numbers. For any system of linear equations $A\bar{x} = \bar{b}$, exactly one of the following is true:

a. There is no solutions

b. There is a unique solution

c. There are infinitely many solutions

Proof: We have to prove that if (a) and (b) are not true, then (c) is the only other possibility. See book, pag. 195.

Basis

A basis for a subspace S of $\mathbb{R}^n$ is a set of vectors in S that

1. spans S and
2. is linearly independent

A subspace can have more than one basis. For examples, the vectors $\left\{\begin{bmatrix} 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \end{bmatrix}\right\}$ form a basis for $\mathbb{R}^2$ as do the canonical basis, $\left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right\}$

Standard or canonical basis: Unit vectors $\bar{e}_1, \cdots, \bar{e}_n$ in $\mathbb{R}^n$ are LI and span $\mathbb{R}^n$. They are called standard (canonical) basis.

Procedure to find the basis of a matrix.

We are interested to find the row basis, the column basis and the null basis of a given matrix.

Let us consider the following matrix $A = \begin{bmatrix} 1 & 1 & 3 & 1 & 6 \\ 2 & -1 & 0 & 1 & -1 \\ -3 & 2 & 1 & -2 & 1 \\ 4 & 1 & 6 & 1 & 3 \end{bmatrix}$

Basis for row(A)

Let us reduce the matrix to its echelon form (even when we reduce the pivot to one it is not necessary). The reduced matrix is

$$R = \begin{bmatrix} 1 & 0 & 1 & 0 & -1 \\ 0 & 1 & 2 & 0 & 3 \\ 0 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{151}$$

By Theorem 3.20, row(A)=row(R). From the reduce matrix R one can see that the first three rows are LI. Then a basis for the row space of A is

$$\{(1\;0\;1\;0\;-1), (0\;1\;2\;0\;3), (0\;0\;0\;1\;4)\} \tag{152}$$

Basis for the column space of A

One method is to transpose the matrix and repeat the procedure used for finding the row basis. By transposing the resulting basis we get a column basis for col(A).
Alternatively, we obtain the column basis by taken the vectors from the matrix A which correspond to the column of the reduced matrix R which correspond to the heads (pivots). En el ejemplo anterior ellos corresponden a las columnas: 1ra, 2da y 4ta de A.
The justification of this procedure is as follows: considering the system $A\bar{x} = 0$, the reduction from A to R represents a dependence relation among the columns of A. Since the elementary row operations do not affect the solution set, if A is row equivalent to R, the columns of A have the same dependence relationships as the columns of R.
Let as called $\bar{a}_i$ the columns of A and $\bar{r}_i$ the columns of R. By inspection we find that

$$\bar{r}_3 = \bar{r}_1 + 2\bar{r}_2 \tag{153}$$
$$\bar{r}_5 = -\bar{r}_1 + 3\bar{r}_2 + 4\bar{r}_4 \tag{154}$$

From the previous argument the same relation exist for the columns of A (verify this)

$$\bar{a}_3 = \bar{a}_1 + 2\bar{a}_2 \tag{155}$$
$$\bar{a}_5 = -\bar{a}_1 + 3\bar{a}_2 + 4\bar{a}_4 \tag{156}$$

and then the columns $\bar{a}_1, \bar{a}_2, \bar{a}_4$ form a basis for the col(A),

$$\left\{ \begin{bmatrix} 1 \\ 2 \\ -3 \\ 4 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -2 \\ 1 \end{bmatrix} \right\} \tag{157}$$

Notice that col(A) and col(R) do not expand the same space (see this by inspection). Then col(A)$\neq$col(R) which implies that $\bar{r}_1, \bar{r}_2, \bar{r}_4$ can no be a basis for col(A), i.e. **elementary row operations change the column space**.

