$$U \; n\times n \text{ unitary}, \; [U^T\mathbf{e}_1]_i \neq 0 \;\forall\, i,$$
$$\mathcal{L}(\mathbf{z}) = U d(U^T\mathbf{z}) d(U^T\mathbf{e}_1)^{-1} U^H, \; \mathbf{z}\in\mathbb{C}^n \; (\,\mathbf{e}_1^T\mathcal{L}(\mathbf{z}) = \mathbf{z}^T\,),$$
$$\mathcal{L}(\mathbf{x}) = \mathcal{L}(\mathbf{z})^2,$$
$$U^T\mathbf{x} = d(U^T\mathbf{e}_1)^{-1} d(U^T\mathbf{z}) U^T\mathbf{z},$$
$$\mathbf{x}^T = \mathbf{z}^T\mathcal{L}(\mathbf{z}) \text{ (Gianluca)},$$
$$\mathcal{L} = \{\mathcal{L}(\mathbf{z}) : \mathbf{z}\in\mathbb{C}^n\} \text{ is a commutative matrix algebra},$$
$$\mathcal{L}(\mathbf{x})\mathcal{L}(\mathbf{y}) = \mathcal{L}(\mathcal{L}(\mathbf{x})^T\mathbf{y}) = \mathcal{L}(\mathbf{y})\mathcal{L}(\mathbf{x}) = \mathcal{L}(\mathcal{L}(\mathbf{y})^T\mathbf{x}).$$

Given $\mathbf{z}^T$, the first row of $\mathcal{L}(\mathbf{z})$, compute the first row of $\mathcal{L}(\mathbf{z})^2$, $\mathcal{L}(\mathbf{z})^4$, $\mathcal{L}(\mathbf{z})^8$, ..., $\mathcal{L}(\mathbf{z})^{2^k}$. Cost = one U^T transform $+kn$ a.o.+ one U transform.

Given $\mathbf{z}$, the first row of $\mathcal{L}(\mathbf{z})$, the eigenvalues λ of $\mathcal{L}(\mathbf{z})$ can be computed by performing a U^T transform, and the eigenvalues of $\mathcal{L}(\mathbf{z})^s$ are simply λ^s.

Circulant, τ, η and μ matrix algebras are of type $\mathcal{L}$.

Given A $n\times n$ and its first row, say $[z_1 \; z_2 \; \cdots \; z_{n-1} \; z_n]$, one can show that $A\in\mathcal{L}$, $\mathcal{L}=\tau,\eta,\mu$, iff

$$a_{i,j-1} + a_{i,j+1} = a_{i-1,j} + a_{i+1,j}, \; 1\leq i,j\leq n,$$

where, for $s=1,\ldots,n$,

$$a_{s,0} = a_{0,s} = a_{s,n+1} = a_{n+1,s} = 0, \text{ if } \mathcal{L}=\tau,$$
$$a_{s,0} = a_{0,s} = a_{1,n-s+1}, \; a_{s,n+1} = a_{n+1,s} = a_{1,s}, \text{ if } \mathcal{L}=\eta,$$
$$a_{s,0} = a_{0,s} = -a_{1,n-s+1}, \; a_{s,n+1} = a_{n+1,s} = -a_{1,s}, \text{ if } \mathcal{L}=\mu.$$

Examples. Let us write the 4×4 τ, η, μ matrices with first row [0 1 0 1]:

$$\begin{bmatrix} 0&1&0&1\\ 1&0&2&0\\ 0&2&0&1\\ 1&0&1&0 \end{bmatrix}, \begin{bmatrix} 0&1&0&1\\ 1&0&1&0\\ 0&1&0&1\\ 1&0&1&0 \end{bmatrix}, \begin{bmatrix} 0&1&0&1\\ 1&0&3&0\\ 0&3&0&1\\ 1&0&1&0 \end{bmatrix},$$

Let us write the 5×5 τ, η, μ matrices with first row [0 1 0 0 1]:

$$\begin{bmatrix} 0&1&0&0&1\\ 1&0&1&1&0\\ 0&1&1&1&0\\ 0&1&1&0&1\\ 1&0&0&1&0 \end{bmatrix}, \begin{bmatrix} 0&1&0&0&1\\ 1&0&1&0&0\\ 0&1&0&1&0\\ 0&0&1&0&1\\ 1&0&0&1&0 \end{bmatrix}, \begin{bmatrix} 0&1&0&0&1\\ 1&0&1&2&0\\ 0&1&2&1&0\\ 0&2&1&0&1\\ 1&0&0&1&0 \end{bmatrix}.$$

Stochastic by columns $\cap$ *matrix algebras*

3×3 stochastic by columns (symmetric, symmetric and persymmetric) matrices:

$$\mathcal{M} = \begin{bmatrix} a & c & e\\ b & d & f\\ 1-a-b & 1-c-d & 1-e-f \end{bmatrix}$$

$$(\mathcal{M}^S = \begin{bmatrix} a & b & 1-a-b\\ b & c & 1-b-c\\ 1-a-b & 1-b-c & -1+a+2b+c \end{bmatrix}, \mathcal{M}^{SP} = \begin{bmatrix} a & b & 1-a-b\\ b & 1-2b & b\\ 1-a-b & b & a \end{bmatrix})$$

$$p_{\mathcal{M}^S}(\lambda) = (1-\lambda)[\lambda^2 - 2\lambda(a+b+c-1) + 3ac - a - 3b^2 + 2b - c],$$
$$eig \;=\; a+b+c-1 \pm \sqrt{a^2+4b^2+c^2+2ab+2bc-ac-a-4b-c+1}$$

3×3 stochastic by columns circulant (symmetric) matrices:

$$\mathcal{C}\cap\mathcal{M} = \begin{bmatrix} a & 1-a-b & b\\ b & a & 1-a-b\\ 1-a-b & b & a \end{bmatrix} \; \left(\mathcal{C}\cap\mathcal{M}^S = \begin{bmatrix} a & \frac{1-a}{2} & \frac{1-a}{2}\\ \frac{1-a}{2} & a & \frac{1-a}{2}\\ \frac{1-a}{2} & \frac{1-a}{2} & a \end{bmatrix} \right)$$

$$\begin{array}{rcl} p_{\mathcal{C}\cap\mathcal{M}}(\lambda) & = & (1-\lambda)[\lambda^2-\lambda(3a-1)+b^3+a^3-3ab(1-a-b)+(1-a-b)^3] \\ & = & (1-\lambda)[\lambda^2-\lambda(3a-1)+3(a^2+b^2-a-b+ab)+1] \end{array}$$

$$b=1-a-b \;\Rightarrow\; p_{\mathcal{C}\cap\mathcal{M}^S}(\lambda)=(1-\lambda)[\lambda^2-\lambda(3a-1)+(\frac{3a-1}{2})^2]=(1-\lambda)(\lambda-\frac{3a-1}{2})^2$$

3×3 stochastic by columns tau matrices:

$$\tau\cap\mathcal{M}=\begin{bmatrix} a & 0 & 1-a \\ 0 & 1 & 0 \\ 1-a & 0 & a \end{bmatrix},\; p_{\tau\cap\mathcal{M}}(\lambda)=(1-\lambda)[(2a-1-\lambda)(1-\lambda)]$$

3×3 stochastic by columns eta matrices:

$$\eta\cap\mathcal{M}=\begin{bmatrix} a & b & 1-a-b \\ b & 1-2b & b \\ 1-a-b & b & a \end{bmatrix}=\mathcal{M}^{SP}\;!\;\; p_{\eta\cap\mathcal{M}}(\lambda)=(1-\lambda)[\lambda^2\;\ldots]$$

3×3 stochastic by columns mu matrices:

$$\mu\cap\mathcal{M}=\begin{bmatrix} a & 0 & 1-a \\ 0 & 1 & 0 \\ 1-a & 0 & a \end{bmatrix},\; p_{\mu\cap\mathcal{M}}(\lambda)=(1-\lambda)[(2a-1-\lambda)(1-\lambda)]$$

4×4 stochastic by columns symmetric and persymmetric matrices:

$$\mathcal{M}^{SP}=\begin{bmatrix} a & b & c & 1-a-b-c \\ b & d & 1-b-c-d & c \\ c & 1-b-c-d & d & b \\ 1-a-b-c & c & b & a \end{bmatrix}$$

