实验：Hill 密码的加密、解密与破译

学院：电子信息与电气工程学院　　姓名：王煌基　　学号：519030910100

一、实验目的

通过线性代数知识、Hill 密码的加密、解密、破译技术的应用来加深自身数学建模的能力以及 MATLAB 熟练度的提升。

二、实验内容

任务一：在问题（2）中，若已知密文的前 4 个字母 OJWP 分别代表 TACO，问能否将此段密文破译？

问题（2）：甲方截获密文：OJWPISWAZUXAUUISEABAUCRSIPLBHAAMMLPJJOTENH 经分析这段密文是用 $Hill_2$ 密码编译的，且这段密文的字母 UCRS 依次代表字母 TACO，问能否破译这段密文的内容？

【解】

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

这里我们也同样利用上课讲过的方法进行分析，可以得到：

$$\begin{cases} \binom{O}{J} \leftrightarrow \beta_1 \leftrightarrow \binom{15}{10} \leftrightarrow A\alpha_1 \\ \binom{W}{P} \leftrightarrow \beta_2 \leftrightarrow \binom{23}{16} \leftrightarrow A\alpha_2 \end{cases}, \text{且} \begin{cases} \alpha_1 \leftrightarrow \binom{20}{1} \leftrightarrow \binom{T}{A} \\ \alpha_2 \leftrightarrow \binom{3}{15} \leftrightarrow \binom{C}{O} \end{cases}$$

计算 β 构造出的矩阵的行列式的值我们可以得到：

$$\det(\beta_1 \quad \beta_2) = \begin{vmatrix} 15 & 23 \\ 10 & 16 \end{vmatrix} = 15 * 16 - 23 * 10 = 10 \equiv 10(mod\ 26)$$

但是我们注意到，模 26 倒数表内：

a	1	3	5	7	9	11	15	17	19	21	23	25
$a^{-1}(mod\ 26)$	1	9	21	15	3	19	7	23	11	5	17	25

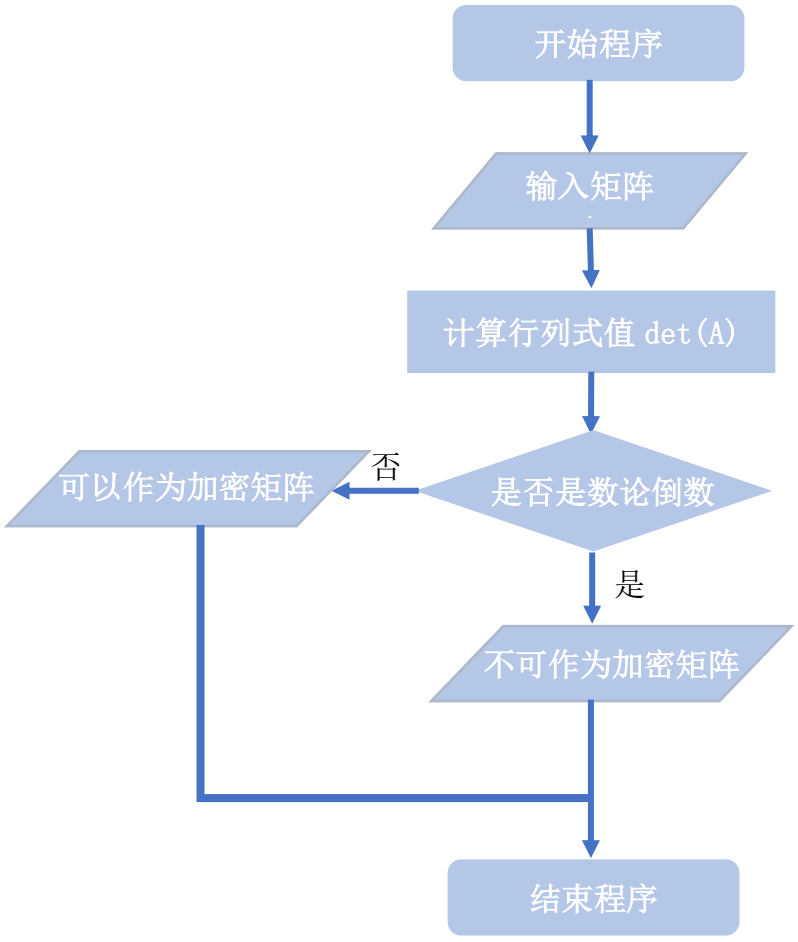
并不存在 10 的情况，即没有数论倒数，因此无法得到在 mod 26 条件下的加密矩阵，故无法进行破译。

任务二：利用所介绍的 $Hill_2$ 密码体制的原理，根据给定的 26 个英文字母的乱序表值，设计与建立 $Hill_4$ 密码体制的加密、解密与破译框图并建立必要的计算机程序. 设英文 26 个字母以下面的乱序表与 Z_{26} 中的整数对应：

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

（1）设 $A=\begin{pmatrix} 8 & 6 & 9 & 5 \\ 6 & 9 & 5 & 10 \\ 5 & 8 & 4 & 9 \\ 10 & 6 & 11 & 4 \end{pmatrix}$，验证矩阵能否作为 $Hill_4$ 密码体制的加密矩阵.用框图画出你的验算过程，并编写相应的计算机程序.

