WEEK 2: NONSINGULAR MATRICES

PADRAIC BARTLETT

1. Last Week's HW

The average was 46/50; consequently, there's not much to say. The proofs in questions #2 and #3 could have been more rigorous, but Prof. Wilson's comments on those questions made it clear that he wasn't expecting them to be terribly tight, so I wasn't too concerned.

2. Random Questions

Question 2.1. *How many rooks can you place on a* 8×8 *chessboard, so that no two pieces can capture each other? How many distinct ways are there of placing this number of rooks on a board?*

Question 2.2. *How many queens can you place on a* 8×8 *chessboard, so that no two pieces can capture each other? How many distinct ways are there of placing this number of queens on a board?*

3. Nonsingular Matrices - Theory

Definition 3.1. A $n \times n$ matrix A is called **invertible** or **nonsingular** if there is a matrix B such that $AB = BA = I_n$. (Here, we denote the $n \times n$ matrix with 1's on its diagonal and 0's everywhere else I_n, and call it the $n \times n$ identity matrix.)

Remark 3.2. Why do we only talk about $n \times n$ matrices being nonsingular? Why not $n \times m$ matrices? The answer is because we consider things to be nonsingular, or "invertible," if there is an inverse to multiply them by to get the identity, and there is in fact no identity element for $n \times m$ matrices – i.e. there is no matrix $I_{n \times m}$ such that for any $n \times m$ matrix A, $AI_{n\times m} = I_{n\times m}A = A$.

The reason for this is elementary: by the rules of matrix multiplication, we know that the only matrices which we can multiply A on the right by are $m \times k$ matrices; if such a matrix then yielded A, we know that k must be equal to m, because A is a $n \times m$-dimensional matrix. As well, the only matrices we can multiply A on the left by are $l \times n$ matrices. So, any matrix that we can multiply A on the left and the right by and get A back has to be both a $m \times m$ and a $l \times n$ matrix at the same time; this is only possible when $m = n$.

Remark 3.3. The above notion gives us the idea of $n \times n$ matrices possibly forming a "group" – i.e. a collection of elements that has some defined operation (multiplication) that has the properties we associate multiplication with in "real life" – i.e. inverses, identity, and associativity. More formally, consider the following definition:

Definition 3.4. A **group** is a set G endowed with a binary operation $\bullet : G \times G \to G$, that obeys the following properties:

- **Associativity**: For all $a, b, c \in G$, $(a \cdot b) \cdot c = a \cdot (b \cdot c)$
- **Identity**: There is a distinguished element $id \in G$ such that for every $g \in G$, $id \cdot G = G \cdot id = G$.
- **Inverses**: For every $g \in G$ there is a distinguished element $g^{-1} \in G$ such that $g \cdot g^{-1} = g^{-1} \cdot g = id$.

Examples 3.5.

- The integers with respect to addition, $(\mathbb{Z}, +)$, form a group: this is because every element has an inverse, and there is an identity element – specifically, 0.
- The integers with respect to multiplication, $(\mathbb{Z}, \cdot)$, do not form a group: this is because many elements do not have inverses (2, for example.)
- The collection of all real-valued $n \times n$ matrices with respect to matrix addition, $(M_n(\mathbb{R}, +)$, forms a group; here, the all-zero matrix is the identity element and the inverse of a matrix $A = \{a_{ij}\}_{i,j=1}^{n,n}$ is the obvious choice $\{-a_{ij}\}_{i,j=1}^{n,n}$.
- The collection of all real-valued $n \times n$ matrices with respect to matrix multiplication, $(M_n(\mathbb{R}, \cdot)$, does not form a group; this is because while an identity exists (I_n), there are many matrices without inverses – i.e. the all-zero matrix, or more generally any singular matrix.
- The collection of all real-valued *nonsingular* $n \times n$ matrices with respect to matrix multiplication, $(GL_n(\mathbb{R}, \cdot)$, trivially forms a group, as these are just the matrices defined to have inverses. This is a very interesting mathematical object to study; over the duration of the course, I'll try to introduce more asides on the mathematical structure on the objects we are studying and try to explain why they have the properties they do.

4. Nonsingular Matrices - Applications

Question 4.1. *When is a $n \times n$ matrix A invertible?*

Answer 4.2. When it has rank n!

