

Systems of Equations

04/03/2025

What is the formula of a linear equation?

$$y = w \cdot x + b$$

↑
weight ↑ bias

wind speed → power output

$$y = w_1 \cdot x_1 + w_2 \cdot x_2 + b$$

↓
wind speed + temperature → power output

How do we formulate linear regression in ML? (as multiple terms)

$$w_1 \cdot x_1 + w_2 \cdot x_2 + \dots + w_n \cdot x_n + b = y$$

A dataset will have lots of rows, therefore ⇒

$$\begin{aligned} w_1 \cdot x_1^{(1)} + w_2 \cdot x_2^{(1)} + \dots + w_n \cdot x_n^{(1)} + b^{(1)} &= y^{(1)} \\ w_1 \cdot x_1^{(2)} + w_2 \cdot x_2^{(2)} + \dots + w_n \cdot x_n^{(2)} + b^{(2)} &= y^{(2)} \\ w_1 \cdot x_1^{(3)} + w_2 \cdot x_2^{(3)} + \dots + w_n \cdot x_n^{(3)} + b^{(3)} &= y^{(3)} \\ \vdots &\quad \vdots &\quad \vdots \\ w_1 \cdot x_1^{(m)} + w_2 \cdot x_2^{(m)} + \dots + w_n \cdot x_n^{(m)} + b^{(m)} &= y^{(m)} \end{aligned}$$

what are the values we aim to find so that we can get as close as possible to the best fitting line?

Now we have a lot of equations, a system of equations, and we aim to find values for weights and bias that gets us as close as possible to solve all these equations at once.

How do we formulate linear regression in ML (as vectors and matrices)?

$$W \cdot X + b = \hat{y}$$

Vector of weights Matrix of features bias Vector of Target Variables

$$[w_1 \ w_2 \ w_3 \dots \ w_n]$$
$$\begin{bmatrix} x_1^{(1)} & x_2^{(1)} & x_3^{(1)} & \dots & x_n^{(1)} \\ x_1^{(2)} & x_2^{(2)} & x_3^{(2)} & \dots & x_n^{(2)} \\ \vdots & & & & \\ x_1^{(m)} & x_2^{(m)} & x_3^{(m)} & \dots & x_n^{(m)} \end{bmatrix}$$
$$[y^{(1)} \ y^{(2)} \ y^{(3)} \dots \ y^{(m)}]$$

When we apply a linear regression model, we actually try to solve a system of equations in a way that our formula ($W \cdot X + b = \hat{y}$) would give us the best solution.

How do we convert a system of equations into vectors and matrices?

Example:

- ① Linear algebra score added to your calculus score minus your probability score was 6.
- ② Your algebra score minus your calculus score plus double your probability score was 4.
- ③ Four times your linear algebra score minus double your calculus score added to your probability score was 10

$$+1a +1c -1p = 6$$

$$+1a -1c +2p = 4$$

$$+4a -2c +1p = 10$$

$$W = [a \ c \ p]$$

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

$$\vec{y} = [6 \ 4 \ 10]$$

$$b = 0$$

What is a system of sentences?

System 1

- The dog is black
- The cat is orange

"Complete,"

↓
Non-Singular
System

System 2

- The dog is black
- The dog is black

"Redundant,"

↓
Singluar
System

System 3

- The dog is black
- The dog is white

"Contradictory,"

What type of systems are there based on completeness and singularity?

"System of equations is basically a system of sentences with numbers"

► A non-singular system carries as many pieces of information as sentences.

Sentences with numbers \Rightarrow equations

► You bought an apple and a banana for \$10. Your wife bought an apple and two bananas for \$12. How much each fruit cost?

$$\begin{aligned} a + b &= 10 \\ a + 2b &= 12 \end{aligned}$$

Non-singular systems are complete. They carry as many pieces of information as equations, and they have one solution. Singular systems carry either redundant information (infinite number of solutions) or contradictory information (no solutions).

- ① You bought an apple and a banana for \$10.
 - ② You bought two apples and two bananas for \$20.
- How much does each cost?
- Any two numbers that add to 10 are solutions.
- Redundant
↓
INFINITE NUMBER OF SOLUTIONS
-
- ① You bought an apple and a banana for \$10.
 - ② You bought two apples and two bananas for \$24.
- How much each cost?
- Contradiction
↓
NO SOLUTIONS

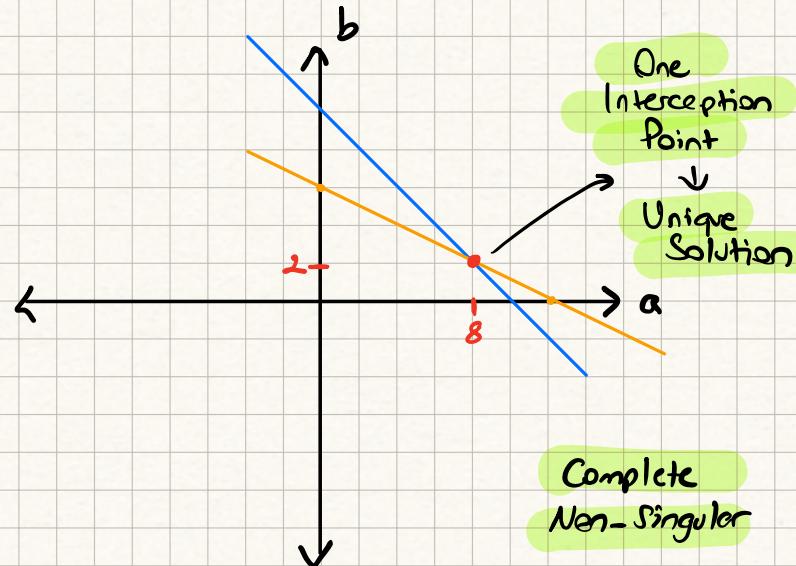
Graphical representation of system of equations

