Hill Cipher Project

`K80TTQ1EP-??,VO.L,XUOH5BY,_71ZVPKOE678_X,N2Y-8HI4VS,,6Z28DDW5N7ADY013`

Directions:

- Answer all numbered questions completely.
- Show non-trivial work in the space provided.
- Non-computational answers should be given in complete sentences.

Introduction

The affine cipher encrypts one letter at a time, resulting in a simple permutation of the alphabet. For the casual observer, messages are unintelligible. But crypto-analysts can easily break the affine cipher by observing letter frequencies. For example, the most commonly occurring letter in the ciphertext is likely to be 'E' in the plaintext.

In this project, we will develop the Hill Cipher, which encrypts several letters at a time, making frequency analysis much more difficult. The method uses modular arithmetic, as well as the basic linear algebra of matrices and vectors.

Vectors and Matrices

For all practical purposes, a vector is just a list of numbers. We will take the convention of writing the vector as a column. For example

$$\vec{u} = \begin{bmatrix} 2 \\ -3 \end{bmatrix} \qquad \vec{v} = \begin{bmatrix} 7 \\ 1 \\ -4 \end{bmatrix} \qquad \vec{w} = \begin{bmatrix} 3 \\ 0 \\ -1 \\ 6 \end{bmatrix}$$

The dimension of a vector is the same as its length. So $\vec{u}$ is two-dimensional, $\vec{v}$ is three-dimensional, and $\vec{w}$ is four-dimensional. A regular number (or a one-dimensional vector), is called a scalar.

If vectors have the same dimension, you can perform basic arithmetic on them: addition/sutraction and scalar multiplication. For example,

$$\begin{bmatrix} 1 \\ 4 \end{bmatrix} - 2 \begin{bmatrix} 3 \\ -2 \end{bmatrix} = \begin{bmatrix} 1 - 2 \cdot 3 \\ 4 - 2 \cdot (-2) \end{bmatrix} = \begin{bmatrix} -5 \\ 8 \end{bmatrix}$$

Often, vectors are identified by name. If we define $\vec{u} = \begin{bmatrix} -4 \\ 1 \\ 2 \end{bmatrix}$, $\vec{v} = \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix}$, and $\vec{w} = \begin{bmatrix} -2 \\ 0 \\ 7 \end{bmatrix}$ then

$$\vec{u} + 5\vec{v} - 2\vec{w} = \begin{bmatrix} -4 \\ 1 \\ 2 \end{bmatrix} + 5 \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} - 2 \begin{bmatrix} -2 \\ 0 \\ 7 \end{bmatrix} = \begin{bmatrix} -4 + 5 \cdot 1 - 2 \cdot -2 \\ 1 + 5 \cdot 3 - 2 \cdot 0 \\ 2 + 5 \cdot 5 - 2 \cdot 7 \end{bmatrix} = \begin{bmatrix} 5 \\ 16 \\ 13 \end{bmatrix}$$

This result is referred to as a linear combination of the three vectors.

When there are many vectors to deal with, they can be concatenated into a matrix. In the previous example, we could define a matrix called A that contains $\vec{u}, \vec{v}$, and $\vec{w}$, and a vector $\vec{x}$ that contains the coefficients in the expression $\vec{u} + 5\vec{v} - 2\vec{w}$.

$$A = \begin{bmatrix} -4 & 1 & -2 \\ 1 & 3 & 0 \\ 2 & 5 & 7 \end{bmatrix} \qquad \vec{x} = \begin{bmatrix} 1 \\ 5 \\ -2 \end{bmatrix}$$

Then the linear combination is written succinctly as matrix-vector multiplication $A\vec{x}$ where :

$$A\vec{x} = \begin{bmatrix} -4 & 1 & -2 \\ 1 & 3 & 0 \\ 2 & 5 & 7 \end{bmatrix} \begin{bmatrix} 1 \\ 5 \\ -2 \end{bmatrix} = \begin{bmatrix} 5 \\ 16 \\ 13 \end{bmatrix}$$

Here is another example:

$$\begin{bmatrix} 3 & 1 \\ 5 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \end{bmatrix} = 2\begin{bmatrix} 3 \\ 5 \end{bmatrix} + 4\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2\cdot 3 + 4\cdot 1 \\ 2\cdot 5 + 4\cdot 2 \end{bmatrix} = \begin{bmatrix} 10 \\ 18 \end{bmatrix}$$

In general, matrices don't have to be square. For example,

$$\begin{bmatrix} 3 & 1 & -2 \\ 1 & 0 & 7 \end{bmatrix}$$

has two rows and three columns, so it's size is 2×3. We could multiply it by a 3 dimensional vector, e.g.

