On the Drazin inverse of the rate matrix

Tech. Note 011 v3 beta
http://threeplusone.com/drazin

Gavin E. Crooks

2018-11-07

Contents

1 The Drazin inverse **1**
1.1 Representation via Core-nilpotent decomposition . . . 1
1.2 Projection operator . . . 2
1.3 Group inverse . . . 2
1.4 Diagonalizable matices . . . 3

2 Transition Rate matrix **3**
2.1 Drazin inverse of the rate matrix . . . 3
2.2 Relaxation times . . . 4
2.3 Instantaneous distribution . . . 4
2.4 Time reversal . . . 4

3 Thermodynamic geometry **4**
3.1 Fisher Information . . . 5
3.2 Currents . . . 5

1 The Drazin inverse

The inverse of a square matrix A is the unique matrix X (if it exists) such that

$$AX = XA = I. \tag{1}$$

A generalized or pseudo-inverse [1, 2] is an operation that has some of the properties of the usual inverse, is the same as the usual inverse if the matrix is nonsingular, but that can be applied to a larger class of singular matrixes.

Different generalized inverses are traditionally characterized by some subset of the follow conditions:

$$\begin{aligned}
AXA &= A && \{1\}\\
XAX &= X && \{2\}\\
(AX)^* &= AX && \{3\}\\
(XA)^* &= XA && \{4\}\\
AX &= XA && \{5\}\\
\\
A^kXA &= A^k && \{1^k\}
\end{aligned}$$

The most commonly encountered variant is the Moore-Penrose pseudo-inverse, which is sometimes denoted the $\{1,2,3,4\}$-inverse since it satisfies conditions 1, 2, 3, and 4. Herein, we're interested in Drazin's $\{1^k,2,5\}$-inverse [3], which satisfies conditions 2 and 5, and the alternative first condition 1^k, where k is some positive integer. The $\{1,2,5\}$-inverse is a common special case of the Drazin inverse, often called the group-inverse.

Although the Drazin inverse is a general construction, we'll restrict our attention to its realization in the linear algebra of matrixes (Since that's where applications of interest to us are to be found). Given a matrix A, we'll denote the $\{1^k,2,5\}$ Drazin inverse as $A^{\mathbf{D}}$, and the $\{1,2,5\}$ group inverse (if we are so restricted) as $A^{\times}$. (Note that the Moore-Penrose inverse can be applied to rectangular matrices, but condition $\{5\}$ (commutation) restricts the Drazin inverse to square matrices.)

The Drazin inverse is unique. Suppose X and Y are both Drazin inverses of A. Then [1]

$$\begin{aligned}
X &= XAX = XAXAX = \ldots = X(AX)^k = A^kX^{k+1} && \text{(2a)}\\
&= A^k(YA)X^{k+1} = \ldots = A^k(YA)^{k+1}X^{k+1} && \text{(2b)}\\
&= Y^{k+1}A^{2k+1}X^{k+1} && \text{(2c)}\\
&= Y^{k+1}A^k(XA)^{k+1} = \ldots = Y^{k+1}A^k(XA) && \text{(2d)}\\
&= Y^{k+1}A^k = (YA)^kY = \ldots = YAYAY = YAY = Y && \text{(2e)}
\end{aligned}$$

(a) We use property $\{2\}$ to append k copies of (XA), then use commutation $\{5\}$ to gather terms. (b) We can then use property $\{1^k\}$ to append $k+1$ copies of (YA). (c) Apply commutation again, we get an expression that is symmetric between X and Y. (d) Reversing the operations, we now remove $k+1$ copies of (XA) (e) And then remove k copies of (YA) to yield Y.