Two-variable linear equations can be visualized as lines in the coordinate plane. (Three-variables → planes in space, more variables → high-dimensional things we won't worry about now!)

Linear Equations → Lines

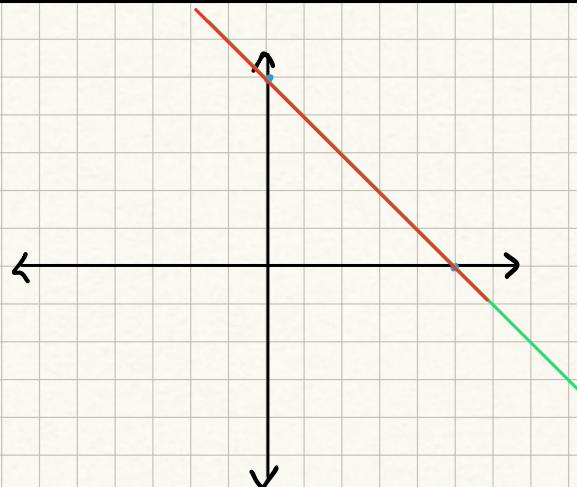
System of Linear Equations → Arrangements of lines

$$\begin{aligned} a + b &= 10 \\ a + 2b &= 12 \end{aligned}$$



$$a + b = 10$$

$$2a + 2b = 20$$



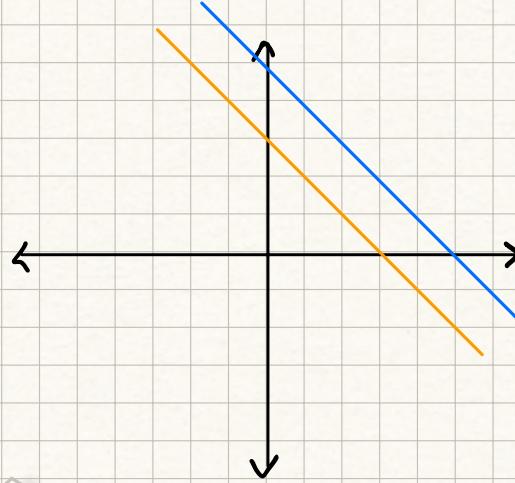
Overlapping
Lines
↓

Ininitely many
Solutions

Redundant Singular

$$a + b = 10$$

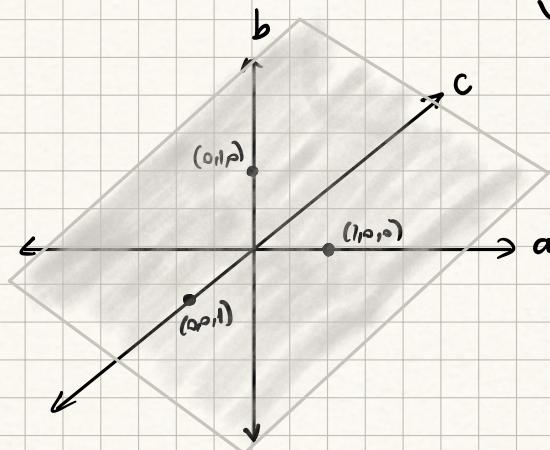
$$2a + 2b = 12$$



Parallel Lines
↓

No Solutions

Contradictory Singular



$$a + b + c = 1 \Rightarrow$$

$$1 + 0 + 0 = 1$$

$$0 + 1 + 0 = 1$$

$$0 + 0 + 1 = 1$$

Do constants matter
for singularity?

In ML, we care about if a system is non-singular or not.
If not, we don't care much about the reason (redundant/contradict)
therefore we can neglect the constants in the equations.

What is "linear dependency" in the context of matrices?

SYSTEM 1

$$\begin{array}{l} a + b = 0 \\ a + 2b = 0 \end{array}$$

Non-Singular System

* No equation is a multiple of the other one

* No row is a multiple of the other one

Rows are linearly independent

SYSTEM 2

$$\begin{array}{l} a + b = 0 \\ 2a + 2b = 0 \end{array}$$

Singular System

* Second equation is a multiple of the first one

* Second row is a multiple of the first one

Rows are linearly dependent

$$\begin{array}{l} a = 1 \\ b = 2 \\ a+b = 3 \end{array} \quad \begin{array}{l} 1a + 0b + 0c = 1 \\ 0a + 1b + 0c = 2 \\ 1a + 1b + 0c = 3 \end{array} \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix}$$

Row 3 depends on rows 1 and 2.

$$\begin{array}{l} a + b + c = 0 \\ a + b + 2c = 0 \\ a + b + 3c = 0 \end{array}$$

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \\ 1 & 1 & 3 \end{bmatrix}$$

Row 2 is the average of the rows 1 and 3.
 \Rightarrow Row 2 depends on rows 1 and 3.

$$\begin{array}{l} a + b + c = 0 \\ a + 2b + c = 0 \\ a + b + 2c = 0 \end{array}$$

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$$

No linear relations between rows
 \Rightarrow Non-Singular System

What is the determinant and how do we calculate it?