$$\begin{bmatrix} 3 & 1 & -2 \\ 1 & 0 & 7 \end{bmatrix} \begin{bmatrix} 1 \\ -3 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix} - 3\begin{bmatrix} 1 \\ 0 \end{bmatrix} + 2\begin{bmatrix} -2 \\ 7 \end{bmatrix} = \begin{bmatrix} -4 \\ 15 \end{bmatrix}$$

---

1. Compute $4\begin{bmatrix} 2 \\ -5 \end{bmatrix} - 2\begin{bmatrix} 8 \\ -12 \end{bmatrix}$.

2. If the linear combination in the previous problem were written as $A\vec{x}$,

   (a) write the matrix A, and state its size

   (b) write the vector $\vec{x}$, and state its dimension.

3. How many rows and colums does this matrix have?

$$A = \begin{bmatrix} 4 & 23 \\ 17 & -2 \\ 4 & 5 \\ 0 & 1 \end{bmatrix}$$

4. Using the A from the previous problem, compute $A \begin{bmatrix} 3 \\ -1 \end{bmatrix}$.

5. If you are multiplying a matrix A by a vector $\vec{x}$, what is the relationship between the the number of columns in A and the dimension of $\vec{x}$?

Matrix-matrix multiplication

Remember that a matrix is just the concatenation of several vectors. We have already defined how to multiply a matrix by a vector. To multiply a matrix by another matrix, just use one vector at a time from the matrix on the right. For example, you can check that

$$\begin{bmatrix} 1 & 3 \\ 2 & 4 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} -1 \\ 7 \end{bmatrix} = \begin{bmatrix} 20 \\ 26 \\ 37 \end{bmatrix} \qquad \text{and} \qquad \begin{bmatrix} 1 & 3 \\ 2 & 4 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 9 \\ 14 \\ 27 \end{bmatrix}$$

Combine these to get

$$\begin{bmatrix} 1 & 3 \\ 2 & 4 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} -1 & 3 \\ 7 & 2 \end{bmatrix} = \begin{bmatrix} 20 & 9 \\ 26 & 14 \\ 37 & 27 \end{bmatrix}$$

Consider another example. To compute

$$\begin{bmatrix} 5 & 7 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 3 & -7 \\ -2 & 5 \end{bmatrix}$$

we may catenate the results of two separate matrix-vector multiplications:

$$\begin{bmatrix} 5 & 7 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 3 \\ -2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad \begin{bmatrix} 5 & 7 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} -7 \\ 5 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

and therefore

$$\begin{bmatrix} 5 & 7 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 3 & -7 \\ -2 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

6. Multiply these matrices:

$$\begin{bmatrix} -1 & 3 \\ 2 & 7 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -3 & 2 \end{bmatrix}$$

7. Multiply these matrices:

$$\begin{bmatrix} 1 & 3 \\ 2 & 4 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ 5 & 3 \end{bmatrix}$$

8. Matrix multiplication is not generally commutative. If $A = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 4 \\ 3 & 5 \end{bmatrix}$, compute AB and BA and show they are not the same.

Identity Matrix

An identity matrix is square (same number of rows and columns), with ones on the main diagonal and zeros elsewhere. Identity matrices go by the label 'I', sometimes with a subscript to indicate the size. For example,

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Verify these two matrix-vector multiplications:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ -7 \end{bmatrix} = \begin{bmatrix} 3 \\ -7 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \\ 8 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ 8 \end{bmatrix}$$

Identity matrices have the special property, illustrated above, that $I\vec{x} = \vec{x}$, no matter what $\vec{x}$ is. Thus I acts just like the number one - multiplying by it doesn't change anything.

---

9. Write I_4 and compute I_4 times the vector $\vec{x} = \begin{bmatrix} 2 \\ -1 \\ 3 \\ 9 \end{bmatrix}$.

Inverses

In mathematics there are many types of inverses:

- The inverse of a number in $\mathbb{R}$.
  $2^{-1} = \frac{1}{2}$ since $2 \cdot \frac{1}{2} = 1$.
- The inverse of a number in $\mathbb{Z}_n$.
  $2^{-1} \pmod 7 = 4$ since $2 \cdot 4 \pmod 7 = 1$
- The inverse of a function.
  If $f(x) = e^x$ and $g(x) = \ln(x)$, then $f^{-1} = g$ since $(g \circ f)(x) = \ln(e^x) = x$.

The common thread here is that two objects are inverses if they *undo* each other and leave you with an identity. Like the antidote to a poison, or offsetting football penalties, the effects of an operation and its inverse will cancel each other out.

