Matrix representation of equation system

① Equation system

$$\left\{\begin{array}{ccccc|cc} a_{11}x_1 & + & a_{12}x_2 & + & a_{13}x_3 & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & a_{23}x_3 & = & b_2 \\ a_{31}x_1 & + & a_{32}x_2 & + & a_{33}x_3 & = & b_3 \end{array}\right.$$

② Equivalent matrix representation:

$$\left[\begin{array}{ccc|c} a_{11} & a_{12} & a_{13} & b_1 \\ a_{21} & a_{22} & a_{23} & b_2 \\ a_{31} & a_{32} & a_{33} & b_3 \end{array}\right]$$

or $[A \mid \vec{b}]$

③ Equivalent matrix-vector multiplication representation:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \times \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

or $A\,\vec{x} = \vec{b}$

- numbers $a_{i\,j}$ and b_i are constants
- x_i are variables to solve for

Vector representation of equation system

Equivalent representation in column-vector format:

$$x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ a_{32} \end{bmatrix} + x_3 \begin{bmatrix} a_{13} \\ a_{23} \\ a_{33} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

Vector $\vec{b}$ is formed as a weighted sum of the columns of A
Each column of the matrix represents a direction in space
Scalars x_1, x_2, ... tell us how much of each column-direction to take

- Take column 1 of A and scale it by x_1
- Take column 2 of A and scale it by x_2

...

- Add these scaled columns together to obtain vector $\vec{b}$

Takeaway:
Solving $A\vec{x} = \vec{b}$ is the same as asking:
Can the vector $\vec{b}$ be built by combining the columns of A,
and if so, with what weights?

If such weights x_1, x_2, ... exist, the system is consistent
The columns of A span the given vector $\vec{b}$

The vector $\vec{x}$ records how the columns are combined

One-sentence summary:
The equation $A\vec{x} = \vec{b}$ means that $\vec{b}$ is a linear combination
of the columns of A, with coefficients given by $\vec{x}$

Solving [A | $\vec{b}$] using RREF

To solve the equation A $\vec{x}$ = $\vec{b}$,
we transform [A | $\vec{b}$] into reduced echelon form R

This transformation simplifies the system by eliminating some variables from certain equations, while preserving the set of all solutions

Eliminating variables does not mean removing them
Instead, their influence is absorbed into simpler equations

The transformed system has the same solutions as the original one:

1. Swapping two equations does not affect which vectors $\vec{x}$ satisfy them
2. Multiplying an equation by a nonzero number does not change its solutions
3. Replacing an equation by the sum of itself and a multiple of another equation does not add or remove solutions

Because each row operation preserves solutions, any sequence of such operations does as well

What matrix R reveals:

- Presence or absence of solution

- Presence & location of free variables
- Isolates leading variables as functions of free variables

Case 1: unique solution

$A\vec{x} = \vec{b}$ has a unique solution for the given $\vec{b}$ if
A is an n×n matrix with rank r = n

⇔

Columns of A are linearly independent

⇔

every column of A produces a pivot column in the A-part of R:
reduced echelon form of $[A \mid \vec{b}]$

⇓

R has a form $[I \mid \vec{x}]$ where
- I is an n×n identity matrix
- $\vec{x}$ is the unique solution

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & x_1 \\ 0 & 1 & 0 & x_2 \\ 0 & 0 & 1 & x_3 \end{array}\right]$$

or equivalently

$$\left\{\begin{array}{ccccc|cc} 1x_1 & + & 0x_2 & + & 0x_3 & = & x_1 \\ 0x_1 & + & 1x_2 & + & 0x_3 & = & x_2 \\ 0x_1 & + & 0x_2 & + & 1x_3 & = & x_3 \end{array}\right.$$

Geometric example of unique solution in 2D

Consider particular equation system of form A $\vec{x} = \vec{b}$:

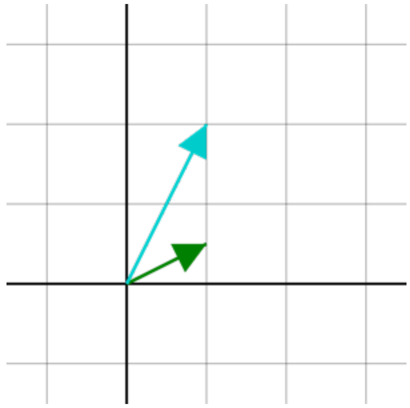
$$\begin{bmatrix} 1 & 1 \\ 0.5 & 2 \end{bmatrix} \times \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$$

with solution $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$

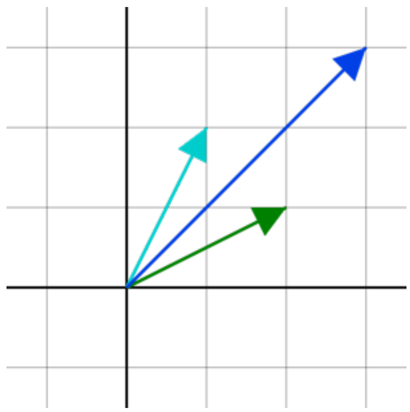
① Column 'perspective' (vector equation):
Vector $\vec{b}$ is composed out of columns of A as follows:

$$x_1 \begin{bmatrix} 1 \\ 0.5 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$$

$\vec{a}1$ (green) & $\vec{a}2$ (cyan)



$$2 \times \vec{a}1 + 1 \times \vec{a}2 = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$$



② Row 'perspective' (equations as lines):
rows of augmented matrix [A | $\vec{b}$]

$$\left[\begin{array}{cc|c} 1 & 1 & 3 \\ 0.5 & 2 & 3 \end{array}\right]$$

or, equivalently

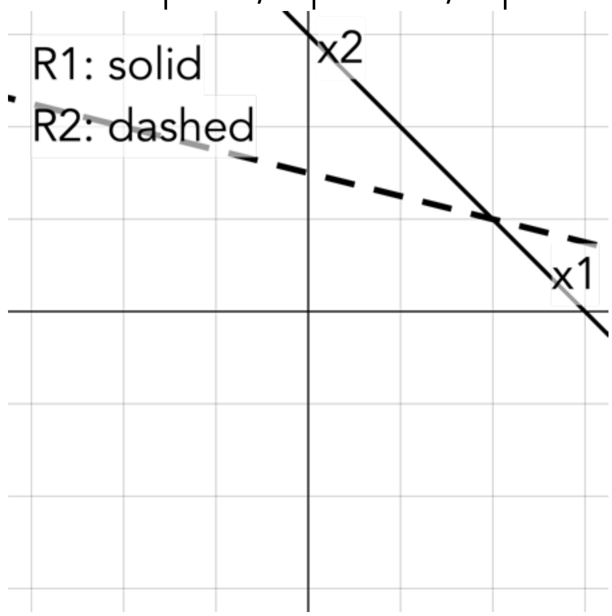
$$\begin{cases} 1x_1 + 1x_2 = 3 \\ 0.5x_1 + 2x_2 = 3 \end{cases}$$