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \underbrace{a \cdot d - b \cdot c}_{\text{the determinant of the matrix}} = 0 \Rightarrow \text{singular}$$

$$= 0 \Rightarrow \text{non-singular}$$

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \quad (a \cdot e \cdot i) + (b \cdot f \cdot g) + (c \cdot d \cdot h) - (c \cdot e \cdot g) - (f \cdot h \cdot a) - (i \cdot b \cdot d)$$

$$= 0 \Rightarrow \text{singular}$$

$$\neq 0 \Rightarrow \text{non-singular}$$

$$\#1 \downarrow \#2 \downarrow \#3 + \#2 \downarrow \#3 \downarrow \#1 + \#3 \downarrow \#1 \downarrow \#2$$

$$- \#3 \downarrow \#2 \downarrow \#1 - \#2 \downarrow \#1 \downarrow \#3 - \#1 \downarrow \#3 \downarrow \#2$$

How do we present and solve linear systems as matrices in NumPy?

$$\begin{aligned} -x_1 + 3x_2 &= 7 \\ 3x_1 + 2x_2 &= 1 \end{aligned} \quad A = \begin{bmatrix} -1 & 3 \\ 3 & 2 \end{bmatrix} \quad b = \begin{bmatrix} 7 \\ 1 \end{bmatrix}$$

`A = np.array([-1, 3], [3, 2]), dtype = np.dtype(float))`

`b = np.array([7, 1], dtype = np.dtype(float))`

`x = np.linalg.solve(A, b)`

`print(x) # Output: [-1, 2]`

`d = np.linalg.det(A)`

`print(d) # Output: -11.00`

`A-system = np.hstack((A, b.reshape(2,1)))`

we need this because:

`b.shape` was (2,) we need (2,1)

* Determinant tells us if the matrix is non-singular (`np.linalg.det(A) ≠ 0`) or singular (`np.linalg.det(A) = 0`)

* `np.linalg.solve(A, b)` returns an array of solutions for x_1 and x_2 (or throws an `LinAlgError` if A is singular.)

Solving Systems of Linear Equations

05/03/2025

How do we solve
a system of
equations with
two variables?

Method 1: Manipulating equations so that we can get rid of one of the variables.

$$\begin{aligned} 3(5a + b = 17) \Rightarrow & \quad 15a + 3b = 51 \\ 4a - 3b = 6 & \quad + \quad 4a - 3b = 6 \\ & \quad \underline{\quad \quad \quad \quad \quad \quad \quad} \\ & \quad 19a + 0b = 57 \\ & \quad a = 3 \end{aligned}$$

$\rightarrow 5a + b = 17$
 $5.3 + b = 17$
 $b = 2$

Method 2: Re-organize one of the equations so that you can define one of the variables in terms of other.

$$\begin{aligned} 5a + b = 17 \Rightarrow b = 17 - 5a \\ 4a - 3b = 6 & \quad \rightarrow 4a - 3(17 - 5a) = 6 \\ & \quad 4a - 51 + 15a = 6 \\ & \quad 19a = 57 \\ & \quad a = 3 \end{aligned}$$

$\rightarrow 5a + b = 17$
 $5.3 + b = 17$
 $b = 2$

What if the system
is redundant?

$$\begin{aligned} (a + b = 10) \cdot -2 & \quad -2a - 2b = -20 \\ 2a + 2b = 20 & \quad + 2a + 2b = 20 \\ & \quad \underline{\quad \quad \quad \quad \quad \quad \quad} \\ & \quad 0 = 0 \end{aligned}$$

Solved System:
 $a = x$
 $b = 10 - x$

"Solution has 1 degree
of freedom, which is
 x . The solutions form
a line."

$$\begin{aligned} (a + b = 10) \cdot -2 & \quad -2a - 2b = -20 \\ 2a + 2b = 20 & \quad 2a + 2b = 24 \\ & \quad \underline{\quad \quad \quad \quad \quad \quad \quad} \\ & \quad 0 = -4 \end{aligned}$$

CONTRADICTION!

Solved System
N/A

Original System

$$5a + b = 17$$

$$4a - 3b = 6$$

Intermediate System

$$a + 0.2b = 3.4$$

$$b = 2$$

Solved System

$$a = 3$$

$$b = 2$$

Original Matrix

$$\begin{bmatrix} 5 & 1 \\ 4 & -3 \end{bmatrix}$$

Upper Diagonal Matrix

$$\begin{bmatrix} 1 & 0.2 \\ 0 & 1 \end{bmatrix}$$

Row Echelon Form

Diagonal Matrix

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

Reduced Row Echelon Form

Row Echelon Form Rules

- * Main diagonal can be 0, or 1s.
- * Below the main diagonal, everything must be 0.
- * Any number can exist on the right-side of 1s.
- * Only zeros are allowed on the right side of zeros.