【解】（1）此处框图如下图：



因此，我们可以通过这个原理得到相应的 m 文件（test.m）：

```
function test(n)
    A = zeros(n,n);
    for i = 1:n
        for j = 1:n
            A(i,j) = input('');
        end
    end
    d = mod(det(A),26);
    flag1 = mod(d,2);
    flag2 = mod(d,13);
    if flag1 == 0 || flag2 == 0
        disp('A 不能作为加密矩阵');
    else
        disp('A 能作为加密矩阵');
    end
```

此处的输入是通过换行输入（一个一个输入的），运行结果如下：

```
>> test(4)
8
6
9
5
6
9
5
10
5
8
4
9
10
6
11
4
A 能作为加密矩阵
```

故 A 矩阵能作为 $Hill_4$ 密码体制的加密矩阵。

（2）设明文为

HILL CRYPTOGRAPHIC SYSTEM IS TRADITIONAL

利用上面的表值与加密矩阵给此明文加密，并将得到的密文解密.画出加密与解密过程的框图并编写相应的计算机程序。

【解】根据题意以及所学知识，可以容易得到相应的 m 文件（encrypt.m 以及 decrypt.m）：

```
function encrypt()
    %Change1 是字母转数字，Change2 是数字转字母
    Change1 = [5,23,2,20,10,15,8,4,18,25,0,16,13,7,3,1,19,6,12,24,21,17,14,22,11,9];
    Change2 = ['K','P','C','O','H','A','R','N','G','Z','E','Y','S','M','W','F','L','V','I'
,'Q','D','U','X','B','T','J'];
    str=input('','s');
```

```
    l = size(str);%l(2)表示 str 的长度（包括空格）
    blank = [];
    tmp = "";
    tot = 1;
    %首先去除空格，并且把空格区域记录下来，便于之后还原的时候利用
    for i = 1:l(2)
        if str(i)==' '
            blank(tot) = i;
            tot = tot + 1;
        else
            tmp = tmp + str(i);
        end
    end
    blank(tot) = 1000000;
    %然后就得到了一组无空格字符串并且方便进行操作
    l = length(tmp{1});
    A = [8,6,9,5;6,9,5,10;5,8,4,9;10,6,11,4];
    B = zeros(1,l);
    %然后我们将其转换成数字
    for i = 1:l
        t = double(tmp{1}(i))-64;
        B(1,i) = Change1(t);
    end
    %然后我们乘上矩阵 A 后进行加密
    Key = zeros(1,l);
    for i = 1:4:l
        temp1 = B(i:i+3)';
        temp2 = mod(A*temp1,26);
        Key(i:i+3) = temp2';
    end
    %然后转回字母，此时同时进行加空格的操作
    Asw = "";
    total = 1;
    x = 1;
    for i = 1:l
        Asw = Asw + Change2(Key(i)+1);
        total = total + 1;
        if(total == blank(x))
            Asw = Asw + ' ';
            x = x + 1;
        end
    end
    %最后输出加密后的字符串
    disp(Asw);

function decrypt()
```

```
2.      %Change1 是字母转数字，Change2 是数字转字母
3.      Change1 = [5,23,2,20,10,15,8,4,18,25,0,16,13,7,3,1,19,6,12,24,21,17,14,22,11,9];
4.      Change2 = ['K','P','C','O','H','A','R','N','G','Z','E','Y','S','M','W','F','L','V','I'
    ,'Q','D','U','X','B','T','J'];
5.      prime = [1,0,9,0,21,0,15,0,3,0,19,0,0,0,7,0,23,0,11,0,5,0,17,0,25,0];
6.      str=input('','s');
7.      l = size(str);%l(2)表示 str 的长度（包括空格）
8.      blank = [];      tmp = "";      tot = 1;
9.      %首先去除空格，并且把空格区域记录下来，便于之后还原的时候利用
10.     for i = 1:l(2)
11.         if str(i)==' '
12.             blank(tot) = i;
13.             tot = tot + 1;
14.         else
15.             tmp = tmp + str(i);
16.         end
17.     end
18.     blank(tot) = 1000000;
19.     %然后就得到了一组无空格字符串并且方便进行操作
20.     l = length(tmp{1});      A = [8,6,9,5;6,9,5,10;5,8,4,9;10,6,11,4];
21.     B = zeros(1,l);
22.     %然后我们将其转换成数字
23.     for i = 1:l
24.         t = double(tmp{1}(i))-64;
25.         B(1,i) = Change1(t);
26.     end
27.     detA = det(A);                    d = round(mod(detA,26));
28.     r1 = prime(d);                    invA = inv(A);
29.     An = detA*invA;                   An1 = round(mod(r1*An,26));
30.     Key = zeros(1,l);
31.     for i = 1:4:l
32.         temp1 = B(i:i+3)';
33.         temp2 = mod(An1*temp1,26);
34.         Key(i:i+3) = temp2';
35.     end
36.     %然后转回字母，此时同时进行加空格的操作
37.     Asw = "";
38.     total = 1;
39.     x = 1;
40.     for i = 1:l
41.         Asw = Asw + Change2(Key(i)+1);
42.         total = total + 1;
43.         if(total == blank(x))
44.             Asw = Asw + ' ';
45.             total = total+1;
46.             x = x + 1;
```

```
47.         end
48.     end
49.     %最后输出解密后的字符串
50.     disp(Asw);
```