Basis for the null space

We have to find the solution of the homogeneous system $A\bar{x} = 0$ from the augmented matrix of A, $[A|\bar{0}] = \left[\begin{array}{rrrrr|r} 1 & 1 & 3 & 1 & 6 & 0 \\ 2 & -1 & 0 & 1 & -1 & 0 \\ -3 & 2 & 1 & -2 & 1 & 0 \\ 4 & 1 & 6 & 1 & 3 & 0 \end{array}\right]$

From the previous calculation we have

$$[R|\bar{0}] = \left[\begin{array}{ccccc|c} 1 & 0 & 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 & 3 & 0 \\ 0 & 0 & 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right] \tag{158}$$

then,

$$x_1 + x_3 - x_5 = 0 \tag{159}$$

$$x_2 + 2x_3 + 3x_5 = 0 \tag{160}$$

$$x_4 + 4x_5 = 0 \tag{161}$$

Since the leading 1s are in columns 1, 2 and 4, we solve for x_1, x_2 and x_4. Let us renamed $x_3 = s$ and $x_5 = t$, then,

$$\bar{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = s \begin{bmatrix} -1 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} 1 \\ -3 \\ 0 \\ -4 \\ 1 \end{bmatrix} \tag{162}$$

Then, the vectors

$$\left\{ \begin{bmatrix} -1 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -3 \\ 0 \\ -4 \\ 1 \end{bmatrix} \right\} \tag{163}$$

span null(A), and since they are LI, they form a basis for null(A).

Exercise for the student in class: Find the row, column and null basis for the following matrix

$$\begin{bmatrix} 3 & 2 & 0 \\ -1 & 1 & -5 \\ 5 & 3 & 1 \end{bmatrix} \tag{164}$$

Dimension and Rank

Theorem 3.23: The basis theorem. Let S be a subspace of $\mathbb{R}^n$. Then any two bases for S have the same number of vectors.
Proof: See book, pag. 200.

Dimension: if S is a subspace of $\mathbb{R}^n$, then the number of vectors in a basis for S is called the dimension of S, denoted dim$S$.

Comment: The zero vector $\bar{0}$ by itself is always a subspace of $\mathbb{R}^n$. It is defined dim$\bar{0}$ to be 0.

Theorem 3.24: The row and column spaces of a matrix A have the same dimension.
Proof: Let R be the reduced row echelon form of A. By theorem 3.20, row(A)=row(R), then dim(row(A))=dim(row(R))=number of nonzero rows of R=number of leading 1s of R=r.
Now col(A)$\neq$col(R), but the columns of A and R have the same dependence relationships.

Therefore, dim(col(A))=dim(col(R)).
Since, there are r leading 1s, R has r columns that are standard unit vectors, $\bar{e}_1, \cdots, \bar{e}_r$. They are LI. Thus, dim(col(R))=r.
It follows that dim(row(A))=r=dim(col(A)).

Rank: The rank of a matrix A is the dimension of its row and column spaces and is denoted by rank(A).

Theorem 3.25: For any matrix A, rank(A^T)=rank(A).
Proof: We have rank(A^T)=dim(col(A^T))=dim(row(A^T))=rank(A).

Nullity: The nullity of a matrix A is the dimension of its null space and is denoted by nullity(A).

Theorem 3.26: Rank Theorem If A is an $m \times n$ matrix, then rank(A)+nullity(A)=n.
This theorem rephrasing the Rank Theorem 2.2.
Proof: Let R be the reduced row echelon form of A, and suppose that rank(A)=r. Then R has r leading 1s, so there are r leading variables and $n-r$ free variables in the solution to $A\bar{x} = \bar{0}$. Since dim(null(A))=n-r, we have rank(A)+nullity(A)=r+(n-r)=n.

Exercise for the student in class: (Example 3.51, pag. 283) Fin de nullity of the matrices M, M^T, N, and N^T where $M = \begin{bmatrix} 2 & 3 \\ 1 & 5 \\ 4 & 7 \\ 3 & 6 \end{bmatrix}$ and $N = \begin{bmatrix} 2 & 1 & -2 & -1 \\ 4 & 4 & -3 & 1 \\ 1 & 7 & 1 & 8 \end{bmatrix}$
Solution:

- rank(M)=2 $\Rightarrow$ nullity(M)=2-rank(M)=0;
- rank(M^T)=2 $\Rightarrow$ nullity(M^T)=4-rank(M^T)=2;
- rank(N)=2 $\Rightarrow$ nullity(N)=4-rank(N)=2;
- rank(N^T)=2 $\Rightarrow$ nullity(N^T)=3-rank(N^T)=1.