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

The first non-zero element of a row is called the pivot. For a matrix to be in row echelon form:

(1) All rows without pivots must be at the bottom

(2) All pivots have to be at the right-side of the pivot of the row above.

* Some texts say that the pivots also have to be reduced to 1. Even if this is not a "must" it's very handy because we must do it for the reduced row echelon form anyways.

What are elementary row operations?

① Switching rows

$$\begin{bmatrix} 5 & 1 \\ 4 & 3 \end{bmatrix} \rightarrow \begin{bmatrix} 4 & 3 \\ 5 & 1 \end{bmatrix}$$

② Multiplying a row by a non-zero scalar

$$\begin{bmatrix} 5 & 1 \\ 4 & 3 \end{bmatrix} \rightarrow \begin{bmatrix} 50 & 10 \\ 4 & 3 \end{bmatrix}$$

③ Adding a row to another row

$$\begin{array}{r} \begin{bmatrix} 5 & 1 \\ 4 & 3 \end{bmatrix} \\ + \\ \hline \begin{bmatrix} 9 & 4 \\ 4 & 3 \end{bmatrix} \end{array}$$

System 1

The dog is black.

The cat is orange.

Two sentences

Two pieces of information.

Rank = 2

System 2

The dog is black.

The dog is black.

Two sentences.

One piece of info.

Rank = 1

System 3

The dog

The dog

Two sentences

Zero pieces of info*

Rank = 0

*About the color.

When we want to turn a matrix into a row echelon form and/or reduced row echelon form, we can perform the following elementary operations:

- ① Switch the order of rows
- ② Multiply the elements of a row by a non-zero scalar.
- ③ Add a row to another.

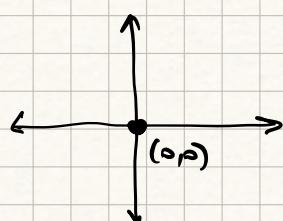
System 1

$$a+b=0$$

$$a+2b=0$$

Rank = 2

$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$$



Dimension of solution space = 0
(It's a point)

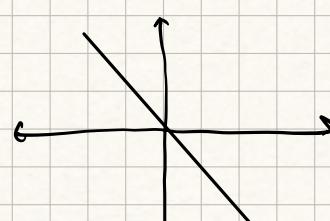
System 2

$$a+b=0$$

$$2a+2b=0$$

Rank = 1

$$\begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}$$



Dimension of solution space = 1
(It's a line)

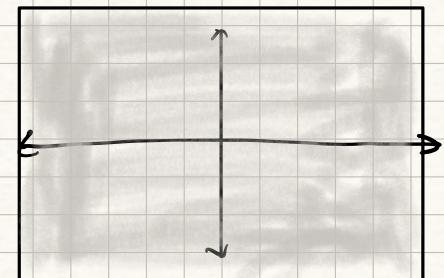
System 3

$$0a+0b=0$$

$$0a+0b=0$$

Rank = 0

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



Dimension of solution space = 2
(It's a plane)

$$\text{Rank}^* = 2 - (\text{Dimension of solution space})$$

*for 2×2

If the rank is equal to the number of rows, the matrix is non-singular. B/C this means it carries as many as information as the number of equations it has.

System 1

$$a+b+c=0$$

$$a+2b+c=0$$

$$a+b+2c=0$$

Rank 3

System 2

$$a+b+c=0$$

$$a+b+2c=0$$

$$a+b+3c=0$$

Rank 2

System 3

$$a+b+c=0$$

$$2a+2b+2c=0$$

$$3a+3b+3c=0$$

Rank 1

System 4

$$0a+0b+0c=0$$

$$0a+0b+0c=0$$

$$0a+0b+0c=0$$

Rank 0

The number of pivots in a row echelon matrix is called its rank. If the rank equals to the number of rows, the matrix is non-singular.

$$\begin{bmatrix} 5 & 1 \\ 4 & -3 \end{bmatrix} \rightarrow \begin{array}{l} \text{let's turn 5 into} \\ 1 \text{ b/c main diag.} \\ \text{can't have 5.} \end{array} \begin{bmatrix} 1 & 0.2 \\ 4 & -3 \end{bmatrix}$$

* 4 has to be converted into zero b/c bottom of the main diag.
has to be zero. How?

$$\rightarrow \begin{bmatrix} 1 & 0.2 \\ 1 & -0.75 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0.2 \\ 0 & -0.95 \end{bmatrix}$$

"first, make it 1", then subtract the first
row from the second

* Now we need 1 on the bottom right. So we divide the row
by -0.95

$$\rightarrow \begin{bmatrix} 1 & 0.2 \\ 0 & 1 \end{bmatrix} \Rightarrow \text{Row Echelon Form} \quad \text{Rank} = 2$$

$$\textcircled{2} \begin{bmatrix} 5 & 1 \\ 10 & 2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0.2 \\ 1 & 0.2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0.2 \\ 0 & 0 \end{bmatrix} \quad \text{Rank} = 1$$

$$\textcircled{3} \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{Rank} = 0$$

* The rank of the matrix is the sum of the values in
the main diagonal of the row echelon form!

