Alex Hawks
Matthew Westberg
Linear Algebra 2270

Cracking the Hill-Cipher: how to break an encrypted message

Abstract:

Password encryption is used today in storing passwords in databases, and even to pass messages. While encrypting messages can be useful when attempting to pass a message, this project focuses more on the cracking of these encrypted messages. When enough sample sizes are gathered from both the message, and the coded message, the key matrix can be solved. There are many ways to crack messages, but this project will focus primarily on cracking the Hill-Cipher.

Background information:

The Hill-Cipher is an encrypted method based on linear algebra that takes a message then converts it to numbers assigned to each letter in that message. This numerical message is then transformed into a matrix, which is then multiplied by a key matrix that is invertible. After doing so, this matrix undergoes modulo operation of the size of alphabet. For instance, if the alphabet is standard from A-Z, one would modulo 26 for the 26 letters. The decryption process is the opposite. One will find the inverse of the key matrix, continuing to apply the modulo value, to keep the numbers within the range of the alphabet. After doing so, one multiplies the inverse matrix with the coded numerical values. After applying the modulo operation, this will give the original message.

Cracking the Hill-Cipher:

Cracking the code means obtaining the key matrix so that if any other encrypted message is given one would be able to decode the message.

When attempting to crack a code, it is important that enough information is gathered. If the key matrix is a N x N matrix, it will be necessary to obtain N^2 messages and encryptions. For example, if the key matrix is a 2x2 matrix, it will be necessary to obtain 2 messages with 2 entries each and their corresponding encryptions, or 1 message with 4 entries and its corresponding encryption.

The following key will be used in assigning numbers with letters.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	.	?	_
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28

For example, we will choose a key matrix A and its inverse to encode and decode messages.

Suppose a key matrix (whose inverse matrix modulo 29 exists) is chosen to be:

$$A = \begin{vmatrix} 1 & 2 \\ 2 & 7 \end{vmatrix} \qquad A^{-1} \textbf{ modulo } 29 = \begin{vmatrix} 12 & 9 \\ 9 & 10 \end{vmatrix}$$

The chosen message is UWIN with the corresponding number vectors (20, 22), and (8, 13). The number vectors multiplied by A (**modulo 29**) =

$$\begin{vmatrix} 1 & 2 \\ 2 & 7 \end{vmatrix} \begin{vmatrix} 20 & 8 \\ 22 & 13 \end{vmatrix} \textbf{ mod } 29 = \begin{vmatrix} 6 & 5 \\ 20 & 20 \end{vmatrix} = \text{GUFU}$$

Now, consider the other team who has intercepted this message along with the coded message. The other team now has the UWIN message along with cyphertext GUFU. Using only this information, they will be able to find the inverse to the key matrix, which will allow them to decode all future messages.

The matrices for the plaintext message and cyphertext messages are:

$$\text{Plaintext} = \begin{vmatrix} 20 & 8 \\ 22 & 13 \end{vmatrix} \qquad \text{Cyphertext} = \begin{vmatrix} 6 & 5 \\ 20 & 20 \end{vmatrix}$$

The trick to breaking the encryption is knowing that we may form an augmented matrix: $(C^T \mid P^T)$, and by row-reducing can find the transpose of the inverse of the key matrix. Row reducing $(C^T \mid P^T)$ will result in $(I \mid (A^{-1})^T)$. (The source of the theorem can be found in the website below). With this, one can transpose the matrix to get the inverse. This can then be used for decoding future messages.

PROOF: https://www.apprendre-en-ligne.net/crypto/hill/Hillciph.pdf PAGE 17-19

Example:

$$P = \begin{vmatrix} 20 & 8 \\ 22 & 13 \end{vmatrix} \qquad C = \begin{vmatrix} 6 & 5 \\ 20 & 20 \end{vmatrix}$$

$$C^T \mid P^T = \begin{vmatrix} 6 & 20 \\ 5 & 20 \end{vmatrix} \begin{vmatrix} 20 & 22 \\ 8 & 13 \end{vmatrix}$$

Row reduction modulo 29:

$$C^T \mid P^T = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} \begin{vmatrix} 12 & 9 \\ 9 & 10 \end{vmatrix}$$

$$(A^{-1})^T = \begin{vmatrix} 12 & 9 \\ 9 & 10 \end{vmatrix} = (A^{-1})^{TT} = A^{-1}$$

And with the inverse, one can crack all future coded messages.

