A Sinkhorn-Knopp Fixed Point Problem: Some Cases and Results Involving Structured and Patterned Matrices

Brent Hancock

Dr. David Strong, Advisor

PEPPERDINE UNIVERSITY

The Problem at a Glance

$$\vec{x} = \left(N^T (N\vec{x})^{(-1)}\right)^{(-1)} \qquad \vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

- N is an $n \times n$ matrix
- $\vec{x}$ is an $n \times 1$ vector
- $\vec{v}^{(-1)}$ is the entrywise reciprocal of a vector $\vec{v}$

Classification of Solutions

It can be shown that if $\vec{x}$ is a solution for a given matrix, then $c\vec{x}$ is also a solution.

So, if $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ is a solution,

$\frac{1}{x_3}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1/x_3 \\ x_2/x_3 \\ 1 \end{bmatrix} = \begin{bmatrix} x_1' \\ x_2' \\ 1 \end{bmatrix}$ is also a solution

Reminder: For the entrywise reciprocal operator to be defined, all entries of $\vec{x}$ must be nonzero (thus $x_3 \neq 0$)

General 2x2 Matrix

$$N = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \qquad \vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\vec{x} = \left(N^T (N\vec{x})^{(-1)}\right)^{(-1)} = \begin{bmatrix} \frac{a}{ax_1+bx_2} + \frac{c}{cx_1+dx_2} \\ \frac{b}{ax_1+bx_2} + \frac{d}{cx_1+dx_2} \end{bmatrix}^{(-1)} = \begin{bmatrix} \frac{(ax_1+bx_2)(cx_1+dx)}{2acx_1+(ad+bc)x_2} \\ \frac{(ax_1+bx_2)(cx_1+dx)}{(ad+bc)x_1+2bdx_2} \end{bmatrix}$$

$$\Rightarrow \quad \vec{x} = \begin{bmatrix} \pm\sqrt{\frac{bd}{ac}}x_2 \\ x_2 \end{bmatrix}$$

Alternating w/ 2 Parameters (4x4)

$$N = \begin{bmatrix} a & b & a & b \\ b & a & b & a \\ a & b & a & b \\ b & a & b & a \end{bmatrix} = N^T \qquad \vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ 1 \end{bmatrix}$$

$$\vec{x} = (N^T (N\vec{x})^{(-1)})^{(-1)} = \left(\begin{bmatrix} a & b & a & b \\ b & a & b & a \\ a & b & a & b \\ b & a & b & a \end{bmatrix}\left(\begin{bmatrix} a & b & a & b \\ b & a & b & a \\ a & b & a & b \\ b & a & b & a \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ 1 \end{bmatrix}\right)^{(-1)}\right)^{(-1)}$$

$$\text{so } \vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ 1 \end{bmatrix} = \begin{pmatrix} \frac{2a}{ax_1+bx_2+ax_3+b} + \frac{2b}{bx_1+ax_2+bx_3+a} \\ \frac{2b}{ax_1+bx_2+ax_3+b} + \frac{2a}{bx_1+ax_2+bx_3+a} \\ \frac{2a}{ax_1+bx_2+ax_3+b} + \frac{2b}{bx_1+ax_2+bx_3+a} \\ \frac{2b}{ax_1+bx_2+ax_3+b} + \frac{2a}{bx_1+ax_2+bx_3+a} \end{pmatrix}^{(-1)}$$

$$\Longrightarrow 2 \text{ Solutions}: \quad \vec{x}^{(1)} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \quad \text{and} \quad \vec{x}^{(2)} = \begin{bmatrix} -1 \\ 1 \\ -1 \\ 1 \end{bmatrix}$$

• Note: For all matrices $N_{n\times n}$ of this form such that n is even, these appear to be the only 2 solutions.

SUMMARY

I consider the fixed point problem $\vec{x} = \left(N^T (N\vec{x})^{(-1)}\right)^{(-1)}$, where $^{(-1)}$ is the entrywise inverse of each vector. This problem arises when using the Sinkhorn-Knopp Algorithm for transforming a square matrix into a doubly-stochastic matrix. I investigate the existence and classification of solutions to this problem in various cases involving matrices with specific structures. Beginning with some simple cases, including low-dimensions and simple patterns, e.g. circulant matrices, I make some conjectures about solutions in higher-dimensional and more complex or generalized cases. I will also explore relationships between solutions and eigenvectors of N, and prove a few simple theorems in the case that N is circulant.