4×4 stochastic by columns τ matrices:

$$\mathcal{M}\cap\tau=\begin{bmatrix} a & b & -b & 1-a \\ b & a-b & 1-a+b & -b \\ -b & 1-a+b & a-b & b \\ 1-a & -b & b & a \end{bmatrix},\;(\mathcal{M}\cap\tau\geq 0)=\begin{bmatrix} a & 0 & 0 & 1-a \\ 0 & a & 1-a & 0 \\ 0 & 1-a & a & 0 \\ 1-a & 0 & 0 & a \end{bmatrix},\lambda=1,1,2a-1,2a-1$$

$$\mathcal{M}\cap\eta=\begin{bmatrix} a & b & c & 1-a-b-c \\ b & a & 1-b-c-a & c \\ c & 1-b-c-a & a & b \\ 1-a-b-c & c & b & a \end{bmatrix},\;(\mathcal{M}\cap\eta\geq 0)=\begin{bmatrix} \; \end{bmatrix}\;\lambda=$$

$$\mathcal{M}\cap\mu=\begin{bmatrix} a & \frac{a-d}{2} & \frac{d-a}{2} & 1-a \\ \frac{a-d}{2} & d & 1-d & \frac{d-a}{2} \\ \frac{d-a}{2} & 1-d & d & \frac{a-d}{2} \\ 1-a & \frac{d-a}{2} & \frac{a-d}{2} & a \end{bmatrix},\;(\mathcal{M}\cap\mu\geq 0)=\begin{bmatrix} \; \end{bmatrix}\;\lambda=$$

2×2 stochastic by columns matrices:

$$\mathcal{M}=\begin{bmatrix} a & b \\ 1-a & 1-b \end{bmatrix},\;\lambda\;=\;1,\,a-b$$

2×2 stochastic by columns circulant matrices:

$$\mathcal{C}\cap\mathcal{M}=\begin{bmatrix} c & 1-c \\ 1-c & c \end{bmatrix},\;\lambda\;=\;1,\,2c-1$$

Assuming $a, b, c \in \mathbb{R}$, the Frobenius norm of $A - X$, $A \in \mathcal{M}$, X varying in $\mathcal{M} \cap \mathcal{C}$, is minimum for $c = \frac{a+1-b}{2}$, i.e. when A and X have the eigenvalues different from 1 equal ($a - b = 2c - 1$).

Fixed $A \in \mathcal{M}$, where $\mathcal{M}$ is the set of all non negative stochastic by columns $n \times n$ matrices, such that $|\lambda_2(A)| < 1$, there exist $X \in \mathcal{C} \cap \mathcal{M}$ such that $|\lambda_2(A)| \leq |\lambda_2(X)| < 1$?

Note that the eigenvalues of X are easily computable.

6×6 stochastic by columns symmetric and persymmetric matrices:

$$\mathcal{M}^{SP} = \begin{bmatrix} a & b & c & d & e & \begin{matrix}1-a-b\\-c-d-e\end{matrix} \\ b & f & g & h & \begin{matrix}1-b-f\\-g-h-e\end{matrix} & e \\ c & g & i & \begin{matrix}1-c-g\\-i-h-d\end{matrix} & h & d \\ d & h & \begin{matrix}1-c-g\\-i-h-d\end{matrix} & i & g & c \\ e & \begin{matrix}1-b-f\\-g-h-e\end{matrix} & h & g & f & b \\ \begin{matrix}1-a-b\\-c-d-e\end{matrix} & e & d & c & b & a \end{bmatrix}$$

6×6 stochastic by columns η matrices:

$$\mathcal{M} \cap \eta = \begin{bmatrix} a & b & c & d & e & \begin{matrix}1-a-b\\-c-d-e\end{matrix} \\ b & a+c-e & b & e & \begin{matrix}-2b-a\\-c-e+1\end{matrix} & e \\ c & b & a & \begin{matrix}-a-b-c\\-d-e+1\end{matrix} & e & d \\ d & e & \begin{matrix}-a-b-c\\-d-e+1\end{matrix} & a & b & c \\ e & \begin{matrix}-2b-a\\-c-e+1\end{matrix} & e & b & a+c-e & b \\ \begin{matrix}1-a-b\\-c-d-e\end{matrix} & e & d & c & b & a \end{bmatrix}$$

- 3×3 singular stochastic by columns matrices:

$$\begin{bmatrix} a & a & a \\ b & b & b \\ 1-a-b & 1-a-b & 1-a-b \end{bmatrix}^n = \begin{bmatrix} a & a & a \\ b & b & b \\ 1-a-b & 1-a-b & 1-a-b \end{bmatrix}$$

$$\begin{bmatrix} a & a & c \\ b & b & d \\ 1-a-b & 1-a-b & 1-c-d \end{bmatrix}^n =?$$

- 3×3 singular stochastic by columns symmetric and persymmetric matrices:

$$A = \begin{bmatrix} a & 1-2a & a \\ 1-2a & -1+4a & 1-2a \\ a & 1-2a & a \end{bmatrix} \in \eta! \;\; \lambda = 0,\, 1,\, \in \mathbb{R}$$

A has rank 1 iff $a = \frac{1}{3}$ (in such case $\lambda = 0, 1, 0$). $A \geq 0$ iff $\frac{1}{4} \leq a \leq \frac{1}{2}$. A is semi positive definite iff $a \geq \frac{1}{3}$. $A \in \tau$ iff $a = \frac{1}{2}$ (in such case $\lambda = 1, 0, 1$).

$$A^n = \begin{bmatrix} a_n & b_n & a_n \\ & & \\ & & \end{bmatrix}, \quad \begin{bmatrix} a_n \\ b_n \end{bmatrix} = \begin{bmatrix} 2a & 1-2a \\ 2(1-2a) & 4a-1 \end{bmatrix} \begin{bmatrix} a_{n-1} \\ b_{n-1} \end{bmatrix}$$

EXAMPLE. Let A be the following $n \times n$ stochastic by columns matrix

$$A = \begin{bmatrix} 0 & b_1 & & & & & \\ 1 & 0 & b_2 & & & & \\ & 1-b_1 & 0 & & & & \\ & & 1-b_2 & & & & \\ & & & & b_{n-2} & & \\ & & & & 0 & 1 & \\ & & & & 1-b_{n-2} & 0 & \end{bmatrix}, \ b_i \in [0,1].$$

Note that the eigenvalues of A are real (even if A is not hermitian), and in the interval $[-1,1]$. They are distinct if $b_i \in (0,1)$. Obviously, 1 is eigenvalue. Moreover, also -1 is eigenvalue (prove it!). The remaining eigenvalues are not known (for generic values of the b_j).

Let $\mathcal{C}$ be the space of $n \times n$ circulant matrices. Let us compute $\mathcal{C}_A$, the minimizer of $\|A - X\|_F$, $X \in \mathcal{C}$, with the aim to compare its eigenvalues with those of A. Let $\{J_1, J_2, \ldots, J_n\}$ be a basis of $\mathcal{C}$. Then $\mathcal{C}_A = \sum_{k=1}^n \alpha_k J_k$, where $B\alpha = \mathbf{c}$, $B_{rs} = (J_r, J_s)$, $c_r = (J_r, A)$, $1 \le r, s \le n$. If $J_k = J_2^{k-1}$ where

$$J_2 = \begin{bmatrix} 0 & 1 & & \\ & & \ddots & \\ & & & 1 \\ 1 & & & 0 \end{bmatrix}$$

then $(J_r, J_s) = n\delta_{rs}$, $(J_1, A) = (J_r, A) = 0$, $r = 3, \ldots, n-1$, and $(J_2, A) = 1 + \sum b_j$, $(J_n, A) = n - 1 - \sum b_j$. Thus

$$\mathcal{C}_A = \begin{bmatrix} & p & & & q \\ q & & p & & \\ & q & & \ddots & \\ & & \ddots & & p \\ p & & & q & \end{bmatrix}, \ p = \frac{1 + \sum b_j}{n}, \ q = \frac{n - 1 - \sum b_j}{n} = 1 - p.$$

The eigenvalues of $\mathcal{C}_A$ can be easily computed. In fact, recalling that

$$\mathcal{C}_A = F d(F \mathcal{C}_A^T \mathbf{e}_1) d(F\mathbf{e}_1)^{-1} F^H, \quad [F]_{ij} = \frac{1}{\sqrt{n}} \omega_n^{(i-1)(j-1)}, \quad 1 \le i, j \le n, \quad \omega_n = e^{\mathbf{i}\frac{2\pi}{n}},$$

first write the vector $\sqrt{n} F \mathcal{C}_A^T \mathbf{e}_1$:

$$\sqrt{n} F \begin{bmatrix} 0 \\ p \\ 0 \\ \cdot \\ 0 \\ q \end{bmatrix} = \sqrt{n}(pF\mathbf{e}_2 + qF\mathbf{e}_n) = \sqrt{n}(pF + qF^H)\mathbf{e}_2, \ F = F^H Q, \ Q = \begin{bmatrix} 1 & & & \\ & & & 1 \\ & & \cdot^{\cdot^{\cdot}} & \\ & 1 & & \end{bmatrix},$$

and then observe that the eigenvalues of $\mathcal{C}_A$ are its entries:

$$p\omega_n^{i-1} + (1-p)\overline{\omega}_n^{i-1} = p\omega_n^{i-1} + (1-p)\omega_n^{n-i+1}, \quad i = 1, \ldots, n.$$

Note that $\frac{1}{n} \le p \le \frac{n-1}{n}$, and that it is sufficient to study the eigenvalues of $\mathcal{C}_A$ for $\frac{1}{2} \le p \le \frac{n-1}{n}$.