By analogy, define two matrices to be inverses if they multiply together (in either order) to be the identity matrix I. For example, consider these two matrices:

$$A = \begin{bmatrix} 5 & 7 \\ 2 & 3 \end{bmatrix} \qquad B = \begin{bmatrix} 3 & -7 \\ -2 & 5 \end{bmatrix}$$

Earlier we saw that

$$AB = \begin{bmatrix} 5 & 7 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 3 & -7 \\ -2 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$$

Therefore these two matrices are inverses of each other. In particular we could write

$$A = B^{-1} \qquad\qquad B = A^{-1}$$

Check that you can reverse the order of multiplication to get

$$BA = \begin{bmatrix} 3 & -7 \\ -2 & 5 \end{bmatrix} \begin{bmatrix} 5 & 7 \\ 2 & 3 \end{bmatrix} = I_2$$

Not all matrices have inverses, but many do. Later in this project we will learn how to compute the inverse of a matrix.

10. Verify that $A = \begin{bmatrix} 12 & 7 \\ 5 & 3 \end{bmatrix}$ and $B = \begin{bmatrix} 3 & -7 \\ -5 & 12 \end{bmatrix}$ are inverses by calculating both AB and BA. Show your work.

11. Show that these matrices are inverses by computing (show your work)

$$\begin{bmatrix} 1 & -2 & 3 \\ 2 & -5 & 10 \\ -1 & 2 & -2 \end{bmatrix} \begin{bmatrix} 10 & -2 & 5 \\ 6 & -1 & 4 \\ 1 & 0 & 1 \end{bmatrix}$$

We'll be using matrices to encrypt messages, so inverses will be essential to guarantee the intended recipient can decipher the message back to its original form.

We will be working in $\mathbb{Z}_n$, in which case two matrices are inverses if AB reduces to I mod n.

$$AB \equiv I \pmod{n}$$

For example, if $A = \begin{bmatrix} 5 & 1 \\ 2 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 5 \\ 3 & 3 \end{bmatrix}$ then in base 7,

$$AB = \begin{bmatrix} 5 & 1 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} 1 & 5 \\ 3 & 3 \end{bmatrix} = \begin{bmatrix} 8 & 28 \\ 14 & 22 \end{bmatrix} \equiv \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \pmod{7}$$

12. Verify that $A = \begin{bmatrix} 3 & 2 \\ 0 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 4 \\ 0 & 4 \end{bmatrix}$ are inverses $\pmod 5$, but they are not inverses $\pmod 7$.

13. Verify that $A = \begin{bmatrix} 21 & 15 & 6 \\ 2 & 23 & 1 \\ 7 & 9 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} 32 & 13 & 0 \\ 17 & 0 & 31 \\ 37 & 16 & 25 \end{bmatrix}$ are inverses $\pmod{41}$.

Encryption

We will now explain the Hill Cipher using the language of modular arithmetic and linear algebra.

Character-Numeric Conversion

First we will map each character in our alphabet to a corresponding number. In order to unambiguously encrypt/decrypt messages, we need those numbers to have inverses modulo the total number of characters. Therefore the length of the alphabet should be prime.

The English alphabet has 26 characters, but 26 is not prime. We also want to include other symbols besides letters. Our map will include a total of 41 characters - a good choice since 41 is prime.

_	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20

U	V	W	X	Y	Z	0	1	2	3	4	5	6	7	8	9	.	?	,	-
21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40

Lower and upper case letters are not distinguished. The phrase “DO IT” would be:

$$(4, 15, 0, 9, 20)$$

---

14. The key to decoding the quote at the beginning of this project is (2, 4, 9, 19, 3, 18, 5, 20, 5). Translate this back to characters.

Encoding Matrix

We will use a combination of matrix multiplication and modular arithmetic to encrypt our messages. First we must create an encoding matrix, which we’ll call A. Given the key (code word), use the table above to convert it into numbers, which are placed in a square matrix. We’ll fill the entries in reading order, padding any extra space with numbers beginning with one.

For example, suppose the key is “MATH”. This converts to $(13, 1, 20, 8)$. Since it has only four characters, we can fit it in a 2×2 matrix as follows

$$A = \begin{bmatrix} 13 & 1 \\ 20 & 8 \end{bmatrix}$$

Notice we begin at the top left and fill row-by-row. Now let’s use the key “CNEAGLE”, which corresponds to $(3, 14, 5, 1, 7, 12, 5)$, and the encoding matrix

$$A = \begin{bmatrix} 3 & 14 & 5 \\ 1 & 7 & 12 \\ 5 & 1 & 2 \end{bmatrix}$$

Note that there were 7 characters, so the first square matrix that fits must be 3×3. The extra spots are padded with ’1’ and ’2’.

