Binary Magic Squares

Alain Riou

October 30, 2023

1 Problem formulation

We aim at generating Binary Magic Squares, i.e. random binary matrices $M = (m_{ij})_{i,j} \in \{0,1\}^{n\times n}$ such that the sum of all lines and columns is equal to a constant k, i.e.

$$\forall i \in \{0,\dots,n-1\},\ \sum_{j=0}^{n-1} m_{ij} = \sum_{j=0}^{n-1} m_{ji} = k \tag{1}$$

2 Intuition

The idea is to compute the magic square column by column. To do so, we start from a matrix $M \in \{0,1\}^{n\times n}$ full of zeros and then, we successively pick k indices $(i_0,\dots,i_{k-1})$ per column t and put a 1 in the corresponding cells $(m_{i_l,t})_{0\le l<k}$, while perserving the following constraints on the sum over the lines through time:

$$\forall i \in \{0,\dots,n-1\}, \forall t \in \{0,\dots,n-1\},\ k+t-n \le \sum_{j=0}^{t} m_{ij} \le k \tag{2}$$

If equation (2) holds, in particular for $t = n-1$ the sum of each line i is $\sum_{j=0}^{n-1} m_{ij} = k$. Moreover, since we pick exactly k indices per column, the sum of each column is also k by construction, so M is actually a Binary Magic Square.

3 Algorithm

We now detail how to pick the indices at each step so that equation (2) is satisfied at each time step. The idea is that at each step t we partition the candidate indices into three subsets A_1, A_2 and A_3 depending on whether the sum of the corresponding line is equal to $k+t-n$, equal to k or strictly in between, and pick the right indices accordingly.

Here, E is the set of k indices that is picked at each time step. We can easily prove that at the end of each iteration t, we have

$$s_i = \sum_{j=0}^{t} m_{ij} \tag{3}$$

Algorithm 1: Binary Magic Square generation

```
for i ∈ {0, ..., n − 1} do
    s_i = 0
for t = 0 to n − 1 do
    A_1 := {i ∈ {0, ..., n − 1} | s_i = k + t − n};
    A_2 := {i ∈ {0, ..., n − 1} | k + t − n < s_i < k};
    A_3 := {i ∈ {0, ..., n − 1} | s_i = k};
    E := A_1 ∪ random_subset(A_2, k − |A_1|);
    for i ∈ E do
        m_it := 1;
        s_i := s_i + 1;
```

For each variable x in algorithm 1, define $x(t)$ its t-th value in the algorithm. In particular, the value of s_i may increase only 1 by 1, i.e. for all t

$$s_i(t) \leq s_i(t+1) \leq s_i(t) + 1 \tag{4}$$

4 Correction

4.1 Sum of the lines

We show by induction that equation (2) is verified at each iteration t.

- $\forall i \in \{0, \ldots, n-1\}, s_i(0) = 0$ so equation (2) holds for $t = 0$.
- Assume equation (2) holds for one $0 \leq t < n-1$. We show that it holds as well for $t+1$.

  By hypothesis, for all $i \in \{0, \ldots, n-1\}$, $k + t - n \leq s_i(t) \leq k$, so $(A_1(t), A_2(t), A_3(t))$ is a partition of $\{0, \ldots, n-1\}$.

  Then, for all $i \in \{0, \ldots, n-1\}$,

  - $i \in A_1(t)$
  $$\begin{aligned} i \in A_1(t) &\Rightarrow s_i(t) = k + t - n \text{ and } i \in E(t) \\ &\Rightarrow s_i(t+1) = s_i(t) + 1 = k + t - n + 1 \end{aligned}$$
  - $i \in A_2(t)$
  $$\begin{aligned} i \in A_2(t) &\Rightarrow k + t - n + 1 \leq s_i(t) \leq k - 1 \\ &\Rightarrow k + t - n + 1 \leq s_i(t+1) \leq k \quad \text{by equation (4)} \end{aligned}$$
  - $i \in A_3(t)$
  $$\begin{aligned} i \in A_3(t) &\Rightarrow s_i(t) = k \text{ and } i \notin E(t) \\ &\Rightarrow s_i(t+1) = s_i(t) = k \end{aligned}$$

  So for all $i \in \{0, \ldots, n-1\}$, $k + t - n + 1 \leq s_i(t+1) \leq k$.