1.1 Representation via Core-nilpotent decomposition

Any square matrix A (or, more generally, any linear operator) can be decomposed into the Jordan canonical form, $A = SJS^{-1}$, where S is some invertible matrix, and J is a block diagonal matrix

$$J = J_{n_1,\lambda_1} \oplus J_{n_2,\lambda_2} \oplus \cdots \oplus J_{n_m,\lambda_m} \tag{3}$$

Each $J_{n,\lambda}$ is a Jordan block, an $n \times n$ matrix with λ on the diagonal, and ones on the super-diagonal. For instance,

$$J_{3,2} = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix} \tag{4}$$

An invertible matrix is nonsingular – it has no zero eigenvalues. The eigenvalues of a Jordan block are all λ, and thus the block is invertible if λ isn't zero. The singular blocks are nilpotent: raised to the nth (or higher) power they yield the zero matrix.

$$(J_{n,0})^n = \mathbb{O}_{n\times n} \tag{5}$$

Core-nilpotent decomposition: Any square matrix can therefore be expressed as

$$A = S[N \oplus C]S^{-1} \,, \tag{6}$$

where S and C are invertible, and N is nilpotent. The *index* k of a square matrix[1] is zero if the matrix in nonsingular, else the size of the largest singular block. Thus a singular matrix raised to kth (or higher) power A^k will have index 1, and the entire nilpotent sector N will be zero.

In this representations the Drazin inverse is

$$A^{\mathbf{D}} = S[\mathbb{O} \oplus C^{-1}]S^{-1} \tag{7}$$

which satisfies the Drazin requirements

$$\begin{aligned} A^k A^{\mathbf{D}} A &= A^k && \{1^k\} \\ A^{\mathbf{D}} A A^{\mathbf{D}} &= A^{\mathbf{D}} && \{2\} \\ A^{\mathbf{D}} A &= A A^{\mathbf{D}} && \{5\} \end{aligned}$$

Intuitively, the Drazin inverse inverts the invertible part of the matrix, and zeros the non-invertible part. The Jordan block decomposition is unique, up to permutation of the blocks. It follows that the Drazin inverse exists for all square matrices, and is unique.

1.2 Projection operator

Perhaps the most important property of the Drazin inverse is that $P_0 = I - AA^{\mathbf{D}}$ is a projection operator ($(I - AA^{\mathbf{D}})^2 = I - AA^{\mathbf{D}}$) that projects onto the null-space of A. (The *null-space* (or *kernel*) is the set of vectors x for which $Ax = 0$.)

$$\begin{aligned} P_0 = I - AA^{\mathbf{D}} &= I - S[N \oplus C]S^{-1}\, S[\mathbb{O} \oplus C^{-1}]S^{-1} && \text{(9a)} \\ &= I - S[\mathbb{O} \oplus I]S^{-1} = S[I \oplus \mathbb{O}]S^{-1} && \text{(9b)} \end{aligned}$$

[1]The index of a matrix is the smallest positive integer for which $\operatorname{rank} A^k = \operatorname{rank} A^{k+1}$

The projector onto the complimentary non-null-space is $AA^{\mathbf{D}}$.

If we know the null-space projection operator (which we often do), we can calculate the Drazin inverse as [4]

$$A^{\mathbf{D}} = (I - P_0)(A - P_0)^{-1} \tag{10}$$

Proof:

$$\begin{aligned} A^{\mathbf{D}} &= (I - P_0)(A - P_0)^{-1} && \text{(11a)} \\ &= \big(S[\mathbb{O} \oplus I]S^{-1}\big)\big(S[N \oplus C]S^{-1} - S[I \oplus \mathbb{O}]S^{-1}\big)^{-1} && \text{(11b)} \\ &= \big(S[\mathbb{O} \oplus I]S^{-1}\big)\big(S[(N - I) \oplus C]S^{-1}\big)^{-1} && \text{(11c)} \\ &= \big(S[\mathbb{O} \oplus I]S^{-1}\big)\big(S[(N - I)^{-1} \oplus C^{-1}]S^{-1}\big) && \text{(11d)} \\ &= S[\mathbb{O} \oplus I][(N - I)^{-1} \oplus C^{-1}]S^{-1} && \text{(11e)} \\ &= S[\mathbb{O} \oplus C^{-1}]S^{-1} && \text{(11f)} \end{aligned}$$