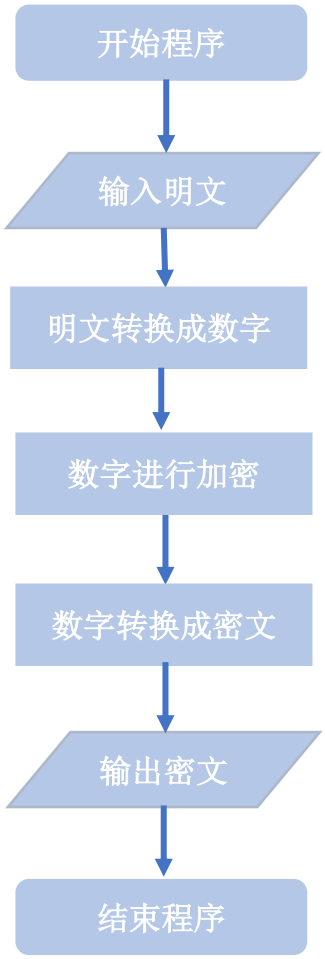
且二者的运行结果如下：

```
>> encrypt
HILL CRYPTOGRAPHIC SYSTEM IS TRADITIONAL
KEGT KPNJKYXRLAOLM ZTPVYT UN HZSCEGDZRPZ
```

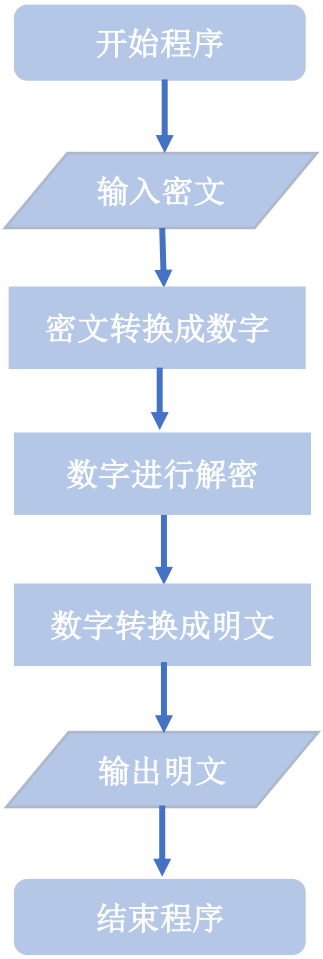
```
>> decrypt
KEGT KPNJKYXRLAOLM ZTPVYT UN HZSCEGDZRPZ
HILL CRYPTOGRAPHIC SYSTEM IS TRADITIONAL
```

流程图如下（其中省略了具体的计算过程以及具体的空格删减增补过程）：

加密过程流程图



解密过程流程图



（3）已知在上述给定表值下的一段 Hill_4 密码的密文为

JCOW ZLVB DVLE QMXC

对应的明文为

DELA YOPE RATI ONSU

能否确定对应的加密矩阵？给出你的判断过程.