If you forgot what rank was, review the definition below:

Definition 4.3. A matrix A has rank n if its row-echelon form has precisely n nonzero rows. Specifically, if A is a $n \times n$ matrix, it has rank n if it can be reduced through elementary row operations to the identity matrix I_n.

These notions are in fact compatible; if a $n \times n$ matrix has n nonzero rows and is in row-echelon form, we know that those rows are all of the form $(0 \dots 1 \dots 0)$, by the definition of row-echelon form; by permuting rows through elementary row operations, we can get the identity matrix.

Similarly, if a matrix A can be transformed to the identity matrix via a series of elementary row operations, we know that $rank(A) = n$ because rank is invariant under elementary row operations (and, I guess, because $rank(I_n) = n$.)

Definition 4.4. Alternately, we say that a $n \times n$ matrix A has "full" rank (i.e. rank n) if you cannot write any of its rows as a linear combination of the other rows (i.e. you couldn't write the second row as three times the first row minus twelve times the third.)

This notion is, again, equivalent to the above definitions: if a $n \times n$ matrix A had rank $< n$, then it has at least one row made entirely of 0's after it is reduced to row-echelon form. Furthermore, we got this "zero row" by adding copies of other rows to this row (and possibly multiplying it by a nonzero constant); so there is some way of writing this row as a combination of the other rows!

The converse is equally fast: if we have some such combination, then the matrix A can be reduced to a matrix A' with one all-zero row through elementary row operations – specifically, the operations that subtract copies of other rows from this "special row" to reduce it to 0. Because elementary row operations do not change the row-echelon form of the ending matrix, we can see that $rank(A) = rank(A') < n$.

So, this is all well and good; but how do we actually compute an inverse?

Idea 4.5. In transforming a matrix to row-echelon form, you're just multiplying it by elementary matrices.

Idea 4.6. We can invert elementary matrices.

Idea 4.7. If we can write (for A an $n \times n$ identity matrix and E_1 elementary matrices)

$$A \cdot E_1 \cdot \ldots E_n = I,$$

then we can write A^{-1} as just

$$E_1 \cdot \ldots E_n.$$

So: a practical example: take a matrix

$$A = \begin{pmatrix} 2 & 1 & 3 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 2 & 3 & 0 \\ 2 & 0 & 0 & 0 \end{pmatrix}$$

How can we find its inverse?

Well: let's write the matrix

$$[A|I_4] = \left(\begin{array}{cccc|cccc} 2 & 1 & 3 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 2 & 3 & 0 & 0 & 0 & 1 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right)$$

and reduce the left-hand portion of the matrix to the identity matrix. Then, the elementary operations we use to do this will appear as a product on the right hand side of the matrix: by our ideas above, this product of elementary matrices will in fact be A^{-1}!

So: we calculate blindly.

$$
\begin{aligned}
[A|I_4] &= \left(\begin{array}{cccc|cccc} 2 & 1 & 3 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 2 & 3 & 0 & 0 & 0 & 1 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right) \\
&\sim_R \left(\begin{array}{cccc|cccc} 1 & 1/2 & 3/2 & 0 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 3/2 & 3/2 & 0 & -1/2 & 0 & 1 & 0 \\ 0 & -1 & -3 & 0 & -1 & 0 & 0 & 1 \end{array}\right) \\
&\sim_R \left(\begin{array}{cccc|cccc} 1 & 1/2 & 3/2 & 0 & 1/2 & 0 & 0 & 0 \\ 0 & -1 & -3 & 0 & -1 & 0 & 0 & 1 \\ 0 & 3/2 & 3/2 & 0 & -1/2 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \end{array}\right) \\
&\sim_R \left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 \\ 0 & 1 & 3 & 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & -3 & 0 & -2 & 0 & 1 & 3/2 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \end{array}\right) \\
&\sim_R \left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 \\ 0 & 1 & 0 & 0 & -1 & 0 & 1 & 1/2 \\ 0 & 0 & 1 & 0 & 2/3 & 0 & -1/3 & -1/2 \\ 0 & 0 & 0 & 1 & -2/3 & 1 & 1/3 & 1/2 \end{array}\right).
\end{aligned}
$$

We thus conclude that

$$
A^{-1} = \begin{pmatrix} 0 & 0 & 0 & 1/2 \\ -1 & 0 & 1 & 1/2 \\ 2/3 & 0 & -1/3 & -1/2 \\ -2/3 & 1 & 1/3 & 1/2 \end{pmatrix}.
$$