Circulant Matrix (3x3) - Solutions

$$\vec{x}^{(1)} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \qquad \vec{x}^{(2)} = \begin{bmatrix} -\frac{1}{2}+\frac{\sqrt{3}}{2}i \\ -\frac{1}{2}-\frac{\sqrt{3}}{2}i \\ 1 \end{bmatrix} \qquad \vec{x}^{(3)} = \begin{bmatrix} -\frac{1}{2}-\frac{\sqrt{3}}{2}i \\ -\frac{1}{2}+\frac{\sqrt{3}}{2}i \\ 1 \end{bmatrix}$$

$$\vec{x}^{(4)} = \begin{bmatrix} \frac{ac^2-b^2c}{ab^2-bc^2} \\ \frac{a^2c-bc^2}{ac^2-a^2b} \\ 1 \end{bmatrix} \qquad \vec{x}^{(5)} = \begin{bmatrix} \frac{ac^2-a^2b}{a^2c-bc^2} \\ \frac{ab^2-a^2c}{a^2b-b^2c} \\ 1 \end{bmatrix} \qquad \vec{x}^{(6)} = \begin{bmatrix} \frac{a^2b-b^2c}{ab^2-a^2c} \\ \frac{ab^2-bc^2}{ac^2-b^2c} \\ 1 \end{bmatrix}$$

OR

$$\vec{x}^{(4)} = \begin{bmatrix} a-\frac{bc}{a} \\ c-\frac{ab}{c} \\ b-\frac{ac}{b} \end{bmatrix} \qquad \vec{x}^{(5)} = \begin{bmatrix} b-\frac{ac}{b} \\ a-\frac{bc}{a} \\ c-\frac{ab}{c} \end{bmatrix} \qquad \vec{x}^{(6)} = \begin{bmatrix} c-\frac{ab}{c} \\ b-\frac{ac}{b} \\ a-\frac{bc}{a} \end{bmatrix}$$

• Note: If these solutions are treated as the columns of another matrix N_2 then N_2 is also circulant.

$$N_2 = \begin{bmatrix} a-\frac{bc}{a} & b-\frac{ac}{b} & c-\frac{ab}{c} \\ c-\frac{ab}{c} & a-\frac{bc}{a} & b-\frac{ac}{b} \\ b-\frac{ac}{b} & c-\frac{ab}{c} & a-\frac{bc}{a} \end{bmatrix}$$

If we treat N_2 as our matrix N in our original equation:

$$\vec{x} = (N^T (N\vec{x})^{(-1)})^{(-1)}$$

$$= \left(\begin{bmatrix} a-\frac{bc}{a} & c-\frac{ab}{c} & b-\frac{ac}{b} \\ b-\frac{ac}{b} & a-\frac{bc}{a} & c-\frac{ab}{c} \\ c-\frac{ab}{c} & b-\frac{ac}{b} & a-\frac{bc}{a} \end{bmatrix}\left(\begin{bmatrix} a-\frac{bc}{a} & b-\frac{ac}{b} & c-\frac{ab}{c} \\ c-\frac{ab}{c} & a-\frac{bc}{a} & b-\frac{ac}{b} \\ b-\frac{ac}{b} & c-\frac{ab}{c} & a-\frac{bc}{a} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ 1 \end{bmatrix}\right)^{(-1)}\right)^{(-1)}$$

$$= \ldots$$

which gives us the following non-eigenvector solutions:

$$\vec{x}^{(4)} = \begin{bmatrix} \frac{ab-c^2}{ac-b^2} \\ \frac{ab-c^2}{cb-a^2} \\ 1 \end{bmatrix} \qquad \vec{x}^{(5)} = \begin{bmatrix} \frac{a^2-cb}{c^2-ab} \\ \frac{a^2-cb}{b^2-ac} \\ 1 \end{bmatrix} \qquad \vec{x}^{(6)} = \begin{bmatrix} \frac{ac-b^2}{cb-a^2} \\ \frac{ac-b^2}{ab-c^2} \\ 1 \end{bmatrix}$$

Notice: the reciprocal pattern between entries of subsequent solutions is still present.