The case $p = \frac{n-1}{n}$ ($b_j = 1\ \forall\, j$).
In this case the eigenvalues of A are obviously known, they are -1, 0 with algebraic multiplicity $n-2$, and 1. The eigenvalues of $\mathcal{C}_A$ are

$$\frac{n-1}{n}\omega_n^{i-1} + \frac{1}{n}\overline{\omega}_n^{i-1}, \ \ i = 1, \dots, n.$$

(draw them!). They are all inside the set $\{z : |z| \le 1\}$, except 1 ($i = 1$) and, for even n, -1 ($i - 1 = \frac{n}{2}$).

The case $p = \frac{1}{2}$ ($\sum b_j + 1 = \frac{n}{2}$).
In this case the eigenvalues of A are not known (?, perhaps are known if $b_j = \frac{1}{2}\ \forall\, j$, and in other particular cases). The eigenvalues of $\mathcal{C}_A$ are $\Re(\omega_n^{i-1}) = \cos\frac{2\pi(i-1)}{n}$, $i = 1, \dots, n$. They are all inside the set $[-1, 1]$, except 1 ($i = 1$) and, for even n, -1 ($i - 1 = \frac{n}{2}$). ...

RESULT. Let $A \in \mathbb{C}^{n\times n}$ be a stochastic by columns (or by rows) $n \times n$ matrix. So, 1 is eigenvalue of A. Let U be a unitary matrix such that $U\mathbf{e}_i = \frac{1}{\sqrt{n}}\mathbf{e}e^{\mathbf{i}\theta}$, for some i and θ ($i = 1, \theta = 0$ if $U = F$), and set $\mathcal{L} = \{Ud(\mathbf{z})U^H : \mathbf{z} \in \mathbb{C}^n\}$. Note that $\mathcal{L}$ is a n-dimensional subspace of $\mathbb{C}^{n\times n}$, i.e. there exist $J_k \in \mathcal{L}$, $k = 1, \dots, n$, linearly independent such that $\mathcal{L} = \mathrm{Span}\{J_k\}$. Let $\mathcal{L}_A$ be the minimizer of $\|A - X\|_F$ in $\mathcal{L}$,

$$\begin{array}{l} \mathcal{L}_A = U\,\mathrm{diag}\,((U^HAU)_{jj})U^H = Ud(U^T\mathbf{z}_A^r)d(U^T\mathbf{v})^{-1}U^H, \\ \mathcal{L}_A = \sum_{k=1}^n \alpha_k J_k,\ \alpha = B^{-1}\mathbf{c},\ B_{rs} = (J_r, J_s),\ c_r = (J_r, A) \end{array}$$

where $\mathbf{v}$ is chosen such that $(U^T\mathbf{v})_j \ne 0\ \forall\, j$ ($\mathbf{v} = \mathbf{e}_1$ if $U = F$).

Then $\mathcal{L}_A$ is stochastic by columns and by rows (SEE the second Theorem in the next pages). In particular, 1 is eigenvalues of $\mathcal{L}_A$. All eigenvalues of $\mathcal{L}_A$ are particular points of the convex set $\{\frac{\mathbf{z}^HA\mathbf{z}}{\mathbf{z}^H\mathbf{z}} : \mathbf{z} \in \mathbb{C}^n\}$. So, when A is normal (hermitian) they are in the minimum polygon (real interval) containing the eigenvalues of A. When alternatively $A \ge 0$ (?$A^k \ge 0$ for some k?) they are in the set $\{z : |z| \le 1\}$ whenever $\mathcal{L}_A \ge 0$, but even in the latter case they can be either inside or outside the minimum polygon containing the eigenvalues of A, SEE the above example (however, if A is also normal, they are inside).

[Question: there are matrices A simultaneously normal, non negative and stochastic by columns (or by rows) which are not real symmetric and not in $\mathcal{C}$?]

Proposition.
If A is a non negative $n \times n$ matrix, then its best approximation in $\mathcal{C}$ is also non negative. (Proof: We know that $\mathcal{C}_A = \sum_k \alpha_k J_k$ with $J_k = J_2^{k-1} \ge 0$ and $\alpha_k = \frac{1}{n}(J_k, A)$. If A is non negative then also $\alpha_k \ge 0$, so $\mathcal{C}_A \ge 0$).

Question: Given A non negative, is $\mathcal{L}_A$ non negative for other spaces $\mathcal{L}$? Is τ_A non negative ? Recall that

$$\tau = \{Ud(\mathbf{z})U^H : \mathbf{z} \in \mathbb{C}^n\},\ \ U = \sqrt{\frac{2}{n+1}}\sin\frac{ij\pi}{n+1},\ 1 \le i, j \le n.$$

Elements of a basis of τ are obtained by choosing $J_k \in \tau$ such that $\mathbf{e}_1^T J_k = \mathbf{e}_k^T$, $k = 1, \dots, n$. They are matrices made up of zeros and ones only. Moreover, for such J_k, we have

$$\tau_A = \sum_k \alpha_k J_k,\ \alpha = B^{-1}\mathbf{c},\ B^{-1} = \frac{1}{2n+2}(3J_1 - J_3),\ c_r = (J_r, A).$$

[$\cdot$]. If A is non negative, then $\mathbf{c} \ge \mathbf{0}$, but $B^{-1}\mathbf{c}$ may have negative entries, but perhaps τ_A is yet non negative (investigate!).

THREE THEOREMS on s-stochastic matrix algebras $\mathcal{L}$ and on the best approximation in $\mathcal{L}$ of A (each more general than the previous one):

First theorem

Set $\mathcal{L} = \{Ud(\mathbf{x})U^H : \mathbf{x} \in \mathbb{C}^n\}$ where U is a unitary matrix. Choose $\mathbf{v}$ such that $[U^T\mathbf{v}]_j \neq 0 \;\forall\, j$. Note that the choice $\mathbf{v} = \mathbf{e}_1$ works for $\mathcal{L} = \mathcal{C}, \tau, \eta, \mu, \ldots$ but not for all low complexity matrix spaces $\mathcal{L}$ [...].

Then $\mathcal{L} = \{\mathcal{L}(\mathbf{z}) : \mathbf{z} \in \mathbb{C}^n\}$ where we have set $\mathcal{L}(\mathbf{z}) = Ud(U^T\mathbf{z})d(U^T\mathbf{v})^{-1}U^H$. Note that $\mathbf{v}^T\mathcal{L}(\mathbf{z}) = \mathbf{z}^T$, and that $\mathbf{x}^T\mathcal{L}(\mathbf{z}) = \mathbf{z}^T\mathcal{L}(\mathbf{x})$, $\forall\, \mathbf{x}, \mathbf{z} \in \mathbb{C}^n$. Moreover, $\mathcal{L}(\mathbf{v}) = I$.

Observe that if $\mathcal{L}(\mathbf{e}) = \mathbf{w}\mathbf{e}^T$ for some $\mathbf{w} \in \mathbb{C}^n$, then $\mathcal{L}(\mathbf{z})$ is $\mathbf{z}^T\mathbf{w}$-stochastic by columns, i.e.

$$\mathbf{e}^T\mathcal{L}(\mathbf{z}) = \mathbf{z}^T\mathcal{L}(\mathbf{e}) = (\mathbf{z}^T\mathbf{w})\mathbf{e}^T$$

[we have observed this first for $\mathcal{L} = \eta$ where $\mathbf{w} = \mathbf{e}$, $\mathbf{v} = \mathbf{e}_1$ (see previous pages)]. When, in general, $\mathcal{L}(\mathbf{e}) = \mathbf{w}\mathbf{e}^T$? Iff $\mathbf{v}^T\mathbf{w} = 1$ and $\mathbf{w}\mathbf{e}^T \in \mathcal{L}$. Assume $\mathbf{w}$ such that $\mathbf{v}^T\mathbf{w} = 1$, so $\mathbf{w}$ is in particular non null. Then $\mathbf{w}\mathbf{e}^T \in \mathcal{L}$ iff

$$\mathbf{w}\mathbf{e}^T = U \begin{bmatrix} & & \\ & \mathbf{e}^T\mathbf{w} & \\ & & \end{bmatrix} U^H = (\mathbf{e}^T\mathbf{w})(U\mathbf{e}_i)(U\mathbf{e}_i)^H, \;\; \mathbf{e}^T\mathbf{w} \neq 0. \qquad (*)$$

