Group Inverses of Singular M–matrices and Applications

Michael Neumann, University of Connecticut

Novi Sad, May, 2010

Research Supported in part by NSA Grant No. 06G–232

From joint works with: Iddo Ben–Ari, Olga Pryporova, and Raymod Sze

This talk is dedicated to Hans Schneider

Introduction

We meet group (generalized) inverses of M–matrices in the following applications:

- **Analysis of convergence of iterative methods for singular systems** – through monotonicity conditions.
- **Markov chains** – expected times of return to a given state, analysis of stability of computations of distribution vectors.
- **Convexity and concavity of the Perron root and vector** – applications to population models.
- **Computations of nonnegative bases for Perron eigenspace**.
- **Graph theory** – estimation of algebraic connectivity of undirected graphs.

Brief Group Inverses

Any matrix $A \in \mathbb{C}^{n,n}$ admits the representation:

$$A = P \begin{bmatrix} C & 0 \\ 0 & N \end{bmatrix} P^{-1}, \quad \text{with} \quad \det(C) \neq 0 \text{ and } N^n = 0.$$

Drazin generalized inverse

$$A^D := P \begin{bmatrix} C^{-1} & 0 \\ 0 & 0 \end{bmatrix} P^{-1}$$

is the unique matrix satisfying the matrix equations:

$$XA^{k+1} = A^k, \quad XAX = X, \quad \text{and} \quad AX = XA,$$

where $k = \text{index}_0(N) = \text{index}_0(A)$.

$$\Rightarrow \quad AA^D = \mathcal{P}_{\mathcal{R}(A^k),\mathcal{N}(A^k)} \text{ and } I - AA^D = \mathcal{P}_{\mathcal{N}(A^k),\mathcal{R}(A^k)}$$

$$\Rightarrow \quad Ax = \lambda x \Leftrightarrow A^D x = \frac{1}{\lambda} x, \quad \lambda \neq 0.$$

When $k = 1$:

$$XA^2 = A \text{ and } AX = XA \Rightarrow AXA = A.$$

Then $X = A^D$ is the unique generalized $\{1,2\}$–inverse of A which commutes with A. X is then called the **group (generalized) inverse of** A and denoted by $A^{\#}$.

Theorem [Ben–Israel & Greville 1973, Campbell & Meyer 1981]:

Let $A \in \mathbb{C}^{n,n}$. Then TFAE:

(i) $A^{\#}$ exists.

(ii) The Jordan blocks of A corresponding to 0 are all 1×1.

(iii) $\mathcal{R}(A) \oplus \mathcal{N}(A) = \mathbb{C}^n$.

□

Comment: If $A \in \mathbb{C}^{n,n}$ is **Hermitian**, then $A^{\#}$ exists and $A^{\#} = A^{\dagger}$, the **Moore–Penrose generalized inverse** of A.

A Bit of History:

In a the celebrated 1975 paper, "The role of the group inverse in the theory of finite Markov chains", Meyer wrote in the abstract:

For an n–state homogeneous Markov chain whose one step transition matrix is T, the group–inverse, $A^{\#}$ of $A = I - T$, is shown to play a central role. For an ergodic chain, virtually everything that one would want to know about the chain can be determined from $A^{\#}$...

Examples:

- Let $\pi \in \mathbb{R}^{n,n}$ be the **stationary distribution vector for the chain**,

$\pi^t T = \pi^t$ and $\|\pi\|_1 = 1$. Then the rows of

$$W = I - AA^{\#}$$

are all equal to π^t.