The solutions $x^{(4)}$, $x^{(5)}$, and $x^{(6)}$ can be re-written as follows:

$$\vec{x}^{(4)} = \begin{bmatrix} \frac{1}{ac-b^2} \\ \frac{1}{bc-a^2} \\ \frac{1}{ab-c^2} \end{bmatrix} \qquad \vec{x}^{(5)} = \begin{bmatrix} \frac{1}{ab-c^2} \\ \frac{1}{ac-b^2} \\ \frac{1}{bc-a^2} \end{bmatrix} \qquad \vec{x}^{(6)} = \begin{bmatrix} \frac{1}{bc-a^2} \\ \frac{1}{ab-c^2} \\ \frac{1}{ac-b^2} \end{bmatrix}$$

• Once again, if these solutions are treated as the columns of another matrix N_3 then N_3 is also circulant.

$$N_3 = \begin{bmatrix} \frac{1}{ac-b^2} & \frac{1}{ab-c^2} & \frac{1}{bc-a^2} \\ \frac{1}{bc-a^2} & \frac{1}{ac-b^2} & \frac{1}{ab-c^2} \\ \frac{1}{ab-c^2} & \frac{1}{bc-a^2} & \frac{1}{ac-b^2} \end{bmatrix}$$

If we treat N_3 as our matrix N in our original equation:

$$\vec{x} = (N^T (N\vec{x})^{(-1)})^{(-1)}$$

$$= \left(\begin{bmatrix} \frac{1}{ac-b^2} & \frac{1}{bc-a^2} & \frac{1}{ab-c^2} \\ \frac{1}{ab-c^2} & \frac{1}{ac-b^2} & \frac{1}{bc-a^2} \\ \frac{1}{bc-a^2} & \frac{1}{ab-c^2} & \frac{1}{ac-b^2} \end{bmatrix}\left(\begin{bmatrix} \frac{1}{ac-b^2} & \frac{1}{ab-c^2} & \frac{1}{bc-a^2} \\ \frac{1}{bc-a^2} & \frac{1}{ac-b^2} & \frac{1}{ab-c^2} \\ \frac{1}{ab-c^2} & \frac{1}{bc-a^2} & \frac{1}{ac-b^2} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ 1 \end{bmatrix}\right)^{(-1)}\right)^{(-1)}$$

which gives us the following non-eigenvector solutions:

$$\vec{x}^{(4)} = \begin{bmatrix} a \\ c \\ b \end{bmatrix} \qquad \vec{x}^{(5)} = \begin{bmatrix} b \\ a \\ c \end{bmatrix} \qquad \vec{x}^{(5)} = \begin{bmatrix} c \\ b \\ a \end{bmatrix} \qquad \Longrightarrow N_4 = \begin{bmatrix} a & b & c \\ c & a & b \\ b & c & a \end{bmatrix}$$

(Our original circulant matrix!)

(Note: Support for this project came from the National Science Foundation via BYU's Center for Undergraduate Research in Mathematics)

Other Misc. Results (Shown w/out Proof)

N	Solutions
$\begin{bmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & f \end{bmatrix}$	$\vec{x} = \begin{bmatrix} \frac{c^2d-bce}{abe} \\ \frac{-c}{b} \\ 1 \end{bmatrix}$
$\begin{bmatrix} a & b & 0 \\ c & d & 0 \\ 0 & 0 & e \end{bmatrix}$	$\vec{x} = \begin{bmatrix} \pm\sqrt{\frac{bd}{ac}}x_2 \\ x_2 \\ 1 \end{bmatrix}$
$\begin{bmatrix} a & b & b \\ b & a & b \\ b & b & a \end{bmatrix}$	3x3 Circulant solutions $\vec{x}^{(1)}$ thru $\vec{x}^{(3)}$ and $\vec{x}^{(4)} = \begin{bmatrix} -\frac{a+b}{a} \\ 1 \\ 1 \end{bmatrix}$, $\vec{x}^{(5)} = \begin{bmatrix} 1 \\ -\frac{a+b}{a} \\ 1 \end{bmatrix}$, $\vec{x}^{(6)} = \begin{bmatrix} 1 \\ 1 \\ -\frac{a+b}{a} \end{bmatrix}$

