MATH 3311, FALL 2025: LECTURE 34, NOVEMBER 19

Video: `https://youtu.be/wKsvtfyWcFM`
Smith Normal Form

Definition 1. A matrix $A \in M_{m\times n}(\mathbb{Z})$ is in **Smitn Normal Form** or **SNF** if:

(1) All its non-zero entrie are along the diagonal.
This means that the matrices
$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \end{pmatrix}$$
are *not* in SNF.

(2) If a diagonal entry is 0 then all following diagonal entries are also 0.
This means that the matrices
$$\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$
are not in SNF.

(3) If d_i is the i-th diagonal entry and d_{i+1} is the $(i+1)$-th entry, then $d_i \mid d_{i+1}$.
This means that the matrices
$$\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}, \begin{pmatrix} 4 & 0 & 0 \\ 0 & 6 & 0 \end{pmatrix}$$
are not in SNF.

(4) All the non-zero entries are *positive*.
This means that the matrix
$$\begin{pmatrix} -2 & 0 \\ 0 & 4 \end{pmatrix}$$
is not in SNF.

Example 1. The matrices
$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 4 & 0 \end{pmatrix}$$
are in SNF.

Theorem 1. *Given $A \in M_{m\times n}(\mathbb{Z})$, there exist $B \in \mathrm{GL}_m(\mathbb{Z})$ and $C \in \mathrm{GL}_n(\mathbb{Z})$ such that BAC is in Smith Normal Form.*

So now we can ask: what is the concrete meaning of multiplying on the left and the right by invertible matrices?

(i) For any integer c and any pair of indices $i \neq j$, we can consider the $r \times r$-matrix B with 1s along the diagonal and only one non-zero entry c in the (i,j)-th entry. This will always be an invertible matrix, and the inverse is explicitly given in the same form except with $-b$ as the only off-diagonal entry in the (i,j)-th place.
For example, if $r = 4$ and $(i,j) = (2,3)$, we have
$$B = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & c & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \; ; \; B^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -c & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Then for any $r \times s$-matrix A, BA will be the matrix obtained from A by adding c-times the j^{th}-row to the i^{th}-row.

(i') For any integer b and any pair of indices $i \neq j$, we can consider the $s \times s$-matrix C with 1s along the diagonal and only one non-zero entry b in the (i,j)-th entry. This will always be an invertible matrix, and the inverse is explicitly given in the same form except with $-b$ as the only off-diagonal entry in the (i,j)-th place.

For example, if $r = 4$ and $(i,j) = (2,3)$, we have

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & b & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \; ; \; C^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -b & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Then for any $r \times s$-matrix A, AC will be the matrix obtained from A by adding b-times the i^{th}-column to the j^{th}-column.

Here is another kind of operation that is helpful: swapping rows and columns. More precisely:

(ii) If $i \neq j$, we can take B to be the $r \times r$-matrix that switches the basis vectors $\overrightarrow{\mathbf{e}}_i$ and $\overrightarrow{\mathbf{e}}_j$, but fixes all the other basis vectors. Such a matrix will be invertible (in fact, it is its own inverse) with determinant -1. As an example, we can take

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Then BA will be obtained from A by switching the i^{th} and j^{th}-rows.

(ii') If $i \neq j$, we can take C to be the $s \times s$-matrix that switches the basis vectors $\overrightarrow{\mathbf{e}}_i$ and $\overrightarrow{\mathbf{e}}_j$, but fixes all the other basis vectors. Such a matrix will be invertible (in fact, it is its own inverse) with determinant -1. As an example, we can take

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Then AC will be obtained from A by switching the i^{th} and j^{th}-columns.

Example 2. Let us look at the matrix

$$\begin{pmatrix} 4 & 0 \\ 0 & 6 \end{pmatrix}$$

Given that we know an SNF exists, we can actually figure out what it is by 'pure thought'. The key point is that multiplying by invertible integer matrices doesn't affect the determinant up to sign. Therefore, the SNF will have to have determinant 24. The possibilities are

$$\begin{pmatrix} 1 & 0 \\ 0 & 24 \end{pmatrix}, \begin{pmatrix} 2 & 0 \\ 0 & 12 \end{pmatrix}$$

Now, if we multiply our given matrix by another integer matrix, we can only get even entries (why?). Therefore, there is no way to get 1 via such a process. This means that the second guy must be the SNF.