- Let $M = (m_{i,j})$ be the mean first passage matrix. Then

$$M = \left[I - A^{\#} + JA^{\#}_{\rm dg}\right](\mathrm{dg}(\pi))^{-1} \Rightarrow m_{i,j} = \frac{A^{\#}_{j,j} - A^{\#}_{i,j}}{\pi_j}, \ \forall\, i \neq j.$$

- **The mean first passage time from state $\mathcal{S}_i$ to state $\mathcal{S}_j$ is the expected number of time–steps for reaching state $\mathcal{S}_j$ for the first time, given that initially the chain was in state $\mathcal{S}_i$.**

Formally, for $1 \leq i, j \leq n$:

$$m_{i,j} = E(F_{i,j}) = \sum_{k=1}^{\infty} kPr(F_{i,j} = k).$$

Here $F_{i,j}$ is the random variable representing the smallest number of time–steps for reaching state $\mathcal{S}_j$ for the first time, given that the system was initially in state $\mathcal{S}_i$. That is:

$$F_{i,j} = \min\{\ell \geq 1 : X_\ell = \mathcal{S}_j | X_0 = \mathcal{S}_i\}.$$

The matrix $M = (m_{i,j})$ is called the *mean first passage matrix* **of the chain.**

In contrast to the stationary distribution vector, the mean first passage times give us information about the short range behavior of the chain. For example, suppose that we go to a holiday destination and we find that it is raining. Our interest then is not the average number of rainy days vs sunny days throughout the year, but rather how long we can expect it to turn sunny, given that it is now raining ?

Another quantity which we can compute using the group inverse is **time to mixing of the chain**, namely, **the time until the Markov chain is "close" to its steady state distribution**, which is given by the quantity:

Kemeny constant:

$$K(T) = \sum_{\mathbf{i=1}}^{\mathbf{n}} \pi_{\mathbf{i}} \sum_{\mathbf{j=1}}^{\mathbf{n}} \mathbf{m_{i,j}} \pi_{\mathbf{j}} = \mathrm{trace}(A^{\#}) + 1 = \sum_{i=2}^{n} \frac{1}{1-\lambda_i} + 1,$$

where $\lambda_2, \ldots, \lambda_n$ are the eigenvalues of T other than 1.

Here we'll focus mainly on MFP matrices & connections to the inverse M–matrix problem.

But before we get deeper into this talk, let me devote a little time to where my work coincided with Hans' work. Hans, the teacher and a mentor of us all, to whom this conference is dedicated.

Iterative methods for regular M–splittings – Hans' contribution

Suppose that $A \in \mathbb{C}^{n,n}$ is **nonsingluar** and consider the problem of solving it by iteration the system $Ax = b$. We split

$$A = M - N, \quad \det(M) \neq 0,$$

and perform the iteration

$$x_i = M^{-1}Nx_{i-1} + M^{-1}b.$$

It is well known that the solution converges from every x_0 if and only if $\rho(M^{-1}N) < 1$.

Recall Varga's celebrated regular splitting result:

Theorem ([Varga 1958]) Let $A = M - N$ be a regular splitting for $A \in \mathbb{R}^{n,n}$, viz. $\exists\ M^{-1} \geq 0$ and $N \geq 0$. Then $\rho(M^{-1}N) < 1$ if and only if $\exists\ A^{-1} \geq 0$.

What if A is **singular?** Then

$$A = M - N = M(I - M^{-1}N) \Rightarrow \rho(M^{-1}N) \geq 1.$$

Examples:

$$A = \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix} \Rightarrow M^{-1}N = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}.$$

$$A = \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \Rightarrow M^{-1}N = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Theorem: ([N–Plemmons 1978])

Let $A = M - N$ be a regular splitting for $A \in \mathbb{R}^{n,n}$. Then $\rho(M^{-1}N) \leq 1$ and index$(M^{-1}N)_1 \leq 1$ if and only if

$$x \in \mathcal{R}(A),\ Ax \geq 0 \ \Rightarrow\ x \ \geq\ 0.$$

or, equivalently,

$$x \ \geq\ 0,\ x \in \mathcal{R}(A) \ \ \Rightarrow A^{\#}x \ \geq\ 0.$$

Discussion:

- The result used heavily the existence of a nonnegative basis for Perron eigenspace, established until then by Rothblum and, independently, by Richman and Schneider [1970s], but can be said to have been initiated by Schneider in his PhD.
- But the result above was not combinatorial. This lead me to seek a non–combinatorial proof of the existence of a nonnegative basis.
- (Hartwig, N., and Rose 1990): For sufficiently samll $\epsilon > 0$, the columns of

$$(A + \epsilon I)^{-1}[I - AA^{D}] \ \geq\ 0$$

contain a nonnegative basis. Here $T \geq 0$, $\rho(T) = 1$; $A = I - T$, and A^{D} is the Drazin inverse of A.
- Together with Hans we refined in 1994 the connection between the combinatorial approach and the analytic approach to the question of the existence of different nonnegative bases.

