NOTES FOR 6 AUG (TUESDAY)

1. Recap

(1) Defined fields and gave examples and counterexamples. (Basically, the only new example was $\mathbb{Z}_n$.)
(2) Showed that row equivalent coefficient matrices give rise to homogeneous equations with the same solution sets.

2. Row operations

Here is a definition : An $m \times n$ matrix A is called row-reduced if :

(1) The first non-zero entry of every row is 1.
(2) The other entries below and above these $1'$s are zero.

Here is a useful proposition : Every matrix over a field is row equivalent to a row-reduced matrix.

Proof. A $1 \times n$ matrix can obviously be row reduced in 1 step. Inductively, consider an $(m+1) \times n$ matrix. First make the pivot of the first row 1. Then row-reduce the the $m \times n$ submatrix (leaving the first row intact). Use the leading $1'$s (the pivots) to make the corresponding entries in the first row zero. □

Here is another definition : A matrix A is said to be in row echelon form if it is row reduced and if the pivot in row k occurs in the k_jth column, then $k_1 < k_2 < \ldots$. It is easy to come up with examples and counterexamples to these concepts.
Every matrix is row equivalent to a row echelon matrix. Indeed, row reduce the matrix and use a finite number of row interchanges to bring it to the row echelon form.
Here is a proposition about solutions of homogeneous equations : If A is an $m \times n$ matrix and $m < n$, then $AX = 0$ has a non-trivial solution.

Proof. Bring A to its row echelon form B. There are at most m pivots and n variables. Hence, we can solve for at most m variables with at least $n - m$ free variables. Therefore there is a non-trivial solution. □

Here is a related result : If A is an $n \times n$ square matrix, then A is row-equivalent to the identity matrix iff $AX = 0$ has only the trivial solution.

Proof. The "if" direction is trivial. If $AX = 0$ has only the trivial solution, then after row-reducing to the echelon form $BX = 0$, the last pivot has to occur in the last row (if the last row is identically zero, we have non-trivial solutions). Hence $B = I$. □

Now we shall deal with the real problem : Solving $AX = Y$. Note that if $AX_0 = Y$, then any other solution of $AX = Y$ is of the form $X_0 + Z$ where $AZ = 0$. So the problem is to know whether there is even a single solution or not. (If there is none, such a system is said to be inconsistent.) When dealing with such equations, row operations should be performed on the right hand side as well. Therefore, we consider the "augmented" matrix $A'_{ij} = A_{ij}\ \forall\ 1 \leq i \leq m, 1 \leq j \leq n$ and $A'_{in+1} = Y_i\ \forall\ 1 \leq i \leq m$. Clearly, row-equivalent augmented matrices give rise to equivalent equations. In the row-echelon form, it is easy to see that if in any row, the first n columns are zero and the last entry is not, then we have an inconsistent system. Otherwise, we can simply solve for the pivots and set the other variables to arbitrary values. It is once again easy to look at examples.

Terminology : The number of non-zero rows in the row echelon form is called the row rank of A.

3. Matrix multiplication and invertible matrices

Since taking linear combinations of rows is so important, given a bunch of $m \times n$ elements of the field A_{ij}, and a bunch of rows $\beta_1, \ldots, \beta_n$, new rows $\gamma_i = \sum_j A_{ij}\beta_j$ define a matrix C such that $C_{ik} = (\gamma_i)_k$. Note that if $B_{jk} = (\beta_j)_k$, then $C_{ik} = \sum_j A_{ij}B_{jk}$. We product of an $m \times n$ matrix A and an $n \times p$ matrix B to be the $m \times p$ matrix $C = AB$ given as above. Clearly, products of arbitrary matrices do not make sense. Moreover, matrix multiplication is not commutative (even over $\mathbb{Z}_2$). It is crucial to note that the rows of C are linear combinations of the rows of B with the coefficients being A. Here are important examples.

(1) $AI = IA = A$ for square matrices.
(2) $0A = 0$ for appropriately sized zero matrices.
(3) $AX = Y$ is honest matrix multiplication.
(4) Matrix multiplication is associative : Indeed, $(A(BC))_{il} = \sum_j A_{ij}(BC)_{jl} = \sum_j A_{ij} \sum_k B_{jk}C_{kl} = \sum_k (\sum_j A_{ij}B_{jk})C_{kl} = ((AB)C)_{il}$.
In particular, A^n is unambiguously defined.
(5) An elementary matrix is an $m \times m$ matrix that can be obtained from the $m \times m$ identity matrix by means of an elementary row operation. Let e be an elementary row operation and $E = e(I)$. Then $e(A) = EA$.

Proof. We shall do this case-by-case.
(a) $R_r \to cR_r$. Then $(E)_{rj} = c\delta_{rj}$, $E_{ij} = \delta_{ij}$ $i \neq r$. Also, $e(A)_{rj} = cA_{rj}$, $e(A)_{ij} = A_{ij}$ $i \neq r$. Hence $e(A) = EA$.
(b) $R_r \leftrightarrow R_s$. In this case $(E)_{rj} = \delta_{sj}$, $E_{sj} = \delta_{rj}$, $E_{ij} = \delta_{ij}$ $\forall\ i \neq r, s$. Clearly, $(EA)_{rk} = \sum_j E_{rj}A_{jk} = A_{sk}$ and likewise for $(EA)_{sk}$. $(EA)_{ik} = \sum_j \delta_{ij}A_{jk} = A_{ik}$.
(c) $R_r \to R_r + cR_s$. $(E)_{ij} = \delta_{ij}$ $\forall\ i \neq r$. $E_{rj} = \delta_{rj} + c\delta_{sj}$. Note that $(EA)_{ik} = A_{ik}$ $\forall\ i \neq r$. $(EA)_{rk} = \sum_j (\delta_{rj} + c\delta_{sj})A_{jk} = A_{rk} + cA_{sk}$.

□