1.3 Group inverse

If the index of a matrix is 1, then the nil-potent sector is the zero matrix ($A = S[\mathbb{O} \oplus C]S^{-1}$), and the Drazin inverse reduces to the $\{1, 2, 5\}$ *group inverse* [5].

$$\begin{aligned} AA^{\times}A &= A && \{1\} \\ A^{\times}AA^{\times} &= A^{\times} && \{2\} \\ A^{\times}A &= AA^{\times} && \{5\} \end{aligned}$$

The name "group inverse" arises because the positive powers of A and $A^{\times}$ form an Abelian group with $AA^{\times}$ as the identity element. Both the notations $A^{\#}$ and A^{+} are commonly encountered, although the latter is also commonly used for the Moore-Penrose pseudoinverse.

The group inverse, if it exists, is unique. Suppose X and Y are both group inverses of A. Then [2]

$$\begin{aligned} X &= XAX = AXX = AYAXX = YAXAX \\ &= YAX \\ &= YAYAX = YYAXA = YYA = YAY = Y \end{aligned} \tag{13}$$

This is just a simplification of the proof used for the Drazin inverse (2).

For the group inverse we also have [0]

$$A^{\times} = P_0 + (A - P_0)^{-1} \tag{14}$$

Proof:

$$\begin{aligned} A^{\times} &= S[I \oplus \mathbb{O}]S^{-1} + \big(S[\mathbb{O} \oplus C]S^{-1} - S[I \oplus \mathbb{O}]S^{-1}\big)^{-1} && \text{(15a)} \\ &= S[I \oplus \mathbb{O}]S^{-1} + \big(S[(-I) \oplus C]S^{-1}\big)^{-1} && \text{(15b)} \\ &= S[I \oplus \mathbb{O}]S^{-1} + S[(-I) \oplus C^{-1}]S^{-1} && \text{(15c)} \\ &= S[\mathbb{O} \oplus C^{-1}]S^{-1} && \text{(15d)} \end{aligned}$$

1.4 Diagonalizable matices

If a matrix is diagonalizable then we can decompose the matrix as $A = \sum_{\lambda \neq 0} \lambda P_\lambda$, where λ are the eigenvalues, and the P's are an orthonormal collection of projection operators $P_\lambda P_\lambda = P_\lambda$, $P_\lambda P_{\lambda'} = 0$ (if $\lambda \neq \lambda'$), $\sum_\lambda P_\lambda = I$. For a diagonalizable matrix, all the Jordan blocks are of size 1, and the index of the matrix is zero (if invertable) or one. Thus the Drazin inverse of a diagonalizable matrix reduces to the group inverse.

The group inverse has a simple realization for diagonalizable matrices.

$$A = \sum_{\lambda \neq 0} \lambda P_\lambda \tag{16a}$$

$$A^\times = \sum_{\lambda \neq 0} \frac{1}{\lambda} P_\lambda \tag{16b}$$

Intuitively, we simple invert all of the invertible eigenvalues.

$$AA^\times A = \sum_{\lambda \neq 0} \lambda P_\lambda = A \tag{17a}$$

$$A^\times A A^\times = \sum_{\lambda \neq 0} \frac{1}{\lambda} P_\lambda = A^\times \tag{17b}$$

$$A^\times A = \sum_{\lambda \neq 0} P_\lambda = I - P_0 = AA^\times \tag{17c}$$

2 Transition Rate matrix

The dynamics of a continuous time, discrete state Markov process can be conveniently described by a transition rate matrix, $R(t)$. The evolution of the probability density is given by a first order differential equation

$$\frac{\partial}{\partial t}\rho(t) = R(t)\rho(t) \tag{18}$$

or alternatively for a temporally homogeneous dynamics, the evolution of the probability distributed during an interval of duration τ can be written as

$$\rho(t+\tau) = e^{\tau R}\rho(t) \tag{19}$$

Note that the rate matrix acts from the left. Be wary, since the opposite convention is also common, particularly in the Markov process literature.