Partially motivated by the above Hans created in 1984 his beautiful notions of a graph compatible splittings and an M–splitting:

Defintion: $A \ = \ M - N$ is an M–splitting if M is an M–matrix and $N \geq 0$.

Theorem: Let A be a singular M–matrix and let $A = M - N$ be an M–splitting. Then:

- $\rho(M^{-1}N) \ = \ 1.$
- $\mathrm{mult}_1(M^{-1}N) \ = \ \mathrm{mult}_0(A).$
- $\mathrm{index}_1(M^{-1}N) \ = \ \mathrm{index}_0(A).$

Returning to Mean First Passage Times

Meyer 1975, Theorem 3.3

The MFP matrix for a Markov chain on n states whose transition matrix is $T \in \mathbb{R}^{n,n}$ is the unique solution of the matrix equation:

$$AX = J - TX_{\mathrm{dg}}$$

and the solution is given by

$$M = \left[I - A^{\#} + JA^{\#}_{\mathrm{dg}}\right] \Pi^{-1},$$

where $A = I - T$, $A^{\#}$ *is the group generalized inverse of* A, $\pi \in \mathbb{R}^{n,n}$ is **stationary distribution vector for the chain**, and where $\Pi = \mathrm{dg}(\pi)$.

Kemeny and Snell show [1960] :

If M is an MFP matrix induced by the transition matrix T, then the matrix $N := M - M_{\mathrm{dg}}$ is **invertible** and T can be represented in terms of M as follows:

$$T = I + (M_{\mathrm{dg}} - J)N^{-1}$$

A First Characterization

Theorem: Giving necessary and sufficient conditions for $M \gg 0$ to be an MFP matrix

Let $M \in \mathbb{R}^{n,n}$ be a positive matrix and set $N := M - M_{\rm dg}$. Then M is the MFP matrix for some Markov chain $\mathcal{C}$ whose transition matrix is T if and only if N is invertible and

$$\hat{T} \;=\; I + (M_{\rm dg} - J)N^{-1}$$

is a nonnegative stochastic matrix. In this case, $T = \hat{T}$.

An corollary showing the structure of the entries of N^{-1} is the following:

Corollary:

Soppose that $N \in \mathbb{R}^{n,n}$ is a nonnegative invertible matrix with zero diagonal entries. Let $N^{-1} = (p_{i,j})$. Then $N = M - M_{\rm dg}$ for some MFP matrix M of a Markov chain $\mathcal{C}$ on n states if and only if

$$\sum_{k=1}^{n} p_{i,k} \;>\; 0 \quad \text{and} \quad p_{i,j} \;\geq\; \frac{\sum_{k=1}^{n} p_{i,k} \sum_{k=1}^{n} p_{k,j}}{\sum_{1\leq k,\ell\leq n} p_{k,\ell}}, \quad \text{for all} \quad i \neq j.$$

Comments on the Corollary

We already mentioned that if we know that M is an MFP matrix, and hence $N = M - M_{\rm dg}$ is invertible, then we could compute the transition matrix T via the formula,

$$T \;=\; I + (M_{\rm dg} - J)N^{-1}.$$

We know that $M_{\rm dg}$ consists of the reciprocals of the stationary probabilities.

Thus it would seem that if we only have $N = M - M_{\rm dg}$ we would not be able to reconstruct from it T without the knowing π.

Fortunately this is not the case. For already in Meyer 1975 paper he shows that:

$$\pi^t \;=\; {\rm trace}(A^{\#})(N^{-1}e)^t,$$

and

$${\rm trace}(A^{\#}) \;=\; e^t N^{-1} e,$$

that is, the sum of all the entries in N^{-1}.