Theorem 3.27: Fundamental Theorem (FT) of Invertible Matrices. Version 2 of 5 . Let A be an $n \times n$ matrix. The following statements are equivalent:

From Version 1

a. A is invertible.

b. $A\bar{x} = \bar{b}$ has a unique solution for every $\bar{b}$ in $\mathbb{R}^n$.

c. $A\bar{x} = 0$ has only the trivial solution.

d. The reduced row echelon form of A is I_n.

e. A is a product of elementary matrices.

New statements

f. rank(A)=n

g. nullity(A)=0

h. The column vectors of A are LI

i. The column vectors of A span $\mathbb{R}^n$

j. The column vectors of A form a basis for $\mathbb{R}^n$

k. The row vectors of A are LI

l. The row vectors of A span $\mathbb{R}^n$

m. The row vectors of A form a basis for $\mathbb{R}^n$

Proof: See book, pag. 204.

This theorem is a labor-saving device. Version 1 allowed us to cut in half the work needed to check that two square matrices are inverses. It also simplifies the task of showing that certain sets of vectors are bases for $\mathbb{R}^n$. Indeed, when we have a set of n vectors in $\mathbb{R}^n$, that set will be a basis for $\mathbb{R}^n$ if either of the necessary properties of linear independence or spanning set is true.

Example 3.52: Let us checks that the three following vectors from a basis for $\mathbb{R}^3$

$$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 4 \\ 9 \\ 7 \end{bmatrix} \tag{165}$$

From (f) or (j) of the FT, the vectors will form a basis if and only if a matrix with these vectors as its columns (or rows) has rank 3. Making row reduction we get

$$\begin{bmatrix} 1 & -1 & 4 \\ 2 & 0 & 9 \\ 3 & 1 & 7 \end{bmatrix} \to \begin{bmatrix} 1 & -1 & 4 \\ 0 & 2 & 1 \\ 0 & 0 & -7 \end{bmatrix} \tag{166}$$

then rank(A)=3. Then the original vectors for a basis for $\mathbb{R}^3$.

Theorem 3.28: Let A be an $m \times n$ matrix. Then

a. rank(A^TA)=rank(A)

b. Then $n \times n$ matrix A^TA is invertible if and only if rank(A)=n.

Proof:
(a) Since (A^TA) is $n \times n$ matrix, it has the same number of columns as A. The Rank Theorem states

$$rank(A) + nullity(A) = n = rank(A^TA) + nullity(A^TA) \tag{167}$$

Hence, to show that rank(A)=rank(A^TA), it is equivalent to show that

$$nullity(A) = nullity(A^TA) \tag{168}$$

We will do so by establishing that the null spaces of A and A^T are the same.
To this end, let $\bar{x}$ be in null(A) so that $A\bar{x} = 0$. Then $A^TA\bar{x} = 0$, then $\bar{x}^TA^TA\bar{x} = \bar{x}^T\bar{0} = 0$. Then,

$$(A\bar{x}) \cdot (A\bar{x}) = (A\bar{x})^T(A\bar{x}) = \bar{x}^TA^TA\bar{x} = 0 \tag{169}$$

and hence $A\bar{x} = \bar{0}$, by theorem 1.2(d). Therefore, $\bar{x}$ is in null(A), so null(A)=null(A^TA), as required.

(b) By the FT, the $n \times n$ matrix A^TA is invertible if and only if rank(A^TA)=n. But, by (a) this is so if and only if rank(A)=n.

Theorem 3.29: Let S be a subspace of $\mathbb{R}^n$ and let $\mathcal{B} = \{\bar{v}_1, \cdots, \bar{v}_k\}$ be a basis for S. For every vector $\bar{v}$ in S, there is exactly one way to write $\bar{v}$ as a linear combination of the basis vectors in $\mathcal{B}$:

$$\bar{v} = c_1\bar{v}_1 + \cdots + c_k\bar{v}_k \tag{170}$$

Proof: See book, pag. 206.