* For 2×2

$$\begin{matrix} 2 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 3 & 1 & 1 \\ 0 & 0 & 0 & -5 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{matrix}$$

Rank 5

① Rows with only zeros must go to the bottom.

② The left-most non-zero entry of a row is called a pivot. Every pivot must be to the right of the pivots on the rows above.

③ Rank of the matrix is the number of pivots. (a general rule, not just for 2×2)

$$\begin{matrix} 3 & 1 & 1 & 1 & 1 \\ 0 & 0 & -1 & 1 & 1 \\ 0 & 0 & 0 & -4 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{matrix} \quad \begin{matrix} \div 3 \\ \div -1 \\ \div -4 \end{matrix}$$

Rank = 3

* We can make all pivots ones if we want.

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

Rank = 3

} 1's are now different of course, but the number of pivots is the same.

Reduced Row Echelon form

- * The matrix must be in row echelon form
- * Pivots must be converted to 1s.
- * Any number above a pivot must be converted to 0.

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

* We need this 2 to be 0. So we multiply the second row by 2 and subtract it from the first row

$$\begin{bmatrix} 1 & 0 & -5 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{bmatrix}$$

Now we must get rid of this -5. Let's multiply 3rd row by 5 and add it to the 1st row.

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{bmatrix}$$

Now let's multiply 3rd row by 4 and subtract from the second.

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

This is our reduced row echelon form.

The Gaussian Elimination Algorithm

$$2a - b + c = 1$$

$$2a + 2b + 4c = -2$$

$$4a + b + 0c = -1$$

2	-1	1	1
2	2	4	-2
4	1	0	-1

Augmented Matrix

① Turn R1 into a 1

$$\text{by } R1 = R1/2$$

1	$-\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$
2	2	4	-2
4	1	0	-1

② Turn R2 into a 0 by :

$$R2 = R2 - (R1 \cdot 2)$$

1	$-\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$
0	3	3	-3
4	1	0	-1

③ Turn R3 into a 0 by:

$$R3 = R3 - (R1 \cdot 4)$$

1	$-\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$
0	3	3	-3
0	3	-2	-3

④ Turn R2 into 1 by

$$R2 = R2/3$$

1	$-\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$
0	1	1	-1
0	3	-2	-3

⑤ Turn R3 into 0 by

$$R3 = R3 - (R2 \cdot 3)$$

1	$-\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$
0	1	1	-1
0	0	-5	0



⑥ Turn R_3 into 1 by

$$R_3 = R_3 / -5$$

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

$\brace{}$

Row Echelon Form

\Rightarrow Back substitution

⑦ Turn R_2 into 0 by

$$R_2 = R_2 - R_3$$

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

⑧ Turn R_1 into 0 by

$$R_1 = R_1 - (1/2 R_3)$$

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

⑨ Turn R_1 into 0 by

$$R_1 = R_1 + (1/2 R_2)$$

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{array} \right] = a$$

$\brace{}$

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{array} \right] = b$$

$\brace{}$

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

$\brace{}$

Identity matrix
(only 1s in the)
diagonal

$$\left[\begin{array}{ccc|c} 1 & 2 & -1 & 5 \\ 0 & 0 & -7 & 9 \\ 0 & 0 & 0 & X \end{array} \right]$$

$X=0 \Rightarrow$ Infinite # of solutions
 $X \neq 0 \Rightarrow$ No solutions

When we have a reduced row echelon of an augmented matrix, the last column is the vector of constants.

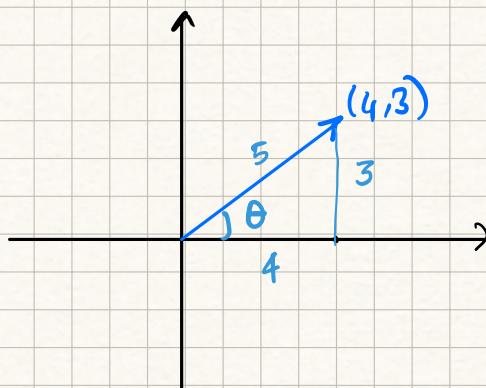
If we have a column with all-zero elements:

- ① If the augmented column is also zero, the system has infinite solutions.
- ② If the augmented column is not zero, the system has no solutions.

Vector Algebra

11.03.2025

Norm and direction



$$l_1 = \|(a,b)\|_1 = |a| + |b| = 7$$

$$l_2 = \|(a,b)\|_2 = \sqrt{a^2 + b^2} = 5$$

$$\tan(\theta) = \frac{3}{4}$$

$$\theta = \arctan(3/4) = 0.64 = 36.87^\circ$$

Notation

Row Vector

$$x = [x_1 \ x_2 \ \dots \ x_n]$$

Column Vector

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$\|x\|_1 = |x_1| + |x_2| + \dots + |x_n|$$

$$\|x\|_2 = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}$$

Vector Operations

$$u = (4, 1)$$

$$v + w = (5, 4)$$

$$v = (1, 3)$$

$$u - v = (3, -2)$$

$$x = [x_1 \ x_2 \ \dots \ x_n]$$

$$y = [y_1 \ y_2 \ \dots \ y_n]$$

$$x + y = [x_1 + y_1 \ x_2 + y_2 \ \dots \ x_n + y_n]$$

$$x - y = [x_1 - y_1 \ x_2 - y_2 \ \dots \ x_n - y_n]$$

$$\lambda \cdot x = [\lambda x_1 \ \lambda x_2 \ \dots \ \lambda x_n]$$

* Every vector must have a magnitude and direction. Magnitude can be calculated in 2 different ways, ℓ_1 norm (aka taxicab) and ℓ_2 norm (Euclidean).