The off-diagonal elements of a rate matrix are are non-negative, and all columns sum to zero.

$$\begin{aligned} R_{fi} &\geqslant 0 && \text{for all } i \neq f \\ \sum_f R_{fi} &= 0 && \text{for all } f \end{aligned}$$

The off-diagonal elements R_{fi} give the rate of going from state i to state f. When a jump occurs, the probability of transiting from state i to state f is proportional to R_{fi}. The waiting time between jumps follows an exponential distribution with a rate given by the negative of the diagonal elements $-R_{ii}$.

If we fix the rate matrix and allow the system to evolve for a long time then the ensemble will reach a stationary distribution which no longer changes with time. We will denote a stationary distribution of the dynamics by the vector π.

$$e^{\tau R}\pi = \pi, \qquad R\pi = 0 \tag{20}$$

We'll generally assume that the rate-matrix is *ergodic* and therefore that the stationary distribution is unique. (In this context *ergodic* means that the system can eventually reach any state from any starting state.) It follows that the rate matrix has a single zero eigenvalue, and that the index of the matrix is one.

Given a transition rate matrix R we can also construct a stochastic matrix for any time interval τ.

$$e^{\tau R} = M^\tau \tag{21}$$

In the infinite time limit the stochastic matrix relaxes the system to the stationary state.

$$\lim_{\tau \to \infty} e^{\tau R} \equiv M^\infty = \pi \mathbf{1}^T \tag{22}$$

Here $\mathbf{1}$ is a vector of ones.

Given that the stationary distribution is unique, then R has one zero eigenvalue, with corresponding projection $P_0 = \pi\mathbf{1}^T = M^\infty$.

2.1 Drazin inverse of the rate matrix

Since the rate matrix has a zero eigenvalue, it cannot be inverted. However, we can construct the *Drazin inverse* of the rate matrix. And since the index of a ergotic rate matrix is one, the Drazin inverse reduces to the $\{1, 2, 5\}$ group inverse.

The most important property here is that $R^\times R$ projects onto the non-null space of R. In other words

$$R^\times R = I - \pi\mathbf{1}^T \tag{23}$$

$$R^\times R\rho = \rho - \pi = \delta\rho. \tag{24}$$

Applying $R^\times R$ to any probability returns the difference of that probability from the stationary distribution. (Note again that we assume that R has only one zero eigenvalue, and therefore a unique stationary state.)

Also

$$R^\times \pi = \mathbf{0} \tag{25}$$

$$\mathbf{1}^T R^\times = \mathbf{0}^T \tag{26}$$

These properties follow from $R^\times R = R\,R^\times = I - \pi\mathbf{1}^T$

$$R^\times R\, R^\times = R^\times(I - \pi\mathbf{1}^T) = (I - \pi\mathbf{1}^T)R^\times \tag{27a}$$
$$R^\times = R^\times - R^\times\pi\mathbf{1}^T = R^\times - \pi\mathbf{1}^T R^\times \tag{27b}$$
$$0 = R^\times\pi\mathbf{1}^T = \pi\mathbf{1}^T R^\times \tag{27c}$$
$$R^\times\pi = \mathbf{0}, \qquad \mathbf{1}^T R^\times = \mathbf{0}^T \tag{27d}$$

(Note that Mandal and Jarzynksi (2015) [6] use these properties, plus commutation, to define the Drazin inverse.)