Coordinates. Let S be a subspace of $\mathbb{R}^n$ and let $\mathcal{B} = \{\bar{v}_1, \cdots, \bar{v}_k\}$ be a basis for S. Let $\bar{v}$ be a vector in S, and write $\bar{v} = c_1\bar{v}_1 + \cdots + c_k\bar{v}_k$. Then $c_1, \cdots, c_k$ are called the coordinates of $\bar{v}$ with respect to $\mathcal{B}$, and the column vector

$$[\bar{v}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_k \end{bmatrix} \tag{171}$$

is called the coordinate vector of $\bar{v}$ with respect to $\mathcal{B}$.

Example 3.53 Let $\mathcal{E} = \{\bar{e}_1, \bar{e}_2, \bar{e}_3\}$ be the standard basis for $\mathbb{R}^3$. The coordinate vector of $\bar{v} = \begin{bmatrix} 2 \\ 7 \\ 4 \end{bmatrix}$ with respect to the basis $\mathcal{E}$ is the same vector:

$$\bar{v} = \begin{bmatrix} 2 \\ 7 \\ 4 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 7 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 4 \end{bmatrix} = 2\bar{e}_1 + 7\bar{e}_2 + 4\bar{e}_3 \tag{172}$$

Example 3.54 Let the vectors $\bar{v}_1$, $\bar{v}_2$ and $\bar{w}$ be in the subspace S of $\mathbb{R}^3$, with $\mathcal{B} = \{\bar{v}_1, \bar{v}_2\}$ a basis for S, where

$$\bar{v}_1 = \begin{bmatrix} 3 \\ -1 \\ 5 \end{bmatrix}, \bar{v}_2 = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} \tag{173}$$

calculate the coordinate vector for the vector $\bar{w} = \begin{bmatrix} 0 \\ -5 \\ 1 \end{bmatrix}$ in the basis $\mathcal{B}$.

Solution: From $\bar{w} = c_1\bar{v}_1 + c_2\bar{v}_2$ we have from the first and second lines

$$3c_1 + 2c_2 = 0 \tag{174}$$

$$-c_1 + c_2 = -5 \tag{175}$$

then $c_2 = -3$ and $c_1 = 2$. The coordinate vector results $[\bar{w}]_{\mathcal{B}} = \begin{bmatrix} 2 \\ -3 \end{bmatrix}$

Introduction to linear transformation

Matrices can be used to transform vectors, acting as a type of function of the form $\bar{w} = T\bar{v}$. Such matrix transformation leads to the concept of a linear transformation. We are only interested in transformations that are compatible with the vector operations of addition and scalar multiplication.

Example: Let $A = \begin{bmatrix} 1 & 0 \\ 2 & -1 \\ 3 & 4 \end{bmatrix}$ $\bar{v} = \begin{bmatrix} x \\ y \end{bmatrix}$ then, we can says that A transform $\bar{v}$ in $\mathbb{R}^2$ into $\bar{w}$ in $\mathbb{R}^3$:

$$A\bar{v} = \begin{bmatrix} 1 & 0 \\ 2 & -1 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ 2x - y \\ 3x + 4y \end{bmatrix} = \bar{w} \tag{176}$$

We denote this transformation by T_A, such that

$$T_A \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ 2x - y \\ 3x + 4y \end{bmatrix} \tag{177}$$

Transformation: More generally, a transformation (or mapping or function) T from $\mathbb{R}^n$ to $\mathbb{R}^m$ is a rule that assigns to each vector $\bar{v}$ in $\mathbb{R}^n$ a unique vector $T\bar{v}$ in $\mathbb{R}^m$.

Domain-Codomain: The domain of T is $\mathbb{R}^n$, and the codomain of T is $\mathbb{R}^m$. This is indicated by writing $T : \mathbb{R}^n \to \mathbb{R}^m$.

Image: For a vector $\bar{v}$ in the domain of T, the vector $T(\bar{v})$ in the codomain is called the image of $\bar{v}$ under (the action of) T.