Message Matrix

Suppose you want to encode the secret message: “RENDEZVOUS AT 9”. First, we convert it to

$$(18, 5, 14, 4, 5, 26, 22, 15, 21, 19, 0, 1, 20, 0, 36)$$

The message matrix M should have the same number of rows as the encoding matrix A. For example, if the key is “MATH”, then both matrices should have two rows, and in particular

$$M = \begin{bmatrix} 18 & 5 & 14 & 4 & 5 & 26 & 22 & 15 \\ 21 & 19 & 0 & 1 & 20 & 0 & 36 & 0 \end{bmatrix}$$

However, if the key is “CNEAGLE”, then we need three rows and

$$M = \begin{bmatrix} 18 & 5 & 14 & 4 & 5 \\ 26 & 22 & 15 & 21 & 19 \\ 0 & 1 & 20 & 0 & 36 \end{bmatrix}$$

Any extra spots in the message matrix may be filled in with zeros.

15. Write the encoding matrix for the key “MATHEMATICS”.

16. If the key is “TIP”, write the message matrix for the message “SELL APPLE”.
What if the key were “HOT TIP” instead ?

17. If the key is k characters long, find a formula for the number of rows in A and M.

Encryption

Matrix multiplication will be used to jumble the message matrix, so that nobody could decipher the result without the key. To encrypt the message with a key:

- Create a square encoding matrix A from the key.
- Create the message matrix M with the same number of rows as A.
- Multiply the encoding matrix by the message matrix to form AM. This creates linear combinations of the encoding matrix columns using coefficients from the message matrix.
- Reduce modulo 41 (or generally the # of characters in your alphabet).
- Convert back to characters.

Let's illustrate by encoding "RENDEZVOUS AT 9" with the key "MATH". We already found A and M:

$$A = \begin{bmatrix} 13 & 1 \\ 20 & 8 \end{bmatrix} \qquad M = \begin{bmatrix} 18 & 5 & 14 & 4 & 5 & 26 & 22 & 15 \\ 21 & 19 & 0 & 1 & 20 & 0 & 36 & 0 \end{bmatrix}$$

Now multiply A by M:

$$\begin{aligned} AM &= \begin{bmatrix} 13 & 1 \\ 20 & 8 \end{bmatrix} \begin{bmatrix} 18 & 5 & 14 & 4 & 5 & 26 & 22 & 15 \\ 21 & 19 & 0 & 1 & 20 & 0 & 36 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 255 & 84 & 182 & 53 & 85 & 338 & 322 & 195 \\ 528 & 252 & 280 & 88 & 260 & 520 & 728 & 300 \end{bmatrix} \\ &\equiv \begin{bmatrix} 9 & 2 & 18 & 12 & 3 & 10 & 35 & 31 \\ 36 & 6 & 34 & 6 & 14 & 28 & 31 & 13 \end{bmatrix} \pmod{41} \end{aligned}$$

Here is a detailed working of the first entry calculation:

$$13 \cdot 18 + 1 \cdot 21 = 255 \equiv 9 \pmod{41}$$

Converting the result matrix back to characters ($9 \to I,\ 2 \to B$, etc), "RENDEZVOUS AT 9" translates to `"IBRLCJ849F7FN14M"`, which can transmitted securely.

Now use the key "CNEAGLE" instead. The corresponding matrices are:

$$A = \begin{bmatrix} 3 & 14 & 5 \\ 1 & 7 & 12 \\ 5 & 1 & 2 \end{bmatrix} \qquad M = \begin{bmatrix} 18 & 5 & 14 & 4 & 5 \\ 26 & 22 & 15 & 21 & 19 \\ 0 & 1 & 20 & 0 & 36 \end{bmatrix}$$

$$\begin{aligned} AM &= \begin{bmatrix} 3 & 14 & 5 \\ 1 & 7 & 12 \\ 5 & 1 & 2 \end{bmatrix} \begin{bmatrix} 18 & 5 & 14 & 4 & 5 \\ 26 & 22 & 15 & 21 & 19 \\ 0 & 1 & 20 & 0 & 36 \end{bmatrix} \\ &= \begin{bmatrix} 418 & 328 & 352 & 306 & 461 \\ 200 & 171 & 359 & 151 & 570 \\ 116 & 49 & 125 & 41 & 116 \end{bmatrix} \\ &\equiv \begin{bmatrix} 8 & 0 & 24 & 19 & 10 \\ 36 & 7 & 31 & 28 & 37 \\ 34 & 8 & 2 & 0 & 34 \end{bmatrix} \pmod{41} \end{aligned}$$

This time the same message gets converted to `"H_XSJ9G41.7HB_7"`.

You should convince yourself of the calculations above. They are quite tedious to do by hand. On the course web page is an app called hillcipher.php that creates the matrices and does the calculations for you.

