Problem

Let

$$\mathbf{M} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

Find $\mathbf{M}^{25}$ (source).

Solution:

My solution is slightly different than the one in the video in the sense that it uses elementary matrix operations.

Matrix $\mathbf{M}$ can be written as the sum of the unit matrix $\mathbf{I}_3$ and another "complement" matrix $\mathbf{C}$:

$$\mathbf{M} = \mathbf{I}_3 + \mathbf{C}, \quad \mathbf{C} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

First, to expand $(\mathbf{I}_3 + \mathbf{C})^{25}$, we use the binomial theorem (see A.1):

$$(\mathbf{I} + \mathbf{C})^{25} = \sum_{k=0}^{25} \frac{25!}{k!(25-k)!} \mathbf{I}^{25-k} \mathbf{C}^k = \sum_{k=0}^{25} \frac{25!}{k!(25-k)!} \mathbf{C}^k \tag{1}$$

Now $\mathbf{C}^k$ needs to be computed for each k. $\mathbf{C}$ resembles with the $\mathbf{I}_3$ unit matrix; we can transform $\mathbf{I}_3$ to $\mathbf{C}$ if we swap rows 1, 3 (operation $R_1 \leftrightarrow R_3$) and then swap rows 2, 3 (operation $R_2 \leftrightarrow R_3$):

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_3} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \xrightarrow{R_2 \leftrightarrow R_3} \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

Operation R_{13} can be performed by matrix multiplication; by multiplying $\mathbf{I}_3$ with a 3×3 so called "elementary" matrix $\mathbf{E}_{13}$ such that it swaps rows 1 and 3. So does R_{23}, but with $\mathbf{E}_{23}$. Therefore, $\mathbf{C}$ can be expressed with matrix multiplication as

$$\mathbf{C} = \mathbf{E}_{23}\mathbf{E}_{13}\mathbf{I}_3 = \mathbf{E}_{23}\mathbf{E}_{13} \tag{2}$$

One can find out (see A.2) that the sought row-swapping matrices $\mathbf{E}_{23}$, $\mathbf{E}_{13}$ are:

$$\mathbf{E}_{23} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \quad \mathbf{E}_{13} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

Computing the next higher power:

$$\begin{aligned} \mathbf{C}^2 = \mathbf{E}_{23}\mathbf{E}_{13}\mathbf{C} = \mathbf{E}_{23}\mathbf{E}_{13} \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} = \mathbf{E}_{23} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \end{aligned} \tag{3}$$

Raising to the next higher power, we obtain a nice recursion:

$$\begin{aligned} \mathbf{C}^3 &= \mathbf{C}\mathbf{C}^2 \\ &= \mathbf{E}_{23}\mathbf{E}_{13}\mathbf{C}^2 = \mathbf{E}_{23}\mathbf{E}_{13} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \\ &= \mathbf{E}_{23} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \mathbf{I}_3 \end{aligned} \tag{4}$$

In general, as derived from Eq. (2), Eq. (3), Eq. (4) powers of $\mathbf{C}$ are repeated every 3:

$$\mathbf{C}^n = \begin{cases} \mathbf{I}_3, & n = 3k, \quad k \in \mathbb{N} \\ \mathbf{C}, & n = 3k+1, \quad k \in \mathbb{N} \\ \mathbf{C}^2, & n = 3k+2, \quad k \in \mathbb{N} \end{cases}$$

Both $\mathbf{C}$ and $\mathbf{C}^2$ have been computed so the sum in Eq. (1) can be calculated in terms of known matrices. Notice that the term $\mathbf{C}^n$ of the sum will equal to $\mathbf{I}_3$ for $n = 0, 3, \ldots, 24$, to $\mathbf{C}$ for $n = 1, 4, \ldots, 23$, and to $\mathbf{C}^2$ for $n = 2, 5, \ldots, 25$. Therefore we break the sum in 3 components, each one running for each indexing mentioned above:

$$\sum_{k=0}^{25} \frac{25!}{k!(25-k)!}\mathbf{C}^k = \underbrace{\sum_{\substack{k=0,\\ k+=3}}^{24} \frac{25!}{k!(25-k)!}}_{\alpha_1}\mathbf{I}_3 + \underbrace{\sum_{\substack{k=1,\\ k+=3}}^{23} \frac{25!}{k!(25-k)!}}_{\alpha_2}\mathbf{C} + \underbrace{\sum_{\substack{k=2,\\ k+=3}}^{25} \frac{25!}{k!(25-k)!}}_{\alpha_3}\mathbf{C}^2$$

So we know in advance that the final will look like in terms of α_1, α_2, α_3 as follows:

$$\sum_{k=0}^{25} \frac{25!}{k!(25-k)!}\mathbf{C}^k = \begin{bmatrix} \alpha_1 & \alpha_3 & \alpha_2 \\ \alpha_2 & \alpha_1 & \alpha_3 \\ \alpha_3 & \alpha_2 & \alpha_1 \end{bmatrix} \tag{5}$$

α_1, α_2, α_3 can easily be computed e.g. in Python3:

```
from math import factorial as fac
a1 = sum([fac(25)/(fac(k)*fac(25-k)) for k in list(range(0,26,3))])
a2 = sum([fac(25)/(fac(k)*fac(25-k)) for k in list(range(1,26,3))])
a3 = sum([fac(25)/(fac(k)*fac(25-k)) for k in list(range(2,26,3))])
```