Since $U\mathbf{e}_i \neq \mathbf{0}$, the equation (*) times $\mathbf{e}_j$, with $\mathbf{e}_j \mid (U\mathbf{e}_i)^H\mathbf{e}_j \neq 0$, implies $\mathbf{w} = \alpha U\mathbf{e}_i$ $\alpha \neq 0$, and, since $(U^T\mathbf{v})_i \neq 0$, $\mathbf{v}^T$ times the equation (*) implies $U\mathbf{e}_i = \beta\mathbf{e}$ $\beta \neq 0$. Thus $\mathbf{w} = \gamma\mathbf{e}$ $\gamma \neq 0$, and, since $\mathbf{v}^T\mathbf{w} = 1$, we have $\gamma = \frac{1}{\mathbf{v}^T\mathbf{e}}$. So, if $\mathcal{L}(\mathbf{e}) = \mathbf{w}\mathbf{e}^T$, then $\mathbf{v}^T\mathbf{e}$ must be non zero and $\mathbf{w}$ must be equal to $\frac{\mathbf{e}}{\mathbf{v}^T\mathbf{e}}$. Now, provided that $\mathbf{v}^T\mathbf{e} \neq 0$, the matrix $\frac{\mathbf{e}\mathbf{e}^T}{\mathbf{v}^T\mathbf{e}}$ is in $\mathcal{L}$ iff

$$\frac{\mathbf{e}\mathbf{e}^T}{\mathbf{v}^T\mathbf{e}} = U \begin{bmatrix} & & \\ & \frac{\mathbf{e}^T\mathbf{e}}{\mathbf{v}^T\mathbf{e}} & \\ & & \end{bmatrix} U^H = \frac{n}{\mathbf{v}^T\mathbf{e}}(U\mathbf{e}_i)(U\mathbf{e}_i)^H. \qquad (**)$$

If $U\mathbf{e}_i = \frac{1}{\sqrt{n}}\mathbf{e}e^{\mathbf{i}\theta}$, then $\mathbf{v}^T\mathbf{e} \neq 0$ and (**) holds. So, we have proved the following theorem:

Theorem.
If $U\mathbf{e}_i = \frac{1}{\sqrt{n}}\mathbf{e}e^{\mathbf{i}\theta}$, then $\mathcal{L}(\mathbf{e}) = \frac{\mathbf{e}\mathbf{e}^T}{\mathbf{v}^T\mathbf{e}}$. It follows that, for any $\mathbf{z} \in \mathbb{C}^n$, the matrix $\mathcal{L}(\mathbf{z})$ is $\frac{\mathbf{z}^T\mathbf{e}}{\mathbf{v}^T\mathbf{e}}$-stochastic by columns, in particular the best approximation of A in $\mathcal{L}$, $\mathcal{L}_A = \mathcal{L}(\mathbf{z}_A) = U\,\mathrm{diag}\,((U^HAU)_{jj})U^H$, is $\frac{\mathbf{z}_A^T\mathbf{e}}{\mathbf{v}^T\mathbf{e}}$-stochastic by columns, i.e. $\mathbf{e}^T\mathcal{L}_A = \frac{\mathbf{z}_A^T\mathbf{e}}{\mathbf{v}^T\mathbf{e}}\mathbf{e}^T$. Finally note that one of the eigenvalues of $\mathcal{L}_A$, $(U^HAU)_{ii}$, is equal to $\frac{1}{n}\mathbf{e}^TA\mathbf{e}$, and that $\frac{\mathbf{z}_A^T\mathbf{e}}{\mathbf{v}^T\mathbf{e}} = \frac{1}{n}\mathbf{e}^TA\mathbf{e}$.
[For example, if $\mathcal{L} = \mathcal{C}, \eta, \ldots$ (where $\mathbf{v}$ can be chosen equal to $\mathbf{e}_1$), then $\mathbf{e}^T\mathcal{L}(\mathbf{z}) = \mathbf{z}^T\mathcal{L}(\mathbf{e}) = (\mathbf{z}^T\mathbf{e})\mathbf{e}^T = (\sum z_i)\mathbf{e}^T$, and $\mathbf{e}^T\mathcal{L}_A = (\frac{1}{n}\mathbf{e}^TA\mathbf{e})\mathbf{e}^T$].
In particular, if A is stochastic by columns ($\mathbf{e}^TA = \mathbf{e}^T$) or by rows ($A\mathbf{e} = \mathbf{e}$), then $\mathcal{L}_A$ is stochastic by columns.

Second theorem

Let U be a $n\times n$ unitary matrix, and set $\mathcal{L}=\{Ud(\mathbf{z})U^H\,:\,\mathbf{z}\in\mathbb{C}^n\}$. Choose $\mathbf{v}\in\mathbb{C}^n$ such that $(U^T\mathbf{v})_j\neq 0\;\forall\, j$ ($\mathbf{v}=\mathbf{e}_1$ if $\mathcal{L}=\mathcal{C},\tau,\eta,\mu,\ldots$; $\mathbf{v}\in\mathbb{R}^n$ whenever possible, f.i. if $U\in\mathbb{R}^{n\times n}$). Then, if we set

$$\begin{array}{l}\mathcal{L}_r(\mathbf{z})=Ud(U^T\mathbf{z})d(U^T\mathbf{v})^{-1}U^H\ \ (\mathbf{v}^T\mathcal{L}_r(\mathbf{z})=\mathbf{z}^T),\\ \mathcal{L}_c(\mathbf{z})=Ud(U^H\overline{\mathbf{v}})^{-1}d(U^H\mathbf{z})U^H\ \ (\mathcal{L}_c(\mathbf{z})\overline{\mathbf{v}}=\mathbf{z}),\end{array}$$

$\mathcal{L}$ can be also represented as $\mathcal{L}=\{\mathcal{L}_r(\mathbf{z})\,:\,\mathbf{z}\in\mathbb{C}^n\}=\{\mathcal{L}_c(\mathbf{z})\,:\,\mathbf{z}\in\mathbb{C}^n\}$. Note that $\mathbf{x}^T\mathcal{L}_r(\mathbf{y})=\mathbf{y}^T\mathcal{L}_r(\mathbf{x})$, $\mathcal{L}_c(\mathbf{y})\mathbf{x}=\mathcal{L}_c(\mathbf{x})\mathbf{y}$, $\mathcal{L}_r(\mathbf{v})=I$, $\mathcal{L}_c(\overline{\mathbf{v}})=I$.

Theorem.

If one of the columns of U has all entries equal each other, i.e. $\exists\, i$ and θ such that $U\mathbf{e}_i=\frac{1}{\sqrt{n}}\mathbf{e}e^{\mathbf{i}\theta}$, then

(1) $\mathcal{L}_r(\mathbf{e})=\frac{\mathbf{e}\mathbf{e}^T}{\mathbf{e}^T\mathbf{v}}$, $\mathcal{L}_c(\mathbf{e})=\frac{\mathbf{e}\mathbf{e}^T}{\mathbf{e}^T\overline{\mathbf{v}}}$ ($\Rightarrow\mathbf{e}\mathbf{e}^T\in\mathcal{L}$) and therefore $\mathcal{L}_r(\mathbf{z})$ is $\frac{\mathbf{z}^T\mathbf{e}}{\mathbf{e}^T\mathbf{v}}$-stochastic by columns, and $\mathcal{L}_c(\mathbf{z})$ is $\frac{\mathbf{z}^T\mathbf{e}}{\mathbf{e}^T\overline{\mathbf{v}}}$-stochastic by rows; in other words, $X\in\mathcal{L}\Rightarrow X$ is s_X-stochastic by rows and by columns for some s_X. [Note that $\mathbf{v}^T\mathbf{e}\neq 0$ because $(\mathbf{v}^TU)_i\neq 0$ ($(\mathbf{v}^TU)_j\neq 0\;\forall\, j$)].

(2) Given $A\in\mathbb{C}^{n\times n}$, the matrix $\mathcal{L}_A=\mathcal{L}_r(\mathbf{z}_A^r)=\mathcal{L}_c(\mathbf{z}_A^c)=U\,\mathrm{diag}\,((U^HAU)_{jj})U^H$, defined as the minimizer on $\mathcal{L}$ of $\|A-X\|_F$, has $(U^HAU)_{ii}=\frac{1}{n}\mathbf{e}^TA\mathbf{e}$ as eigenvalue, and is $\big(\frac{1}{n}\mathbf{e}^TA\mathbf{e}\big)$-stochastic by rows and by columns, i.e. $\mathcal{L}_A\mathbf{e}=\frac{1}{n}(\mathbf{e}^TA\mathbf{e})\mathbf{e}$, $\mathbf{e}^T\mathcal{L}_A=\frac{1}{n}(\mathbf{e}^TA\mathbf{e})\mathbf{e}^T$.

