

Topic 3: Tuples, Vectors, and Matrices

02-680: Essentials of Mathematics and Statistics

August 26, 2024

1 Tuples

Unlike sets, **tuples** (also called **sequences** or **lists**) are an *ordered* list of objects. Think of the position on chess board (or a 2D plane), a color in RGB, etc.

When these have small cardinality we can use terms like (ordered) pair [2], triple [3], quadruple [4], or more generically an “ n -tuple”.

If we’re being precise, we normally use angle brackets (“ \langle ”, “ \rangle ”) but a lot of times we will be lazy and just use parentheses (“(”, “)”)

Cartesian Product. A very useful way to construct **set of** tuples is using the **cartesian product** operator, which in essence creates all possible pairs of elements from two sets.

$$S \times T = \{\langle x, y \rangle \mid x \in S \wedge y \in T\}$$

As an example, lets remember the first two sets from our examples last topic:

$$A = \{\text{“Welcome”, “to”, “02-680”}\} \text{ and}$$

$$B = \{x^2 \mid x = 2 \vee x = 3\}.$$

In this case the cartesian product is

$$A \times B = \{\langle \text{“Welcome”, } 4 \rangle, \langle \text{“to”, } 4 \rangle, \langle \text{“02-680”, } 4 \rangle, \langle \text{“Welcome”, } 9 \rangle, \langle \text{“to”, } 9 \rangle, \langle \text{“02-680”, } 9 \rangle\}$$

It doesn’t have to be different sets in the cartesian product though, we can have the product with a set and itself. In fact this is performed so often it has its own notation:

$$B \times B = B^2 = \{\langle 4, 4 \rangle, \langle 4, 9 \rangle, \langle 9, 4 \rangle, \langle 9, 9 \rangle\}.$$

This notation also generalizes, so $S^3 = S \times S \times S$, $S^4 = S \times S \times S \times S$ and so on. Notice that order matters in tuples (unlike sets) so $\langle 4, 9 \rangle \neq \langle 9, 4 \rangle$.

A note about this notation: Sometimes we want to have a set of tuples of different lengths (remember sets don't need to be over objects of the same type) so something like

$$B^2 \cup B^3 = \left\{ \begin{array}{l} \langle 4, 4 \rangle, \langle 4, 9 \rangle, \langle 9, 4 \rangle, \langle 9, 9 \rangle, \\ \langle 4, 4, 4 \rangle, \langle 4, 4, 9 \rangle, \langle 4, 9, 4 \rangle, \langle 4, 9, 9 \rangle, \\ \langle 9, 4, 4 \rangle, \langle 9, 4, 9 \rangle, \langle 9, 9, 4 \rangle, \langle 9, 9, 9 \rangle \end{array} \right\}$$

If we wanted to enumerate all binary numbers up to 8 digits (while omitting leading 0s):

$$\{1\} \times \bigcup_{i=1}^7 \{0, 1\}^i$$

But that leads to the notation we saw last time for strings: Σ^* . Sometimes we want the set of all tuples of any length, then we use the **Kleene star** (or Kleene operator); for some set S ,

$$S^* = \bigcup_{i=0}^{\infty} S^i.$$

We will define $S^0 = \langle \rangle$ (the empty tuple) for any S , in the case of Σ^0 we often call it the empty string.

2 Vectors

When the set used to define a tuple is the set of real \mathbb{R} (or complex \mathbb{C}^1) numbers, we call the tuple a **vector**. Specifically an n -vector x is defined as an element in

$$x \in \mathbb{R}^n.$$

If we want to reference the i -th element of x we will write

$$x_i$$

(or sometimes $x[i]$, this is true of tuples as well).

2.1 Simple Operations

Graphically we can think of vectors (in \mathbb{R}^2) in two way, which are... somewhat equivalent: as a point on the plane, or as an arrow from the origin. The second will be useful in this section, but the latter is sometimes useful as well.

¹It will generally be true throughout the class that the properties we're discussing also apply to complex numbers, but for simplicity we will usually only directly discuss reals.

Vector length. If we think of the vector as an arrow, we can say the *length* of the vector (arrow) is the same as the hypotenuse right triangle with each leg having the same length as each one of the elements. In that case we know that for vector $x = \langle x_1, x_2 \rangle \in \mathbb{R}^2$, the length is $\sqrt{x_1^2 + x_2^2}$. To generalize this we say the length of a vector $x \in \mathbb{R}^n$

$$\|x\| = \sqrt{\sum_{i=1}^n (x_i)^2}.$$

Often times we call this the L_2 -norm of a vector.

Vector addition and scalar multiplication. Both of these operations are element-wise. For vectors $x, y \in \mathbb{R}^n$

$$z = x + y \rightarrow z_i = x_i + y_i \quad \forall 1 \leq i \leq n.$$

Thus $z \in \mathbb{R}^n$ as well.