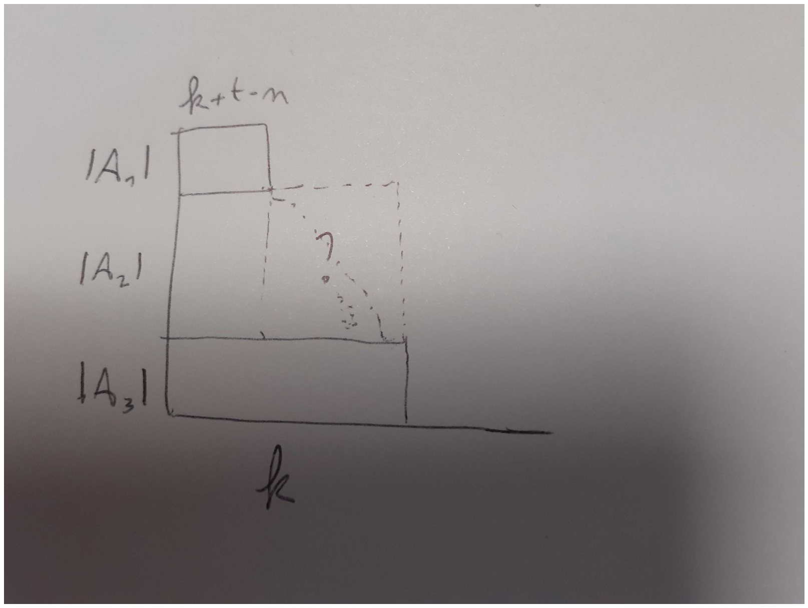

Figure 1: Overview of the algorithm technique. According to equation (2), lines indices are partitioned at each time step t between A_1 (if their sums equal $k+t-n$), A_3 (if their sums equal k) or A_2 (if their sums are strictly between $k+t-n$ and k).

- By induction,

$$\forall t \in \{0,\ldots,n-1\}, \forall i \in \{0,\ldots,n-1\}, k+t-n \leq s_i(t) \leq k \tag{5}$$

  In particular, for all $i \in \{0,\ldots,n-1\}$, $s_i(n-1) = k$, i.e. the sum of each line equals k.

4.2 Sum of the columns

For each row i and column t, $m_{it} = 1$ if $i \in E(t)$ and 0 otherwise, by design. Therefore for all t

$$\sum_{i=0}^{n-1} m_{it} = |E(t)| \tag{6}$$

We show by induction that for all $t \in \{0,\ldots,n-1\}$, $|E(t)| = k$.

- Assume $t = 0$.
  - If $k = 0$, then $A_1(0) = A_2(0) = \emptyset$ so $E(0) = \emptyset$ and $|E(0)| = 0 = k$.
  - If $k = n$, then $k+t-n = 0$, i.e. $A_1(0) = \{0,\ldots,n-1\}$ and $A_2(0) = A_3(0) = \emptyset$. Then $E(0) = A_1(0) = \{0,\ldots,n-1\}$ and $|E(0)| = n = k$.

– If $0 < k < n$, then $A_1(0) = \emptyset$ and $A_2(0) = \{0, \dots, n-1\}$.
  Then, $E(0)$ is a subset of $A_2(0)$ of size $k - A_1(0)$, i.e. a subset of $\{0, \dots, n-1\}$ of size k. So $|E(0)| = k$.

- Assume that exists $0 < t \leq n-1$ such that for all $0 \leq j < t$, $|E(j)| = k$. We show that $|E(t)| = k$ as well.

  At time step t, the total number of ones in M is then $\sum_{j=0}^{t-1} |E(j)| = tk$ (i.e. the sum of the ones in each column).

  Also, the total number of ones in M can be expressed as the sum of the number of ones in each line, i.e. $\sum_{i=0}^{n-1} s_i(t)$.