2.2 Relaxation times

The rate matrix encodes transition rates between states. Conversely, the pseudoinverse matrix $R^\times$ tells us the characteristic relaxation times for the process. First note the following integral [7, Eq. 6.3],

$$\begin{aligned} R^\times R \int_0^\infty e^{tR} dt &= R^\times \int_0^\infty R e^{tR} dt \\ &= R^\times (e^{tR}|_0^\infty) \\ &= R^\times(M^\infty - I) = -R^\times R R^\times \\ &= -R^\times \end{aligned} \tag{28}$$

As $R^\times R$ projects onto the non-null space we can write the relaxation of any initial probability to the stationary probability as

$$e^{tR}\ \delta p(0) = e^{tR} R^\times R\ \delta p(0) = \delta p(t) \tag{29}$$

The exponential is just a polynomial in R, and R and $R^\times$ commute, so we can move the projector to the left of the exponential. We then integrate w.r.t. time

$$\int_0^\infty dt\ R^\times R\ e^{tR}\ \delta p(0) = \int_0^\infty dt\ \delta p(t) \tag{30}$$

The left side simplifies due to the previous remarked integral.

$$\int_0^\infty dt\ R^\times R\ e^{tR}\ p(0) = R^\times R \int_0^\infty e^{tR} dt\ \delta p(0) = R^\times\ \delta p(0) \tag{31}$$

and on the right with have a vector $t^{\text{eff}}\delta p(0)$, where τ^{eff} are the effective relaxation times of each microstate. Thus

$$R^\times \delta p(0) = \tau^{\text{eff}} \delta p(0) \tag{32}$$

(Kudos: This basic idea is in the Supplementary Materials of Mandal2015 [6].)

2.3 Instantaneous distribution

Suppose we have a non-temporally homogeneous dynamics where the rate matrix R(t) is a function of time. Lets assume continuously differentiable control. Then we can express the instantaneous distribution in terms of the Drazin inverse of the rate matrix.

$$\tfrac{d}{dt} p(t) = R p(t) \tag{33a}$$
$$R^\times \tfrac{d}{dt} p(t) = R^\times R p(t) = p(t) - \pi(t) \tag{33b}$$
$$(I - R^\times \tfrac{d}{dt}) p(t) = \pi(t) \tag{33c}$$
$$\begin{aligned} p(t) &= \left(I - R^\times \tfrac{d}{dt}\right)^{-1} \pi(t) \\ &= \sum_{n=0}^{\infty} \left(R^\times \tfrac{d}{dt}\right)^n \pi(t) \end{aligned} \tag{33d}$$

(a) Definition (b) multiple both sides by $R^\times$. On the right the projection $R^\times R$ yeilds δp. (c) gather like terms (d) This follows by treating $(1-x)^{-1} = 1 + x + x^2 + \cdots$ as a formal power series generating function. This relation is formally exact.

The first term in the power series is the identity, so we can also write

$$\delta p(t) = \sum_{n=1}^{\infty} \left(R^\times \tfrac{d}{dt}\right)^n \pi \tag{34}$$

(Kudos: This is a slightly simplified derivation and result from what is found in Mandal and Jarzynski (2015) [6])

2.4 Time reversal

Let R be an ergotic rate matrix with steady state π. The time reversal of the rate matrix is $\widetilde{R} = D_\pi^{-1} R^T D_\pi$, where D_π is a diagonal matrix with the vector π along the diagonal [8, 9].

The Drazin inverse commutes with time reversal [10].

$$\widetilde{R}^\times = \left(D_\pi^{-1} R^T D_\pi\right)^\times = \left(D_\pi^{-1} (R^\times)^T D_\pi\right) = \widetilde{(R^\times)} \tag{35}$$

3 Thermodynamic geometry

Let $D_\rho = \text{diag}(\rho)$ be the diagonal matrix with the vector ρ along the diagonal. Then we can compactly express the relative entropy between the instantaneous and equilibrium distributions as

$$D(\rho\|\pi) = \mathbf{1}^T(\ln D_p - \ln D_\pi)\rho \tag{36}$$

This relative entropy is the free energy difference between the instantaneous and equilibrium ensembles [0, 0].

