MysteryTwister C3

THE CRYPTO CHALLENGE CONTEST

Hill Cipher

Author: Peter Uelkes

September 2022

The Hill cipher (1/5)

The Hill cipher as described in [1] was introduced by Lester S. Hill in 1929, see also [2] and [3]. It uses methods from linear algebra to implement a polygraphic substitution of a message.
In this challenge, we will use the plaintext alphabet

$$\texttt{ABCDEFGHIJKLMNOPQRSTUVWXYZ}$$

i.e. the uppercase English letters in ordinary sequence. We assign numerical values to them:

$$\mathrm{A} = 0, \mathrm{B} = 1, \ldots, \mathrm{Z} = 25$$

We choose an integer $n \geqslant 2$ and an $n \times n$ matrix M. The matrix must be invertible modulo the length of the alphabet, so for this challenge M will be invertible modulo 26.

The Hill cipher (2/5)

A matrix M over $\mathbb{Z}$ is invertible modulo some integer n iff* there exists a matrix M^{-1} with the property

$$M \cdot M^{-1} = M^{-1} \cdot M = I_n$$

where I_n is the $n \times n$ identity matrix, i.e. the $n \times n$ matrix with ones on the main diagonal and zeroes everywhere else.

It can be shown that such an inverse matrix exists iff the determinant $\det(M)$ of M is co-prime to n. In our case the determinant must therefore be co-prime to 26. This is the case iff $\det(M)$ is co-prime to 2 as well as to 13 because $26 = 2 \cdot 13$ is the prime factorization of 26 [4].

*"iff"is short for "if and only if".

The Hill cipher (3/5)

The chosen M is the secret key which must be agreed upon by the involved communication partners.
As an example, let $n = 2$ and choose M as

$$M = \begin{pmatrix} 19 & 17 \\ 10 & 1 \end{pmatrix}$$

To encrypt a block of text of length n, each character of the block is converted to integers:
For example, we want to encrypt "TOBEORNOTTOBE". The first block is "TO", correspondig to the vector $\begin{pmatrix} 19 \\ 14 \end{pmatrix}$ which is then multiplied (from the right) into the matrix M, the result is taken modulo 26:

$$\begin{pmatrix} 19 & 17 \\ 10 & 1 \end{pmatrix} \cdot \begin{pmatrix} 19 \\ 14 \end{pmatrix} = \begin{pmatrix} 599 \\ 204 \end{pmatrix} \equiv \begin{pmatrix} 1 \\ 22 \end{pmatrix} \pmod{26}$$

The Hill cipher (4/5)

The vector $\begin{pmatrix} 1 \\ 22 \end{pmatrix}$ is now converted back to letters, resulting in "BW", this is the beginning of the ciphertext. The remaining plaintext is treated in the same way. Because the number of letters is odd and therefore not an integer multiple of the matrix dimension n, we attach an X to the plaintext before encrypting.

The resulting ciphertext is "BWJOJBROIBXLZL". To decrypt a ciphertext, we calculate the inverse matrix M^{-1} (modulo 26). This can be achieved e.g. via the Gaussian algorithm[†]. In our example, we have

$$M^{-1} = \begin{pmatrix} 21 & 7 \\ 24 & 9 \end{pmatrix}$$

[†]As we chose M to be invertible modulo 26.

The Hill cipher (5/5)

For deciphering we now take blocks of length n (=2 for the example case) of the ciphertext, convert them to numbers and multiply them from the right into M^{-1} (modulo 26).
E.g. for the first two ciphertext symbols "BW", corresponding to $\begin{pmatrix} 1 \\ 22 \end{pmatrix}$:

$$M^{-1} \cdot \begin{pmatrix} 1 \\ 22 \end{pmatrix} = \begin{pmatrix} 21 & 7 \\ 24 & 9 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 22 \end{pmatrix} = \begin{pmatrix} 175 \\ 22 \end{pmatrix} \equiv \begin{pmatrix} 19 \\ 14 \end{pmatrix} \pmod{26}$$

which converted back into letters is "TO" (not much of a surprise there!). Continuing in blocks of length 2 we get back our plaintext "TOBEORNOTTOBEX" (with the additional "X").

Challenge description (1/2)

This is a ciphertext-only challenge.
For this challenge, the above alphabet of length 26 is used. An 8×8 matrix M was chosen which is invertible modulo 26.

A text P (only consisting of characters A-Z) of length 512 was chosen from the Brown Corpus ([5], [6]), converted to upper case and using only letters from the alphabet above.
The text P was enciphered with the chosen matrix M and is given in the file `ciphertext.txt` in the additional files.

Your task is to find the plaintext P and the encryption matrix M.
Please submit the last word of the plaintext, the determinant of M and the trace[‡] of M, separated by single spaces.

[‡]The trace of a square matrix is the sum of its main diagonal elements.

Challenge description (2/2)

The determinant and the trace both have to be submitted modulo 26. So, for example, if the last word of the plaintext is "SECRET", the determinant is 9 (modulo 26) and the trace is 17 (modulo 26), you would submit: "SECRET 9 17" (without the quotation marks).

Good luck!

Hint 1: Because everything is done modulo 26, all entries of M are integers in the range $0, 1, \ldots, 25$.
Hint 2: The number of invertible matrices modulo 26 is of $\mathcal{O}(26^{n^2}) \approx 10^{90}$ for $n = 8$ [7], so brute forcing the matrices won't solve this challenge.

Resources (1/2)

1. Wikipedia article: en.wikipedia.org/wiki/Hill_cipher
2. David Kahn: "The Code-Breakers", revised revision 1996, pages 404-410
3. Friedrich L. Bauer: "Entzifferte Geheimnisse", 2nd edition, pages 86, 129, 223
4. Murray Eisenberg: apprendre-en-ligne.net/crypto/hill/Hillciph.pdf

Resources (2/2)

5. Brown Corpus, Wikipedia article: en.wikipedia.org/wiki/Brown_Corpus
6. Brown Corpus, download: www.nltk.org/nltk_data
7. K. Pommerening: www.staff.uni-mainz.de/pommeren/Cryptology/Classic/9_Lin/NumbInv.pdf

Additional files

→ `ciphertext.txt`: The ciphertext of the challenge.