  As $A_1(t)$, $A_2(t)$ and $A_3(t)$ partition $\{0, \dots, n-1\}$, we have therefore

$$\begin{aligned} tk &= \sum_{i \in A_1(t)} s_i(t) + \sum_{i \in A_2(t)} s_i(t) + \sum_{i \in A_3(t)} s_i(t) \\ &= |A_1(t)|(k+t-n) + \sum_{i \in A_2(t)} s_i(t) + |A_3(t)|k \quad \text{by definition of } A_1 \text{ and } A_3 \end{aligned} \tag{7}$$

  By definition of A_2,

$$|A_2(t)|(k+t-n) < \sum_{i \in A_2(t)} s_i(t) < |A_2(t)|k \tag{8}$$

  So, by injecting (8) into (7),

$$(|A_1(t)| + |A_2(t)|)(k+t-n) + |A_3(t)|k < tk < |A_1(t)|(k+t-n) + (|A_2(t)| + |A_3(t)|)k \tag{9}$$

  As $A_1(t)$, $A_2(t)$ and $A_3(t)$ partition $\{0, \dots, n-1\}$,

$$|A_1(t)| + |A_2(t)| + |A_3(t)| = n \tag{10}$$

  Therefore

$$\begin{aligned} &(n - |A_3(t)|)(k+t-n) + |A_3(t)|k < tk \\ &nk + nt - n^2 - |A_3(t)|k - |A_3(t)|t + |A_3(t)|n + |A_3(t)|k < tk \\ &|A_3(t)|(n-t) < n^2 - nk - nt + tk \\ &|A_3(t)|(n-t) < (n-k)(n-t) \\ &|A_3(t)| < n-k \end{aligned}$$

  and

$$\begin{aligned} &tk < |A_1(t)|(k+t-n) + (n - |A_1(t)|)k \\ &tk < |A_1(t)|k + |A_1(t)|t - |A_1(t)|n + nk - |A_1(t)|k \\ &|A_1(t)|(n-t) < nk - tk \\ &|A_1(t)| < k \end{aligned}$$

  Then, $E(t) = A_1(t) \cup \textbf{random\_subset}(A_2(t), k - |A_1(t)|)$ by definition.

$$\begin{aligned} |A_2(t)| &= n - |A_1(t)| - |A_3(t)| \\ &> k - |A_1(t)| && \text{since } |A_3(t)| < n-k \\ &> 0 && \text{since } |A_1(t)| \end{aligned}$$

so $|\textbf{random\_subset}(A_2(t), k - |A_1(t)|)| = k - |A_1(t)|$ and $|E(t)| = k$ since $A_1(t) \cap A_2(t) = \emptyset$.

- By induction, for all $t \in \{0, \dots, n-1\}$, $|E(t)| = k$.

We proved that the sum of every line and column of a matrix M generated by algorithm 1 is equal to k, and is therefore a Binary Magic Square.

5 Complexity

Without any parallelization trick, the overall complexity of algorithm 1 is

$$C(n) = \mathcal{O}\left(n^2\right) \tag{11}$$

However, all operations inside the **for** loop can be done using vectorized operations in practice. If we have p processes, the overall complexity of algorithm 1 then becomes

$$C(n) = \mathcal{O}\left(n\left\lceil\frac{n}{p}\right\rceil\right) \tag{12}$$

6 Towards non-square Binary Magic Squares

One can extend the definition of Binary Magic Squares to non-square matrices, by defining it as a matrix $M = (m_{ij})_{i,j} \in \{0,1\}^{m \times n}$ such that

$$\begin{cases} \exists a \in \{0, \dots, n\},\ \forall i \in \{0, \dots, m-1\}, & \sum_{j=0}^{n-1} m_{ij} = a \\ \exists b \in \{0, \dots, m\},\ \forall j \in \{0, \dots, n-1\}, & \sum_{i=0}^{m-1} m_{ij} = b \end{cases} \tag{13}$$

However, we show that not all combinations of a, b, m, n can lead to valid magic squares.

Let $M = (m_{ij})_{i,j} \in \{0,1\}^{m \times n}$ be a BMS whose sum of every line (resp. column) is a (resp. b).

$$\sum_{i=0}^{m-1}\sum_{j=0}^{n-1} m_{ij} = \sum_{j=0}^{n-1}\sum_{i=0}^{m-1} m_{ij} \tag{14}$$

which leads to

$$am = bn \tag{15}$$

and in particular $m|bn$ and $n|am$.