can be presented as 2D image where

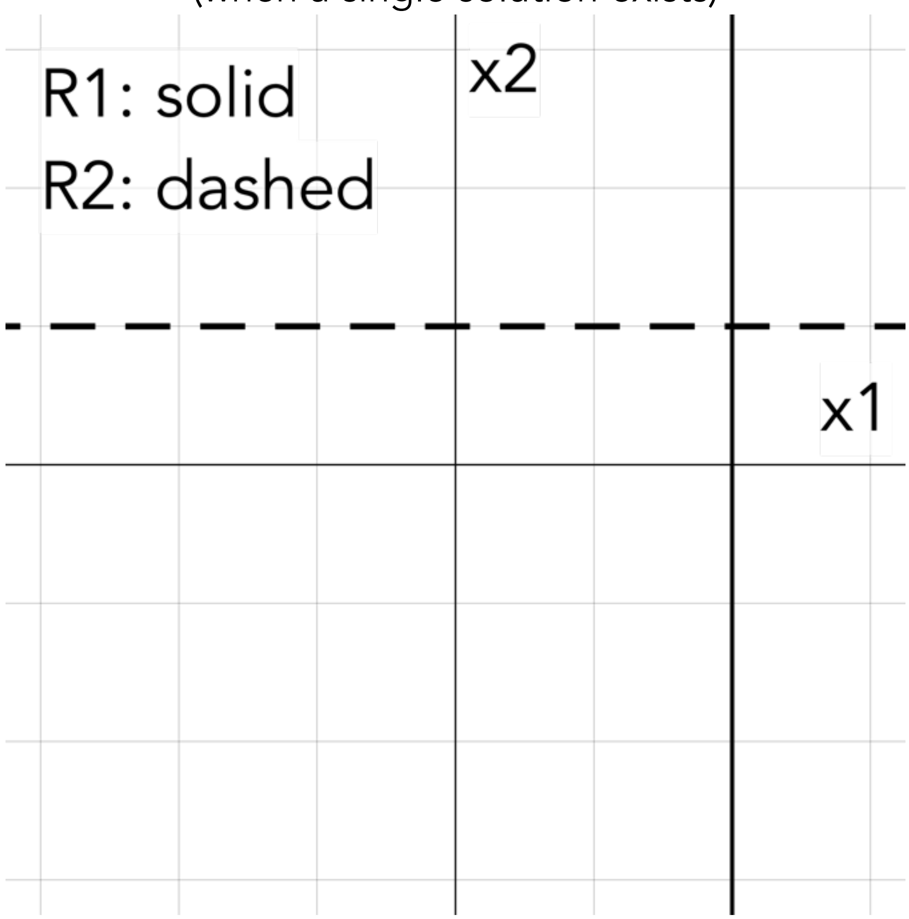
- each row represents one linear equation of type

$$a_{i1}x_1 + a_{i2}x_2 = b_i$$

- common intersection point, if present, represents the solution



After row reduction,
equations become aligned with coordinate axes,
so the solution can be read directly
(when a single solution exists)



Summary:

- Column view explains how the right-hand side is built
- Row view explains the geometric constraints the solution must satisfy
- Row reduction transforms those conditions into a form where their existence and their set are visible

---

Case 2: no solution

$A\vec{x} = \vec{b}$ has no solution for the given $\vec{b}$ if
$\vec{b}$ cannot be presented as a linear combination of columns of A
$\Leftrightarrow$
there is no choice of x_1, x_2, … that satisfies the following:

$$x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ a_{32} \end{bmatrix} + x_3 \begin{bmatrix} a_{13} \\ a_{23} \\ a_{33} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

Columns of A span a certain subspace & vector $\vec{b}$ lies outside of it

When $\vec{b}$ is added as an extra column, it introduces a new independent direction
$\Downarrow$
$[A \mid \vec{b}]$ has higher rank than A itself

If R is reduced echelon form of $[A \mid \vec{b}]$, it has a form

$$\left[\begin{array}{ccc|c} 1 & 0 & r_{13} & r_{14} \\ 0 & 1 & r_{23} & r_{24} \\ 0 & 0 & 0 & r_{34} \neq 0 \end{array}\right]$$

or

$$\left\{\begin{array}{ccccc|cc} 1x_1 & + & 0x_2 & + & r_{13}x_3 & = & r_{14} \\ 0x_1 & + & 1x_2 & + & r_{23}x_3 & = & r_{24} \\ 0x_1 & + & 0x_2 & + & 0x_3 & = & r_{34} \neq 0 \end{array}\right.$$

As in $[A \mid \vec{b}]$, the last column introduces an extra independent direction that cannot be spanned by the columns of A, producing the inconsistent bottom row

(one or more inconsistent rows may be present)

Looking ahead: ⏭️

Even though the system $A\vec{x} = \vec{b}$ has no exact solution,
we can still ask a meaningful question:

among all vectors that can be produced by A,
which one comes closest to $\vec{b}$?