【解】此处我们先找到明文与密文之间的对应关系：

$$\begin{pmatrix} J \\ C \\ O \\ W \end{pmatrix} \leftrightarrow \beta_1 = \begin{pmatrix} 25 \\ 2 \\ 3 \\ 14 \end{pmatrix} = A\alpha_1 \text{，且} \begin{pmatrix} D \\ E \\ L \\ A \end{pmatrix} \leftrightarrow \alpha_1 = \begin{pmatrix} 20 \\ 10 \\ 16 \\ 5 \end{pmatrix}$$

$$\begin{pmatrix} Z \\ L \\ V \\ B \end{pmatrix} \leftrightarrow \beta_2 = \begin{pmatrix} 9 \\ 16 \\ 17 \\ 23 \end{pmatrix} = A\alpha_2 \text{，且} \begin{pmatrix} Y \\ O \\ P \\ E \end{pmatrix} \leftrightarrow \alpha_2 = \begin{pmatrix} 11 \\ 3 \\ 1 \\ 10 \end{pmatrix}$$

$$\begin{pmatrix} D \\ V \\ L \\ E \end{pmatrix} \leftrightarrow \beta_3 = \begin{pmatrix} 20 \\ 17 \\ 16 \\ 10 \end{pmatrix} = A\alpha_3 \text{，且} \begin{pmatrix} R \\ A \\ T \\ I \end{pmatrix} \leftrightarrow \alpha_3 = \begin{pmatrix} 6 \\ 5 \\ 24 \\ 18 \end{pmatrix}$$

$$\begin{pmatrix} Q \\ M \\ X \\ C \end{pmatrix} \leftrightarrow \beta_4 = \begin{pmatrix} 19 \\ 13 \\ 22 \\ 2 \end{pmatrix} = A\alpha_4 \text{，且} \begin{pmatrix} O \\ N \\ S \\ U \end{pmatrix} \leftrightarrow \alpha_4 = \begin{pmatrix} 3 \\ 7 \\ 12 \\ 21 \end{pmatrix}$$

%这里的所有计算均写入了 m 文件中（solve2_3.m），在本题结束处展示。

因此，我们通过 MATLAB 计算可以得到，β 的列向量构成的矩阵的行列式的值为：

$$\det(\beta_1 \quad \beta_2 \quad \beta_3 \quad \beta_4) = \det\begin{pmatrix} 25 & 9 & 20 & 19 \\ 2 & 16 & 17 & 13 \\ 3 & 17 & 16 & 22 \\ 14 & 23 & 10 & 2 \end{pmatrix} = 77881 \equiv 11(mod26)$$

显然，5 是其中一个数论倒数，说明确实存在加密矩阵 A。

下面通过 MATLAB 计算加密矩阵 A。

$$A\alpha = \beta \Leftrightarrow A = \beta\alpha^{-1}$$

其中，α、β 的矩阵满足：

$$\beta = \begin{matrix} 25 & 9 & 20 & 19 \\ 2 & 16 & 17 & 13 \\ 3 & 17 & 16 & 22 \\ 14 & 23 & 10 & 2 \end{matrix}, \quad \alpha = \begin{matrix} 20 & 11 & 6 & 3 \\ 10 & 3 & 5 & 7 \\ 16 & 1 & 24 & 12 \\ 5 & 10 & 18 & 21 \end{matrix}$$

```
>> solve2_3

ans =

   11.0000


ans =

   16.0000   12.0000    5.0000   23.0000
   12.0000    5.0000   23.0000   20.0000
   23.0000   16.0000    8.0000    5.0000
   20.0000   12.0000    9.0000    8.0000
```

将其带入相应 m 文件后我们得到上图所示的结果，因此得到结论：存在加密矩阵 A 为：

$$A=\begin{pmatrix}16 & 12 & 5 & 23\\ 12 & 5 & 23 & 20\\ 23 & 16 & 8 & 5\\ 20 & 12 & 9 & 8\end{pmatrix}$$

下面是相应的 m 文件代码（solve2_3.m）：

```
function solve2_3()
    A = [25,9,20,19;2,16,17,13;3,17,16,22;14,23,10,2];
    mod(det(A),26)
    beta = [25,9,20,19;2,16,17,13;3,17,16,22;14,23,10,2];
    alpha= [20,11,6,3;10,3,5,7;16,1,24,12;5,10,18,21];
    inva = det(alpha)*inv(alpha);
    ans = mod(beta*inva,26)
```

任务三：设已知一份密文为 Hi112 密码体系，其中出现频数最高的双字母是 RH 和 NI，而在明文语言中，出现频数最高的双字母为 TH 和 HE 由这些信息按表 10.5 给出的表值能得到什么样的加密矩阵？

表 10.5 明文字母的表值

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

$$\begin{cases}\binom{R}{H}\leftrightarrow\beta_1\leftrightarrow\binom{17}{7}\leftrightarrow A\alpha_1\\ \binom{N}{I}\leftrightarrow\beta_2\leftrightarrow\binom{13}{8}\leftrightarrow A\alpha_2\end{cases},\text{且}\begin{cases}\alpha_1\leftrightarrow\binom{19}{7}\leftrightarrow\binom{T}{H}\\ \alpha_2\leftrightarrow\binom{7}{4}\leftrightarrow\binom{H}{E}\end{cases}$$