Then, unless $a = b = 0$,

- If $m \wedge n = 1$, then $m|b$ by Euclid's lemma and finally $b = m$. Symmetrically, $a = n$.
- If $m|n$, then there exists $k \geq 1$ such that $n = km$.
  Equation (15) then becomes $am = bkm$, which leads to $a = bk$.
- Symmetrically, if $n|m$ then $b = a\frac{m}{n}$.

In other words, if $m \wedge n = 1$ then the only BMS are the trivial ones. Otherwise, the ratio between a and b has to be the same as between m and n.

Algorithm 2: Non-square Binary Magic Square generation

for $i \in \{0, \ldots, m-1\}$ **do**
    $s_i = 0$

for $t = 0$ **to** $n - 1$ **do**
    $A_1 := \{i \in \{0, \ldots, m-1\} \mid s_i = a + t - n\}$;
    $A_2 := \{i \in \{0, \ldots, m-1\} \mid a + t - n < s_i < a\}$;
    $A_3 := \{i \in \{0, \ldots, m-1\} \mid s_i = a\}$;
    $E := A_1 \cup$ **random_subset**$(A_2, b - |A_1|)$;
    for $i \in E$ **do**
        $m_{it} := 1$;
        $s_i := s_i + 1$;

6.1 Sum of the lines

We show by induction that equation (2) is verified at each iteration t.

- $\forall i \in \{0, \ldots, m-1\}, s_i(0) = 0$ so equation (2) holds for $t = 0$.
- Assume equation (2) holds for one $0 \leq t < n - 1$. We show that it holds as well for $t + 1$.

  By hypothesis, for all $i \in \{0, \ldots, m-1\}$, $a + t - n \leq s_i(t) \leq a$, so $(A_1(t), A_2(t), A_3(t))$ is a partition of $\{0, \ldots, m-1\}$.

  Then, for all $i \in \{0, \ldots, m-1\}$,

  - $i \in A_1(t)$
  $$\begin{aligned} i \in A_1(t) &\Rightarrow s_i(t) = a + t - n \text{ and } i \in E(t) \\ &\Rightarrow s_i(t+1) = s_i(t) + 1 = a + t - n + 1 \end{aligned}$$
  - $i \in A_2(t)$
  $$\begin{aligned} i \in A_2(t) &\Rightarrow a + t - n + 1 \leq s_i(t) \leq a - 1 \\ &\Rightarrow a + t - n + 1 \leq s_i(t+1) \leq k \quad \text{by equation (4)} \end{aligned}$$
  - $i \in A_3(t)$
  $$\begin{aligned} i \in A_3(t) &\Rightarrow s_i(t) = a \text{ and } i \notin E(t) \\ &\Rightarrow s_i(t+1) = s_i(t) = a \end{aligned}$$

  So for all $i \in \{0, \ldots, n-1\}$, $a + t - n + 1 \leq s_i(t+1) \leq a$.
- By induction,
$$\forall t \in \{0, \ldots, n-1\}, \forall i \in \{0, \ldots, m-1\},\ a + t - n \leq s_i(t) \leq a \tag{16}$$

In particular, for all $i \in \{0, \ldots, m-1\}$, $s_i(n-1) = a$, i.e. the sum of each line equals a.

6.2 Sum of the columns

For each row i and column t, $m_{it} = 1$ if $i \in E(t)$ and 0 otherwise, by design. Therefore for all t

$$\sum_{i=0}^{n-1} m_{it} = |E(t)| \tag{17}$$

We show by induction that for all $t \in \{0, \ldots, n-1\}$, $|E(t)| = b$.