(3) If A is stochastic by columns or by rows, then $\mathcal{L}_A$ is stochastic by rows and by columns.

proof. (1): Note that $M_j:=U\begin{bmatrix} & & \\ & \mathbf{e}^T\mathbf{e} & \\ & & \end{bmatrix}U^H\in\mathcal{L},\ \forall\, j,\ \ M_j=(\mathbf{e}^T\mathbf{e})(U\mathbf{e}_j)(\overline{U}\mathbf{e}_j)^T$.

Moreover, by the assumption $U\mathbf{e}_i=\frac{1}{\sqrt{n}}\mathbf{e}e^{\mathbf{i}\theta}$, we have $M_i=\mathbf{e}\mathbf{e}^T$. So, $\mathbf{e}\mathbf{e}^T\in\mathcal{L}$ and, obviously, $\mathcal{L}_r(\mathbf{e})=\frac{\mathbf{e}\mathbf{e}^T}{\mathbf{v}^T\mathbf{e}}$ ($\mathbf{v}^T\mathcal{L}_r(\mathbf{e})=\mathbf{e}^T!$), $\mathcal{L}_c(\mathbf{e})=\frac{\mathbf{e}\mathbf{e}^T}{\overline{\mathbf{v}}^T\mathbf{e}}$ ($\mathcal{L}_c(\mathbf{e})\overline{\mathbf{v}}=\mathbf{e}!$). Thus, $\forall\,\mathbf{z}\in\mathbb{C}^n$ we have $\mathbf{e}^T\mathcal{L}_r(\mathbf{z})=\mathbf{z}^T\mathcal{L}_r(\mathbf{e})=\frac{\mathbf{z}^T\mathbf{e}}{\mathbf{v}^T\mathbf{e}}\mathbf{e}^T$, $\mathcal{L}_c(\mathbf{z})\mathbf{e}=\mathcal{L}_c(\mathbf{e})\mathbf{z}=\frac{\mathbf{e}^T\mathbf{z}}{\mathbf{e}^T\overline{\mathbf{v}}}\mathbf{e}$.

(2): It is enough to observe that $(U^HAU)_{ii}=\frac{1}{n}\mathbf{e}^TA\mathbf{e}$ and use the formula $\mathcal{L}_A=U\,\mathrm{diag}\,((U^HAU)_{jj})U^H$. However, let us obtain the thesis from (1). As a consequence of (1), the matrix

$$\mathcal{L}_A=Ud(U^T\mathbf{z}_A^r)d(U^T\mathbf{v})^{-1}U^H=Ud(U^H\overline{\mathbf{v}})^{-1}d(U^H\mathbf{z}_A^c)U^H$$

is both $\frac{\mathbf{e}^T\mathbf{z}_A^r}{\mathbf{v}^T\mathbf{e}}$-stochastic by columns and $\frac{\mathbf{e}^T\mathbf{z}_A^c}{\overline{\mathbf{v}}^T\mathbf{e}}$-stochastic by rows. Let us prove that $\frac{\mathbf{e}^T\mathbf{z}_A^r}{\mathbf{v}^T\mathbf{e}}=\frac{\mathbf{e}^T\mathbf{z}_A^c}{\overline{\mathbf{v}}^T\mathbf{e}}=\frac{1}{n}\mathbf{e}^TA\mathbf{e}$. Since $U^H\mathbf{e}=\sqrt{n}e^{-\mathbf{i}\theta}\mathbf{e}_i$, we have

$$\begin{array}{rcl}(\mathbf{z}_A^r)^T\mathbf{e}&=&\mathbf{v}^TU\,\mathrm{diag}\,((U^HAU)_{jj})U^H\mathbf{e}=\sqrt{n}e^{-\mathbf{i}\theta}\mathbf{v}^TU\mathbf{e}_i(U^HAU)_{ii}\\ &=&\mathbf{v}^T\mathbf{e}(U^HAU)_{ii}=\mathbf{v}^T\mathbf{e}(U\mathbf{e}_i)^HA(U\mathbf{e}_i)=\mathbf{v}^T\mathbf{e}\frac{1}{n}\mathbf{e}^TA\mathbf{e}\end{array}$$

and, since $\mathbf{e}^TU=\sqrt{n}e^{\mathbf{i}\theta}\mathbf{e}_i^T$, we have

$$\begin{array}{rcl}\mathbf{e}^T\mathbf{z}_A^c&=&\mathbf{e}^TU\,\mathrm{diag}\,((U^HAU)_{jj})U^H\overline{\mathbf{v}}=\sqrt{n}e^{\mathbf{i}\theta}(U^HAU)_{ii}\mathbf{e}_i^TU^H\overline{\mathbf{v}}\\ &=&\mathbf{e}^T\overline{\mathbf{v}}(U^HAU)_{ii}=\mathbf{e}^T\overline{\mathbf{v}}\frac{1}{n}\mathbf{e}^TA\mathbf{e}.\end{array}\qquad\square$$

Question: when $\mathcal{L}_c(\mathbf{z})=\mathcal{L}_r(\mathbf{x})$? Iff $U^T\mathbf{x}=d(\mathbf{u})U^H\mathbf{z}$, $\mathbf{u}=d(U^H\overline{\mathbf{v}})^{-1}U^T\mathbf{v}$ [$\mathbf{u}=\mathbf{e}$ if $U\in\mathbb{R}^{n\times n}$ ($\mathbf{v}\in\mathbb{R}^n$) or $U=F$ ($\mathbf{v}=\mathbf{e}_1$); $|u_i|=1\;\forall\, i$].

Third Theorem

Let U, V be two $n \times n$ unitary matrices. Choose $\mathbf{v}, \mathbf{u} \in \mathbb{C}^n$ such that $(U^T\mathbf{v})_i \neq 0$, $(V^T\mathbf{u})_i \neq 0$, $\forall\, i$. Given $\mathbf{z} \in \mathbb{C}^n$, set

$$\mathcal{L}_r(\mathbf{z}) = Ud(V^T\mathbf{z})d(U^T\mathbf{v})^{-1}V^H, \quad \mathcal{L}_c(\mathbf{z}) = Ud(V^H\overline{\mathbf{u}})^{-1}d(U^H\mathbf{z})V^H.$$

Note that $\mathbf{v}^T\mathcal{L}_r(\mathbf{z}) = \mathbf{z}^T$, $\mathcal{L}_c(\mathbf{z})\overline{\mathbf{u}} = \mathbf{z}$.

Theorem.
If there exists i such that $V\mathbf{e}_i = \frac{1}{\sqrt{n}}\mathbf{e}e^{\mathbf{i}\theta}$, $U\mathbf{e}_i = \frac{1}{\sqrt{n}}\mathbf{e}e^{\mathbf{i}\varphi}$, then $Vd(U^T\mathbf{e})d(U^T\mathbf{v})^{-1}V^H = \frac{\mathbf{e}\mathbf{e}^T}{\mathbf{e}^T\mathbf{v}}$, $Ud(V^H\overline{\mathbf{u}})^{-1}d(V^H\mathbf{e})U^H = \frac{\mathbf{e}\mathbf{e}^T}{\mathbf{e}^T\overline{\mathbf{u}}}$, and therefore

$$\begin{aligned} &\mathbf{e}^T\mathcal{L}_r(\mathbf{z}) = \mathbf{z}^TVd(U^T\mathbf{e})d(U^T\mathbf{v})^{-1}V^H = (\tfrac{\mathbf{z}^T\mathbf{e}}{\mathbf{e}^T\mathbf{v}})\mathbf{e}^T, \\ &\mathcal{L}_c(\mathbf{z})\mathbf{e} = Ud(V^H\overline{\mathbf{u}})^{-1}d(V^H\mathbf{e})U^H\mathbf{z} = \mathbf{e}(\tfrac{\mathbf{e}^T\mathbf{z}}{\mathbf{e}^T\overline{\mathbf{u}}}). \end{aligned}$$