Let's do one more example, this time by hand and via computer. The key is “NERD” and the message is “GOBBLE”. Check that the encoding and message matrices are:

$$A = \begin{bmatrix} 14 & 5 \\ 18 & 4 \end{bmatrix} \qquad M = \begin{bmatrix} 7 & 15 & 2 \\ 2 & 12 & 5 \end{bmatrix}$$

When you multiply them, you should get

$$AM = \begin{bmatrix} 108 & 270 & 53 \\ 134 & 318 & 56 \end{bmatrix} \equiv \begin{bmatrix} 26 & 24 & 12 \\ 11 & 31 & 15 \end{bmatrix} \pmod{41}$$

Finally read off the encoded message to be `"ZXLK4O"`. Use the web app to check these calculations.

---

18. What is the encoding matrix for the key “HOKIES” ?

19. Reduce the matrix $\begin{bmatrix} 48 & 15 & 138 & 209 \\ 410 & 92 & 97 & 43 \end{bmatrix}$ (mod 41) and read off the message.

20. Use the web app to encrypt the message “THE CAT IN THE HAT” using the key “SEUSS”.

    How do the two occurrences of “THE” translate as ciphertext?

21. Use the web app to encrypt your favorite quotation. Email me (kmassey@cn.edu) the key and ciphertext.

22. By hand, show the steps to encrypt the message "RU READY" using the key "NERD". You can check your answer with the computer.

Decryption

Only the intended recipient should be able to read your message. So your key should be privately shared, and not easy to guess. Given the key and the encrypted message, how can we *undo* the encryption? The answer is to use a matrix inverse.

When decrypting, the hillcipher web app will display both the encoding matrix A, and a decoding matrix called B. This is calculated so that $BA \equiv I \pmod{41}$, and thus B and A are inverses. In other words, multiplying by B reverses the effect of multiplying by A. When both operations are appled to a message, you end up with the original.

For example, if the key is "MATH", then

$$A = \begin{bmatrix} 13 & 1 \\ 20 & 8 \end{bmatrix} \qquad B = \begin{bmatrix} 4 & 20 \\ 31 & 27 \end{bmatrix}$$

You should check that the decoding and encoding matrices multiply to give the identity.

$$BA = \begin{bmatrix} 4 & 20 \\ 31 & 27 \end{bmatrix} \begin{bmatrix} 13 & 1 \\ 20 & 8 \end{bmatrix} = \begin{bmatrix} 452 & 164 \\ 943 & 247 \end{bmatrix} \equiv \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \pmod{41}$$

A message is decrypted just like it is encrypted, except that we use the decoding matrix B instead of the encoding matrix A. For example, to decrypt `"IBRLCJ849F7FN14M"` using the key "MATH":

$$\begin{aligned} B\hat{M} &= \begin{bmatrix} 4 & 20 \\ 31 & 27 \end{bmatrix} \begin{bmatrix} 9 & 2 & 18 & 12 & 3 & 10 & 35 & 31 \\ 36 & 6 & 34 & 6 & 14 & 28 & 31 & 13 \end{bmatrix} \\ &= \begin{bmatrix} 756 & 128 & 752 & 168 & 292 & 600 & 760 & 384 \\ 1251 & 224 & 1476 & 534 & 471 & 1066 & 1922 & 1312 \end{bmatrix} \\ &\equiv \begin{bmatrix} 18 & 5 & 14 & 4 & 5 & 26 & 22 & 15 \\ 21 & 19 & 0 & 1 & 20 & 0 & 36 & 0 \end{bmatrix} \pmod{41} \end{aligned}$$

As hoped, this procedure has restored the original message "RENDEZVOUS AT 9".

For another example, let's decrypt `"F3RM7VZUJ6SE"` using the key "NERD". The web app tells us that the decoding matrix is

$$B = \begin{bmatrix} 24 & 11 \\ 15 & 2 \end{bmatrix}$$

The encrypted message matrix is

$$\hat{M} = \begin{bmatrix} 6 & 30 & 18 & 13 & 34 & 22 \\ 26 & 21 & 10 & 33 & 19 & 5 \end{bmatrix}$$

Multiplying them gives:

$$\begin{aligned} B\hat{M} &= \begin{bmatrix} 24 & 11 \\ 15 & 2 \end{bmatrix} \begin{bmatrix} 6 & 30 & 18 & 13 & 34 & 22 \\ 26 & 21 & 10 & 33 & 19 & 5 \end{bmatrix} \\ &= \begin{bmatrix} 430 & 951 & 542 & 675 & 1025 & 583 \\ 142 & 492 & 290 & 261 & 548 & 340 \end{bmatrix} \\ &\equiv \begin{bmatrix} 20 & 8 & 9 & 19 & 0 & 9 \\ 19 & 0 & 3 & 15 & 15 & 12 \end{bmatrix} \pmod{41} \end{aligned}$$

which translates to `"THIS_IS_COOL"`. Notice that plaintext "OO" appeared as "6S" in ciphertext.

