Polytope of Matrices with Invariant Row- and Column- Sums

Pramod Mathai *

February 26, 2019

In this note, we answer the following two questions and discuss their relevance to a hidden variable theory (HVT) of Aaronson's (the discrete counterpart of Schrödinger's HVT) [1]:
(1) *Construction*: Given $\mathbf{r}, \mathbf{c} \in \mathbb{R}_+^{n\times 1}$, with $\mathbf{1}^{\mathrm{T}}\mathbf{r} = \mathbf{1}^{\mathrm{T}}\mathbf{c} = \gamma$, how does one construct $\mathbf{P} \in \mathbb{R}_+^{n\times n}$ with row and column sums satisfying $\mathbf{P1} = \mathbf{r}, \mathbf{P}^{\mathrm{T}}\mathbf{1} = \mathbf{c}$? Here, $\mathbf{1}$ is the vector of all ones. The case $\gamma = 1$ corresponds to requiring that $\mathbf{P}$ be a transition matrix - the instance specific to Aaronson's HVT.
(2) *Robustness*: How robust are the entries of $\mathbf{P}$ to perturbations in $\mathbf{r}$ and $\mathbf{c}$? The perturbed vectors $\mathbf{r}'$ and $\mathbf{c}'$ are subject to the same constraints as $\mathbf{r}$ and $\mathbf{c}$ in (1).

Construction
$\mathcal{S}_{(\mathbf{r},\mathbf{c})}$, the set of of all solutions to (1), is convex. We construct $\mathcal{V}_{(\mathbf{r},\mathbf{c})}$, the set of vertices of a polytope $\mathcal{P}_{(\mathbf{r},\mathbf{c})} \subseteq \mathcal{S}_{(\mathbf{r},\mathbf{c})}$.

Without loss of generality, we assume that the elements of $\mathbf{r}$ and $\mathbf{c}$ are arranged in increasing order: $r_i \leq r_{i+1}$ for $i \in [n-1]$, likewise for $\mathbf{c}$ ($[k]$ denotes the set of positive integers not exceeding k). If one solves for this case, then the rows and columns of $\mathbf{P}$ can always be permuted back to their original order without changing the values of the respective sums.

Since $\mathbf{r}$ and $\mathbf{c}$ have a finite number of elements, they can be ordered simultaneously [2]. For example, when $n = 3$, such an ordering could be $r_1 \leq c_1 \leq c_2 \leq r_2 \leq r_3 \leq c_3$.

We define $m_i := \min(r_i, c_i)$ and $\mathbf{e}_i$ as the column vector with 1 in the i^{th} position and zeros in the other $(n-1)$ positions.

The polytope $\mathcal{P}_{(\mathbf{r},\mathbf{c})}$ is then defined as the convex hull of the set

$$\mathcal{V}_{(\mathbf{r},\mathbf{c})} = \mathbf{P}^* \bigcup_{\substack{i,j\in[n],\\ i<j}} \mathbf{P}_{(i,j)} \tag{❖}$$

of, in general, $1 + \frac{n(n-1)}{2}$ distinct matrices where $\mathbf{P}_{(i,j)} = \mathbf{P}^* + m_1 \cdot (\mathbf{e}_i\mathbf{e}_j^{\mathrm{T}} + \mathbf{e}_j\mathbf{e}_i^{\mathrm{T}} - \mathbf{e}_i\mathbf{e}_i^{\mathrm{T}} - \mathbf{e}_j\mathbf{e}_j^{\mathrm{T}})$ and $\mathbf{P}^*$ has entries $p^*_{i,j}$ equal to zero except for the following elements:
$p^*_{i,i} = m_i,\ \ p^*_{n,i} = c_i - m_i,\ \ p^*_{i,n} = r_i - m_i$ for $i \in [n-1]$ and
$p^*_{n,n} = r_n + c_n - \gamma + \sum_{i=1}^{n-1} m_i$.

For example, in the case $n = 3$ with ordering $r_1 \leq c_1 \leq c_2 \leq r_2 \leq r_3 \leq c_3$, every element in the polytope defined by these four matrices will be a solution to (1):

*pramod.m@gmail.com
[1]See p.183 of "Quantum Computing since Democritus" by Scott Aaronson, *Cambridge University Press*, 2013.
[2]without invoking the Axiom of Choice, which may be necessary to transfer the present approach to a continuous setting.

$$\mathbf{P}^* = \begin{bmatrix} r_1 & 0 & 0 \\ 0 & c_2 & r_2 - c_2 \\ c_1 - r_1 & 0 & r_3 + c_3 - \gamma + (r_1 + c_2) \end{bmatrix}, \mathbf{P}_{(1,2)} = \begin{bmatrix} 0 & r_1 & 0 \\ r_1 & c_2 - r_1 & r_2 - c_2 \\ c_1 - r_1 & 0 & r_3 + c_3 - \gamma + (r_1 + c_2) \end{bmatrix},$$

$$\mathbf{P}_{(1,3)} = \begin{bmatrix} 0 & 0 & r_1 \\ 0 & c_2 & r_2 - c_2 \\ c_1 & 0 & r_3 + c_3 - \gamma + c_2 \end{bmatrix}, \mathbf{P}_{(2,3)} = \begin{bmatrix} r_1 & 0 & 0 \\ 0 & c_2 - r_1 & r_2 - c_2 + r_1 \\ c_1 - r_1 & r_1 & r_3 + c_3 - \gamma + c_2 \end{bmatrix}.$$

