1 Classical Cryptosystems

(Continued from previous lecture.)

1.1 Interlude: Modular Linear Algebra

Before going into polygraphic ciphers, let us first discuss how *linear algebra* interacts with modular arithmetic. We'll just work on 2×2 matrices for now.

1.1.1 2×2 Matrices

Definition 1.1

A 2×2 integer **matrix** (or just *matrix* for short) is a 2×2 box of numbers $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ where $a, b, c, d \in \mathbb{Z}$.

- The **determinant** of A is the integer $\det(A) = ad - bc$.
- The **identity matrix** is the matrix $I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.
- Suppose $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and $B = \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix}$ are two matrices. Their product AB is defined to be

$$AB = \begin{bmatrix} aa' + bc' & ba' + db' \\ ca' + dc' & cb' + dd' \end{bmatrix}.$$

(Example.) Let $A = \begin{bmatrix} 3 & 2 \\ 1 & 7 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix}$. We know that

$$\det(A) = 3 \cdot 7 - 2 \cdot 1 = 19.$$

We also know that

$$AB = \begin{bmatrix} 7 & 18 \\ 15 & 25 \end{bmatrix}$$

and

$$BA = \begin{bmatrix} 7 & 30 \\ 9 & 25 \end{bmatrix}.$$

Remark: It should be clear from the above example that $AB \neq BA$. That is, matrix multiplication is not commutative.

(Exercise.) Let A be a 2×2 integer matrix. Show that

$$AI = IA = A.$$

Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Then,

$$IA = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

and

$$AI = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Theorem 1.1: Multiplicativity of Determinant

If A and B are matrices, then $\det(I) = 1$ and

$$\det(AB) = \det(A)\det(B).$$

Definition 1.2

A **vector** v is a vertical column

$$v = \begin{bmatrix} x \\ y \end{bmatrix},$$

where $x, y \in \mathbb{Z}$.

Definition 1.3

If $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is a matrix, then the product $Ab = \begin{bmatrix} ax + by \\ cx + dy \end{bmatrix}$.

1.1.2 Congruences and Inversion for Matrices

Definition 1.4

Fix a positive integer n and suppose A and B are both matrices:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad B = \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix}.$$

We say that $A \equiv B \pmod{n}$ if all four of the entries of the two matrices are congruent mod n, i.e., if all of the following are true:

$$a \equiv a' \pmod{n}$$
$$b \equiv b' \pmod{n}$$
$$c \equiv c' \pmod{n}$$
$$d \equiv d' \pmod{n}$$

Definition 1.5

A matrix A is *invertible mod* n if there exists a matrix X such that $AX \equiv I \pmod{n}$. In this case, X is called an inverse of $A \pmod{n}$. In symbols, we write $X \equiv A^{-1} \pmod{n}$.

Theorem 1.2: Modular Inversion Theorem

Suppose $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is a matrix. Then, A is invertible if and only if $\det(A)$ is invertible mod n. Moreover, if $e \equiv \det(A)^{-1} \pmod{n}$, then

$$X = \begin{bmatrix} ed & -eb \\ -ec & ea \end{bmatrix}$$

is an inverse of $A \pmod{n}$.

(Example.) Suppose we have $A = \begin{bmatrix} 3 & 2 \\ 1 & 7 \end{bmatrix}$. We know that $\det(A) = 19$ is invertible mod 26, so A is also invertible mod 26. We have

$$19^{-1} \equiv 11 \pmod{26},$$

so the formula for the inverse from the Matrix Inversion Theorem tells us that

$$A^{-1} \equiv \begin{bmatrix} 11 \cdot 7 & -11 \cdot 2 \\ -11 \cdot 1 & 11 \cdot 3 \end{bmatrix} \equiv \begin{bmatrix} 77 & -22 \\ -11 & 33 \end{bmatrix} \equiv \begin{bmatrix} 25 & 4 \\ 15 & 7 \end{bmatrix} \pmod{26}.$$

In other words,

$$X = \begin{bmatrix} 15 & 4 \\ 15 & 7 \end{bmatrix}$$

is an inverse of A mod 26. It follows that $AX = I$.