Entire Process Summary

Imagine Alice and Bob want to communicate secret messages. If any third party intercepts a message, they shouldn't be able to read it.

1. Alice and Bob agree on a key word or phrase.
2. Alice uses the key to create the encryption matrix A.
3. Bob uses the key to create the decryption matrix B, so that $BA = I$.
4. Alice creates a message matrix M, pre-multiplies it by A to get AM, and sends the corresponding ciphertext to Bob.
5. If anybody intercepts the ciphertext, it will look like gibberish.
6. Bob receives the ciphertext, converts it to $\hat{M}$, knowing that $\hat{M} = AM$. Then Bob pre-multiplies by B

   $$B\hat{M} = B(AM) = (BA)M = IM = M$$

   which is the original plaintext message.

---

23. Use the web app with the key from problem #14 to decode the quote at the beginning of this project.

24. Use the web app to find the encoding and decoding matrices for the key "EASY", then verify that $AB \equiv I_2 \pmod{41}$.

25. Use the web app to find the decoding matrix for the key "HOKIES".

26. By hand, show the steps to decode `"II?N4B''` using the key "HOKIES".

Computing the Inverse Matrix

Suppose Alice and Bob will communicate using the key “MATH”. It is easy for Alice to find the encoding matrix A. Just fill the 2×2 matrix with the numeric equivalents of M-A-T-H.

$$A = \begin{bmatrix} 13 & 1 \\ 20 & 8 \end{bmatrix}$$

Bob’s task is more difficult. He must find the decoding matrix B so that $BA = I$; thus B is the inverse of A. We can easily check that

$$B = \begin{bmatrix} 4 & 20 \\ 31 & 27 \end{bmatrix}$$

is the inverse, because:

$$BA = \begin{bmatrix} 4 & 20 \\ 31 & 27 \end{bmatrix} \begin{bmatrix} 13 & 1 \\ 20 & 8 \end{bmatrix} = \begin{bmatrix} 452 & 164 \\ 943 & 247 \end{bmatrix} \equiv \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \pmod{41}$$

But inverses commute, so AB is also the identity:

$$AB = \begin{bmatrix} 13 & 1 \\ 20 & 8 \end{bmatrix} \begin{bmatrix} 4 & 20 \\ 31 & 27 \end{bmatrix} = \begin{bmatrix} 83 & 287 \\ 328 & 616 \end{bmatrix} \equiv \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \pmod{41}$$

It turns out that checking the solution is easy, but finding it is not. We will now describe an algorithm that Bob can use to find B.

Linear Equations

Assume that we are working in $\mathbb{Z}_{41}$, so all equations are true $\pmod{41}$. Let’s label the unknown entries of B, and consider the equation $AB = I$.

$$AB = \begin{bmatrix} 13 & 1 \\ 20 & 8 \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

By the definition of matrix multiplication, this breaks into two matrix-vector equations:

$$\begin{bmatrix} 13 & 1 \\ 20 & 8 \end{bmatrix} \begin{bmatrix} b_{11} \\ b_{21} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad \begin{bmatrix} 13 & 1 \\ 20 & 8 \end{bmatrix} \begin{bmatrix} b_{12} \\ b_{22} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Let’s focus on the first of these:

$$\begin{bmatrix} 13 & 1 \\ 20 & 8 \end{bmatrix} \begin{bmatrix} b_{11} \\ b_{21} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

which corresponds to the following system of simultaneous linear equations

$$\begin{aligned} 13b_{11} + b_{22} &= 1 \\ 20b_{11} + 8b_{22} &= 0 \end{aligned}$$

Row Operations

By convention, let's refer to equation i as row i, denoted R_i. Any linear system can be manipulated in three ways without changing the solution, as illustrated here with shorthand descriptions of the row operations:

- swap two equations (e.g. $R_1 \leftrightarrow R_2$)

$$\begin{aligned} 20b_{11} + 8b_{22} &= 0 \\ 13b_{11} + b_{22} &= 1 \end{aligned}$$

- multiply one equation by a non-zero constant (e.g. $R_2 \leftarrow 3R_2$)

$$\begin{aligned} 20b_{11} + 8b_{22} &= 0 \\ 39b_{11} + 3b_{22} &= 3 \end{aligned}$$

- add a multiple of one equation to another equation (e.g. $R_2 \leftarrow R_2 + 2R_1$)

$$\begin{aligned} 20b_{11} + 8b_{22} &= 0 \\ 79b_{11} + 19b_{22} &= 3 \end{aligned}$$

Gauss Elimination

We will design a sequence of row operations that progressively simplify the system of equations, eventually revealing the solution.