These ingredients used in proving the above corollary also lead to the result:

Observation: The diagonal entries of N^{-1} are all negative.

Several inequalities derive from this observation. For example: for each $i = 1, \ldots, n$,

$$\pi_i \;<\; \sum_{k=1}^{n} A^{\#}_{k,k}\left(a_{i,i} - a_{k,i}\right).$$

Connection Between MFP Problems and M–matrices

Meyer [1975], Cho & Meyer [2000], Ditzenbacher [1988]) show that:

If T is a transition matrix for a Markov chain, $A = I - T$, and M is its MFP matrix, then

$$\overline{M}_j \; := \; [m_{1,j}, \ldots, m_{j-1,j}, m_{j+1,j} \ldots, m_{n,j}]^T \; = \; A_j^{-1} e,$$

where A_j the the $(n-1) \times (n-1)$ principal submatrix of A obtained by deleting its j–th row and column. We know that A_j is a nonsingular diagonally dominant M–matrix.

But most suggestively:

Theorem: Tetali [1994]

Let $T = (t_{i,j}) \in \mathbb{R}^{n,n}$ be a transition matrix for a Markov chain with $T_{\rm dg} = 0$. Let and $A = I - T$ and $\Pi_n = {\rm dg}(\pi_1, \ldots, \pi_{n-1})$. Then

$$(\Pi_n A_n) H \; = \; I_{n-1},$$

where $H \; = \; (h_{i,j})$ is given by:

$$\mathbf{h_{i,j}} \; = \; \begin{cases} m_{i,n} + m_{n,i}, & \text{if } i = j, \\ m_{i,n} + m_{n,j} - m_{i,j}, & \text{if } i \neq j, \end{cases}$$

Connection Between MFP Problems and M–matrices, Continued

Note: $A_n \in \mathbb{R}^{n,n}$ is a row diagonally dominant M–matrix $\Rightarrow \Pi_n A_n$ is a row diagonally dominant M–matrix $\Rightarrow H^{-1} = \Pi_n A_n$ is a row diagonally dominant M–matrix.

Claim: H^{-1} is (also) a column diagonally dominant M–matrix.

Proof: Partition π as $\pi = [\bar{\pi}^t \mid \pi_n]^t$, where $\bar{\pi} \in R^{n-1}$. Observe first that because $0 = [\bar{\pi}^t \mid \pi_n] \left[\begin{array}{c|c} A_n & * \\ \hline a_{n,1}, \ldots, a_{n,n-1} & * \end{array}\right]$ and A is an M–matrix, we can write that:

$$\bar{\pi}^t A_n = -\pi_n [a_{n,1} \ldots a_{n,n-1}] \geq 0.$$

But $\Pi_n A_n = H^{-1}$ and so:

$$e^T H^{-1} = e^t \Pi_n A_n = \bar{\pi}^t A_n \geq 0.$$

Towards a Second Characterization for Being an MFP Matrix

Given a positive matrix $M = (m_{i,j}) \in \mathbb{R}^{n,n}$, define the matrix $H \in \mathbb{R}^{n-1,n-1}$ by:

$$\mathbf{h_{i,j}} = \begin{cases} m_{i,n} + m_{n,i}, & \text{if } i = j, \\ \\ m_{i,n} + m_{n,j} - m_{i,j}, & \text{if } i \neq j. \end{cases}$$

One can immediately show that a Tetali–type theorem holds here if M is assumed to be an MFP matrix, without the restriction that $T_{\mathrm{dg}} = 0$.