计算 β 构造出的矩阵的行列式的值我们可以得到：

$$\det(\beta_1 \quad \beta_2) = \begin{vmatrix} 17 & 13 \\ 7 & 8 \end{vmatrix} = 17 * 8 - 7 * 13 = 45 \equiv 19(mod\ 26)$$

而 19 是存在数论倒数的，故存在相应的加密矩阵 A 使得加密关系成立。

下面通过 MATLAB 计算加密矩阵 A，用与 2(3)类似的方法可以解得，原理为：

$$A\alpha = \beta \Leftrightarrow A = \beta\alpha^{-1}$$

解得加密矩阵 A 为：

$$A = \begin{pmatrix} 3 & 24 \\ 24 & 25 \end{pmatrix}$$

相应的 m 文件如下(Hill2.m)：

```
1.  function Hill2()
2.      alpha = [19,7;7,4];
3.      beta =  [17,13;7,8];
4.      detb = mod(det(beta),26)
5.      inva = det(alpha).*inv(alpha);
6.      A = mod(beta*inva,26)
```

任务四：如下的密文据表 10.1 以 $Hill_2$ 加密，密文为

VIKYNOTC**LK**YRJQETIRECVUZLNOJTUYDIMHRCFITQ

已获知其中相邻字母 LK 表示字母 KE，试破译这份密文.

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

$$\begin{pmatrix} L \\ K \end{pmatrix} \leftrightarrow \beta_1 \leftrightarrow \begin{pmatrix} 12 \\ 11 \end{pmatrix} \leftrightarrow A\alpha_1, \text{且 } \alpha_1 \leftrightarrow \begin{pmatrix} 11 \\ 5 \end{pmatrix} \leftrightarrow \begin{pmatrix} K \\ E \end{pmatrix}$$

但是此时我们需要注意到，这样的条件似乎是不足够的，还少了一组明文密文的转换方式，因此这里可以考虑采取先设出相应的加密矩阵（因为这里的规模比较小，只是 2X2 的规模），再通过计算寻找相应的性质，设矩阵 A 为：$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$，我们可以通过$\beta_1 = A\alpha_1$来寻找 A 矩阵的某些性质后进行枚举这些可能的 A 矩阵，判断其中语句恰当的情况：

$$\because \beta_1 = \begin{pmatrix} 12 \\ 11 \end{pmatrix} = A\alpha_1 = \begin{pmatrix} 11a + 5b \\ 11c + 5d \end{pmatrix} \ \therefore \begin{cases} 11a + 5b = 12 \\ 11c + 5d = 11 \end{cases} (mod\ 26), \text{且} a, b, c, d \in Z.$$

因此在这里我们只需要寻找a, b, c, d从 0 至 25 的所有可能情况即可，而且，有趣的是，$\gcd(11,5) = 1$，这意味着a与b,c与d是互质的，则当前者确定的时候，后者也一定是确定

的，因此我们只需要枚举相应的前者，就能够得到相应的后者，然后我们将所有可能的语句进行简单阅读后再判断哪句话是有意义的即可，这里我们只需要对 decrypt.m 文件进行简单的修改后就可以得到适应本题的新程序(new_decrypt.m)，且将相应的输出结果存放于 txt 文件中(密文转明文.txt)，由于数量级在三位数，故可以通过认为扫描来排除，而当数量级更大的时候引发了我在后文的新思考。这里的扫描其实也很方便，因为很多语句的开头就不是单词的前缀或者部分前缀，例如

A = 1 21

9 24

WUCNOUWKKEJWMKPWRGKVYKBPOUTBKXSCJOKIEEUU

A = 1 21

10 1

KCTGOUQOKEOXIEYDETFUCQRNOUTBJCCEYRSUAYYA

很明显，前面的三个字母在英文表述中毫无章法，因此这为我的检索带来很大的方便。

然后我注意到其中可以组成句子的特殊语句有且仅有一句，即：

A = 5 7

2 3

CANYOUMAKEANOMELETTEWITHOUTBREAKINGEGGSS

这句话是英文谚语：有失才有得。在这里出现是非常巧合的，故我认为这个是一个合理的加密矩阵 A。为什么后面是 2 个 S 呢？因为这里要求两个两个一组，而在这个句子中最后的 S 是单个的，故需要多写一个 S，从而使得程序能正常进行。