Range: The set of all possible images $T(\bar{v})$ is called the range of T.

Example: For the transformation $T_A \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ 2x - y \\ 3x + 4y \end{bmatrix}$ we have

- Domain: $\mathbb{R}^n$
- Codomain: $\mathbb{R}^m$
- Image of $\bar{v}$: $T_A(\bar{v})$
- Range: $T_A \begin{bmatrix} x \\ y \end{bmatrix} = x \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + y \begin{bmatrix} 0 \\ -1 \\ 4 \end{bmatrix}$ then, image is the plane span$(\bar{v}_1, \bar{v}_2)$=col(A), i.e. the column space of A (dibujar diagrama) by $\bar{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, $\bar{v}_2 = \begin{bmatrix} 0 \\ -1 \\ 4 \end{bmatrix}$

Linear Transformation (LT): A transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is called a linear transformation if

1. $T(\bar{u} + \bar{v}) = T(\bar{u}) + T(\bar{v})$ for all $\bar{u}$ and $\bar{v}$ in $\mathbb{R}^n$ and
2. $T(c\bar{v}) = cT(\bar{v})$ for all $\bar{v}$ in $\mathbb{R}^n$ and all scalar c.

Example: Let us shows that the transformation $T_A \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ 2x - y \\ 3x + 4y \end{bmatrix}$ is a linear transformation.

Exercise for the student in class: Show that the previous transformation is linear using instead $T(a\bar{u} + b\bar{v}) = aT(\bar{u}) + bT(\bar{v})$.

Getting the matrix transformation from a transformation: From the transformation $T \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ 2x - y \\ 3x + 4y \end{bmatrix}$ we can get a matrix which makes the same job as the transformation:

$$T \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ 2x - y \\ 3x + 4y \end{bmatrix} = x \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + y \begin{bmatrix} 0 \\ -1 \\ 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & -1 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \tag{178}$$

then, $T = T_A$, where $A = \begin{bmatrix} 1 & 0 \\ 2 & -1 \\ 3 & 4 \end{bmatrix}$

Theorem 3.30: Let A be an $m \times n$ matrix. Then the matrix transformation $T_A : \mathbb{R}^n \to \mathbb{R}^m$ defined by $T_A(\bar{x}) = A\bar{x}$, with $\bar{x} \in \mathbb{R}^n$ is a linear transformation.
Prof: Let $\bar{u}$ and $\bar{v}$ be vectors in $\mathbb{R}^n$ and let c be a scalar. Then

$$\begin{aligned} T_A(\bar{u} + \bar{v}) &= A(\bar{u} + \bar{v}) = A\bar{u} + A\bar{v} = T_A\bar{u} + T_A\bar{v} && (179) \\ T_A(c\bar{v}) &= A(c\bar{v}) = cA\bar{v} = cT_A(\bar{v}) && (180) \end{aligned}$$

Example: 90^o **counterclockwise rotation.** Let $R : \mathbb{R}^2 \to \mathbb{R}^2$ be the transformation that rotates each point 90^o counterclockwise about the origin (hacer diagrama). Show that R is a LT: $R \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -y \\ x \end{bmatrix}$. Its associate matrix is $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, then R is a matrix transformation and from the Theorem 3.30 R is a LT.

Exercise for the student in class: The transformation which transform each point in $\mathbb{R}^2$ into its reflection with respect to x axis (hacer diagrama) is a LT: $F \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ -y \end{bmatrix}$. Its associate matrix is $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, then F is a matrix transformation and from the Theorem 3.30 F is a LT.