1.2 Hill Cipher

The *Hill Cipher* is the first polygraphic cipher we'll talk about. We'll focus on the digraphic case, which replaces 2 letters of plaintext at a time. Our **key** for this cipher is a matrix that is invertible mod 26.

(Example.) Suppose we want to encrypt the message `You have saved us all`. Begin with the usual encoding process:

Y	O	U	H	A	V	E	S	A	V	E	D	U	S	A	L	L
24	14	20	7	0	21	4	18	0	21	4	3	20	18	0	11	11

(The numbers below the letters represent the ranking of each letter.) Let's suppose our key is

$$A = \begin{bmatrix} 3 & 2 \\ 1 & 7 \end{bmatrix},$$

which has determinant 19 and is thus invertible mod 26. It follows that A is an invertible matrix mod 26, which can thus be used as a key.

For encrypting, the idea is to go through the list of numbers, replacing each pair of numbers with the result of multiplying that pair by the matrix A (mod 26). For example, for the pair 24 and 14, we can make a vector containing these numbers,

$$v = \begin{bmatrix} 24 \\ 14 \end{bmatrix},$$

and then compute

$$Av = \begin{bmatrix} 3 & 2 \\ 1 & 7 \end{bmatrix} \begin{bmatrix} 24 \\ 14 \end{bmatrix} \pmod{26} = \begin{bmatrix} 100 \\ 122 \end{bmatrix} \pmod{26} = \begin{bmatrix} 22 \\ 18 \end{bmatrix}.$$

So, we replace the numbers 24 and 14 with the numbers 22 and 18, respectively. In other words, the first two letters of the message will be replaced by `W` and `S`, respectively.

We can continue this process with the next pair of numbers (20, 7), and so on. Eventually, we'll reach the end. Note that, if you have an odd number of letters, you can add an additional random letter at the end (e.g., `Z`). With this in mind, the net result is the ciphertext

```
WSWRQRWAQRSZSQWZFE
```

As you might expect, to decrypt a message, we just need to multiply the pairs of numbers by the *inverse* of A mod 26.

(Exercise.) Use the matrix

$$A = \begin{bmatrix} 3 & -1 \\ 2 & 5 \end{bmatrix}$$

as the key for a Hill cipher. Encrypt the message `Go to Lake Lerna`.

First, we verify that this matrix can be used as a key by checking the determinant.

$$\det(A) = 15 - 2(-1) = 15 + 2 = 17.$$

Because 17 is invertible mod 26, it follows that we can use A as a key. So, begin by encoding the message:

```
G  O  T  O  L A  K E  L E  R  N A  Z
6 14 19 14 11 0 10 4 11 4 17 13 0 25
```

Note that we put a `Z` at the end so that the length of the plaintext is even (that way, we can do pairwise encryption.) We'll now process each pair of letters.

- For pair $(6, 14)$, we have

$$\begin{bmatrix} 3 & -1 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} 6 \\ 14 \end{bmatrix} \pmod{26} = \begin{bmatrix} 4 \\ 82 \end{bmatrix} \pmod{26} = \begin{bmatrix} 4 \\ 4 \end{bmatrix},$$

  which corrsponds to `E` and `E`.

- For pair $(19, 14)$, we have

$$\begin{bmatrix} 3 & -1 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} 19 \\ 14 \end{bmatrix} \pmod{26} = \begin{bmatrix} 43 \\ 108 \end{bmatrix} \pmod{26} = \begin{bmatrix} 17 \\ 4 \end{bmatrix} \pmod{26},$$

  which corresponds to `R` and `E`.