Instead of trying to solve $A\vec{x} = \vec{b}$,
we may replace $\vec{b}$ by a closest vector that lies in Col(A) by 'flattening' $\vec{b}$ onto that space

We return to this idea later

Case 3: infinitely many solutions

System $A\,\vec{x} = \vec{b}$ has infinitely many solutions when both hold:

- $\vec{b}$ can be presented as a linear combination of columns of A
- columns of A are not linearly independent

Mechanism:

there is one or more non-zero vector(s) $\vec{y}$ such that $A\,\vec{y} = \vec{0}$

$\Leftrightarrow$

$A(\vec{x} + \vec{y}) = A\,\vec{x}$

Vectors $\vec{y}$ are directions that allow freedom in $\vec{x}$ while producing same $\vec{b}$

Computation:

R (reduced echelon form of $[A \mid \vec{b}]$)

will allow us to determine all solutions $\vec{x}$, including

- particular solution $\vec{x}_p$
- all vectors $\vec{y}$ satisfying $A\,\vec{y} = \vec{0}$

so that $A(\vec{x}_p + \vec{y}) = \vec{b}$

Below is an example of R for system with

- m = 4 equations and n = 4 variables
- dependent columns in positions 2, 3

$\Downarrow$

rank r = 2

$$\left[\begin{array}{cccc|c} 1 & r_{12} & r_{13} & 0 & r_{15} \\ 0 & 0 & 0 & 1 & r_{25} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Or in equivalent equation form

$$\left\{\begin{array}{ccccccc|c} 1x_1 & + & r_{12}x_2 & + & r_{13}x_3 & + & 0x_4 & = r_{15} \\ 0x_1 & + & 0x_2 & + & 0x_3 & + & 1x_4 & = r_{25} \\ 0x_1 & + & 0x_2 & + & 0x_3 & + & 0x_4 & = 0 \\ 0x_1 & + & 0x_2 & + & 0x_3 & + & 0x_4 & = 0 \end{array}\right.$$

Recall that

- independent columns in $[A \mid \vec{b}]$ produce pivot columns in R
- dependent columns in $[A \mid \vec{b}]$ produce non-pivot columns in R (shown in blue)
- row rank = col rank

⇓

there are n-r all-zero rows on the bottom of coefficient part of R

- system is not inconsistent

⇔

$\vec{b}$ in $[A \mid \vec{b}]$ does not add dimension

⇓

last column in R does not add a new independent column

⇓

there are n-r all-zero rows on the bottom of R, therefore, no contradiction

'Reading' above equation system produces

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} r_{15} \\ 0 \\ 0 \\ r_{25} \end{bmatrix} + x_2 \begin{bmatrix} -r_{12} \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -r_{13} \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

Image above shows that $\vec{x}$ is composed of following parts:
① Particular solution $\vec{x}_p$ chosen so that A $\vec{x}_p = \vec{b}$ when all $\vec{y} = \vec{0}$
is shown in cyan and composed as follows:

Note that each pivot in R has both a row index and a column index
• each entry of $\vec{x}_p$ corresponds to a variable
• for a pivot variable, the entry of $\vec{x}_p$ is taken from the constant column of R
using the row index of the corresponding pivot
• entries corresponding to free variables are set to 0

② So-called homogeneous solutions, one for each free variable,
shown in blue

Each free variable generates one homogeneous direction vector
and any free variable vector x_i is composed as follows:
• the entry i is set to 1
• for any j≠i, corresponding to another free variable,
entry j is set to 0
• for any k corresponding to a pivot variable,
entry k is taken from the constant column of R, with sign negated
using the row index of the pivot in column k
(note that the homogeneous vectors are constructed so that they are linearly independent)

• we denote free variable vectors as directions $\vec{y}$

• every $\vec{y}$ satisfies $A\vec{y} = \vec{0}$
• every $\vec{y}$ is orthogonal to every row of A

• all homogeneous solutions go through the origin
• the particular solution shifts solution set away from the origin

---

Geometric example of infinitely many solutions in 2D

Consider particular equation system of form $A\vec{x} = \vec{b}$:

$$\begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix} \times \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 9 \end{bmatrix}$$

where col2(A) is a multiple of col1(A) and $\vec{b}$ is within span of those columns

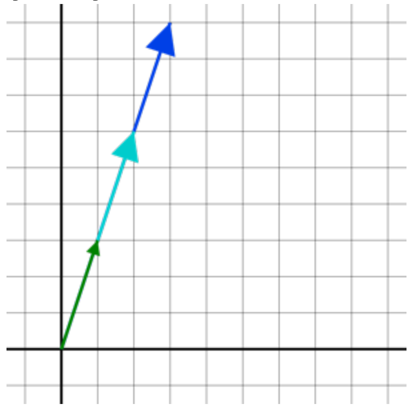
$$\text{R: REF form of } [\, A \mid \vec{b}\,] = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \end{bmatrix}$$