Rate of mean excess work is $\langle\beta\dot{W}^{\text{ex}}\rangle = \frac{d\lambda}{dt}\frac{d}{d\lambda}D(p\|\pi)$.

$$\begin{aligned}
\langle\beta\dot{W}^{\text{ex}}\rangle &= \tfrac{d\lambda}{dt}\tfrac{d}{d\lambda}D(p\|\pi) && \text{(37a)}\\
&= \tfrac{d\lambda}{dt}\tfrac{d}{d\lambda}\mathbf{1}^T(\ln D_p - \ln D_\pi)p && \text{(37b)}\\
&= -\tfrac{d\lambda}{dt}\cdot\mathbf{1}^T(\tfrac{d}{d\lambda}\ln D_\pi)p && \text{(37c)}\\
&= -\tfrac{d\lambda}{dt}\cdot\mathbf{1}^T(\tfrac{d}{d\lambda}\ln D_\pi)\left(I - R^\times\tfrac{d}{dt}\right)^{-1}\pi && \text{(37d)}\\
&= -\tfrac{d\lambda}{dt}\cdot\mathbf{1}^T(\tfrac{d}{d\lambda}\ln D_\pi)\left(I + R^\times\tfrac{d}{dt}+\cdots\right)\pi && \text{(37e)}\\
&\approx \tfrac{d\lambda}{dt}\cdot\mathbf{1}^T(\tfrac{d}{d\lambda}\ln D_\pi)(-R^\times)\tfrac{d}{dt}\pi && \text{(37f)}\\
&\approx \tfrac{d\lambda}{dt}\cdot\mathbf{1}^T(\tfrac{d}{d\lambda}\ln D_\pi)(-R^\times)\tfrac{d\lambda}{dt}\tfrac{d}{d\lambda}\pi && \text{(37g)}\\
&\approx \tfrac{d\lambda}{dt}\cdot\mathbf{1}^T(\tfrac{d}{d\lambda}\ln D_\pi)(-R^\times)(\tfrac{d}{d\lambda}\ln D_\pi)\pi\cdot\tfrac{d\lambda}{dt} && \text{(37h)}\\
&\approx \tfrac{d\lambda}{dt}\cdot\zeta\cdot\tfrac{d\lambda}{dt} && \text{(37i)}
\end{aligned}$$

(a) Definition of mean excess work
(b) Insert definition of relative entropy (36)
(c) Derivative w.r.t. to λ only hits the score of the equilibrium distribution
(d) Insert expression for the instantaneous probability (34)
(e) Expand the generating function …
(f) and truncate at second term. First term is identically zero since $\mathbf{1}^T\frac{d}{d\lambda}\pi = 0$ due to normalization.
(g) chain rule
(h) $(\frac{d}{d\lambda}\ln D_\pi)\pi = (\frac{d}{d\lambda}D_\pi)D_\pi^{-1}\pi, = (\frac{d}{d\lambda}D_\pi)\mathbf{1} = \frac{d}{d\lambda}\pi$.
(i) Define the friction ζ ('zeta')

$$\begin{aligned}
\zeta &= \mathbf{1}^T(\tfrac{d}{d\lambda}\ln D_\pi)(-R^\times)(\tfrac{d}{d\lambda}\ln D_\pi)\pi && \text{(38a)}\\
&= \int_0^\infty dt\ \mathbf{1}^T(\tfrac{d}{d\lambda}\ln D_\pi)(e^{tR}RR^\times)(\tfrac{d}{d\lambda}\ln D_\pi)\pi && \text{(38b)}\\
&= \int_0^\infty dt\ \mathbf{1}^T(\tfrac{d}{d\lambda}\ln D_\pi)\ e^{tR}\ (\tfrac{d}{d\lambda}\ln D_\pi)\ \pi && \text{(38c)}\\
&= \int_0^\infty dt\ \left\langle(\tfrac{d}{d\lambda}\ln D_\pi(t))\ (\tfrac{d}{d\lambda}\ln D_\pi(0))\right\rangle && \text{(38d)}
\end{aligned}$$

(c) Note that $RR^\times = (I-\pi\mathbf{1}^T)$, and that $\mathbf{1}^T\frac{d}{d\lambda}\pi = 0$ by conservation of normalization. Therefore $RR^\times(\frac{d}{d\lambda}\ln D_\pi)\pi = (\frac{d}{d\lambda}\ln D_\pi)\pi$.