Theorem:

Suppose that T is the transition matrix of a Markov chain $\mathcal{C}$ on n states with the MFP matrix M and the stationary vector $\pi = (\pi_1, \ldots, \pi_n)^t$. Let $A = I - T$ and set $\Pi_n = \mathrm{diag}\,(\pi_1, \ldots, \pi_{n-1})$. Then

$$(\Pi_n A_n)H = I.$$

Let M be a positive matrix and let

$$P = [I_{n-1} \;\; -e] \in \mathbb{R}^{(n-1),n}.$$

Then one can check that

$$H = -P(M - M_{\mathrm{dg}})P^t = \begin{cases} m_{i,n} + m_{n,i}, & \text{if } i = j, \\ m_{i,n} + m_{n,j} - m_{i,j}, & \text{if } i \neq j. \end{cases}$$

We already saw that when H is derived from an MFP matrix $M \in R^{n,n}$, then H is the inverse of a row and column diagonally dominant M–matrix.

We are now ready for our main result:

Theorem:

Let $H \in \mathbb{R}^{(n-1),(n-1)}$. Then the following are equivalent:

(a) H is invertible and H^{-1} is an irreducible row and column diagonally dominant M–matrix, and

$$\mathrm{trace}((I + J)H^{-1}) \leq 1.$$

(b) There exists a Markov chain $\mathcal{C}$ on n states with transition matrix $T \in \mathbb{R}^{n,n}$ and a stationary distribution vector $\pi = (\pi_1, \ldots, \pi_n)$ such that

$$(\Pi_n A_n)H = I,$$

where $A = I - T$ and where H^{-1} is irreducible.

(c) There exists an MFP matrix M of a Markov chain $\mathcal{C}$ such that

$$H = -P(M - M_{\mathrm{dg}})P^t \gg 0.$$

Proof that (a) $\Rightarrow$ (b)

$$\operatorname{trace}\left(H^{-1}\right) \;<\; \operatorname{trace}\left(H^{-1}\right) + \operatorname{trace}\left(JH^{-1}\right) \;\le\; \operatorname{trace}\left((I+J)H^{-1}\right) \;\le\; 1.$$

Let $d_1, \ldots, d_{n-1}$ be the diagonal entries of H^{-1}. Then we see that

$$\sum_{j=1}^{n-1} d_j \;<\; 1.$$

We can now choose positive numbers $\pi_1, \ldots, \pi_n$, with $\sum_{j=1}^{n} \pi_j = 1$, such that

$$\pi_j \;\ge\; d_j, \quad \text{for } j = 1, \ldots, n-1.$$

Set $\pi = (\pi_1, \ldots, \pi_n)^t$, $\Pi := \operatorname{diag}(\pi_1, \ldots, \pi_n)$, and

$$\mathbf{T} \;:=\; \mathbf{I} - \mathbf{\Pi}^{-1}\mathbf{P}^t\mathbf{H}^{-1}\mathbf{P} \;=\; \mathbf{I} - \mathbf{\Pi}^{-1} \begin{bmatrix} H^{-1} & -H^{-1}e \\ -e^tH^{-1} & e^tH^{-1}e \end{bmatrix}.$$

One can readily check that T is nonnegative and irreducible, $Te = e$, and $\pi^tT = \pi^t$. Thus T is a transition matrix for some Markov chain whose stationary distribution is the vector π. Furthermore, we have that $\Pi_n A_n = H^{-1}$, where $A = I - T$.

Main Corollary:

Suppose that $A = (a_{i,j}) \in \mathbb{R}^{n \times n}$. Then the following conditions are equivalent:

(a) A is invertible and A^{-1} is an irreducible row and column diagonally dominant M–matrix.

(b) There exists a Markov chain $\mathcal{C}$ on $n+1$ states and a constant $k > 0$ such that

$$0 < a_{i,j} = \begin{cases} k\,(m_{i,n+1} + m_{n+1,j} - m_{i,j})\,, & \text{if} \quad i \neq j, \\ k\,(m_{i,n+1} + m_{n+1,j})\,, & \text{if} \quad i = j, \end{cases}$$

where $M = (m_{i,j}) \in \mathbb{R}^{n+1,n+1}$ is the MFP matrix for the chain.

Extending Also Fiedler's Results

In 1998 Fiedler characterized inverses of symmetric diagonally dominant M–matrices in terms of **resistive electrical networks**.