下面是相应的 m 文件(new_decrypt.m)：

```
function new_decrypt(str)
    %Change1 是字母转数字，Change2 是数字转字母
    Change1 = [1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,0];
    Change2 = ['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'];
    prime = [1,0,9,0,21,0,15,0,3,0,19,0,0,0,7,0,23,0,11,0,5,0,17,0,25,0];
    for ii = 0:25
        for jj = 0:25
            A = zeros(2,2);
            A(1,1)=ii;A(2,1)=jj;
            for kk = 0:25
                if mod(ii*11+kk*5,26)==12
                    A(1,2) = kk;
                end
```

```
                if mod(jj*11+kk*5,26)==11
                    A(2,2) = kk;
                end
            end
            A
            detA = round(det(A));
            if mod(detA,2)~=0&&mod(detA,13)~=0
                l = length(str{1});
                B = zeros(1,l);
                %然后我们将其转换成数字
                for i = 1:l
                    t = double(str{1}(i))-64;
                    B(1,i) = Change1(t);
                end

                d = round(mod(detA,26));
                r1 = prime(d);
                invA = inv(A);
                An = detA*invA;
                An1 = round(mod(r1*An,26));

                Key = zeros(1,l);
                for i = 1:2:l
                    temp1 = B(i:i+1)';
                    temp2 = mod(An1*temp1,26);
                    Key(i:i+1) = temp2';
                end
                %然后转回字母
                Asw = "";
                for i = 1:l
                    if(Key(i)==0)
                        Asw = Asw + Change2(26);
                    else
                        Asw = Asw + Change2(Key(i));
                    end
                end
                %最后输出解密后的字符串
                disp(Asw);
            end
        end
    end
```

任务五：找出元素属于 Z_{26} 的所有可能的 $Hill_2$ 密码加密矩阵. 若截获了如下一段密文

UTCQCVFOYQUVMGMGULFOLEYHDUHOPEASWXTIFBAMWT

且已知它是根据表 10.1 按 $Hill_2$ 密码体制加密的，你能否将其解密？

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

这里其实和任务四是非常类似的，故我们将任务四的一些条件简单修改后就可以得到相应的 m 文件(new2_decrypt.m)了。

```
function new2_decrypt(str)
    fid = fopen('b.txt','a');
    %Change1 是字母转数字，Change2 是数字转字母
    Change1 = [1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,0];
    Change2 = ['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'];
    prime = [1,0,9,0,21,0,15,0,3,0,19,0,0,0,7,0,23,0,11,0,5,0,17,0,25,0];
    for ii = 0:25
        for jj = 0:25
            for kk = 0:25
                for ll = 0:25
                    A = zeros(2,2);
                    A(1,1)=ii;A(2,1)=jj;A(1,2)=kk;A(2,2)=ll;
                    detA = round(det(A));
                    if mod(detA,2)~=0&&mod(detA,13)~=0
                        l = length(str{1});
                        B = zeros(1,l);
                        %然后我们将其转换成数字
                        for i = 1:l
                            t = double(str{1}(i))-64;
                            B(1,i) = Change1(t);
                        end

                        d = round(mod(detA,26));
                        r1 = prime(d);
                        invA = inv(A);
                        An = detA*invA;
                        An1 = round(mod(r1*An,26));

                        Key = zeros(1,l);
                        for i = 1:2:l
                            temp1 = B(i:i+1)';
                            temp2 = mod(An1*temp1,26);
                            Key(i:i+1) = temp2';
                        end
                        %然后转回字母
```

```
                        Asw = "";
                        for i = 1:l
                            if(Key(i)==0)
                                Asw = Asw + Change2(26);
                            else
                                Asw = Asw + Change2(Key(i));
                            end
                        end
                        %最后输出解密后的字符串
                        fprintf(fid,'%d    ',A(1,1));
                        fprintf(fid,'%d    ',A(1,2));
                        fprintf(fid,'%d    ',A(2,1));
                        fprintf(fid,'%d\n',A(2,2));
                        fprintf(fid,'%s\n',Asw);
                    end
                end
            end
        end
    end
    fclose(fid);
end
```

由于运行结果过大，故放在相应的 txt 文件(密文转明文 2.txt)中。

这里由于量实在过大，故我直接查找特殊情况下的加密矩阵（此处以 a=5 的情况来进行查找，因为任务四中是以 a=5 的情况作为加密矩阵），结果并没有找到合适的英文句子，我便思考是否是中文的拼音翻译呢？果不其然，在 A=　5　2; 3　11 处找到了中文句子：

WEIRUANGONGSIJIJIANGTUICHUXINYIDAIBENTENGG

翻译过来为：微软公司即将推出新一代奔腾（句子后的两个 G 是凑数专用的）

当然，这里不排除还有其他的翻译情况，但是由于体量过大且我想不到什么好的方法来进行筛选，故只能做到这里了。

三、问题思考

在实验过程中，我发现有一个问题，就是当我们提供的条件非常少甚至不够充足的情况下实际上是非常难对相应的句子进行破译的（如任务四仅提供一个条件、任务五没有提供条件），而且，当加密矩阵的阶数增加的时候，破译难度可以说是指数级增加，因此密码的安全性就更高了。然而，当我们在进行任务四、五的破译的时候，除了人工查找，有什么其他的方法吗？