If λ^i is a vector of control parameters, then we get a matrix of friction coefficients ζ_{ij}, and the mean excess work is

$$\langle\beta\dot{W}^{\text{ex}}\rangle \approx \tfrac{d\lambda^i}{dt}\cdot\zeta_{ij}\cdot\tfrac{d\lambda^j}{dt} \tag{39}$$

3.1 Fisher Information

Lets define Fisher information via the Cramer-Rao bound using the same notation. Suppose we have some measurement A, that is an estimate for the parameter λ; $\langle A\rangle = \hat{\lambda} = \mathbf{1}^T D_A p$ (A is a vector of measurement outcomes for individual states). For an unbiased estimator $\frac{d}{d\lambda}\langle A\rangle = 1$.

$$\begin{aligned}
1 &= \tfrac{d}{d\lambda}\langle A\rangle && \text{(40a)}\\
1 &= \tfrac{d}{d\lambda}\mathbf{1}^T D_A p && \text{(40b)}\\
1 &= \mathbf{1}^T D_A \tfrac{d}{d\lambda}p && \text{(40c)}\\
1 &= \mathbf{1}^T D_A(\tfrac{d}{d\lambda}D_p)p = \operatorname{cov}(A, \tfrac{d}{d\lambda}\ln p) && \text{(40d)}\\
1 &= \operatorname{cov}(A, \tfrac{d}{d\lambda}\ln p)^2 \leqslant \operatorname{var}(A)\operatorname{var}(\tfrac{d}{d\lambda}\ln D_p) && \text{(40e)}\\
&\text{therefore } \operatorname{var}(A) \geqslant \frac{1}{I_\lambda},\quad I_\lambda = \mathbf{1}^T(\tfrac{d}{d\lambda}\ln D_p)^2\,p && \text{(40f)}
\end{aligned}$$

3.2 Currents

Suppose we wish to measure flows between states. We construct a skew-symmetric matrix $F = -F^T$ containing only $\{-1, 0, 1\}$. The diagonal entries are zero, and each off-diagonal $+1$ records a transition between states. The corresponding -1 skew-symmetric entry counts the backwards transitions between the same states. The operator that measures the instantaneous current is then $J = R\odot F$, where $\odot$ is the Hadamard (entry by entry) product, and the average current is $\langle J\rangle = \mathbf{1}^T J\, p$.

$$\begin{aligned}
\langle J\rangle &= \mathbf{1}^T J\, p && \text{(41a)}\\
&\approx \mathbf{1}^T J\,(I + R^\times\tfrac{d}{dt})\,\pi && \text{(41b)}\\
&\approx \mathbf{1}^T J\,\pi + \mathbf{1}^T J\, R^\times\tfrac{d}{dt}\,\pi && \text{(41c)}\\
&\approx \mathbf{1}^T J\,\pi - \tfrac{d\lambda}{dt}\mathbf{1}^T J\,(-R^\times)(\tfrac{d}{d\lambda}\ln D_\pi)\,\pi && \text{(41d)}
\end{aligned}$$

With this notation we can express the autocorrelation function of the current fluctuations as

$$C_{JJ} = -\mathbf{1}^T J R^\times J\,\pi = -\mathbf{1}^T(R\odot F)R^\times(R\odot F)\pi \tag{42}$$

Proof:

$$\begin{aligned}
C_{JJ} &= -\mathbf{1}^T J R^\times J\,\pi && \text{(43a)}\\
&= \int_0^\infty dt\ \mathbf{1}^T J(e^{tR}RR^\times)J\,\pi && \text{(43b)}\\
&= \int_0^\infty dt\ \mathbf{1}^T J\ e^{tR}\ (I-\pi\mathbf{1}^T)J\,\pi && \text{(43c)}\\
&= \int_0^\infty dt\ \mathbf{1}^T J\ (e^{tR}-\pi\mathbf{1}^T)J\,\pi && \text{(43d)}\\
&= \int_0^\infty \langle J(t)J(0)\rangle - \langle J\rangle^2 && \text{(43e)}
\end{aligned}$$

(Kudos: This is inspired by Eq 16-17 of Baiesi2009 [11]. But we seem to get a slightly different conclusion.)