Theorem ([Fiedler 1998])

Let $B = (b_{i,j}) \in \mathbb{R}^{n,n}$ be a real $n \times n$ matrix. Then the following are equivalent.

(a) B is invertible and B^{-1} is an irreducible diagonally dominant symmetric M–matrix.

(b) There is a connected resistive network $\mathcal{N}(G)$ with $n+1$ nodes $1, \ldots, n+1$ such that the *effective resistances* $R_{i,j}$ satisfy

$$b_{i,j} = \frac{1}{2}[R_{i,n+1} + R_{n+1,j} - R_{i,j}], \quad i,j = 1, \ldots, n.$$

Remark: Here $R_{i,i} = 0$, for all $i = 1, \ldots, n$.

Extending Fiedler's Results Continued

With the connected resistive network $\mathcal{N}(G)$ one can associate a random walk using the probabilities:

$$t_{i,j} \;=\; \frac{c_{i,j}}{\sum_{k\in V} c_{i,k}}, \quad i,j = 1,\ldots,n+1,$$

where $c_{i,j}$ denotes the conductance between node i and node j in the resistive network.

To connect Fiedler's result with our work we need the following 1996 proposition due to Chandra, Raghavan, Ruzzo, Smolensky, and Tiwari:

Proposition: Chandra, Raghavan, Ruzzo, Smolensky, and Tiwari, [1996]

Suppose that $\mathcal{N}(G)$ is a connected resistive network. Set:

$$\hat{C} \;:=\; \sum_{(i,j)\in V\times V} c_{i,j}.$$

Then for any two distinct nodes $i,j \in V$,

$$2m_{i,j} \;=\; \hat{C}R_{i,j},$$

Completing the arguments

$$\text{With} \qquad k \;:=\; \frac{1}{\hat{C}}$$

we find that for $i \neq j$, i,j=1,. . . , n+1,

$$km_{i,j} \;=\; \frac{1}{2}\hat{C}R_{i,j} \;=\; \frac{1}{2}R_{i,j}$$

and so for $i \neq j$, we have that from Fiedler's condition on the $a_{i,j}$'s that:

$$a_{i,j} \;=\; \frac{1}{2}\left(R_{i,n+1} + R_{n+1,j} - R_{i,j}\right) \;=\; k\left(m_{i,n+1} + m_{n+1,j} - m_{i,j}\right),$$

while for $i = j$ we see that

$$a_{i,i} \;=\; \frac{1}{2}\left(R_{i,n+1} + R_{n+1,i} - \underbrace{R_{i,i}}_{=\,0}\right) \;=\; k(m_{i,n+1} + m_{n+1,i}).$$

Hence Fiedler's conditions lead to satisfying the condition of our Main Corollary with $k = \hat{C}$. Moreover, we observe that since A is symmetric in Fieldler's assumptions, it is the inverse of a row and column diagonally dominant M–matrix.

A Question of Kemeny and Snell

Question (Kemeny and Snell [1960]):

Given the M is an MFP for a Markov chain, when does M or $N = M - M_{\rm dg}$ represent a **regular chain**, i.e., ergodic, but not cyclic ?

This is the hard question and is not unlike the question if we just look at the positive inverse of an irreducible M–matrix, when is it the inverse of a primitive or just the inverse of a cyclic M–matrix?

Since we have a formula for T, viz., $T = I + (M_{\rm dg} - J)N^{-1}$, we know that (the irreducible matrix) T will be primitive if one of its diagonal entries are positive. What we can do is begin to answer when will all the diagonal entries of T be zero:

Observation: All the diagonal entries of T are zero if and only if:

$$\mathrm{trace}(N^{-1}) = \frac{1}{\mathrm{trace}(A^{\#})}\sum_{i=1}^{n}(N^{-1}e)_i^2 + \frac{1}{\mathrm{trace}^2(A^{\#})}\sum_{i=1}^{n}\sum_{j=1}^{n} A^{\#}_{i,i}a_{i,j}\left(N^{-1}e\right)_j - 1.$$