- Assume $t = 0$.
  - If $b = 0$, then $a = 0$.
    Consequently, $A_1(0) = A_2(0) = \emptyset$ so $E(0) = \emptyset$ and $|E(0)| = 0 = b$.
  - If $b = m$, then $a = n$, and $a+t-n = 0$, i.e. $A_1(0) = \{0, \ldots, m-1\}$ and $A_2(0) = A_3(0) = \emptyset$.
    Then $E(0) = A_1(0) = \{0, \ldots, m-1\}$ and $|E(0)| = m = b$.
  - If $0 < b < m$, then $0 < a < n$.
    Consequently, $A_1(0) = \emptyset$ and $A_2(0) = \{0, \ldots, m-1\}$.
    Then, $E(0)$ is a subset of $A_2(0)$ of size $b - |A_1(0)|$, i.e. a subset of $\{0, \ldots, m-1\}$ of size b. So $|E(0)| = b$.

- Assume that exists $0 < t \leq n-1$ such that for all $0 \leq j < t$, $|E(j)| = b$. We show that $|E(t)| = b$ as well.

  At time step t, the total number of ones in M is then $\sum_{j=0}^{t-1} |E(j)| = tb$ (i.e. the sum of the ones in each column).

  Also, the total number of ones in M can be expressed as the sum of the number of ones in each line, i.e. $\sum_{i=0}^{n-1} s_i(t)$.
  As $A_1(t)$, $A_2(t)$ and $A_3(t)$ partition $\{0, \ldots, m-1\}$, we have therefore

$$\begin{aligned} tb &= \sum_{i \in A_1(t)} s_i(t) + \sum_{i \in A_2(t)} s_i(t) + \sum_{i \in A_3(t)} s_i(t) \\ &= |A_1(t)|(a+t-n) + \sum_{i \in A_2(t)} s_i(t) + |A_3(t)|a \quad \text{by definition of } A_1 \text{ and } A_3 \end{aligned} \tag{18}$$

  By definition of A_2,

$$|A_2(t)|(a+t-n) < \sum_{i \in A_2(t)} s_i(t) < |A_2(t)|a \tag{19}$$

  So, by injecting (19) into (18),

$$(|A_1(t)| + |A_2(t)|)(a+t-n) + |A_3(t)|a < tb < |A_1(t)|(a+t-n) + (|A_2(t)| + |A_3(t)|)a \tag{20}$$

  As $A_1(t)$, $A_2(t)$ and $A_3(t)$ partition $\{0, \ldots, m-1\}$,

$$|A_1(t)| + |A_2(t)| + |A_3(t)| = m \tag{21}$$

Therefore

$$\begin{aligned}
&(m - |A_3(t)|)(a + t - n) + |A_3(t)|a < tb \\
&am + mt - mn - |A_3(t)|a - |A_3(t)|t + |A_3(t)|n + |A_3(t)|a < tb \\
&|A_3(t)|(n - t) < mn - am - mt + tb \\
&|A_3(t)|(n - t) < mn - bn - mt + tb && \text{by (15)} \\
&|A_3(t)|(n - t) < (m - b)(n - t) \\
&|A_3(t)| < m - b
\end{aligned}$$

and

$$\begin{aligned}
&tb < |A_1(t)|(a + t - n) + (m - |A_1(t)|)a \\
&tb < |A_1(t)|a + |A_1(t)|t - |A_1(t)|n + am - |A_1(t)|a \\
&|A_1(t)|(n - t) < am - tb \\
&|A_1(t)|(n - t) < bn - tb && \text{by (15)} \\
&|A_1(t)| < b
\end{aligned}$$

Then, $E(t) = A_1(t) \cup \textbf{random\_subset}(A_2(t), b - |A_1(t)|)$ by definition.

$$\begin{aligned}
|A_2(t)| &= m - |A_1(t)| - |A_3(t)| \\
&> b - |A_1(t)| && \text{since } |A_3(t)| < m - b \\
&> 0 && \text{since } |A_1(t)| < b
\end{aligned}$$

so $|\textbf{random\_subset}(A_2(t), b - |A_1(t)|)| = b - |A_1(t)|$ and $|E(t)| = b$ since $A_1(t) \cap A_2(t) = \emptyset$.

- By induction, for all $t \in \{0, \ldots, n-1\}$, $|E(t)| = b$.

We proved that the sum of every line (resp. column) of a matrix M generated by algorithm 2 is equal to a (resp. b), and is therefore a Binary Magic Square.