Group Inverse: Theory, Computation, and Applications in Mathematical Biology

Steve Kirkland

Department of Mathematics
University of Manitoba

We have already seen how eigenvalues/eigenvectors play a key role in age/stage classified population models.
E.g. For the desert tortoise, we have the projection matrix

$$A = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1.300 & 1.980 & 2.570 \\ 0.716 & 0.567 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.149 & 0.567 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.149 & 0.604 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.235 & 0.560 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.225 & 0.678 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.249 & 0.851 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.016 & 0.860 \end{bmatrix}$$

with Perron value $r = 0.9581$, right Perron vector $x =$
$\begin{bmatrix} 0.2217 & 0.4058 & 0.1546 & 0.0651 & 0.0384 & 0.0309 & 0.0718 & 0.0117 \end{bmatrix}^\top$.

What happens if one or more of the demographic rates change? That question might arise in considering species management approaches, for example.

Reframe the question more mathematically: If $A \hookrightarrow A + \epsilon E$ where $\epsilon > 0$ is small and E is some fixed matrix, how do r and x behave as a function of ϵ?

Want to find the derivative of r wrt ϵ, evaluated at $\epsilon = 0$, and similarly for the derivative of x.

Setup:
Suppose that A is an irreducible nonnegative matrix with Perron value r, right Perron vector x, and left Perron vector $y^\top$, normalized so that $y^\top x = 1$.

Fix a matrix E such that $A + \epsilon E$ is also irreducible and nonnegative for all ϵ such that $|\epsilon|$ is sufficiently small. Thinking of the corresponding Perron value as $r(\epsilon)$ and right Perron vector $x(\epsilon)$, we want $\frac{dr}{d\epsilon}\Big|_{\epsilon=0}, \frac{dx}{d\epsilon}\Big|_{\epsilon=0}$.

Do these 'derivatives' even make sense? For r, recall that it's a simple root of the characteristic polynomial, whose coefficients are linear in ϵ, so differentiability of r follows from the implicit function theorem.

For differentiability of x, we need to be careful about how x is normalized.

Cautionary example: Consider $t > -1$,
$A = \begin{bmatrix} 1+t & 1+t \\ 1 & 1 \end{bmatrix}$, $r = 2+t$. Here's the Perron vector x, normalized so that $||x||_\infty = 1 : x = \begin{bmatrix} 1+t \\ 1 \end{bmatrix}$, if $0 > t > -1$, and $x = \begin{bmatrix} 1 \\ \frac{1}{1+t} \end{bmatrix}$, if $t \geq 0$. Observe that this x is not differentiable at $t = 0$.

However, writing $A = \left[\begin{array}{c|c} a_{11} & u^\top \\ \hline v & A_{(1)} \end{array}\right]$, we find that x can be written as $x_1 \left[\begin{array}{c} 1 \\ \hline (rI - A_{(1)})^{-1} v \end{array}\right]$ so there are (lots of) normalizations so that x is differentiable. E.g. $x_1 = 1, ||x||_1 = 1, ||x||_2 = 1$.

Choose a good normalization for x, so that $x(\epsilon), r(\epsilon)$ are differentiable at $\epsilon = 0$ for $A + \epsilon E$.

$Ax = rx \implies A'x + Ax' = r'x + rx' \implies Ex + Ax' = r'x + rx' \implies y^\top Ex + y^\top Ax' = r'y^\top x + ry^\top x'$.

Deduce that

$$r' \equiv \frac{dr}{dE} = y^\top Ex.$$

(Side note: yet another reason to care about Perron vectors!)

E.g. Derivatives for the desert tortoise.

$$\begin{bmatrix} * & * & * & * & * & \color{blue}{0.0060} & 0.0140 & 0.0023 \\ 0.0580 & 0.1062 & * & * & * & * & * & * \\ * & 0.2786 & 0.1062 & * & * & * & * & * \\ * & * & 0.2786 & 0.1173 & * & * & * & * \\ * & * & * & 0.1767 & 0.1043 & * & * & * \\ * & * & * & * & 0.1845 & 0.1482 & * & * \\ * & * & * & * & * & 0.1352 & 0.3145 & * \\ * & * & * & * & * & * & \color{red}{0.3678} & 0.0600 \end{bmatrix}$$

Small entry: birth rate for subadults (first stage capable of reproducing).

Large entry: rate of movement from adult 1 to adult 2 (latter has the highest birth rate).

Derivative of the eigenvector? Recall that we had $Ex + Ax' = r'x + rx'$ which rearranges to

$$(rI - A)x' = Ex - r'x = Ex - (y^\top Ex)x.$$

The issue here is that $rI - A$ is singular (with nullity 1).