- For pair $(11, 0)$, we have

$$\begin{bmatrix} 3 & -1 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} 11 \\ 0 \end{bmatrix} \pmod{26} = \begin{bmatrix} 33 \\ 22 \end{bmatrix} \pmod{26} = \begin{bmatrix} 7 \\ 22 \end{bmatrix} \pmod{26},$$

  corresponding to `H` and `W`.

By continuing this process, we end up with the ciphertext

```
EEREHWAODQMVBV
```

(Exercise.) Use the matrix

$$A = \begin{bmatrix} 3 & -1 \\ 2 & 5 \end{bmatrix}$$

as the key for a Hill cipher. Decrypt the message `RNCQYVFRRLZI`.

Note again that $\det(A) = 17$. In order to decrypt the message, we need to find the inverse of A mod 26.

Finding GCD: Recall that the Matrix Inversion Theorem states that A is invertible if and only if $\det(A)$ is invertible mod n. To see if $\det(A)$ is invertible mod n, we need to see if $\gcd(\det(A), n) = 1$. So, let's find $\gcd(17, 26)$.

a	b	$b = aq + r$	q	r
17	26	$26 = 17q + r$	1	9
9	17	$17 = 9q + r$	1	8
8	9	$9 = 8q + r$	1	1
1	8	$8 = 1q + r$	8	0

Therefore, $\gcd(17, 26) = 1$ as desired. Thus, an inverse must exist.

Finding Bezout: Now, we need to find the Bezout coefficients. Labeling each equation, we have

- (Eq. 1) $26 = 17(1) + 9 \implies 9 = 26 + 17(-1)$
- (Eq. 2) $17 = 9(1) + 8 \implies 8 = 17 + 9(-1)$
- (Eq. 3) $9 = 8(1) + 1 \implies 1 = 9 + 8(-1)$

Now that we've labeled each relevant operation, we can find the Bezout coefficients:

$$\begin{aligned}
1 &= 9 + 8(-1) \\
&= 9 + (\underbrace{17 + 9(-1)}_{\text{Eq. 2}})(-1) \\
&= 9 + 17(-1) + 9(-1)(-1) \\
&= 9 + 17(-1) + 9 \\
&= 9(2) + 17(-1) \\
&= (\underbrace{26 + 17(-1)}_{\text{Eq. 1}})(2) + 17(-1) \\
&= 26(2) + 17(-1)(2) + 17(-1) \\
&= 26(2) + 17(-2) + 17(-1) \\
&= 26(2) + 17(-3)
\end{aligned}$$

From this, it follows that $x = -3$, which is the desired inverse.

Decrypting: With this in mind, we have

$$X = \begin{bmatrix} -3(5) & 3(-1) \\ 3(2) & -3(3) \end{bmatrix} = \begin{bmatrix} -15 & -3 \\ 6 & -9 \end{bmatrix} \pmod{26}.$$

Now that we have the matrix needed to decrypt the message, we can proceed. Labeling each character in the message gives us

```
R  N C  Q  Y  V F  R  R  L  Z I
17 13 2 16 24 21 5 17 17 11 25 8
```

Iterating over each pair, we have

- For $(17, 13)$,
$$X \begin{bmatrix} 17 \\ 13 \end{bmatrix} \pmod{26} = \begin{bmatrix} -294 \\ -15 \end{bmatrix} \pmod{26} = \begin{bmatrix} 18 \\ 11 \end{bmatrix} \pmod{26},$$
or `S` and `L`.
- For $(2, 16)$,
$$X \begin{bmatrix} 2 \\ 16 \end{bmatrix} \pmod{26} = \begin{bmatrix} -78 \\ -132 \end{bmatrix} \pmod{26} = \begin{bmatrix} 0 \\ 24 \end{bmatrix} \pmod{26},$$
or `A` and `Y`.

