Linear Algebra and the Hill Cipher

Seth Reelitz

4/24/16

Introduction

In this report we will be examining how a Hill cipher utilizes fundamental concepts from linear algebra to achieve the encryption and decryption of messages.

Although our primary focus is that of linear algebra, there are several other core concepts and definitions that will need to be discussed along the way. Outside of linear algebra, will be looking at the classification of the cipher, the alphabet assignment, and modular arithmetic.

Classification in Cryptography

Within the field of cryptography there are several methods, or ciphers, to encode and decode messages. One of the most basic ciphers is called a basic substitution cipher (BSC), this is where one letter is substituted with another. Although the Hill cipher uses the concept of substitution to change the plaintext to ciphertext, unlike the BSC, it deals with groups of letters instead of individual ones. This difference in grouping the plaintext is classified as a polygraphic substitution cipher (PSC). A PSC is a more or less a BSC, but as mentioned above, it takes a group of letters and encodes them, instead of the individual letters. This grouping of the plaintext is important to help make frequency analysis more difficult on the encrypted text, with a BSC this analysis is much easier.

Alphabet Assignment and Modular Arithmetic

In this paper we will be using the standard 26 letter English alphabet, each letter will correspond to a number from 0-25. The follow chart will be used as our reference.

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

The Hill cipher requires the use of modular arithmetic in order to be able to take the encoded and decoded messages from the numerical value to its corresponding letter. With our alphabet having 26 characters, we will taking our results with Mod 26.

The Hill Cipher

In order to use the Hill cipher we need to take note of a few fundamental concepts to ensure we can use it. The first, for us to be able to encode a message we must take a plain text and create n x 1 vectors with the corresponding numerical value. Secondly, our encryption matrix needs to be an n x n matrix, which is invertible with a determinant ≠ 0. If our encryption matrix does not contain these properties then we will only be able to encode the plaintext and have no method of decryption. In order to be able to decode, we will need to be using a matrix that is the inverse of encryption matrix. This is one of the several aspects of using the Hill cipher that makes it less practical, finding the encryption matrix.

Let us find our encryption matrix A, remember it must be an invertible matrix with a determinant ≠ 0. To save time we will use a matrix we know meets the criteria instead of having to test the properties.

$$\text{Matrix A} = \begin{bmatrix} 1 & 0 & 2 \\ 10 & 20 & 15 \\ 0 & 1 & 2 \end{bmatrix}$$

Now that we have encryption matrix, it is time for us to find the inverse matrix (decryption matrix), the following steps help us achieve this.

$$\text{Matrix B} = \begin{bmatrix} 15 & 22 & 2 \\ 14 & 22 & 3 \\ 6 & 15 & 12 \end{bmatrix}$$

I have chosen the following message to encrypt “Sorry Hill, RSA is the best”, let us look at that in vector form.

$$\begin{bmatrix} S \\ \mathrm{O} \\ R \end{bmatrix}, \begin{bmatrix} R \\ Y \\ H \end{bmatrix}, \begin{bmatrix} \mathrm{I} \\ L \\ L \end{bmatrix}, \begin{bmatrix} R \\ S \\ A \end{bmatrix}, \begin{bmatrix} \mathrm{I} \\ S \\ T \end{bmatrix}, \begin{bmatrix} H \\ E \\ B \end{bmatrix}, \begin{bmatrix} E \\ S \\ T \end{bmatrix}$$

Here are the vectors after the letters are assigned with their corresponding numerical value;

$$\begin{bmatrix}18\\14\\17\end{bmatrix},\begin{bmatrix}17\\24\\7\end{bmatrix},\begin{bmatrix}8\\11\\11\end{bmatrix},\begin{bmatrix}17\\18\\0\end{bmatrix},\begin{bmatrix}8\\18\\19\end{bmatrix},\begin{bmatrix}7\\4\\1\end{bmatrix},\begin{bmatrix}4\\18\\19\end{bmatrix}$$

Now that we have our matrix A to encrypt our message, we will take matrix A and multiply by each one of the 7 vectors above. It is important to remember that after we get our result we convert it using Mod 26, otherwise we cannot convert it to the alphabet.

$$\begin{bmatrix}0\\13\\22\end{bmatrix},\begin{bmatrix}5\\1\\12\end{bmatrix},\begin{bmatrix}4\\23\\7\end{bmatrix},\begin{bmatrix}17\\10\\18\end{bmatrix},\begin{bmatrix}20\\23\\4\end{bmatrix},\begin{bmatrix}9\\9\\6\end{bmatrix},\begin{bmatrix}16\\9\\4\end{bmatrix}$$

The resulting vectors are now encoded.

$$\begin{bmatrix}A\\N\\W\end{bmatrix},\begin{bmatrix}F\\B\\M\end{bmatrix},\begin{bmatrix}E\\X\\H\end{bmatrix},\begin{bmatrix}R\\K\\S\end{bmatrix},\begin{bmatrix}U\\X\\E\end{bmatrix},\begin{bmatrix}J\\J\\G\end{bmatrix},\begin{bmatrix}Q\\J\\E\end{bmatrix}$$

As we can see in our above results our encoded message reads “ANWFBMEXHRKSUXEJJGQJE“.

We will now take our encoded message and perform the same steps that we did with the encryption to now decrypt the message. Using our inverse matrix B times the vector mod 26 we obtain the following results.

$$\begin{bmatrix}18\\14\\17\end{bmatrix},\begin{bmatrix}17\\24\\7\end{bmatrix},\begin{bmatrix}8\\11\\11\end{bmatrix},\begin{bmatrix}17\\18\\0\end{bmatrix},\begin{bmatrix}8\\18\\19\end{bmatrix},\begin{bmatrix}7\\4\\1\end{bmatrix},\begin{bmatrix}4\\18\\19\end{bmatrix}$$

Then once again we take the numerical value and match it up with its corresponding letter, we can see have gone back to the original message. “Sorry Hill, RSA is the best.

$$\begin{bmatrix} S \\ \mathrm{O} \\ R \end{bmatrix}, \begin{bmatrix} R \\ Y \\ H \end{bmatrix}, \begin{bmatrix} \mathrm{I} \\ L \\ L \end{bmatrix}, \begin{bmatrix} R \\ S \\ A \end{bmatrix}, \begin{bmatrix} \mathrm{I} \\ S \\ T \end{bmatrix}, \begin{bmatrix} H \\ E \\ B \end{bmatrix}, \begin{bmatrix} E \\ S \\ T \end{bmatrix}$$

Conclusion

After having gone through the process of using the Hill cipher, there are couple of major drawbacks to its use in real life scenarios. The first would be that of frequency analysis, although it would be more difficult than analyzing a BSC, if determined, it could be achieved. Secondly, finding the encryption key matrix is time consuming, and difficult at times. The Hill cipher is however a great way to see just another one of linear algebras many uses. Bottom line, if you need to encrypt data, R.S.A. is the way to go, but if you want to have an exercise for students learning linear algebra, the Hill cipher is a great resource.