Matrix $\boldsymbol{A}$ times Vector $\boldsymbol{x}$: Column Way and Row Way

Gilbert Strang

$$A\boldsymbol{x} = \begin{bmatrix} 2 & 1 & 3 \\ 3 & 4 & 7 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1 \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ 4 \\ 2 \end{bmatrix} + x_3 \begin{bmatrix} 3 \\ 7 \\ 3 \end{bmatrix} \quad \begin{matrix}\textbf{Combination} \\ \textbf{of} \\ \textbf{columns}\end{matrix}$$

$$= \begin{bmatrix} 2x_1 + 1x_2 + 3x_3 \\ 3x_1 + 4x_2 + 7x_3 \\ 1x_1 + 2x_2 + 3x_3 \end{bmatrix} = \begin{bmatrix} (\text{row } 1) \cdot \boldsymbol{x} \\ (\text{row } 2) \cdot \boldsymbol{x} \\ (\text{row } 3) \cdot \boldsymbol{x} \end{bmatrix} \quad \begin{matrix}\textbf{Dot products} \\ \textbf{with rows}\end{matrix}$$

Example of a vector space = **Column space of $\boldsymbol{A}$** = $\mathbf{C}(A)$

The column space $\mathbf{C}(A)$ contains **all combinations** of the columns: all x_1, x_2, x_3

All multiples of column 1 = **Line** in 3-dimensional space

All combinations of columns 1 and 2 = **Plane** in 3-dimensional space

Notice that column 3 = column 1 + column 2

All combinations of columns $1, 2, 3$ = **Same plane** in 3-dimensions = $\mathbf{C}(A)$

Columns 1 and 2 are **independent**

They are a **basis** for the column space

(only $x_1 = x_2 = 0$ gives combination = 0)

Columns $1, 2, 3$ are **dependent**

They are **not a basis** for $\mathbf{C}(A)$

$(x_1, x_2, x_3) = (1, 1, -1)$ gives combination = 0

The rank of $\boldsymbol{A}$ is $\boldsymbol{r = 2}$

Basis for $\mathbf{C}(A)$ has r vectors

$$\boldsymbol{A} = \boldsymbol{CR} = \begin{bmatrix} 2 & 1 \\ 3 & 4 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} \mathbf{1} & 0 & \mathbf{1} \\ 0 & \mathbf{1} & \mathbf{1} \end{bmatrix}$$

Thoughts on Chapter 1

Most textbooks don't have a place for the author's thoughts. But a lot of decisions go into starting a new textbook. This chapter has intentionally jumped right into the subject, with discussion of independence and rank. There are so many good ideas ahead, and they take time to absorb, so why not get started? Here are two questions that influenced the writing.

What makes this subject easy?	All the equations are linear.
What makes this subject hard?	So many equations and unknowns and ideas.

Book examples are small size. But if we want the temperature at many points of an engine, there is an equation at every point: easily $n = 1000$ unknowns.

I believe the key is to work right away with matrices. $A\boldsymbol{x} = \boldsymbol{b}$ is a perfect format to accept problems of all sizes. The linearity is built into the symbols $A\boldsymbol{x}$ and the rule is $A(\boldsymbol{x}+\boldsymbol{y}) = A\boldsymbol{x} + A\boldsymbol{y}$. Each of the m equations in $A\boldsymbol{x} = \boldsymbol{b}$ represents a flat surface:

$2x + 5y - 4z = 6$ is a plane in three-dimensional space

$2x + 5y - 4z + 7w = 9$ is a 3D plane (*hyperplane*?) in four-dimensional space

Linearity is on our side, but there is a serious problem in visualizing 10 planes meeting in 11-dimensional space. Hopefully they meet along a line: dimension $11 - 10 = 1$. An 11th plane should cut through that line at one point (which solves all 11 equations). What the textbook and the notation must do is to keep the counting simple

Here is what we expect for a random m by n matrix A:

$\boldsymbol{m < n}$ Probably many solutions to the m equations $A\boldsymbol{x} = \boldsymbol{b}$

$\boldsymbol{m = n}$ Probably one solution to the n equations $A\boldsymbol{x} = \boldsymbol{b}$

$\boldsymbol{m > n}$ Probably no solution: too many equations with only n unknowns in $\boldsymbol{x}$

But this count is not necessarily what we get! Columns of A can be combinations of previous columns: nothing new. An equation can be a combination of previous equations. **The rank $\boldsymbol{r}$ tells us the real size of our problem**, from independent columns and rows. The beautiful formula is $\boldsymbol{A = CR} = (m \times r)\,(r \times n)$: three matrices of rank r.

Notice: The columns of A that go into C must multiply the matrix I inside R.

We end with the great **associative law $\boldsymbol{(AB)\,C = A\,(BC)}$**. Suppose C has 1 column:

AB has columns $A\boldsymbol{b}_1, \ldots, A\boldsymbol{b}_n$ and then $(AB)\boldsymbol{c}$ equals $c_1 A\boldsymbol{b}_1 + \cdots + c_n A\boldsymbol{b}_n$.

$B\boldsymbol{c}$ has one column $c_1\boldsymbol{b}_1 + \cdots + c_n\boldsymbol{b}_n$ and then $A(B\boldsymbol{c}) = A\,(c_1\boldsymbol{b}_1 + \cdots + c_n\boldsymbol{b}_n)$.

Linearity gives equality of those two sums. This proves $(AB)\,\boldsymbol{c} = A\,(B\boldsymbol{c})$.

The same is true for every column of C. Therefore $\boldsymbol{(AB)\,C = A\,(BC)}$.

Notice that over and over—for $\boldsymbol{Ax}$ and $\boldsymbol{AB}$ and $\boldsymbol{CR}$—we write about linear combinations of columns of $\boldsymbol{A}$ or $\boldsymbol{C}$.

Not about dot products with the rows!