By continuing this process, we end up with

```
SLAYTHEHYDRA
```

1.3 Playfair Cipher

The **Playfair Cipher** is another digraphic cipher, like the Hill cipher we just discussed above. The key for a Playfair cipher is a 5×5 grid of letters, where each letter appears exactly once. Because there are 26 letters in the English alphabet but 25 letters can fit in a grid, we treat `I` and `J` as the same letter[1].

How do we start constructing a grid? An easy and convenient way of doing this is to start with a secret keyword. For example, suppose `ALPHABET` is our keyword. We can start filling out our grid by writing out the letters of our keyword across the rows, skipping over the letters we've written.

A	L	P	H	B
E	T			

We can then fill out the remaining squares with the remaining letters of the alphabet, skipping over anything we've already written down and remembering that `I` and `J` are the same.

A	L	P	H	B
E	T	C	D	F
G	I	K	M	N
O	Q	R	S	U
V	W	X	Y	Z

We can encode our message by doing the following:

[1] We could also use a variant where we use a 6×6 grid that includes all 26 letters and 10 digits, instead.

1. Remove all non-alphabet characters and capitalize everything.
2. Replace all instances of `J` with `I`.
3. Group the letters into pairs.
4. If there are any pairs where both letters are the same, inser the letter `X` in between the two letters of that pair and regroup into pairs.
5. If there's an unpaired letter at the end, insert the letter `X` after it.

Remark: You may need to apply rule 4 multiple times.

(Example.) Suppose we want to encode the message `hidden jewels in trees.` Here's what will happen after each step described above.

1. `HIDDENJEWELSINTHETREES`
2. `HIDDENIEWELSINTHETREES`
3. `HI DD EN IE WE LS IN TH ET RE ES`
4. `HI DX DE NI EW EL SI NT HE TR EX ES`
5. `HI DX DE NI EW EL SI NT HE TR EX ES`

To encrypt, we need to replace each pair with another pair using the grid by following the rules:

- (Row Rule.) If both letters in the pair occur in the same row, replace each letter of the pair with the letter that appears immediately to its right (wrapping around to the left side of the row if needed).
- (Column Rule.) If both letters in the pair occur in the same column, replace each letter of the pair with the letter that appears immediately below it (wrapping around to the top of the column if needed).
- (Rectangle Rule.) Otherwise, the two letters define a rectangle inside the grid, and we replace each letter with the letter on the same row but the opposite of that rectangle.

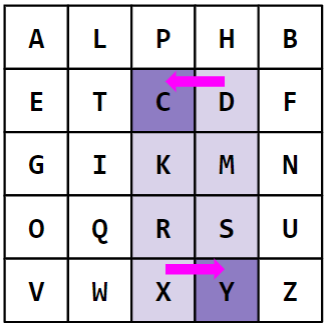
(Example.) Suppose we want to encrypt the message `HI DX DE NI EW EL SI NT HE TR EX ES` (see previous example for encoding). Let's look at each pair.

- For `HI`, notice that `H` and `I` do not appear in the same row or column. Therefore, the rectangle rule applies. Observe the highlighted cells:

A	L	P	H	B
E	T	C	D	F
G	I	K	M	N
O	Q	R	S	U
V	W	X	Y	Z

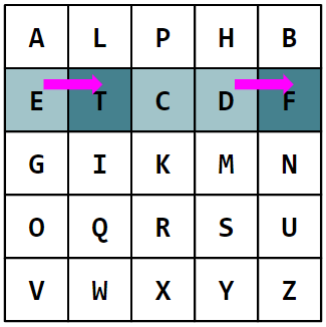
  Here, the letter in the same row as `H` but opposite side is `L`, and the letter in the same row as `I` but the opposite side is `M`. Therefore, `HI` becomes `LM`.

- For `DX`, we also apply the rectangle rule. Observe the highlighted cells:



So, it follows that DX gets replaced with CY.

- For DE, both letters are on the same row so we apply the row rule. Observe that



So, it follows that DE becomes FT.

Continuing this process yields the desired result.