The inverse of $rI - A$ is not available for finding x', so we look for the next best thing.

Suppose that M is a real square matrix of order n. Suppose further that M is singular, with 0 as a semi–simple eigenvalue (i.e. the algebraic and geometric multiplicities of 0 coincide). Of course M is not invertible, but it has a group inverse, which we now define.

The group inverse of M is the unique matrix X satisfying the following three properties: i) $MX = XM$, ii) $MXM = M$ and iii) $XMX = X$. We denote this group X by $M^{\#}$. One way of computing $M^{\#}$ is to work with a full rank factorisation of M: if M has rank k, then there is an $n \times k$ matrix U and a $k \times n$ matrix V such that $M = UV$. In that case, $M^{\#}$ can be written as $M^{\#} = U(VU)^{-2}V$. (Note that $M = UV$ has $n - k$ nonzero eigenvalues; since UV and VU have the same nonzero eigenvalues, VU is invertible.)

Uniqueness: $MXMX = MX$, so MX is a projection matrix. We have $col(MX) \subseteq col(M)$ and
$rank(M) = rank(MXM) \leq rank(MX) \leq rank(M)$, so rank(MX)=rank(M). Deduce that $col(MX) = col(M)$. Similarly we deduce $N(MX) = N(M)$ (here $N(\bullet)$ is the null space). Deduce that MX is the projection matrix with range $col(M)$ and null space $N(M)$. Suppose that X_1 and X_2 are two solutions to i)–iii). Then $MX_1 = MX_2$. But then we have

$$X_1 = X_1 MX_1 = X_1(MX_2) = (X_1 M)X_2 = (X_2 M)X_2 = X_2.$$

Hence there is a unique matrix satisfying i)–iii).

Consider the special case that 0 is a simple eigenvalue of M, say with u and $v^\top$ as right and left null vectors, normalised so that $v^\top u = 1$. In this case, X is the group inverse iff $MX = XM = I - uv^\top, Xu = 0$ and $v^\top X = 0^\top$.

Sketch: Evidently if $MX = XM = I - uv^\top, Xu = 0$ and $v^\top X = 0^\top$, then X satisfies i)–iii). Suppose now that X satisfies i)–iii). Since $M(XM - I) = 0$, each column of $XM - I$ is a scalar multiple of u. Also, $(MX - I)M = 0$, so each row of $MX - I$ is a scalar multiple of $v^\top$. Hence, $XM = I + uw^\top$ for some w and $MX = I + zv^\top$ for some z. But $XM = MX$, so it must be the case that $XM = MX = I + tuv^\top$ for some scalar t. Since $det(XM) = 0$, it follows that $t = -1$.

Back to the issue at hand: $(rI - A)x' = Ex - (y^\top Ex)x$. Multiply by $(rI - A)^\#$ to get

$$(rI - A)^\#(rI - A)x' = (rI - A)^\# Ex - (y^\top Ex)(rI - A)^\# x.$$

In our setting, $(rI - A)^\#(rI - A) = I - xy^\top$ and $(rI - A)^\# x = 0$, so we get $(I - xy^\top)x' = (rI - A)^\# Ex$ so

$$x' \equiv \frac{dx}{dE} = (y^\top x')x + (rI - A)^\# Ex = \mathrm{const} \times x + (rI - A)^\# Ex.$$

How to find the constant? Depends on the normalization of x that we started with.

Suppose that there is a fixed vector z, and that x has been normalized so that $z^\top x = 1$. Then $z^\top x' = 0$, and we find that $0 = \mathrm{const} + z^\top (rI - A)^\# Ex$. Deduce that

$$x' = \frac{dx}{dE} = (-z^\top (rI - A)^\# Ex)x + (rI - A)^\# Ex.$$

Observe that this covers the case $z = \mathbf{1}$, which corresponds to $||x||_1 = 1$.

Fix $p > 0$, and suppose that we normalize x so that $||x||_p = 1$. Set $z_j = x_j^{p-1}, j = 1, \ldots, n$ and notice that $z^\top x' = 0$. As above, we have

$$x' = \frac{dx}{dE} = \left(- \begin{bmatrix} x_1^{p-1} & \ldots & x_n^{p-1} \end{bmatrix} (rI - A)^\# Ex \right) x + (rI - A)^\# Ex.$$