Circulant Matrix (n x n) - Solutions

* **n^{th} roots of unity** (complex solutions to $z^n = 1$) are $\omega_m = e^{\frac{2\pi i}{n}m}$

$= \cos\left(\frac{2\pi m}{n}\right) + i\sin\left(\frac{2\pi m}{n}\right)$, for $m = 0, 1, \ldots, n-1$.

Note: $\omega_m{}^n = 1$ and $\frac{1}{\omega_m} = \omega_{n-m}$

FACT: (Davis, P.J. Circulant Matrices. Wiley, 1979)

The $n \times n$ circulant matrix $N = \begin{bmatrix} C_0 & C_1 & \cdots & C_{n-1} \\ C_{n-1} & C_0 & \cdots & C_{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ C_1 & C_2 & \cdots & C_0 \end{bmatrix}$ has n eigenvectors of the form

$\vec{x}_m = \begin{bmatrix} \omega_m \\ \omega_m{}^2 \\ \vdots \\ \omega_m{}^{n-1} \\ 1 \end{bmatrix}$ corresponding to eigenvalues $\lambda_m = C_0 + C_1\omega_m + C_2\omega_m{}^2 + \cdots + C_{n-1}\omega_m{}^{n-1}$ (for $m = 0, 1, \cdots, n-1$)

* Note: The eigenvalues λ_m depend on the entries of N, but the eigenvectors $\vec{x}_m$ do not.

Lemma: $\vec{x}_m$ is also an eigenvector of N^T, but corresponds to eigenvalue λ_{n-m}, $m = 0, 1, \cdots, n-1$

Proof:

N^T is also circulant, so $\vec{x}_m$ is an eigenvector of N^T. However, the corresponding eigenvalue is:

$$C_0 + C_{n-1}\omega_m + C_{n-2}\omega_m{}^2 + \cdots + C_1\omega_m{}^{n-1}$$

$$= C_0 + C_1\omega_m{}^{n-1} + C_2\omega_m{}^{n-2} + \cdots + C_{n-1}\omega_m$$

$$= C_0 + C_1\left(\omega_m{}^{n-1}\right) + C_2\left(\omega_m{}^{n-1}\right)^2 + \cdots + C_{n-1}\left(\omega_m{}^{n-1}\right)^{n-1}$$

$$= \lambda_{n-m} \quad \text{since } \omega_m{}^{n-1} = \omega_m{}^n\,\omega_m{}^{-1} = 1\left(\frac{1}{\omega_m}\right) = \omega_{n-m}$$

Theorem:

If N is an $n \times n$ circulant matrix, then each eigenvector $\vec{x}_m$ (for $m = 0, 1, \cdots, n-1$) is a solution to our problem.

Proof:

First observe $\vec{x}_m{}^{(-1)} = \vec{x}_{n-m}$ since $\frac{1}{(\omega_m)^k} = \left(\frac{1}{\omega_m}\right)^k = (\omega_{n-m})^k$

$$\text{Thus } (N^T (N\vec{x}_m)^{(-1)})^{(-1)} = (N^T (\lambda_m \vec{x}_m)^{(-1)})^{(-1)} = (N^T \frac{1}{\lambda_m}(\vec{x}_m)^{(-1)})^{(-1)}$$

$$= (\frac{1}{\lambda_m} N^T \vec{x}_{n-m})^{(-1)}$$

$$= (\frac{1}{\lambda_m} \lambda_{n-(n-m)} \vec{x}_{n-m})^{(-1)} \quad \text{(by the lemma)}$$

$$= (\vec{x}_{n-m})^{(-1)} = \vec{x}_{n-(n-m)} = \vec{x}_m$$