In other words, $X \in \mathcal{L} \Rightarrow X$ is s_X-stochastic by rows and by columns for some $s_X \in \mathbb{C}$. Since, moreover, $(U^HAV)_{ii} = \frac{1}{n}e^{\mathbf{i}(\theta-\varphi)}\mathbf{e}^TA\mathbf{e}$, if $\mathcal{L}_A = Ud(V^T\mathbf{z}_A^r)d(U^T\mathbf{v})^{-1}V^H = Ud(V^H\overline{\mathbf{u}})^{-1}d(U^H\mathbf{z}_A^c)V^H = U\,\mathrm{diag}\,((U^HAV)_{jj})V^H$ is the best approximation of A in $\mathcal{L} = \{Ud(\mathbf{z})V^H : \mathbf{z} \in \mathbb{C}^n\}$, then we have that

$$\begin{aligned} &(\mathbf{z}_A^r)^T\mathbf{e} = \mathbf{v}^TU\,\mathrm{diag}\,((U^HAV)_{jj})V^H\mathbf{e} = \tfrac{1}{n}(\mathbf{e}^TA\mathbf{e})\mathbf{v}^T\mathbf{e}, \\ &\mathbf{e}^T(\mathbf{z}_A^c) = \mathbf{e}^TU\,\mathrm{diag}\,((U^HAV)_{jj})V^H\overline{\mathbf{u}} = \tfrac{1}{n}(\mathbf{e}^TA\mathbf{e})\mathbf{e}^T\overline{\mathbf{u}}, \end{aligned}$$

and therefore

$$\mathbf{e}^T\mathcal{L}_A = \frac{1}{n}(\mathbf{e}^TA\mathbf{e})\mathbf{e}^T, \quad \mathcal{L}_A\mathbf{e} = \frac{1}{n}(\mathbf{e}^TA\mathbf{e})\mathbf{e}.$$

It follows that whenever $A \in \mathbb{C}^{n\times n}$ is stochastic by rows ($A\mathbf{e} = \mathbf{e}$) or stochastic by columns ($\mathbf{e}^TA = \mathbf{e}^T$), its better approximation $\mathcal{L}_A$ in $\mathcal{L}$ is stochastic simultaneously by rows and by columns.

proof. $V\mathbf{e}_i = \frac{1}{\sqrt{n}}e^{\mathbf{i}\theta}\mathbf{e}$ $\left[\, U\mathbf{e}_i = \frac{1}{\sqrt{n}}e^{\mathbf{i}\varphi}\mathbf{e} \,\right] \Rightarrow$

$$V\begin{bmatrix} & & \\ & \frac{\mathbf{e}^T\mathbf{e}}{\mathbf{e}^T\mathbf{v}} & \\ & & \end{bmatrix}V^H = \frac{\mathbf{e}^T\mathbf{e}}{\mathbf{e}^T\mathbf{v}}V\mathbf{e}_i(V\mathbf{e}_i)^H = \frac{\mathbf{e}\mathbf{e}^T}{\mathbf{e}^T\mathbf{v}} \;\left[U\begin{bmatrix} & & \\ & \frac{\mathbf{e}^T\mathbf{e}}{\mathbf{e}^T\overline{\mathbf{u}}} & \\ & & \end{bmatrix}U^H = \frac{\mathbf{e}^T\mathbf{e}}{\mathbf{e}^T\overline{\mathbf{u}}}U\mathbf{e}_i(U\mathbf{e}_i)^H = \frac{\mathbf{e}\mathbf{e}^T}{\mathbf{e}^T\overline{\mathbf{u}}}\right].$$

$U\mathbf{e}_i = \frac{1}{\sqrt{n}}e^{\mathbf{i}\varphi}\mathbf{e}$ $\left[\, V\mathbf{e}_i = \frac{1}{\sqrt{n}}e^{\mathbf{i}\theta}\mathbf{e}\right] \Rightarrow$
$\mathbf{e}_i^TU^H = \frac{1}{\sqrt{n}}e^{-\mathbf{i}\varphi}\mathbf{e}^T$ $\left[\mathbf{e}_i^TV^H = \frac{1}{\sqrt{n}}e^{-\mathbf{i}\theta}\mathbf{e}^T\right] \Rightarrow$

$$\begin{matrix} \mathbf{e}_i^T(U^H\mathbf{e}) = \frac{1}{\sqrt{n}}e^{-\mathbf{i}\varphi}\mathbf{e}^T\mathbf{e} = \frac{(U^H\overline{\mathbf{v}})_i}{\mathbf{e}^T\overline{\mathbf{v}}}\mathbf{e}^T\mathbf{e}, \\ \mathbf{e}_j^T(U^H\mathbf{e}) = \mathbf{e}_j^TU^He^{-\mathbf{i}\varphi}\sqrt{n}U\mathbf{e}_i = 0, \; j \neq i \end{matrix} \quad \left[\begin{matrix} \mathbf{e}_i^T(V^H\mathbf{e}) = \frac{1}{\sqrt{n}}e^{-\mathbf{i}\theta}\mathbf{e}^T\mathbf{e} = \frac{(V^H\overline{\mathbf{u}})_i}{\mathbf{e}^T\overline{\mathbf{u}}}\mathbf{e}^T\mathbf{e}, \\ \mathbf{e}_j^T(V^H\mathbf{e}) = \mathbf{e}_j^TV^He^{-\mathbf{i}\theta}\sqrt{n}V\mathbf{e}_i = 0, \; j \neq i \end{matrix} \right].$$

Thus we have

$$d(U^H\mathbf{e})d(U^H\overline{\mathbf{v}})^{-1} = \begin{bmatrix} & & \\ & \frac{\mathbf{e}^T\mathbf{e}}{\mathbf{e}^T\overline{\mathbf{v}}} & \\ & & \end{bmatrix} \;\left[d(V^H\mathbf{e})d(V^H\overline{\mathbf{u}})^{-1} = \begin{bmatrix} & & \\ & \frac{\mathbf{e}^T\mathbf{e}}{\mathbf{e}^T\overline{\mathbf{u}}} & \\ & & \end{bmatrix}\right],$$

and therefore $Vd(U^T\mathbf{e})d(U^T\mathbf{v})^{-1}V^H = \frac{\mathbf{e}\mathbf{e}^T}{\mathbf{e}^T\mathbf{v}}$ $\left[Ud(V^H\overline{\mathbf{u}})^{-1}d(V^H\mathbf{e})U^H = \frac{\mathbf{e}\mathbf{e}^T}{\mathbf{e}^T\overline{\mathbf{u}}}\right]$.

The equalities $V^H\mathbf{e} = e^{-\mathbf{i}\theta}\sqrt{n}\mathbf{e}_i$, $\mathbf{e}^TU = e^{\mathbf{i}\varphi}\sqrt{n}\mathbf{e}_i^T$, let us easily obtain the assertions on $\mathcal{L}_A$.

REMARK. $\mathcal{L}=\{Ud(\mathbf{z})V^H : \mathbf{z}\in\mathbb{C}^n\}$, U,V unitary, $\mathcal{L}_A = U\,\mathrm{diag}\,((U^HAV)_{jj})V^H$:

$$V\mathbf{e}_i = \tfrac{1}{\sqrt{n}}e^{\mathbf{i}\theta}\mathbf{e} \;\Rightarrow\; \mathcal{L}_A\mathbf{e} = ((U\mathbf{e}_i)^H A\mathbf{e})U\mathbf{e}_i,\; (U\mathbf{e}_i)^H\mathcal{L}_A = \tfrac{(U\mathbf{e}_i)^HA\mathbf{e}}{n}\mathbf{e}^T;$$

$$V\mathbf{e}_i = \tfrac{1}{\sqrt{n}}e^{\mathbf{i}\theta}\mathbf{e},\; U\mathbf{e}_i = \tfrac{e^{\mathbf{i}\varphi}}{\|\mathbf{e}_\le\|}\mathbf{e}_\le,\; A\mathbf{e}=\mathbf{e}_\le \;\Rightarrow\; \mathcal{L}_A\mathbf{e}=\mathbf{e}_\le,\; \mathbf{e}_\le^H\mathcal{L}_A = \tfrac{\|\mathbf{e}_\le\|^2}{n}\mathbf{e}^T;$$

$$U\mathbf{e}_i = \tfrac{1}{\sqrt{n}}e^{\mathbf{i}\varphi}\mathbf{e} \;\Rightarrow\; \mathbf{e}^T\mathcal{L}_A = (\mathbf{e}^TAV\mathbf{e}_i)(V\mathbf{e}_i)^H,\; \mathcal{L}_A(V\mathbf{e}_i) = \tfrac{\mathbf{e}^TAV\mathbf{e}_i}{n}\mathbf{e};$$

$$U\mathbf{e}_i = \tfrac{1}{\sqrt{n}}e^{\mathbf{i}\varphi}\mathbf{e},\; V\mathbf{e}_i = \tfrac{e^{\mathbf{i}\theta}}{\|\mathbf{e}_\le\|}\mathbf{e}_\le,\; \mathbf{e}^TA=\mathbf{e}_\le^H \;\Rightarrow\; \mathbf{e}^T\mathcal{L}_A = \mathbf{e}_\le^H,\; \mathcal{L}_A\mathbf{e}_\le = \tfrac{\|\mathbf{e}_\le\|^2}{n}\mathbf{e}.$$

Exercise.
Given $\mathbf{w}\in\mathbb{C}^n$, $\mathbf{w}\neq\mathbf{0}$, set $\mathcal{M}=\{X\in\mathbb{C}^{n\times n} : X\mathbf{w}\mathbf{e}^T=\mathbf{w}\mathbf{e}^TX\}$. Prove that
(i) $\mathcal{M}$ is a matrix algebra ;
(ii) $\mathcal{M}=\{X\in\mathbb{C}^{n\times n} : X\mathbf{w}=c\mathbf{w}$ & $\mathbf{e}^TX=c\mathbf{e}^T$ for some $c\in\mathbb{C}\}$, i.e. $X\in\mathcal{M}$ implies X is s_X-stochastic by columns ;
(iii) if $\mathbf{w}=\mathbf{e}$, then $\mathcal{M}=\{X\in\mathbb{C}^{n\times n} : X\mathbf{e}=c\mathbf{e}$ & $\mathbf{e}^TX=c\mathbf{e}^T$ for some $c\in\mathbb{C}\}$, i.e. $X\in\mathcal{M}$ implies X is s_X-stochastic by rows and by columns.

