COMP0014 Tutorial 2

Yuzuko Nakamura

More Linear Algebra Basics

Norms and magnitude

A **norm** is a function that maps vectors to the set of non-negative real numbers and obeys certain properties. More formally, we can say a norm is a function $f : V \to \mathbb{R}$, where V is a vector space. There are multiple norms that exist, and several of these norms are important in the field of machine learning, but for now we will focus on the most common norm: the L2 norm or *Euclidean norm*:

$$\|v\| = \|v\|_2 = \sqrt{\sum_{i=1}^{n} v_i^2} \tag{1}$$

The **magnitude** of a mathematical object refers to the size of that object. Magnitudes are defined for multiple types of objects other than vectors, but in the case of vectors, the magnitude is just the Euclidean norm, so magnitude and norm are synonyms. They both refer to the length of a vector in Euclidean space, which is calculated by the equation above: the square root of the sum of squares.

Norms obey the following properties:

- Subadditivity, or triangle inequality, i.e. if f is a norm, then for all $v_1, v_2 \in V$

$$f(v_1 + v_2) \leq f(v_1) + f(v_2) \tag{2}$$

- Scalability, i.e. for all $v \in V$ and $\lambda \in \mathbb{R}$

$$f(\lambda v) = \lambda f(v) \tag{3}$$

- The zero vector is (the only vector) assigned a norm of 0.

$$f(\mathbf{0}) = 0 \tag{4}$$

$$f(v) = 0 \implies v = \mathbf{0} \tag{5}$$

Matrices

A matrix is a rectangular 2-D array of numbers, each drawn from a field $\mathbb{F}$. You can think of it as a set of vectors concatenated horizontally (or vertically).

Matrix multiplication

Matrix multiplication is like the vector dot product, except that the dot product is performed for each row vector from the left side matrix paired with each column vector of the right side matrix. That is to say, if matrix A is the product of matrices B and C, A_{ij} is the ith row of B dotted with the jth column of C.

In order for two matrices to be multiplied together, their inner dimensions have to match. The dimensions of matrices are specified in (rows, columns) order. So an $m \times n$ matrix (m rows, n columns) can be multiplied with an $n \times k$ matrix, for any integers m, n, k. The result of the multiplication will be a matrix with the outer dimensions, i.e. an $m \times k$ matrix.

Inverse matrices

Some matrices have an inverse A^{-1}, such that $AA^{-1} = I$ and $A^{-1}A = I$. I is the identity matrix: the matrix consisting of 1's along the diagonal (upper-left to lower-right) and 0's everywhere else. Multiplying the identity matrix by any other vector or matrix will result in the original vector or matrix again, i.e.:

$$Ix = x \qquad \text{for any matrix or vector x} \tag{6}$$

assuming that I's dimensions are such that the above matrix multiplication can be performed.

In other words, if multiplying a matrix by A is seen as applying a linear transformation, then multiplying the result by A^{-1} will undo that linear transformation.

Inverse matrices only exist for square matrices that are full rank. **Rank** refers to the maximal number of linearly independent column vectors in a matrix. If all the column vectors of $n \times n$ matrix A are linearly independent, the rank of the matrix is n and the matrix has full rank and is invertible. If the rank is less than n (i.e. at least one vector can be written as a linear combination of the others i.e. at least one of the vectors is "redundant" and could be removed without changing the spanning set of the column vectors of A), the matrix is said to be **rank deficient**.

Column and row spaces

In a matrix, the columns (or rows) can be thought of as a set of vectors. The spanning set (recall from last week) of the set of column or row vectors is called the **column space** and **row space** respectively. (Note: The column space of matrix A is the row space of matrix A^T and vice versa.)

If you have a matrix with m rows (so each column vector $\in \mathbb{F}^m$ for some field $\mathbb{F}$), then the column space will be a **subspace** of $\mathbb{F}^m$. A (linear) subspace is a portion/subset of another vector space, which is itself a vector space. It is the portion of $\mathbb{F}^m$ reachable through linear combinations of the column vectors, and obeys the addition and scalar multiplication closure properties of vector spaces.

For a full rank matrix, the column space is just $\mathbb{F}^m$. However, the column space of a deficient rank matrix will span only a particular portion of $\mathbb{F}^m$ (which depends on the particular column vector(s) involved).