* We can add and subtract vectors from each other and multiply a vector with a scalar as if we're doing the math operations on regular numbers.

The Dot Product

$$x = [x_1 \ x_2 \ \dots \ x_n] \quad y = [y_1 \ y_2 \ \dots \ y_n]$$

$$x \cdot y = (x_1 \cdot y_1) + (x_2 \cdot y_2) + \dots + (x_n \cdot y_n)$$

$$\langle x, y \rangle = x \cdot y$$

Transposing a vector / matrix

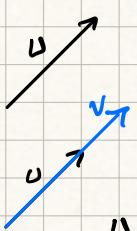
$$x = [1 \ 4 \ 8] \Rightarrow x^T = \begin{bmatrix} 1 \\ 4 \\ 8 \end{bmatrix}$$

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \Rightarrow A^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$$

Geometric Dot Product

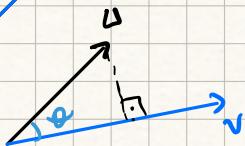
* Orthogonal vectors have dot product 0

$$\langle u, v \rangle = 0$$



$$\langle u, u \rangle = \|u\|^2 = \|u\| \cdot \|u\|$$

$$\langle u, v \rangle = \|u\| \cdot \|v\|$$



$$\langle u, v \rangle = \|u\| \cdot \|v\| \cdot \cos(\theta)$$

$$\langle u, v \rangle > 0 \Rightarrow 0 < \theta < 90^\circ$$

$$\langle u, v \rangle < 0 \Rightarrow 90^\circ < \theta < 180^\circ$$

The dot product is a mathematical operation that takes two vectors and returns a scalar, representing the projection of one vector onto another, or, equivalently, the sum of the products of their corresponding components.

Multiplying a matrix by a vector

Equations as dot product

$$a + b + c = 10$$

$$\begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} a \\ b \\ c \end{bmatrix} = 10$$

$$a + 2b + c = 15$$

$$\begin{bmatrix} 1 & 2 & 1 \end{bmatrix} \cdot \begin{bmatrix} a \\ b \\ c \end{bmatrix} = 15$$

$$a + b + 2c = 12$$

$$\begin{bmatrix} 1 & 1 & 2 \end{bmatrix} \cdot \begin{bmatrix} a \\ b \\ c \end{bmatrix} = 12$$

$$\begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} a \\ b \\ c \end{bmatrix} = 10$$

$$\begin{bmatrix} 1 & 2 & 1 \end{bmatrix} \cdot \begin{bmatrix} a \\ b \\ c \end{bmatrix} = 15$$

$$\begin{bmatrix} 1 & 1 & 2 \end{bmatrix} \cdot \begin{bmatrix} a \\ b \\ c \end{bmatrix} = 12$$

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} \cdot \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 10 \\ 15 \\ 12 \end{bmatrix}$$

$3 \times 3 \leftarrow = 3 \times 1 \rightarrow 3 \times 1$

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \\ 2 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 10 \\ 15 \\ 12 \\ 13 \end{bmatrix}$$

4×3

3×1

4×1

Number of columns of the matrix has to match the number of rows of the vector.

To multiply a matrix and a vector, # of columns of the matrix must be equal to the # of rows of the vector. The i -th element of the resulting vector is the dot product of the i -th row of the matrix and the input vector.

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} ax + by \\ cx + dy \end{bmatrix}$$

Linear Transformation = function
 input vector $\rightarrow [f] \rightarrow$ output vector

1. Lines must remain lines
 2. Origin must remain fixed in place.

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = x \cdot \begin{bmatrix} a \\ c \end{bmatrix} + y \begin{bmatrix} b \\ d \end{bmatrix}$$

$\begin{bmatrix} a \\ c \end{bmatrix}$ is where \hat{i} lands

$$= \begin{bmatrix} ax + by \\ cx + dy \end{bmatrix}$$

$\begin{bmatrix} b \\ d \end{bmatrix}$ is where \hat{j} lands

$\begin{bmatrix} x \\ y \end{bmatrix}$ is the input vector

"We applied a linear transformation on the $\begin{bmatrix} x \\ y \end{bmatrix}$ vector."

The linear transformation we applied can be shown as a matrix:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

$\begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix}$ tells us that:

① $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ goes to $\begin{bmatrix} 3 \\ 1 \end{bmatrix}$

② $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ goes to $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$

Linear transformations as matrix multiplication

$$\begin{bmatrix} 2 & -1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \end{bmatrix} \quad \begin{bmatrix} 2 & -1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 4 \end{bmatrix}$$

* A linear transformation is a function between two vector spaces that preserves the operations of vector addition and scalar multiplication.

$$* T(u+v) = T(u) + T(v)$$

$$* T(c \cdot u) = c \cdot T(u)$$

* Linear transformation maintains the "straightness" of lines and keep the origin fixed.