Acknowledgments

History

v3 (2018-11-01) First public release. Kudos: David Limmer

v2 (2016) Kudos: Lyndon Zhang (Various bug reports)

v1 (2015) Kudos: Input and inspiration from Dibyendu Mandal, Paul Riechers, and Subhaneil Lahiri. I was first introduced to the Drazin inverse by Dibyendu Mandal [6], although I had to figure out the mathematical gadget he was using was the Drazin inverse for myself. Jordan Horowitz's thesis also contains useful material [12], although he never uses the term "Drazin inverse" either. Then I discovered that Subhaneil Lahiri [10] and Paul Riechers [13] were both, independently, using the Drazin for somewhat related problems. These notes grew out of trying to understand these connections and to find further applications of the Drazin inverse to stochastic thermodynamics.

References

[0] [citation needed]. (pages 2, 4, and 4).

[1] Adi Ben-Israel and Thomas N. E. Greville. *Generalized Inverses: Theory and Applications*. Springer-Verlag, New York, 2nd edition (2003). (page 1 and 1).

[2] Robert Pizia and P. L. Odell. *Matrix Theory: From Generalized Inverses to Jordan Form*. Chapman and Hall (2007). (pages 1 and 2).

[3] M. P. Drazin. Pseudo-inverses in associative rings and semigroups. *Am. Math. Monthly*, 65(7):506–514 (1958). doi:10.2307/2308576. (page 1).

[4] Uriel G. Rothblum. A representation of the Drazin inverse and characterization of the index. *SIAM J. Appl. Math.*, 31(4):646–648 (1976). https://www.jstor.org/stable/2100514. (page 2).

[5] M. J. Englefield. The commuting inverses of a square matrix. *Proc. Camb. Phil. Soc.*, 62(4):667–671 (1966). doi:10.1017/S0305004100040317. (page 2).

[6] Dibyendu Mandal and Christopher Jarzynski. Analysis of slow transitions between nonequilibrium steady states. *J. Stat. Mech.: Theor. Exp.*, (6):063204 (2015). doi:10.1088/1742-5468/2016/06/063204. (pages 4, 4, 4, and 6).

[7] Jaromír J. Koliha and Ivan Straškraba. Power bounded and exponentially bounded matrices. *Appl. Math.*, 44:289–308 (1999). doi:10.1023/A:1023032629988. (page 4).

[8] John G. Kemeny, J. Laurie Snell, and Anthony W. Knapp. *Denumerable Markov Chains*. Springer-Verlag, New York, 2nd. edition (1976). (page 4).

[9] James R. Norris. *Markov Chains*. Cambridge University Press, Cambridge, England (1997). (page 4).

[10] Subhaneil Lahiri, Jascha Sohl-Dickstein, and Surya Ganguli. A universal tradeoff between power, precision and speed in physical communication. 1603.07758. (pages 4 and 6).

[11] Marco Baiesi, Christian Maes, and Karel Netočný. Computation of current cumulants for small nonequilibrium systems. *J. Stat. Phys.*, 135:57–75 (2009). doi:10.1007/s10955-009-9723-3. (page 5).

[12] Jordan M. Horowitz. *Controlling molecular-scale motion: Exact predictions for driven stochastic systems*. Ph.D. thesis, University of Maryland, College Park (2010). (page 6).

[13] Paul M. Riechers and James P. Crutchfield. Beyond the spectral theorem: Spectrally decomposing arbitrary functions of nondiagonalizable operators. 1607.06526. (page 6).