→ Investigate low complexity spaces $\mathcal{L}$ of matrices commuting with $\mathbf{w}\mathbf{e}^T$, in particular commutative spaces $\mathcal{L}$ including $\mathbf{w}\mathbf{e}^T$. We have seen examples in the case $\mathbf{w}=\mathbf{e}$.

Let U be a $n\times n$ unitary matrix. Set $\mathcal{L}=\{U(\mu\circ Z)U^H : Z\in\mathbb{C}^{n\times n}\}$ where μ is a fixed matrix whose entries are 0 or 1 and $\circ$ is the entry by entry product. For example

$$\mu=\begin{bmatrix} & & 1\\ 1 & 1 & 1\\ & & 1\end{bmatrix},\ Z=\begin{bmatrix} z_{11} & z_{12} & z_{13}\\ z_{21} & z_{22} & z_{23}\\ z_{31} & z_{32} & z_{33}\end{bmatrix},\ \mu\circ Z=\begin{bmatrix} & & z_{13}\\ z_{21} & z_{22} & z_{23}\\ & & z_{33}\end{bmatrix}.$$

Note that if $\mu=I$, then $\mathcal{L}=\{Ud(\mathbf{z})U^H : \mathbf{z}\in\mathbb{C}^n\}$.

The space of matrices $\mathcal{L}$ is a vector subspace of $\mathbb{C}^{n\times n}$, and is a matrix algebra (i.e. product of matrices from $\mathcal{L}$ are in $\mathcal{L}$) if the matrix μ satisfies the condition

$$[\mu]_{ij}=0 \quad\Rightarrow\quad [\mu^2]_{ij}=0$$

[or $[\mu^2]_{ij}\neq 0\Rightarrow[\mu]_{ij}\neq 0$; or $\mu^2\leq\alpha\mu$ for some $\alpha>0$ (the pattern of μ^2 is enclosed in the pattern of μ)]. Examples of μ satisfying $\mu^2\leq\alpha\mu$:

$$\mu=I,\ \mathbf{e}\mathbf{e}^T,\ \begin{bmatrix}1 & & \\ 1 & & \\ & & \end{bmatrix},\ \begin{bmatrix} & & \\ & 1 & 1\\ & & \end{bmatrix},\ \begin{bmatrix}1 & 1 & \\ 1 & 1 & \\ & & 1\end{bmatrix},\ \begin{bmatrix}1 & & \\ 1 & & \\ 1 & & \end{bmatrix},\ \begin{bmatrix} & & 1\\ 1 & 1 & 1\\ & & 1\end{bmatrix},\ \begin{bmatrix}1 & & \\ 1 & 1 & \\ 1 & 1 & 1\end{bmatrix},$$

$$\begin{bmatrix}1 & & \\ 1 & 1 & \\ & & 1\end{bmatrix},\ \begin{bmatrix}1 & & \\ & 1 & \\ 1 & & 1\end{bmatrix},\ \begin{bmatrix} & & \\ 1 & & \\ 1 & & 1\end{bmatrix},\ \begin{bmatrix}1 & & \\ 1 & 1 & \\ 1 & & 1\end{bmatrix},\ \begin{bmatrix}1 & & \\ 1 & 1 & \\ 1 & & \end{bmatrix},\ \begin{bmatrix}1 & 1 & 1\\ & & 1\\ & & 1\end{bmatrix}.$$

Given $A\in\mathbb{C}^{n\times n}$, and defined $\mathcal{L}_A$ as the minimizer of $\|A-U(\mu\circ Z)U^H\|_F$, $Z\in\mathbb{C}^{n\times n}$, we have

$$\mathcal{L}_A=U(\mu\circ(U^HAU))U^H.$$

Observe that if μ has a triangular structure, then the eigenvalues of $\mathcal{L}_A$ are $\mu_{jj}(U^HAU)_{jj}$, $j=1,\ldots,n$, i.e. null or the same of $U\,\mathrm{diag}\,((U^HAU)_{jj})U^H$.

In the particular case where $U\mathbf{e}_1=\frac{1}{\sqrt{n}}e^{\mathbf{i}\theta}\mathbf{e}$ and $\mu_{11}=1$, the matrix $\mu\circ(U^HAU)$ can be written as follows

$$\mu\circ(U^HAU)=\begin{bmatrix} [\frac{1}{n}\mathbf{e}^TA\mathbf{e}] & \cdot & \mu_{1j}[\frac{1}{\sqrt{n}}e^{-\mathbf{i}\theta}\mathbf{e}^TA(U\mathbf{e}_j)] & \cdot\\ \cdot & & \cdot & \\ \mu_{i1}[\frac{1}{\sqrt{n}}e^{\mathbf{i}\theta}(U\mathbf{e}_i)^HA\mathbf{e}] & \cdot & \mu_{ij}[(U\mathbf{e}_i)^HA(U\mathbf{e}_j)] & \cdot\\ \cdot & & \cdot & \end{bmatrix}.$$

Theorem (stoch by rows).
Assume $U\mathbf{e}_1=\frac{1}{\sqrt{n}}e^{\mathbf{i}\theta}\mathbf{e}$ and $\mu_{11}=1$. If $A\mathbf{e}=\mathbf{e}$, then $\mathcal{L}_A\mathbf{e}=\mathbf{e}$ and

$$\mu\circ(U^HAU)=\begin{bmatrix} 1 & \cdot & \mu_{1j}[\frac{1}{\sqrt{n}}e^{-\mathbf{i}\theta}\mathbf{e}^TA(U\mathbf{e}_j)] & \cdot\\ 0 & & \cdot & \\ \cdot & \cdot & \mu_{ij}[(U\mathbf{e}_i)^HA(U\mathbf{e}_j)] & \cdot\\ 0 & & \cdot & \end{bmatrix}.$$

If moreover $\mu_{1j}=0$, $j=2,\ldots,n$, then

$$\mu\circ(U^HAU)=\begin{bmatrix} 1 & 0 & \cdot & 0\\ 0 & & \cdot & \\ \cdot & \cdot & \mu_{ij}[(U\mathbf{e}_i)^HA(U\mathbf{e}_j)] & \cdot\\ 0 & & \cdot & \end{bmatrix},\quad \mathbf{e}^T\mathcal{L}_A=\mathbf{e}^T.$$

(thus choose $\mu \geq \mathbf{e}_1\mathbf{e}_1^T + \mathbf{e}_1\mathbf{e}_j^T$ for some j in order to have $\mathbf{e}^T\mathcal{L}_A \neq \mathbf{e}^T$).
If alternatively A is quasi-stochastic by rows, $A\mathbf{e} = \mathbf{e}_{\leq}$, with $\mathbf{0} \leq \mathbf{e}_{\leq} \leq \mathbf{e}$, then $\mathcal{L}_A\mathbf{e} = \frac{\mathbf{e}^T\mathbf{e}_{\leq}}{n}\mathbf{e}$ whenever $\mu_{i1} = 0\ \forall\, i \geq 2$.

proof: investigate the first column in the equality $\mathcal{L}_A U = U(\mu \circ (U^H AU))$:

$$\mathcal{L}_A\mathbf{e} = \frac{1}{n}(\mathbf{e}^T A\mathbf{e})\mathbf{e} + \sum_{i=1,\, i\neq 1}^{n} \mu_{i1}((U\mathbf{e}_i)^H A\mathbf{e})U\mathbf{e}_i. \quad \square$$