It is unnecessary to write the variables each time; only the coefficients and right-hand side values are required. Going back to the original system, write the augmented matrix

$$\left[\begin{array}{cc|c} 13 & 1 & 1 \\ 20 & 8 & 0 \end{array}\right]$$

First multiply row 1 by $13^{-1} \pmod{41} = 19$, and reduce $\pmod{41}$.

$$(R_1 \leftarrow 19R_1) \qquad \left[\begin{array}{cc|c} 247 & 19 & 19 \\ 20 & 8 & 0 \end{array}\right] = \left[\begin{array}{cc|c} 1 & 19 & 19 \\ 20 & 8 & 0 \end{array}\right]$$

Now subtract 20 times row 1 from row 2, and reduce.

$$(R_2 \leftarrow R_2 - 20R_1) \qquad \left[\begin{array}{cc|c} 1 & 19 & 19 \\ 0 & -372 & -380 \end{array}\right] = \left[\begin{array}{cc|c} 1 & 19 & 19 \\ 0 & 38 & 30 \end{array}\right]$$

Next multiply row 2 by $38^{-1} \pmod{41} = 27$, and reduce.

$$(R_2 \leftarrow 27R_2) \qquad \left[\begin{array}{cc|c} 1 & 19 & 19 \\ 0 & 1026 & 810 \end{array}\right] = \left[\begin{array}{cc|c} 1 & 19 & 19 \\ 0 & 1 & 31 \end{array}\right]$$

Finally, subtract 19 times row 2 from row 1, and reduce.

$$(R_1 \leftarrow R_1 - 19R_2) \qquad \left[\begin{array}{cc|c} 1 & 0 & -570 \\ 0 & 1 & 31 \end{array}\right] = \left[\begin{array}{cc|c} 1 & 0 & 4 \\ 0 & 1 & 31 \end{array}\right]$$

Since the coefficient matrix is now the identity, it simply reads $b_{11} = 4$ and $b_{21} = 31$.

27. Go through the same exact process to solve the second system

$$\begin{bmatrix} 13 & 1 \\ 20 & 8 \end{bmatrix} \begin{bmatrix} b_{12} \\ b_{22} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Gauss Elimination Algorithm

Since the exact same row operations must be done for both systems, they can be solved together using the augmented matrix:

$$\left[\begin{array}{cc|cc} 13 & 1 & 1 & 0 \\ 20 & 8 & 0 & 1 \end{array}\right]$$

The general procedure for turning the coefficient matrix into I is called Gaussian elimination. Here the steps are illustrated with a 3×3 matrix.

(I) Start with column $j = 1$.

(II) If the diagonal position j, j is zero, swap row j with another row below it.

(III) Multiply row j by a constant to put a '1' in position j, j.

$$\left[\begin{array}{ccc|ccc} 1 & * & * & * & * & * \\ * & * & * & * & * & * \\ * & * & * & * & * & * \end{array}\right]$$

(IV) Add/subtract a multiple of row j to each row below it to put zeros under the j, j entry.

$$\left[\begin{array}{ccc|ccc} 1 & * & * & * & * & * \\ 0 & * & * & * & * & * \\ 0 & * & * & * & * & * \end{array}\right]$$

(V) Move to the next column, and go back to step (II). After all the diagonal entries are '1', with zeros below, your augmented matrix will look like:

$$\left[\begin{array}{ccc|ccc} 1 & * & * & * & * & * \\ 0 & 1 & * & * & * & * \\ 0 & 0 & 1 & * & * & * \end{array}\right]$$

(VI) Now work your way backward, starting with the last column, zeroing out the entries above the diagonal:

$$\left[\begin{array}{ccc|ccc} 1 & * & 0 & * & * & * \\ 0 & 1 & 0 & * & * & * \\ 0 & 0 & 1 & * & * & * \end{array}\right]$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & * & * & * \\ 0 & 1 & 0 & * & * & * \\ 0 & 0 & 1 & * & * & * \end{array}\right]$$

(VII) Your augmented matrix now displays the inverse on the right-hand side.

