Systems of equations and translation into matrices

August 5, 2022

A linear equation in n variables $x_1, \ldots, x_n$ over real numbers (there are other frames of reference) is **typically** of the form

$$a_1x_1 + \cdots + a_nx_n = b_1$$

One could also consider a set of m such equations written on successive lines one below the other forming m rows. We ensure that the same variables are present in the same column in different equations. The coefficient of any variable in any equation is some real number. We have denoted it in the above as a_i for some index i. If a particular variable is missing in some equation, it can be treated as present with coefficient 0.

One can rewrite a system of m linear equations in n variables as a matrix equation

$$AX = b$$

where A is an $m \times n$ matrix, X is a column vector of height n, also describable as an $n \times 1$ matrix and b is a column vector of height m also describable as an $m \times 1$ matrix.

Here the matrix A consists of only the coefficients of the variables from the equations with the variable names removed. The variables appear in the column vextor X. The constants on the right hand side of the equations constitute the column vector b.

When every entry in the column vector b is 0, the system of equations is called homogeneous. A homogeneous system of equations **always** has a solution because the vector with every entry 0, satisfies such a system of equations.

We also observed that one linear equation in any number of variables always has a solution. For inconsistency to arise, we need at least two equations.

Each linear equation in n real valued variables with real coefficients is basically some criterion to partition the points of R^n, the n Euclidean space into points that satisfy the equation and points that do not. So each equation determines a subset of points that conform. So in other words it is a subset of points in R^n that satisfy the equation. When taken across m such equations what we get is m subsets of points in R^n. The solution to the entire system of equations is the intersection of these m subsets. Thus the solution to a system of simultaneous linear equations in a certain numbr of variables can be viewed as a subset of the euclidean space R^n.

Alternatively, one can view any matrix as either the aggregate of its individual column vectors or its individual row vectors. We take the aggregate of column vectors view, together with a solution to the system and see what it indicates. By looking at the expansion, as we did in the lecture we can say that the column vector b of dimension m is being generated as a linear combination of the column vectors of the matrix A, which are also of dimension m. The multipliers of the column vectors to generate the solution vector b is some solution of the system of linear equations.

Thus one can view the solution of a system of m linear equations in n variables as the intersection of m subsets of R^n or alternatively as the multiplies to n column vectors of dimension m each, needed to generate the column vector b of dimension m.

This latter viewing corresponds to the idea of generating vectors of some dimension by linear combinations of given vectors of the same dimension. In general one vector can only generate vectors that are scaled versions of itself. Two vectors may generate more, provided they are not scaled versions of eah other. However they may not generate the whole space. Given a set of vectors in R_m, the set of all vectors that can be generated by linear combinations of these is called the **column space** of this set of vectors. Instead of talkig about some arbitrary set of column vectors of dimension m each we can talk about the n column vectors of dimension m each constituting the n coluns of an $m \times n$ matrix. This then becomes the **column space of that matrix**.

One can view the matrix of coefficients and variables while allowing b to be varying. Under this model, we can talk about spaces of the matrix. The system of equations has a solution if and only if the column vector b is in the column space of the matrix as defined earlier. When we limit ourselves to homogeneous equations, then the solutions (instantiations of constants into the variable column vector X) are called the **null space** of that matrix.

Thus when we move from a system of equations to their matrix counterpart and look only at the coefficient matrix and allow b to vary, we get new notions of the column space and null space of a matrix. All these will be relooked at in the context of subspaces when we look at vecror spaces, later on in the course.