Theorem (stoch by columns).
Assume $U\mathbf{e}_1 = \frac{1}{\sqrt{n}}e^{\mathbf{i}\theta}\mathbf{e}$ and $\mu_{11} = 1$. If $\mathbf{e}^T A = \mathbf{e}^T$, then $\mathbf{e}^T\mathcal{L}_A = \mathbf{e}^T$ and

$$\mu \circ (U^H AU) = \begin{bmatrix} 1 & 0 & \cdot & 0 \\ \cdot & & \cdot & \\ \mu_{i1}[\frac{1}{\sqrt{n}}e^{\mathbf{i}\theta}(U\mathbf{e}_i)^H A\mathbf{e}] & \cdot & \mu_{ij}[(U\mathbf{e}_i)^H A(U\mathbf{e}_j)] & \cdot \\ \cdot & & \cdot & \end{bmatrix}.$$

If moreover $\mu_{i1} = 0$, $i = 2, \ldots, n$, then

$$\mu \circ (U^H AU) = \begin{bmatrix} 1 & 0 & \cdot & 0 \\ 0 & & \cdot & \\ \cdot & \cdot & \mu_{ij}[(U\mathbf{e}_i)^H A(U\mathbf{e}_j)] & \cdot \\ 0 & & \cdot & \end{bmatrix}, \quad \mathcal{L}_A\mathbf{e} = \mathbf{e}.$$

(thus choose $\mu \geq \mathbf{e}_1\mathbf{e}_1^T + \mathbf{e}_i\mathbf{e}_1^T$ for some i in order to have $\mathcal{L}_A\mathbf{e} \neq \mathbf{e}$).
If alternatively A is quasi-stochastic by columns, $\mathbf{e}^T A = \mathbf{e}_{\leq}^T$, with $\mathbf{0} \leq \mathbf{e}_{\leq} \leq \mathbf{e}$, then $\mathbf{e}^T\mathcal{L}_A = \frac{\mathbf{e}^T\mathbf{e}_{\leq}}{n}\mathbf{e}^T$ whenever $\mu_{1j} = 0\ \forall\, j \geq 2$.

proof: investigate the first row in the equality $U^H\mathcal{L}_A = (\mu \circ (U^H AU))U^H$:

$$\mathbf{e}^T\mathcal{L}_A = \frac{1}{n}(\mathbf{e}^T A\mathbf{e})\mathbf{e}^T + \sum_{j=1,\, j\neq 1}^{n} \mu_{1j}(\mathbf{e}^T A(U\mathbf{e}_j))(U\mathbf{e}_j)^H. \quad \square$$

Note. If $\mathcal{L}_A$ and its eigenvalues do not fit our requirements, then we could introduce a perturbation of $\mathcal{L}_A$, yet in $\mathcal{L}$, in place of it, for example the matrix

$$M = \mathcal{L}_A + U(\mu \circ (\,(U^H AU) \circ \varepsilon\,))U^H,\ \varepsilon \in \mathbb{R}^{n\times n},\ |\varepsilon_{ij}| \text{ small}.$$

Note that

$$\|A - M\|_F \leq \|A - \mathcal{L}_A\|_F + \sqrt{\sum_{i,j,\, \mu_{ij}=1} |\varepsilon_{ij}|^2 |(U^H AU)_{ij}|^2}.$$

But the Frobenius norm is the right norm?

Results on $\mathcal{L}_A$ more general than the ones in REMARK, where $\mathcal{L} = \{Ud(\mathbf{z})V^H : \mathbf{z} \in \mathbb{C}^n\}$, and the ones in Theorem stoch by columns, and Theorem stoch by rows, where $\mathcal{L} = \{U(\mu \circ Z)U^H : Z \in \mathbb{C}^{n\times n}\}$:

Let U, V be $n \times n$ unitary matrices. Set $\mathcal{L} = \{U(\mu \circ Z)V^H : Z \in \mathbb{C}^{n\times n}\}$ where μ is a fixed matrix whose entries are 0 or 1 and $\circ$ is the entry by entry product. Note that if $\mu = I$, then $\mathcal{L} = \{Ud(\mathbf{z})V^H : \mathbf{z} \in \mathbb{C}^n\}$.

The space of matrices $\mathcal{L}$ is a vector subspace of $\mathbb{C}^{n\times n}$. In general it is not a matrix algebra.

Given $A \in \mathbb{C}^{n\times n}$, and defined $\mathcal{L}_A$ as the minimizer of $\|A - U(\mu \circ Z)V^H\|_F$, $Z \in \mathbb{C}^{n\times n}$, we have

$$\mathcal{L}_A = U(\mu \circ (U^H AV))V^H.$$

Observe that if μ has a diagonal structure, then the eigenvalues of $\mathcal{L}_A\mathcal{L}_A^H$ are $\mu_{jj}|(U^H AV)_{jj}|^2$, $j = 1, \ldots, n$.

Note that the equalities $\mathcal{L}_A V = U(\mu\circ(U^H AV))$ and $U^H\mathcal{L}_A = (\mu\circ(U^H AV))V^H$ imply, respectively,

$$\mathcal{L}_A(V\mathbf{e}_1) = \mu_{11}(U^H AV)_{11}U\mathbf{e}_1 + \textstyle\sum_{i=1,\, i\neq 1}^{n} \mu_{i1}(U^H AV)_{i1}U\mathbf{e}_i,$$
$$(U\mathbf{e}_1)^H\mathcal{L}_A = \mu_{11}(U^H AV)_{11}(V\mathbf{e}_1)^H + \textstyle\sum_{j=1,\, j\neq 1}^{n} \mu_{1j}(U^H AV)_{1j}(V\mathbf{e}_j)^H.$$

In the following two theorems $\mathbf{e}_\leq$ can be an arbitrary vector with complex entries. However, we think to use the stated results for $\mathbf{e}_\leq = \mathbf{e}$ (stochastic case) or for $\mathbf{e}_\leq$, $0 \leq [\mathbf{e}_\leq]_j \leq 1$ (quasi-stochastic case).

Theorem (stochastic by rows)
$V\mathbf{e}_1 = \frac{1}{\sqrt{n}}e^{\mathbf{i}\theta}\mathbf{e}$ & $\mu_{11} = 1 \;\Rightarrow$

$$\mathcal{L}_A\mathbf{e} = ((U\mathbf{e}_1)^H A\mathbf{e})U\mathbf{e}_1 + \sum_{i=1,\, i\neq 1}^{n} \mu_{i1}((U\mathbf{e}_i)^H A\mathbf{e})U\mathbf{e}_i.$$

Thus
(i) if $A\mathbf{e} = \mathbf{e}_\leq$ & $U\mathbf{e}_1 = \frac{e^{\mathbf{i}\varphi}}{\|\mathbf{e}_\leq\|}\mathbf{e}_\leq$, then $\mathcal{L}_A\mathbf{e} = \mathbf{e}_\leq$. If, moreover, $\mu_{1j} = 0$ $\forall\, j \neq 1$, then $\mathbf{e}_\leq^H\mathcal{L}_A = \frac{\|\mathbf{e}_\leq\|^2}{n}\mathbf{e}^T$.
(ii) if $A\mathbf{e} = \mathbf{e}_\leq$ & $U\mathbf{e}_1 = \frac{e^{\mathbf{i}\varphi}}{\sqrt{n}}\mathbf{e}$, then $\mathcal{L}_A\mathbf{e} = \frac{\mathbf{e}^T\mathbf{e}_\leq}{n}\mathbf{e}$ whenever $\mu_{i1} = 0\; \forall\, i \neq 1$.

Theorem (stochastic by columns)
$U\mathbf{e}_1 = \frac{1}{\sqrt{n}}e^{\mathbf{i}\varphi}\mathbf{e}$ & $\mu_{11} = 1 \;\Rightarrow$

$$\mathbf{e}^T\mathcal{L}_A = (\mathbf{e}^T A(V\mathbf{e}_1))(V\mathbf{e}_1)^H + \sum_{j=1,\, j\neq 1}^{n} \mu_{1j}(\mathbf{e}^T AV\mathbf{e}_j)(V\mathbf{e}_j)^H.$$

Thus
(i) if $\mathbf{e}^T A = \mathbf{e}_\leq^H$ & $V\mathbf{e}_1 = \frac{e^{\mathbf{i}\theta}}{\|\mathbf{e}_\leq\|}\mathbf{e}_\leq$, then $\mathbf{e}^T\mathcal{L}_A = \mathbf{e}_\leq^H$. If, moreover, $\mu_{i1} = 0$ $\forall\, i \neq 1$, then $\mathcal{L}_A\mathbf{e}_\leq = \frac{\|\mathbf{e}_\leq\|^2}{n}\mathbf{e}$.
(ii) if $\mathbf{e}^T A = \mathbf{e}_\leq^H$ & $V\mathbf{e}_1 = \frac{e^{\mathbf{i}\theta}}{\sqrt{n}}\mathbf{e}$, then $\mathbf{e}^T\mathcal{L}_A = \frac{\mathbf{e}_\leq^H\mathbf{e}}{n}\mathbf{e}^T$ whenever $\mu_{1j} = 0$ $\forall\, j \neq 1$.