Example

Find the inverse of:

$$A = \begin{bmatrix} 2 & 5 & 7 \\ 8 & 20 & 1 \\ 3 & 9 & 18 \end{bmatrix}$$

Start with an augmented matrix $[A \mid I]$, and do row operations until it becomes $[I \mid B]$.

$$\left[\begin{array}{ccc|ccc} 2 & 5 & 7 & 1 & 0 & 0 \\ 8 & 20 & 1 & 0 & 1 & 0 \\ 3 & 9 & 18 & 0 & 0 & 1 \end{array}\right]$$

$$(R_1 \leftarrow 2^{-1}R_1) \quad \left[\begin{array}{ccc|ccc} 1 & 23 & 24 & 21 & 0 & 0 \\ 8 & 20 & 1 & 0 & 1 & 0 \\ 3 & 9 & 18 & 0 & 0 & 1 \end{array}\right]$$

$$(R_2 \leftarrow R_2 - 8R_1) \quad \left[\begin{array}{ccc|ccc} 1 & 23 & 24 & 21 & 0 & 0 \\ 0 & 0 & 14 & 37 & 1 & 0 \\ 3 & 9 & 18 & 0 & 0 & 1 \end{array}\right]$$

$$(R_3 \leftarrow R_3 - 3R_1) \quad \left[\begin{array}{ccc|ccc} 1 & 23 & 24 & 21 & 0 & 0 \\ 0 & 0 & 14 & 37 & 1 & 0 \\ 0 & 22 & 28 & 19 & 0 & 1 \end{array}\right]$$

$$(R_2 \leftrightarrow R_3) \quad \left[\begin{array}{ccc|ccc} 1 & 23 & 24 & 21 & 0 & 0 \\ 0 & 22 & 28 & 19 & 0 & 1 \\ 0 & 0 & 14 & 37 & 1 & 0 \end{array}\right]$$

$$(R_2 \leftarrow 22^{-1}R_2) \quad \left[\begin{array}{ccc|ccc} 1 & 23 & 24 & 21 & 0 & 0 \\ 0 & 1 & 5 & 40 & 0 & 28 \\ 0 & 0 & 14 & 37 & 1 & 0 \end{array}\right]$$

$$(R_3 \leftarrow 14^{-1}R_3) \quad \left[\begin{array}{ccc|ccc} 1 & 23 & 24 & 21 & 0 & 0 \\ 0 & 1 & 5 & 40 & 0 & 28 \\ 0 & 0 & 1 & 29 & 3 & 0 \end{array}\right]$$

$$(R_2 \leftarrow R_2 - 5R_3) \quad \left[\begin{array}{ccc|ccc} 1 & 23 & 24 & 21 & 0 & 0 \\ 0 & 1 & 0 & 18 & 26 & 28 \\ 0 & 0 & 1 & 29 & 3 & 0 \end{array}\right]$$

$$(R_1 \leftarrow R_1 - 24R_3) \quad \left[\begin{array}{ccc|ccc} 1 & 23 & 0 & 22 & 10 & 0 \\ 0 & 1 & 0 & 18 & 26 & 28 \\ 0 & 0 & 1 & 29 & 3 & 0 \end{array}\right]$$

$$(R_1 \leftarrow R_1 - 23R_2) \quad \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 18 & 27 & 12 \\ 0 & 1 & 0 & 18 & 26 & 28 \\ 0 & 0 & 1 & 29 & 3 & 0 \end{array}\right]$$

We have found the inverse:

$$A^{-1} = B = \begin{bmatrix} 18 & 27 & 12 \\ 18 & 26 & 28 \\ 29 & 3 & 0 \end{bmatrix}$$

As you can see, finding the inverse can be computationally intense. Mathematicians prefer to write computer programs to implement such algorithms.

28. Use the Gaussian elimination algorithm to find the inverse of $A = \begin{bmatrix} 14 & 5 \\ 18 & 4 \end{bmatrix}$. (the key “NERD”)

29. Use shorthand to describe the first three row operations you would do on this matrix to implement Gaussian elimination. Do not actually do them.

$$\left[\begin{array}{ccc|ccc} 4 & 15 & 3 & 1 & 0 & 0 \\ 21 & 2 & 33 & 0 & 1 & 0 \\ 7 & 11 & 13 & 0 & 0 & 1 \end{array}\right]$$

Conclusion

Will Gaussian elimination always work? The answer is no - some matrices do not have inverses, and the algorithm will get stuck. Take a course in linear algebra to learn the details.

It turns out that the Hill Cipher can be broken. If an enemy knows the plaintext and corresponding ciphertext of a single message, she can reverse engineer to discover the key, with which to decrypt future messages.

Fortunately for modern electronic communication, mathematicians have developed more sophisticated and secure methods of encryption. They all build upon the basic ideas we have learned in discrete math.

30. (Bonus) The plaintext `"BONUS PROBLEM"` gets encrypted to `"QF434Q6.1FJJXTS"`, and the Hill Cipher key is nine characters long. Decrypt the message `"V.8Y_Q063I63.I5"`.