COMP0014 Tutorial 2

Yuzuko Nakamura

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.:

$$Ix = x \qquad \text{for any matrix or vector x} \tag{1}$$

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**.

Matrix transpose

A transpose is an operation that can be done to matrices where the rows and columns are flipped. To be more precise, if A is an $m \times n$ matrix, then A^T is an $n \times m$ matrix where each element a^T_{ij} (the element at row i and column j) of A^T is:

$$a^T_{ij} = a_{ji} \qquad \text{for all i, j} \tag{2}$$

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). The dimension of this portion of a vector space (subspace) will be the number of linearly-independent column vectors.

Null space

The null space of a matrix A is the set of vectors such that $Ax = 0$. The **nullity** refers to the dimension of the null space in the same way that *rank* refers to the dimension of the column space of A.

The nullity-rank theorem states:

$$rank(A) + nullity(A) = ncols(A) \tag{3}$$

This is important because some findings about the rank of a matrix can be proved by proving something about the null space of the matrix instead.

Symmetric matrices

A symmetric matrix is a particular kind of matrix where elements mirrored across the diagonal (that is, with row index and column index transposed) are identical. In other words:

$$a_{ij} = a_{ji} \qquad \text{for any i, j} \tag{4}$$

An equivalent way of stating symmetry is:

$$A = A^T \tag{5}$$

As such, only square matrices can be symmetric.

Some notable things ways to create symmetric matrices (you can prove these are true when doing the notebook):

$$A + A^T \qquad \text{is a symmetric matrix for any matrix } A \tag{6}$$

$$A^T A \qquad \text{is a symmetric matrix for any matrix } A \tag{7}$$

(Same goes for AA^T.)

Positive definiteness

When it comes to symmetric matrices (made up of real numbers/defined on the field $\mathbb{R}$) specifically, we can talk about their positive/negative (semi)definiteness. There are several possible qualities, the most important of which are positive definite and positive semi-definite:

- A is *positive semidefinite (psd)* ($A \succeq 0$) iff.

$$x^T A x \geq 0 \qquad \text{for all non-trivial } x \ (x \neq 0) \tag{8}$$

- A is *positive definite (pd)* ($A \succ 0$) iff.

$$x^T A x > 0 \qquad \text{for all } x \neq 0 \tag{9}$$

- A is *negative semidefinite (nsd)* ($A \preceq 0$) iff.

$$x^T A x \leq 0 \qquad \text{for all } x \neq 0 \tag{10}$$

- A is *negative definite (nd)* ($A \prec 0$) iff.

$$x^T A x < 0 \qquad \text{for all } x \neq 0 \tag{11}$$

- A is *indefinite* otherwise.

Positive definite matrices are of special interest because when combined with a vector variable x in a quadratic form (for example, $x^T A x = c$) they represent a geometric object: an ellipsoid. This is similar to the way the equations $x^2 = r$ or $ax^2 + by^2 = c$ can be used to represent circles and ellipses, but with the principle extended to higher dimensions and linear algebra-style setup. These quadratic forms are useful in various multivariate applications such as linear regression.

Properties of $X^T X$

Because of the way matrix multiplication works, symmetry and quadratic matrix forms (of the type $X^T X$) are strongly connected. The $X^T X$ matrix has some particular properties (you can prove these). For any matrix X, the following are true:

- $rank(X^T X) = rank(X)$
- $X^T X$ is positive semidefinite.
- In addition, $X^T X$ is also positive definite iff. it is full rank.