* We can show linear transformations as matrices. Linear Transformations can be applied on vectors as well as matrices.

This means \Rightarrow

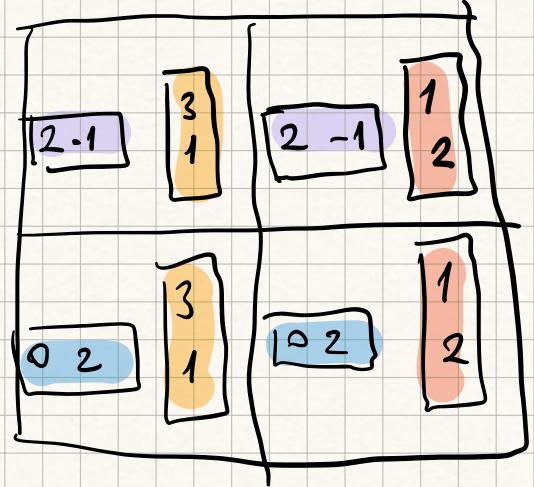
$$\begin{bmatrix} 2 & -1 \\ 0 & 2 \end{bmatrix} \cdot \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 5 & 0 \\ 2 & 4 \end{bmatrix}$$

SECOND!
Transformation Matrix

FIRST!
Transformation Matrix

How do we perform matrix multiplication?

$$\begin{bmatrix} 2 & -1 \\ 0 & 2 \end{bmatrix} \cdot \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix} =$$



of rows don't need to match.
of columns of matrix 1 has to match # of rows of the matrix 2

$$\begin{bmatrix} 3 & 1 & 4 \\ 2 & -1 & 2 \end{bmatrix} \cdot \begin{bmatrix} 3 & 0 & 1 & -2 \\ 1 & 5 & 2 & 0 \\ -2 & 1 & 4 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 9 & 21 & -6 \\ 1 & -3 & 8 & -4 \end{bmatrix}$$

2×3 $=$ 3×4 2×4

What happens if we use an identity matrix as the transformation matrix?

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} a \\ b \\ c \\ d \\ e \end{bmatrix} = \begin{bmatrix} a \\ b \\ c \\ d \\ e \end{bmatrix}$$

What is an inverse matrix?

$$[\text{Matrix}] \cdot [\text{Inverse Matrix}] = [\text{Identity Matrix}]$$

$$\begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix} \cdot \underbrace{\begin{bmatrix} a & b \\ c & d \end{bmatrix}}_{\begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix}^{-1}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$3a + 1c = 1$$

$$3b + 1d = 0$$

$$1a + 2c = 0$$

$$1b + 2d = 1$$

$$\left[\begin{array}{cccc|c} 3 & 0 & 1 & 0 & 1 \\ 0 & 3 & 0 & 1 & 0 \\ 1 & 0 & 2 & 0 & 0 \\ 0 & 1 & 0 & 2 & 1 \end{array} \right]$$

$$a = 2/5$$

$$b = -1/5$$

$$c = -1/5$$

$$d = 3/5$$

Matrices and linear transformations in Natural Language Processing

Spam	Lottery	Win
Yes	1	1
Yes	2	1
No	0	0
Yes	0	2
No	0	1
No	1	0
Yes	2	2
Yes	2	0
Yes	1	2

Scores:

Lottery = points / occurrence

Win = points / occurrence

Rule:

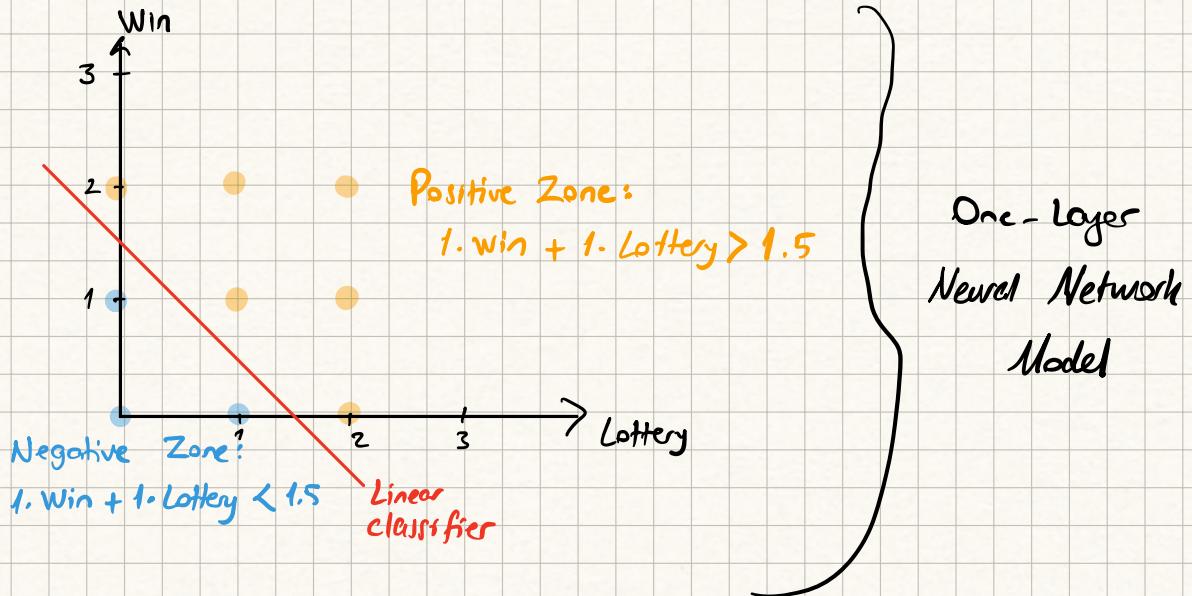
If the number of points of the sentence is bigger than, then the email is spam.