E.g. Desert tortoise, right Perron vector x with $||x||_1 = 1$ is:

$$x = \begin{bmatrix} 0.2217 \\ 0.4058 \\ 0.1546 \\ 0.0651 \\ 0.0384 \\ 0.0309 \\ 0.0718 \\ 0.0117 \end{bmatrix}. \text{ Derivatives of } x\text{: } \underset{\mathrm{wrt}(1,6)}{\underset{\uparrow}{\begin{bmatrix} 0.0070 \\ 0.0066 \\ 0.0001 \\ -0.0011 \\ -0.0012 \\ -0.0016 \\ -0.0078 \\ -0.0020 \end{bmatrix}}}, \quad \underset{\mathrm{wrt}(8,7)}{\underset{\uparrow}{\begin{bmatrix} 0.2657 \\ 0.1048 \\ -0.1055 \\ -0.1120 \\ -0.1016 \\ -0.1222 \\ -0.5308 \\ 0.6015 \end{bmatrix}}}.$$

Suppose that A is an irreducible $n \times n$ nonnegative matrix with Perron value r, and left and right Perron vectors $y^\top, x$, normalized so that $y^\top x = 1$. Write $A, x, y^\top$ as $\left[\begin{array}{c|c} A_{(n)} & * \\ \hline * & * \end{array}\right]$, $\left[\begin{array}{c} \overline{x} \\ \hline x_n \end{array}\right]$, $\left[\begin{array}{c|c} \overline{y}^\top & y_n \end{array}\right]$. Then $(rI - A)^\# =$

$$\left(\overline{y}^\top (rI - A_{(n)})^{-1}\overline{x}\right) xy^\top +$$
$$\left[\begin{array}{c|c} (rI - A_{(n)})^{-1} - (rI - A_{(n)})^{-1}\overline{x}\overline{y}^\top - \overline{x}\overline{y}^\top (rI - A_{(n)})^{-1} & -y_n (rI - A_{(n)})^{-1}\overline{x} \\ \hline -x_n \overline{y}^\top (rI - A_{(n)})^{-1} & 0 \end{array}\right].$$

The formula can be deduced from the eigenequations $Ax = rx, y^\top A = ry^\top$, which yield a full rank factorization for $rI - A$. In particular, we can find the group inverse in roughly $2n^3$ flops.

Suppose that we have the Jordan form available for $rI - A$, say

$$rI - A = S \begin{bmatrix} 0 & & & \\ & J(\lambda_2)_{k_2} & & \\ & & \ddots & \\ & & & J(\lambda_q)_{k_q} \end{bmatrix} S^{-1}$$

for some invertible matrix S. Then

$$(rI - A)^{\#} = S \begin{bmatrix} 0 & & & \\ & J(\lambda_2)_{k_2}^{-1} & & \\ & & \ddots & \\ & & & J(\lambda_q)_{k_q}^{-1} \end{bmatrix} S^{-1}.$$

In particular the spectral properties of $(rI - A)^{\#}$ are closely related to those of $rI - A$.

Suppose that A is an irreducible stochastic matrix with stationary vector $w^\top$. Now perturb A to get $\tilde{A} = A + E$, where $\tilde{A}$ is also irreducible and stochastic. What is the new stationary distribution vector $\tilde{w}^\top$?

We have $\tilde{w}^\top(A+E) = \tilde{w}^\top$, so $\tilde{w}^\top(I-A) = \tilde{w}^\top E$. Hence $\tilde{w}^\top(I-A)(I-A)^\# = \tilde{w}^\top E(I-A)^\#$ so that $\tilde{w}^\top(I - \mathbf{1}w^\top) = \tilde{w}^\top E(I-A)^\#$.

Deduce that $\tilde{w}^\top(I - E(I-A)^\#) = w^\top$. It turns out that $(I - E(I-A)^\#)$ is nonsingular, and hence

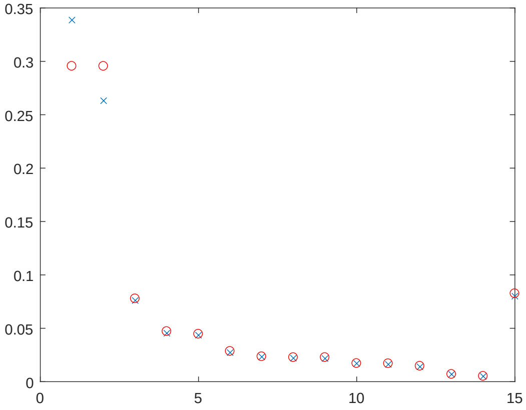
$$\tilde{w}^\top = w^\top(I - E(I-A)^\#)^{-1}.$$

