Lecture 4-6

We now begin the linear algebra portion of the course, starting with the very concrete setting of systems of linear equations. Given such a system of m equations in n unknowns $x_1, \ldots, x_n$, we choose notation so that the ith equation reads $\sum m_{ij}x_j = b_i$ (for $1 \le i \le m$), where the coefficients m_{ij} and constants b_i on the right side are known (and assumed to be real, though we could in fact allow numbers from any field here). We now know that if we define an $m \times n$ matrix M by decreeing that its ij-th entry be m_{ij}, and if we further write X, B for the respective column vectors $(x_1, \ldots, x_n), (b_1, \ldots, b_n)$ of unknowns and right-hand sides of the equations, then we may rewrite the entire system as a single matrix equation $MX = B$. Let M' be the so-called *augmented* matrix obtained from M by adding B as a new column on the right.

We solve this system be repeatedly replacing the equations by simpler ones with the same solutions. More precisely, given two equations $a_1x_1 + \ldots + a_nx_n = a, b_1x_1 + \ldots + b_nx_n = b$ with $a_1 \neq 0$, we can replace the second equation by itself minus $\frac{b_1}{a_1}$ times the first equation to get a second pair of equations with (clearly) exactly the same (possibly empty) solution set as the first pair, in which the coefficient of the first variable x_1 in the second equation is 0. Regarding the coefficients $(a_1, \ldots, a_n, a), (b_1, \ldots, b_n, b)$ of the variables x_i, followed by the constants a, b on the right sides, as vectors R_1, R_2, respectively, we see that we are subtracting $\frac{b_1}{a_1}$ times R_1 from R_2. We call this a *row operation* and sometimes denote it as $R_2 \to R_2 - (b_1/a_1)R_1$. Iterating this idea and starting now from the augmented matrix M', assume that its top left entry m_{11} is nonzero. Label the rows of M' as $R_1, \ldots, R_m$ from top to bottom. Then by subtracting suitable multiples of R_1 from $R_2, \ldots, R_m$ we arrive at the augmented matrix of another system with the same solution set as the original system. Now if some row of this new system takes the form $(0, \ldots, 0, x)$ for some $x \neq 0$, then that row corresponds to an equation $0x_1 + \ldots + 0x_n = x$, which already has no solution; so certainly the original system has no solution.

We are now completely done with both the first row and first column of our system; no entry in either of these will ever change henceforth. We proceed to look at the second column in the new system. If the second entry of this column is nonzero, then by subtracting suitable multiples of the second row from the third and lower rows, we arrive at a new system equivalent to the original one in which all entries below the second one in the second column and all entries below the first one in the first column are 0. Once we have finished with the second column (and thereby also the second row), we proceed to the third column, and so on, until we either reach the last equation of the system or run out of columns (we will not get to all of its columns if it has more unknowns than equations). Along the way, if we come to column i and its ith entry is 0 but some lower entry, say the jth, in the same column is nonzero, then we interchange rows i and j (denoted $R_i \leftrightarrow R_j$) in the system before proceeding further; clearly this new row operation again leads to an equivalent system. Otherwise, if the ith and lower entries in the ith column are all 0, we are done with the ith column and so proceed to the next column.

In the end we arrive at an equivalent system whose augmented matrix has all zeroes below the leftmost nonzero entry (called the *pivot*) in every row and the position of this leftmost nonzero entry moves strictly to the right as we go down the equations in the system. Also if any row of the system consists entirely of zeroes, then all lower rows also consist of zeroes. Such a system is said to be in *echelon*(=staircase) form. We then solve it in reverse order, beginning with the last equation and proceeding to the first one. For example, if the original system had $M = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 10 \end{pmatrix}$ and $B = \begin{pmatrix} 6 \\ 15 \\ 25 \end{pmatrix}$, then after clearing out the first column our augmented matrix is $\begin{pmatrix} 1 & 2 & 3 & 6 \\ 0 & -3 & -6 & -9 \\ 0 & -6 & -11 & -17 \end{pmatrix}$, while after clearing out the second column the matrix becomes $\begin{pmatrix} 1 & 2 & 3 & 6 \\ 0 & -3 & -6 & -9 \\ 0 & 0 & 1 & 1 \end{pmatrix}$. Now the last equation of this last system clearly implies $x_3 = 1$, whence we get $x_2 = 1$ from the second equation, whence finally $x_1 = 1$ from the first equation. The unique solution to the system is $(1 \quad 1 \quad 1)$. Had the last equation in the original system been replaced by $7x_1 + 8x_2 + 9x_3 = 25$, then all steps in solving the system would have been the same, except that the last equation of the last system would have read $0x_1 + 0x_2 + 0x_3 = 1$, which clearly has no solution; so the original system has no solution as well. Had the last equation in the original system been replaced by $7x_1 + 8x_2 + 9x_3 = 24$, then the last equation of the last system would have read $0x_1 + 0x_2 + 0x_3 = 0$. Now of course the situation is very different. We can in fact assign any value we like to x_3, then use the second equation to solve for x_2 and finally use the first one to solve for x_1; thus we get a family of solutions instead of just one solution. We say that x_3 is a *free variable* in this case; it corresponds to the (in this case only) row of the original system that has no pivot in the echelon form. The other variables x_1, x_2 are called *bound* (since they are determined by the values of the free variables). We will observe next time that *a system of linear equations is inconsistent (i.e. has no solution) if and only if some row of its echelon form has just a single nonzero as its rightmost entry.*