We obtain that $\alpha_1 = \alpha_2 = 11184811$ and $\alpha_3 = 11184810$, therefore using Eq. (5) the final answer is

$$\mathbf{M}^{25} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}^{25}$$

$$= \sum_{k=0}^{25} \frac{25!}{k!(25-k)!}\mathbf{C}^k = \begin{bmatrix} 11184811 & 11184810 & 11184811 \\ 11184811 & 11184811 & 11184810 \\ 11184810 & 11184811 & 11184811 \end{bmatrix}$$

References

[1] *Proof of the binomial theorem by mathematical induction*. [Online]. Available: http://amsi.org.au/ESA_Senior_Years/SeniorTopic1/1c/1c_2content_6.html#:~:text=Proof%5C%20of%5C%20the%5C%20binomial%5C%20theorem%5C%20by%5C%20mathematical%5C%20induction,-In%5C%20this%5C%20section&text=(a%5C%2Bb)n%5C%3D,%5C%3D1%5C%20and%5C%20n%5C%3D2.&text=(a%5C%2Bb)k%5C%3D,k%5C%E2%5C%88%5C%921%5C%2Bbk..

A Appendices

A.1 Binomial theorem

THEOREM A.1 (Binomial Theorem)**.** *Let $x, y \in \mathbb{R}$ and $n \geq 0$. Then*

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k, \qquad \binom{n}{k} = \frac{n!}{k!(n-k)!}$$

Proof. It is derived by induction as described in [1]. □

A.2 Elementary row interchange operations

We will observe and derive the elementary row interchange matrices $\mathbf{E}_{ij}$, $i, j \leq 3, i \neq j$ for 3×3 matrices, however they can easily be generalised for $m \times n$. Supposed we are given a

$$\mathbf{M} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

and we want to generate a new matrix $\mathbf{N}$ whose 1st row is $\mathbf{M}$'s 3rd row and is otherwise empty. Consider the row vectors $\mathbf{e}_1 = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}$, $\mathbf{e}_2 = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}$, $\mathbf{e}_3 = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}$. Multiplying $\mathbf{e}_1$ by $\mathbf{M}$ extracts the 1st row of $\mathbf{M}$, and so on for $\mathbf{e}_2$, $\mathbf{e}_3$:

$$\mathbf{e}_1\mathbf{M} = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = \begin{bmatrix} a & b & c \end{bmatrix}, \quad \mathbf{e}_2\mathbf{M} = \begin{bmatrix} d & e & f \end{bmatrix}, \qquad \mathbf{e}_3\mathbf{M} = \begin{bmatrix} g & h & i \end{bmatrix},$$

Therefore if we multiply $\mathbf{E}_{13} := \begin{bmatrix} \mathbf{e}_3 & \mathbf{e}_2 & \mathbf{e}_2 \end{bmatrix}^\top$ with $\mathbf{M}$, the result will consist of $\mathbf{M}$'s 3rd row stacked on top of its 2nd row, stacked on top of the 1st row, i.e. $\mathbf{M}$ with its rows 1 and 3 swapped:

$$\mathbf{E}_{13}\mathbf{M} = \begin{bmatrix} \mathbf{e}_3 \\ \mathbf{e}_2 \\ \mathbf{e}_1 \end{bmatrix} \mathbf{M} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \mathbf{M} = \begin{bmatrix} g & h & i \\ d & e & f \\ a & b & c \end{bmatrix}$$

Similarly, $\mathbf{E}_{23}\mathbf{M}$ swaps rows 2 and 3 of matrix $\mathbf{M}$ and $\mathbf{E}_{23}\mathbf{E}_{13}\mathbf{M}$ swaps row 1 with 3, and then rows 2 with 3:

$$\mathbf{E}_{23}\mathbf{E}_{13}\mathbf{M} = \mathbf{E}_{23} \begin{bmatrix} g & h & i \\ d & e & f \\ a & b & c \end{bmatrix} = \begin{bmatrix} g & h & i \\ a & b & c \\ d & e & f \end{bmatrix}$$

There's a neat property to note about the row interchange matrix $\mathbf{E}_{ij}$; its inverse is itself. That should be intuitively obvious, as multiplying .e.g $\mathbf{E}_{13}$ by $\mathbf{M}$ swaps rows 1, 3 and multiplying again by $\mathbf{E}_{13}$ swaps again rows 1,3, leading back to the original matrix $\mathbf{M}$.

Lemma A.1 (inverse of row interchange matrix)**.** *Let $\boldsymbol{E}_{ij}$ be an elementary row interchange matrix that swaps rows $i, j, i, j \geq m$ of an invertible matrix $\boldsymbol{M} \in \mathbb{R}^{m \times n}$. Then:*

$$\boldsymbol{E}_{ij} = \boldsymbol{E}_{ij}^{-1} \tag{A.1}$$

Proof.

$$\mathbf{E}_{ij}\mathbf{E}_{ij}\mathbf{M} = \mathbf{M} \Rightarrow$$

$$\mathbf{E}_{ij}\mathbf{E}_{ij} = \mathbf{I}_m$$

□

Some other properties that can be readily derived from Eq. (A.2) are:

Lemma A.2 (properties of row interchange matrices)**.** *If $\boldsymbol{E}_{ij}$ is an elementary matrix that interchanges rows i and j, then:*

$$\boldsymbol{E}_{ij}^n = \begin{cases} \boldsymbol{E}_{ij}, & n = 2k, \quad k \in \mathbb{N} \\ \boldsymbol{I}, & n = 2k+1, \quad k \in \mathbb{N} \end{cases} \tag{A.2}$$

Furthermore, $\boldsymbol{E}_{ij}$ is symmetric:

$$\boldsymbol{E}_{ij}^\top = \boldsymbol{E}_{ij} \tag{A.3}$$