The construction (❖) makes it clear that any $\mathbf{P}_{(i,j)}$ has the same row and columns sums as $\mathbf{P}^*$. So it suffices to show that $\mathbf{P}^*$ satisfies (1) and that $p^*_{i,i} \geq m_1$ for $i \in [n]$; the additional condition ensures that the diagonal elements of $\mathbf{P}_{(i,j)}$ are nonnegative. We prove these properties in order:

- *Row sums*: Denote $\mathbf{P}^*_{i,\cdot}$ to be the i^{th} row of $\mathbf{P}^*$. For $i \in [n-1]$, $\mathbf{P}^*_{i,\cdot}\mathbf{1} = m_i + r_i - m_i = r_i$. $\mathbf{P}^*_{n,\cdot}\mathbf{1} = \sum_{i=1}^{n-1}(c_i - m_i) + p^*_{n,n} = r_n + (\sum_{i=1}^{n} c_i) - \gamma = r_n$.
- *Column sums*: The calculation is analogous to the one for row sums.
- *Nonnegativity of off-diagonal terms; diagonal terms* $\geq m_1$: Readily follows from (❖) for all elements except $p^*_{n,n}$.
  Requiring $p^*_{n,n} \geq m_1$ is equivalent to requiring $r_n + c_n + \sum_{i=2}^{n-1} m_i \geq \gamma$. By the inequality of arithmetic and geometric means, the left hand side is minimized when each of the n terms is equal, say, to M. The orderings $r_i \leq r_{i+1}$ and $c_i \leq c_{i+1}$ then imply that $r_i = c_i = M$ for $2 \leq i \leq n$, $M \geq m_1$ and thus $(n-1)M + m_1 = \gamma$. $p^*_{n,n} \geq m_1$ follows.

This shows that all matrices in $\mathcal{P}_{(\mathbf{r},\mathbf{c})}$ are solutions to (1) for any valid choice of $\mathbf{r}, \mathbf{c}$. But $\mathcal{P}_{(\mathbf{r},\mathbf{c})}$ does not necessarily contain all solutions to (1), i.e., $\mathcal{P}_{(\mathbf{r},\mathbf{c})} \subseteq \mathcal{S}_{(\mathbf{r},\mathbf{c})}$; indeed, the construction shows that any matrix in $\mathcal{P}_{(\mathbf{r},\mathbf{c})}$ has a symmetric principal minor of order $n-1$. It is unclear if $\mathcal{P}_{(\mathbf{r},\mathbf{c})}$ is the largest polytope in $\mathcal{S}_{(\mathbf{r},\mathbf{c})}$.

Robustness

A solution $\mathbf{P} \in \mathcal{P}_{(\mathbf{r},\mathbf{c})}$ is not robust, in the infinity norm, to perturbations in $\mathbf{r}, \mathbf{c}$ simply because the diameter of the polytope will, in general, be positive. The perturbed vectors $\mathbf{r}', \mathbf{c}'$ may satisfy $0 \leq \|\mathbf{r} - \mathbf{r}'\|_\infty, \|\mathbf{c} - \mathbf{c}'\|_\infty \leq \delta$ for an arbitrarily small δ. But the construction (❖) shows the existence of solutions $\mathbf{P}$ (to $\mathbf{r}, \mathbf{c}$) and $\mathbf{P}'$ (to $\mathbf{r}', \mathbf{c}'$) that will satisfy $\|\mathbf{P} - \mathbf{P}'\|_\infty \geq \max(m_1, m_1')$, a quantity that is independent of δ. Aaronson described a dynamical process (row/column scaling) to converge onto a solution $\mathbf{P}$. When applied to find $\mathbf{P}'$ this process might converge to a solution differing - in the infinity norm - by $\max(m_1, m_1')$ from $\mathbf{P}$ depending on the choice of the initial point in the dynamical process.

However, if the HVT (now, $\gamma = 1$) is modified to reward a unique solution in $\mathcal{S}_{(\mathbf{r},\mathbf{c})}$ then it is conceivable that this solution is robust. For example, if the HVT asks for the sparsest solution then $\mathbf{P} = \mathbf{P}^*$ is, arguably, the unique choice in $\mathcal{S}_{(\mathbf{r},\mathbf{c})}$ and not just in $\mathcal{P}_{(\mathbf{r},\mathbf{c})}$ [3]. The construction (❖) would then imply that $\|\mathbf{P} - \mathbf{P}'\|_\infty \sim O(\delta n)$. Additionally, the Fannes–Audenaert inequality, applied to the diagonal density matrices in $\mathbb{R}^{n^2 \times n^2}$ with entries from $\mathbf{P}$ and $\mathbf{P}'$, can then be used to conclude that, for small enough δ [4], the entropy difference between those two density matrices is $O(\delta n \log(n))$.

[3] $\mathbf{P}^*$ has at most $2n-1$ nonzero entries and any solution in $\mathcal{S}_{(\mathbf{r},\mathbf{c})}$ must in general have at least as many; but it is not clear if $\mathbf{P}^*$ is uniquely sparsest in $\mathcal{S}_{(\mathbf{r},\mathbf{c})}$.

[4] The $O(\delta n)$ difference should be less than 1 in order to be accommodated in Audenaert's binary-entropy term.