Extracting a single column: Observe that if we multiply a matrix A by standard basis vectors $\bar{e}_i$, we obtain the column i of the matrix A:

$$A\bar{e}_1 = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \end{bmatrix}, \quad A\bar{e}_2 = \begin{bmatrix} a_{12} \\ a_{22} \\ a_{32} \end{bmatrix} \tag{181}$$

Theorem 3.31: Standard matrix Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a LT. Then T is a matrix transformation (MT). More specifically, $T = T_A$, where A is the $m \times n$ matrix $A = [T(\bar{e}_1)|T(\bar{e}_2)| \cdots |T(\bar{e}_n)]$.
Proof: Let $\bar{e}_1, \cdots, \bar{e}_n$ be the standard basis vectors in $\mathbb{R}^n$ and let $\bar{x}$ be a vector in $\mathbb{R}^n$. We can write $\bar{x} = x_1\bar{e}_1 + \cdots + x_n\bar{e}_n$. Then

$$T(\bar{x}) \quad = \quad x_1T(\bar{e}_1) + \cdots + x_nT(\bar{e}_n) \tag{182}$$

because T is a LT. Each term $T(\bar{e}_i)$ is a vector of dimension m. We can interpret the above product as the product of a matrix with n columns $T(\bar{e}_i)$, then

$$\begin{aligned} T(\bar{x}) &= x_1T(\bar{e}_1) + \cdots + x_nT(\bar{e}_n) && (183) \\ &= [T(\bar{e}_1)| \cdots |T(\bar{e}_n)] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = A\bar{x} && (184) \end{aligned}$$

with $A = [T(\bar{e}_1)| \cdots |T(\bar{e}_n)]$ a matrix of order $m \times n$ called standard matrix of the linear transformation T.

Example: Rotation. (Example 3.58, pags. 214-215) Let us build the standard matrix for the rotation about the origin through an angle: $R_\theta : \mathbb{R}^2 \to \mathbb{R}^2$.
First we apply the transformation to $\bar{e}_1$ and $\bar{e}_2$ (hacer diagrama de pag. 215):

$$\begin{aligned} R_\theta(\bar{e}_1) &= R_\theta \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} cos\theta \\ sin\theta \end{bmatrix} && (185) \\ R_\theta(\bar{e}_2) &= R_\theta \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -sin\theta \\ cos\theta \end{bmatrix}, && (186) \end{aligned}$$

then we build the matrix,

$$R_\theta = \begin{bmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{bmatrix} \tag{187}$$

Exercise for the student in class: Applied the rotation transformation to the 90^o counterclockwise rotation and compare with the previous result.

Composition of LT

If $T : \mathbb{R}^m \to \mathbb{R}^n$ and $S : \mathbb{R}^n \to \mathbb{R}^p$ are LT, then we may follow T by S to form the composition of the two transformations, denoted $S \circ T$, where the codomain of T and the domain of S must match. The composite transformation $S \circ T$ goes from the domain of T to the codomain of S, $S \circ T : \mathbb{R}^m \to \mathbb{R}^p$, with $(S \circ T)(\bar{v}) = S(T(\bar{v}))$.

Theorem 3.32: Let $T : \mathbb{R}^m \to \mathbb{R}^n$ and $S : \mathbb{R}^n \to \mathbb{R}^p$ be LT. Then $S \circ T : \mathbb{R}^m \to \mathbb{R}^p$ is a LT with standard matrices given by the product of the standard matrices of each transformation

$$[S \circ T] = [S][T] \tag{188}$$

where we used the notation $[T]$ for the standard matrix of the transformation T.
Proof: Let $[S] = A$ and $[T] = B$ with A of dimension $p \times n$ and B of dimension $n \times m$. If $\bar{v}$ is a vector in $\mathbb{R}^m$, then (mostrar detalles de las dimensiones)

$$(S \circ T)(\bar{v}) = S(T(\bar{v})) = S(B\bar{v}) = A(B\bar{v}) = (AB)\bar{v} = [S][T]\bar{v} = [S \circ T]\bar{v} \tag{189}$$

Example: (Example 3.60, pag. 218) Let us build the composite transformation $S \circ T$ for $T : \mathbb{R}^2 \to \mathbb{R}^3$ and $S : \mathbb{R}^3 \to \mathbb{R}^4$, with

$$T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ 2x_1 - x_2 \\ 3x_1 + 4x_2 \end{bmatrix} \tag{190}$$

$$S\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 2y_1 + y_3 \\ 3y_2 - y_3 \\ y_1 - y_2 \\ y_1 + y_2 + y_3 \end{bmatrix} \tag{191}$$