Goal:

Find the best points and threshold.

Inverse matrix as a linear transformation transforms the plane to its original state. Only non-singular matrices have inverses. They are also called invertible.

- a) Lottery: 1 pt , Win: 2 pts , Threshold: 3 pts.
- b) Lottery: 1 pt , Win: 1 pt , Threshold: 1.5 pts.
- c) Lottery: 2 pts , Win: 3 pts , Threshold: 1.5 pts.



Input Data	$\begin{bmatrix} \#word \\ \text{Lottery} \end{bmatrix}$	$\begin{bmatrix} \#word \\ \text{Win} \end{bmatrix}$	Model	$\begin{bmatrix} \text{pts. for lottery} \\ \text{pts. for win} \end{bmatrix}$	> Threshold ?
	$\begin{bmatrix} 2 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$		= 3 > 1.5 (spam)	
	$\begin{bmatrix} 0 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$		= 1 > 1.5 (not spam)	
Data	$\begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 0 & 0 \\ 0 & 2 \\ 0 & 1 \\ 1 & 0 \\ 2 & 2 \\ 2 & 0 \\ 1 & 2 \end{bmatrix}$	Model	$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$	Product	Target
		.		$\begin{bmatrix} 2 \\ 3 \\ 0 \\ 2 \\ 1 \\ 1 \\ 4 \\ 2 \\ 3 \end{bmatrix}$	> 1.5 ?
					$\begin{bmatrix} \text{YES} \\ \text{YES} \\ \text{NO} \\ \text{YES} \\ \text{NO} \\ \text{NO} \\ \text{YES} \\ \text{YES} \\ \text{YES} \end{bmatrix}$

$$1. \text{Win} + 1. \text{Lottery} > 1.5$$

$$1. \text{Win} + 1. \text{Lottery} - 1.5 > 0$$

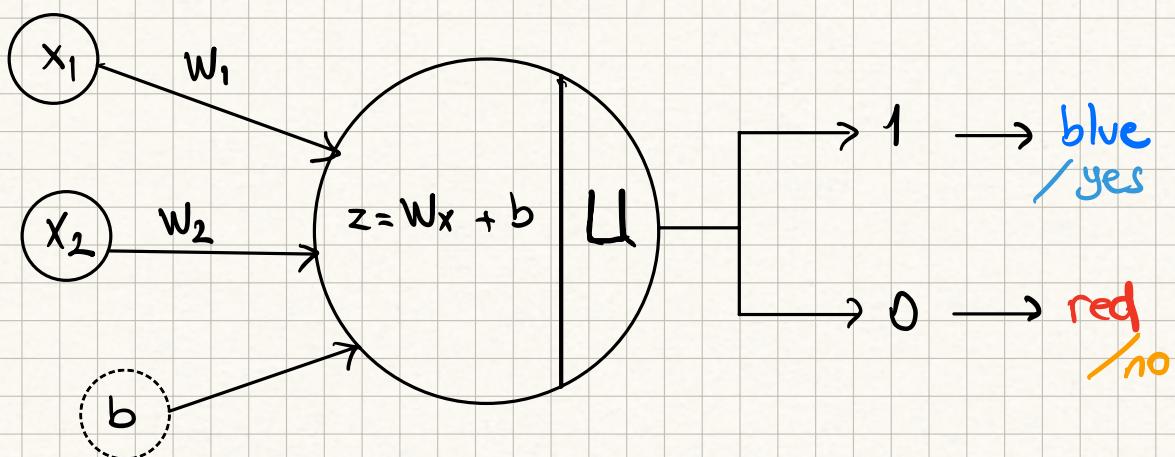
↳ BIAS

"AND" logic with linear transformation.

AND	x	y		Model	=	Dot Product.
No	0	0				0
No	1	0		1		1
No	0	1		1		1
Yes	1	1				2

$> 1.5 ?$

How does a single perceptron neural network work?



A perceptron is the simplest type of artificial neural network, designed to model a single neuron. It takes multiple inputs, each with an associated weight; it calculates a weighted sum of these inputs; it applies an activation function (typically a step function) to the weighted sum; and outputs a binary value indicating the category the input belongs to. Perceptrons use the concept of a linear transformation, which can be expressed as a matrix multiplication, to calculate its output and make decisions.