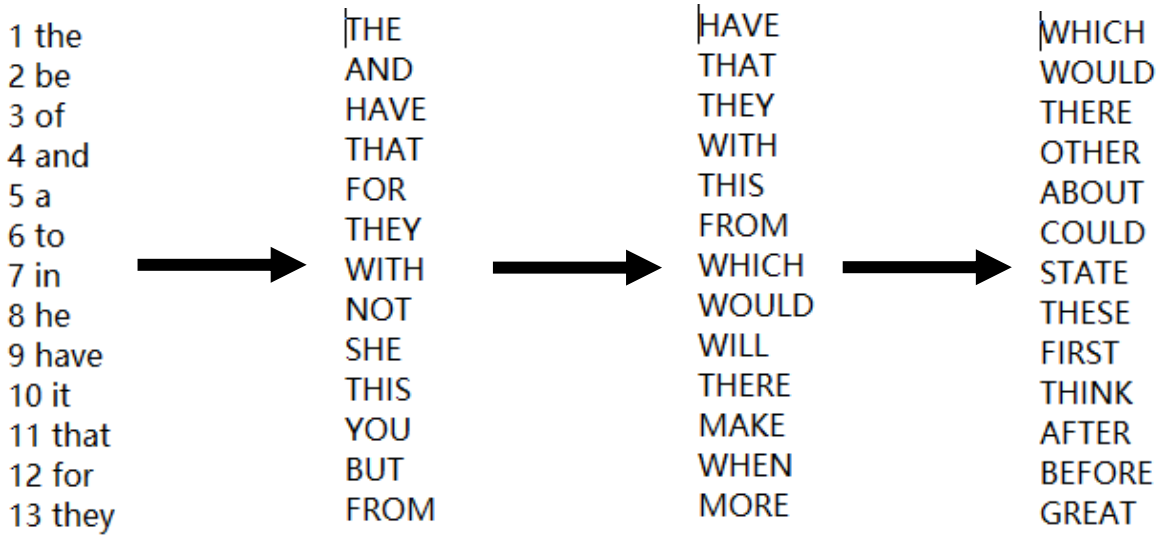
这里我有一个想法（感觉可行，但是不知道怎么实际操作）：将英语中的常用单词全部选出来（大约 3 万个），并且排除其中的 1-3 个字母的单词情况（避免出现过多的无关句

子，加大我们后面辨析的难度）后，在所有我们觉得可能的明文中进行枚举查找，判断这句明文中是否出现了至少一个常用单词，如果有，说明有可能是正确的，无，说明不可能是正确的，以这样的方法就可以大大缩小规模过大的情况（例如任务五可能是可以实现的），但是这个方法的难题在于怎么确定常用的 3 万个单词，这个问题我不知道怎么解决。

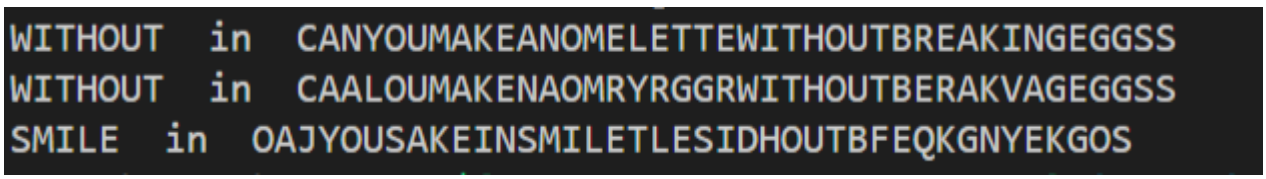
经过一段时间的思考后，我想到一个方法，就是利用字符串的匹配功能，大致原理与我的想法是符合的，只是这个方法存在一些漏洞，后文会提出。（此处想法的所有文件均在 py 文件夹中，因为是利用 python 写的程序代码）

首先我通过百度的方式找到了两千个常用英语单词（均存放于 foo.txt 中），当然，很显然的一点是不能利用长度为 1-3 的单词进行匹配，因为出现的重复率是非常高的，故这里我采取的是利用长度为 4 的单词来进行匹配，当然这里如果匹配不到，则需要考虑长度为 5 的，以此类推，所以这里的方法有点类似于“试根法”，强行找在句子内部的单词。