'Reading' R produces

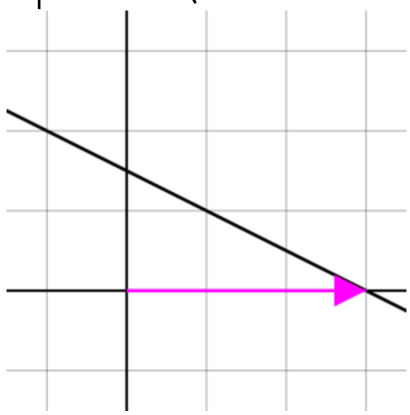
$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

① Column 'perspective' (vector equation):

Vector $\vec{b}$ can be expressed as a linear combination of columns of A in infinitely many ways as shown on the image below:



② Row 'perspective' (non-0 row of R as line):



$A\,\vec{x} = \vec{0}$

forms a line through the origin in the direction of the null-space vector

[-2, 1]

A particular solution $\vec{x}_p$ satisfies

$A\,\vec{x}_p = \vec{b}$

and here $\vec{x}_p = [3, 0]$

The full solution set of

$A\,\vec{x} = \vec{b}$

is obtained by translating the homogeneous solution set by $\vec{x}_p$

Equivalently,

$\vec{x} = \vec{x}_p + x_2\,\vec{y}$, with $A\,\vec{y} = \vec{0}$

so in this example

$\vec{x} = [3, 0] + x_2 [-2, 1]$

All homogeneous solutions pass through the origin
The particular solution $\vec{x}_p$ shifts the solution set away from the origin
The direction of the solution set is unchanged

Summary for case 3 with infinite solutions

$A \vec{x} = \vec{b}$ has infinitely many solutions if and only if both conditions are true:

- $\vec{b}$ can be written as a linear combination of columns of A
- columns of A are not linearly independent

① Components of solution set $\vec{x}$

If the columns of A are not linearly independent, there exists at least one non-zero vector $\vec{y}$ such that

$A \vec{y} = \vec{0}$,

and each vector $\vec{y}$ is associated with one free variable

In addition to the vectors $\vec{y}$, there is one particular solution $\vec{x}_p$ defined as follows: $A \vec{x}_p = \vec{b}$ when all free variables or $\vec{y}$ vectors are set to $\vec{0}$,

Then, every solution of $A \vec{x} = \vec{b}$ can be written as

$\vec{x} = \vec{x}_p + \vec{y}$

- every $\vec{y}$ goes through origin
- $\vec{x}_p$ shifts solution set away from the origin

② Orthogonality concept

If the rows of A are denoted by a_i,
then $A\vec{y} = \vec{0}$ implies
$a_i \cdot \vec{y} = 0$ for every row a_i
so every $\vec{y}$ is orthogonal to every row of A

From infinite solutions case to domain of n×n square matrix A

Domain of A is the set of all vectors $\vec{x}$ such that $A\vec{x}$ is valid: $\mathbb{R}^n$

As we saw, reduced echelon form of A has 2 fundamental numbers:

- r = number of pivot variables
- n - r = number of free variables

These two numbers describe decomposition of the domain $\mathbb{R}^n$
into two complementary subspaces

① row space(A)

- is a subspace of $\mathbb{R}^n$
- has dimension r
- is spanned by r independent rows of A

② null space(A)

- is a subspace of $\mathbb{R}^n$
- has dimension n - r
- consists of all vectors $\vec{y}$ such that $A\vec{y} = \vec{0}$
- each free variable generates one independent direction vector in null(A)

As we saw on the previous page, each $\vec{y}$ is orthogonal to every row of A
This implies that null space and row space do not overlap

Taken together, their dimension sums to n

$\Downarrow$

row space(A) and null space(A) are complementary subspaces of $\mathbb{R}^n$

For a general m×n matrix A, the same reasoning applies to the domain $\mathbb{R}^n$