In the context of this course, especially when talking about vectors, a *scalar* is a single number (i.e. $a \in \mathbb{R}$) rather than a vector. (The implication is that \mathbb{R} and \mathbb{R}^1 are not the same thing.) When multiplying a vector $x \in \mathbb{R}^n$ by a scalar $a \in \mathbb{R}$, we once again apply this element-wise. Thus the result, $z \in \mathbb{R}^n$ can be computed as:

$$z = ax \rightarrow z_i = ax_i \quad \forall 1 \leq i \leq n.$$

2.2 Dot Product

While a scalar times a vector is a vector, it turns out a vector times a vector is a scalar! Lets define it first then we will dive in; for $x, y \in \mathbb{R}^n$

$$x \bullet y = \sum_{i=1}^n x_i \cdot y_i.$$

(note sometimes we will use \bullet vs \cdot to differentiate scalar multiplication and dot product, but generally only the latter is used since the domains of the functions are different and can be extracted from context.)

We often think of dot product as telling us how vectors go “in the same direction”. Consider two cardinal vectors (going directly along an axis) on the plane; intuitively they go in totally different directions (note, this is different from *opposite* directions). WLG one vector must be $\langle a, 0 \rangle$ and the other must be $\langle 0, b \rangle$ (for scalars $a, b \in \mathbb{R}$). Using the definition above $\langle a, 0 \rangle \bullet \langle 0, b \rangle = 0$. Consider a third vector $\langle c, 0 \rangle$ with $c \in \mathbb{R}$, but enforce that $a \geq 0$ and $c \leq 0$. In this case the result is $\langle a, 0 \rangle \bullet \langle c, 0 \rangle = ac$, which we know is negative; they share a lot of direction, but go opposite ways!

We can actually redefine the L_2 -norm of a vector $x \in \mathbb{R}^n$ using the dot product:

$$\|x\| = \sqrt{x \cdot x}.$$

3 Matrices

You can almost think of a **matrix** as a 2-dimension vector. We say that an “ n -by- m ” matrix $M \in \mathbb{R}^{n \times m}$ has n rows and m columns and we usually write it as:

$$M = \begin{bmatrix} M_{1,1} & M_{1,2} & \dots & M_{1,m} \\ M_{2,1} & M_{2,2} & \dots & M_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ M_{n,1} & M_{n,2} & \dots & M_{n,m} \end{bmatrix}$$

3.1 Special Matrices

In a square matrix $N \in \mathbb{R}^{n \times n}$, we define the **main diagonal** as the entries where the horizontal and vertical component are equal; i.e. $\{N_{i,i} \mid 1 \leq i \leq n\}$.

The **identity** matrix $I_n \in \mathbb{R}^{n \times n}$ (sometimes simplified to just I when the size is implied from context) is a special square matrix where the main diagonal values are 1 and all other values are 0.

$$\forall 1 \leq i, j \leq n : I_{i,j} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

3.2 Matrix Operations

Addition and Scalar Multiplication. Like with vectors, addition of two matrices as well as scalar multiplication are element-wise operations, so for matrices $M, N \in \mathbb{R}^{n \times m}$ and scalar $a \in \mathbb{R}$:

$$O = M + N \rightarrow O_{i,j} = M_{i,j} + N_{i,j} \quad \forall 1 \leq i \leq n, 1 \leq j \leq m$$

$$O = aM \rightarrow O_{i,j} = aM_{i,j} \quad \forall 1 \leq i \leq n, 1 \leq j \leq m$$

Matrix Multiplication Just like with vectors, multiplying two matrices is more complicated than scalars. The first question is the size of the result, if we multiply $C \in \mathbb{R}^{n \times p}$ with $D \in \mathbb{R}^{p \times m}$ we get a matrix $E \in \mathbb{R}^{n \times m}$; notice that the *inner* dimensions are the same. And the values in E are defined as follows:

$$E_{i,j} = \sum_{k=1}^m C_{i,k} D_{k,j}$$

We can actually rewrite this using dot product, lets say that $C_{i,*}$ is the i -th column of C , and $D_{*,j}$ is the j -th column of D . In that case

$$E_{i,j} = C_{i,*} \cdot D_{*,j}^T.$$

What can we do with it? Lets define the following:

- G is an n -by- m matrix where $G_{i,j} = 1$ if actor i was in an episode of the show j (and 0 otherwise)
- H be an m -by- p matrix where $H_{j,k} = 1$ if the show j is available to stream on service k (and 0 otherwise)

Useful References

Liben-Nowell, “Connecting Discrete Mathematics and Computer Science, 2e”. §2.4