首先我对于网上的单词进行一定的处理，效果图如下图所示（转换成大写状态后再通过长度进行筛选，筛选出$\geq 3, \geq 4, \geq 5$的单词情况）：



然后就直接进行字符串的匹配操作，这里采取的是任务四中的样例，对于任务四中产生的大规模数据进行处理后我们可以得到如下结果（此处采取的是长度为$\geq 4, \geq 5$的情况，由于版面问题，第一张图是≥ 5的情况，第二张图才是≥ 4的情况）：



```
URGE  in  SOMPOUUUKEPSCIFURGEZIMPXOUTBODEUBCEMUCEW
KISS  in  MSOFOUEWKELMASDERGIFKCXJOUTBUZWIPKISSMGM
WITH  in  CANYOUMAKEANOMELETTEWITHOUTBREAKINGEGGSS
WIPE  in  WIPEOUWEKEWBMGCFETXQYOBFOUTBXWSMWDKQEAUY
WITH  in  CAALOUMAKENAOMRYRGGRWITHOUTBERAKVAGEGGSS
CAKE  in  IAYLOUCAKERAQMTYRGCRUILHOUTBYRIKHACEIGQS
CAKE  in  IALYOUCAKEENQMGLETPEUILHOUTBLEIKUNCEIGQS
SKIN  in  QOEPOUGUKEFSKINURGOZAMJXOUTBQDKUXCOMCCWW
BONE  in  SEMBOUUCKEPUCWFIRGEXIYPTOUTBONEYBIEKUQEI
SURE  in  YSXSOUKWKEGZESURETNSGCHJOUTBVMMIAXASWMCM
HIGH  in  IKYZOUCSKERYQYTKRGCTUWLLOUTBYHIGHUCGISQG
TOWN  in  AKSZOUYSKEDYWYZKRGQTOWNLOUTBGHGGRUQGOSKG
MANY  in  YMKHOUKGKETIEQHCRGAJGEHROUTBIFMANYAWWKCO
WAVE  in  QEROOUGCKESHKWAVETBKAYJTOUTBDAKYKVOKCQWI
SAKE  in  OAWLOUSAKEVASMVYRGYRSIDHOUTBSRQKTAYEKGOS
MILE  in  OAJYOUSAKEINSMILETLESIDHOUTBFEQKGNYEKGOS
SIZE  in  SIZEOUUEKECBCGSFETRQIOPFOUTBBWEMODEQUAEY
JOKE  in  CUNAOUMKKEAJOKEJETTIWKTPOUTBRKACIBGIGESU
MALE  in  QARYOUGAKESNKMALETBEAIJHOUTBDEKKKNOECGWS
KEEP  in  IQLKOUCIKEEPQAGZETPCUULDOUTBLOIOUTCCIUQE
SUIT  in  OYJQOUSMKEIDSUITETLOSADBOUTBFGQQGJYOKOOK
```

我们可以发现，原来数百条甚至上千条的可能语句被我们经过简单的筛选方式就可以缩减了非常大的数据规模，这里从 312 条语句直接缩减成了 21 条（≥ 4的情况）和 3 条（≥ 5的情况），效率是非常高的，我们从这些数据中也可以很方便地看出合理的句子确实是

CANYOUMAKEANOMELETTEWITHOUTBREAKINGEGGSS

同样的，对于任务五的句子我们也可以如此操作，由于任务五的数据规模更大，故这个方法是非常必要的。通过相应的程序，我们成功把原本 9131kb 的文件缩减到 11kb（虽然还是很多，但是相比而言缩减率是非常高的！）。但是，结果却是令人失望的，我并没有找到合理的英语语句，所以大概率这题的答案确实是用拼音写成的句子（这里我也无法确定）。

这个方法存在一些弊端：首先，如果我们的这个常用英语字典库选的不够充足，那么必然存在一些被遗漏的存在；其次，当一个英语语句的单词全由少字母的单词构成的（例如语句：”AREYOUABOY”，最长 3 个字母，最短 1 个字母；再例如：”HEISME”），说明确实可能存在较短单词构成的语句而由于筛选条件的问题被遗漏；最后，当语言不通时（我怀疑任务五确实是中文），只能更新筛选的语言词库，否则将无法找到结果。

四、反思体会

学习 MATLAB 已经快 3 个月了，对于 MATLAB 的很多基本操作更加熟练掌握，例如文件的读写、字符串的处理操作、逻辑运算符等的应用都逐渐熟悉。此外，本次的实验让我将我的程序设计课程的思想与线性代数的思想相结合，我学科知识的掌握又加深了许多，可以说，虽然仅是一位大一的学生，但是在数学实验这门课中我学到的也是相当的丰富，提前接触了很多可能我未来会接触到的课程，这对我而言是大有裨益的！比较可惜的是目前我对于 MATLAB 中小程序的制作还是不是很清楚，这将是我下一个阶段努力的方向。