We have that the standard matrices are

$$[T] = \begin{bmatrix} 1 & 0 \\ 2 & -1 \\ 3 & 4 \end{bmatrix} \tag{192}$$

$$[S] = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 3 & -1 \\ 1 & -1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \tag{193}$$

then

$$[S \circ T] = [S][T] = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 3 & -1 \\ 1 & -1 & 0 \\ 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 2 & -1 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 5 & 4 \\ 3 & -7 \\ -1 & 1 \\ 6 & 3 \end{bmatrix} \tag{194}$$

Then,

$$(S \circ T)\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 5 & 4 \\ 3 & -7 \\ -1 & 1 \\ 6 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{195}$$

$$= \begin{bmatrix} 5x_1 + 4x_2 \\ 3x_1 - 7x_2 \\ -x_1 + x_2 \\ 6x_1 + 3x_2 \end{bmatrix} \tag{196}$$

Homework: Checks that $[S \circ T]\bar{x} = S(T(\bar{x}))$. Notice that the left hand side is a matrix product, while the right h.s. are transformations (not matrix product) as defined in (190) and (191). Make explicit the dimensions in each step.

Lectura Sugerida: Section Robotics, pag. 224. (Explicar)

Inverses of LT

Consider the effect of a 90^o counterclockwise rotation about the origin R_{90} followed by a 90^o clockwise rotation about the origin R_{-90}. This two transformations leave every point in $\mathbb{R}^2$ unchanged. Then, $(R_{-90} \circ R_{90})(\bar{v}) = \bar{v}$ for every $\bar{v}$ in $\mathbb{R}^2$. It is also true $(R_{90} \circ R_{-90})(\bar{v}) = \bar{v}$. Then $(R_{-90} \circ R_{90})(\bar{v})$ and $(R_{90} \circ R_{-90})$ are identity transformation.

Identity transformation: An identity transformation $I_n : \mathbb{R}^n \to \mathbb{R}^n$ is the transformation which satisfies $I_n(\bar{v}) = \bar{v}$ for every $\bar{v}$ in $\mathbb{R}^n$.

Inverse transformation: Let S and T be LT from $\mathbb{R}^n$ to $\mathbb{R}^n$. Then S and T are inverse transformation is $S \circ T = I_n$ and $T \circ S = I_n$.

Invertible transformation: Due to the symmetry of the definition of inverse transformation we will say that S is the inverse of T and T is the inverse of S. We will say that S and T are invertible.

Standard matrix of the identity transformation: If S and T are inverse transformations, then $[S][T] = [S \circ T] = [I] = I$ where I is the identity matrix. It is also true $[T][S] = [T \circ S] = [I] = I$.

Theorem 3.33: Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be an invertible LT. Then its standard matrix $[T]$ is an invertible matrix, and $[T^{-1}] = [T]^{-1}$, where $[T]^{-1}$ means the inverse of the standard matrix of the transformation T!!.

Example 3.62: Let us find the standard matrix which rotates 60^o clockwise about the origin from the previous calculated standard matrix R_θ which rotates counterclockwise about the origin.
From the previous example about the rotation transformation we have

$$[R_{60}] = \begin{bmatrix} cos(60) & -sin(60) \\ sin(60) & cos(60) \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix} \tag{197}$$

then (review how to calculate the inverse of 2×2 matrix)

$$[R_{-60}] = [([R_{60}])^{-1}] = \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}^{-1} = \begin{bmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix} \tag{198}$$

Homework: Check that $[R_{-60}]$ does what it is expected to do.

No inverse: From the Theorem of inversible matrices if the inverse of a matrix associate to a transformation has no inverse, then the transformation has no inverse.

Exercise for the student in class: Show that the transformation which projects onto the x-axis is not invertible.
Solution: it standard matrix is $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$. Then rank of this matrix is less than the number of columns, then, from the Theorem of invertible matrices it is not invertible. From the inverse of two by two matrix we know that it is not invertible because its determinant is nil.