Summary of results:

The results that we found were that if given enough information, it was in fact possible to find the key matrix to the encryption. The trick is to have enough message to find linearly independent column vectors. If column vectors (encrypted numerical values) are linearly dependent, then those column vectors cannot be used in the process, and more data will be needed. We found that for a 2x2 key matrix, 2 encrypted messages and their 2 plaintext messages were sufficient to find the key matrix.

To make the process more complex, we found that if someone were to use a 3x3 key matrix or a 4x4 key matrix, more intercepted messages and their meanings would be necessary to find the key matrix. Based on the results that we found, one strategy to increase the security of the

encrypted code would be to use a larger key matrix. Mathematically speaking, as the key matrix increases in size, the number of intercepted messages and their meanings that are needed to crack the code would rise in a quadratic manner.

Extra example of a 3x3 key matrix:

In order to hack a cypher with a 3 by 3 key matrix, a message of 9 letters with their cyphers are needed.

Key matrix divisible by mod 29:

$$A = \begin{pmatrix} 1 & 7 & 9 \\ 2 & 7 & 7 \\ 3 & 15 & 0 \end{pmatrix} \qquad A^{-1} \textbf{ modulo } 29 = \begin{pmatrix} 14 & 11 & 27 \\ 3 & 21 & 14 \\ 22 & 5 & 28 \end{pmatrix}$$

The chosen message is USSRHACK! with the corresponding number vectors (20, 18, 18), (17, 7, 0) and (2, 10, 28). The number vectors multiplied by A (**modulo 29**) =

$$\begin{pmatrix} 1 & 7 & 9 \\ 2 & 7 & 7 \\ 3 & 15 & 0 \end{pmatrix} \begin{pmatrix} 20 & 17 & 2 \\ 18 & 7 & 10 \\ 18 & 0 & 28 \end{pmatrix} \textbf{ mod } 29 = \begin{pmatrix} 18 & 8 & 5 \\ 2 & 25 & 9 \\ 11 & 11 & 11 \end{pmatrix} = \text{SCLIZLFJL}$$

Now to crack the cypher, and get the key matrix!

$$\text{Plaintext} = \begin{pmatrix} 20 & 17 & 2 \\ 18 & 7 & 10 \\ 18 & 0 & 28 \end{pmatrix} \qquad \text{Cyphertext} = \begin{pmatrix} 18 & 8 & 5 \\ 2 & 25 & 9 \\ 11 & 11 & 11 \end{pmatrix}$$

Transpose plaintext and cyphertext and combine into an augmented matrix:

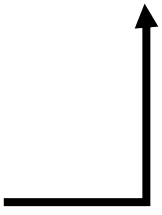
$$C^T \mid P^T = \left(\begin{array}{ccc|ccc} 18 & 2 & 11 & 20 & 18 & 18 \\ 8 & 25 & 11 & 17 & 7 & 0 \\ 5 & 9 & 11 & 2 & 10 & 28 \end{array}\right)$$

Row reduce mod 29:

$$C^T \mid P^T = \begin{pmatrix} 1 & 0 & 0 & 14 & 3 & 22 \\ 0 & 1 & 0 & 11 & 21 & 5 \\ 0 & 0 & 1 & 27 & 14 & 28 \end{pmatrix}$$

$$(A^{-1})^T = \begin{pmatrix} 14 & 11 & 27 \\ 3 & 21 & 14 \\ 22 & 5 & 28 \end{pmatrix} \quad = (A^{-1})^{TT} = A^{-1}$$

$$A^{-1} \textbf{ modulo } 29 = \begin{pmatrix} 14 & 11 & 27 \\ 3 & 21 & 14 \\ 22 & 5 & 28 \end{pmatrix} \; =$$



The matrix has been cracked!