Species succession transition matrix A:

$$\begin{bmatrix}
.7725 & .1022 & .0170 & .0040 & .0150 & .0010 & .0180 & .0120 & .0020 & .0140 & .0030 & .0020 & .0050 & .0030 & .0291 \\
.1450 & .6090 & .0310 & .0110 & .0280 & .0050 & .0220 & .0250 & .0110 & .0150 & .0120 & .0080 & .0050 & .0040 & .0690 \\
.0519 & .0609 & .7093 & .0040 & .0200 & .0040 & .0080 & .0080 & .0250 & .0030 & .0050 & .0070 & .0020 & .0080 & .0839 \\
.0170 & .0541 & .0060 & .8398 & .0050 & 0 & .0040 & .0060 & .0080 & .0040 & .0060 & .0110 & 0 & .0030 & .0360 \\
.1169 & .2178 & .0350 & .0040 & .4036 & .0080 & .0330 & .0320 & .0130 & .0070 & .0060 & .0050 & .0060 & .0050 & .1079 \\
.0090 & .0240 & .0120 & 0 & .0160 & .8647 & .0010 & .0070 & .0160 & .0030 & .0040 & .0070 & 0 & 0 & .0361 \\
.2412 & .2232 & .0511 & .0160 & .0801 & .0240 & .1051 & .0410 & .0140 & .0330 & .0250 & .0050 & .0140 & .0120 & .1151 \\
.1986 & .2345 & .0379 & .0180 & .0888 & .0070 & .0439 & .1537 & .0150 & .0269 & .0160 & .0200 & .0090 & .0090 & .1218 \\
.0559 & .1469 & .0260 & .0110 & .0200 & .0060 & .0110 & .0260 & .5854 & .0210 & .0060 & .0050 & .0010 & .0050 & .0739 \\
.3084 & .2275 & .0309 & .0100 & .0269 & .0060 & .0419 & .0309 & .0100 & .1647 & .0130 & .0080 & .0120 & .0060 & .1038 \\
.0559 & .2216 & .0279 & .0080 & .0359 & 0 & .0250 & .0200 & .0070 & .0070 & .5060 & .0020 & .0050 & .0030 & .0758 \\
.0250 & .0680 & .0180 & .0300 & .0160 & 0 & .0100 & .0160 & .0040 & .0030 & .0010 & .5370 & .0030 & .0030 & .2660 \\
.3210 & .1790 & .0230 & 0 & .0630 & 0 & .0300 & .0200 & .0030 & .0200 & .0170 & 0 & .2480 & 0 & .0760 \\
.1583 & .4489 & .0180 & .0180 & .0852 & .0060 & .0301 & .0180 & .0180 & .0301 & .0060 & .0060 & 0 & .0301 & .1273 \\
.1010 & .3200 & .0250 & .0090 & .0620 & .0050 & .0480 & .0340 & .0130 & .0310 & .0170 & .0170 & .0110 & .0130 & .2940
\end{bmatrix}$$

Perturbation: $a_{11} \to a_{11} - .0517,\ a_{12} \to a_{12} + .0517.$

Stationary distribution plots, before and after the perturbation.



Other uses in the Markov chain context:

For an irreducible stochastic matrix A with stationary vector $w^\top$, the mean first passage matrix is given by

$$M = \left(I - (I-A)^{\#} + J(I-A)^{\#}_{\mathrm{dg}}\right) W^{-1},$$

where $W = \mathrm{diag}(w)$.

Then $Mw = (1 + \mathrm{trace}((I-A)^{\#}))\mathbf{1}$, so Kemeny's constant $\kappa(A)$ is equal to $\mathrm{trace}((I-A)^{\#})$.

Imagine we have an outbreak of a disease – the outbreak occurs heterogeneously in several different geographic locations ("patches"), and there is the possibility of movement between different locations.

Cholera, with communities located along a river, and the pathogen carried between patches by contaminated water.

The disease has characteristics that are patch–dependent: indirect transmission rate from pathogen to host, pathogen decay rate, pathogen shedding rate, and decay rate of infectious host individuals.

Movement is possible between the patches, with parameters $m_{ij} \geq 0$ representing the representing rate of the pathogen/host dispersal from patch j to patch i.

If there are n patches, then we may construct the corresponding movement matrix $M = [m_{ij}]_{i,j=1,\ldots,n}$, as well as the associated Laplacian matrix $L = diag(\mathbf{1}^T M) - M$.

For each patch we keep track of the number of susceptible, infected and recovered individuals, and the water, with patch–dependent rates.

We model these as a system of ODEs.

Will the disease persist or not? The answer is determined by the network basic reproduction number, $\mathcal{R}_0$; if $\mathcal{R}_0 < 1$ the disease dies out, while if $\mathcal{R}_0 > 1$ the disease persists.

$\mathcal{R}_0$ can also be interpreted as the expected number of infections directly generated by one infected individual.

Each patch has its own basic reproductive number $\mathcal{R}_0^{(k)}$, based on the patch–specific parameters. If the dispersal between patches is much faster than the disease dynamics, it turns out that we have $\mathcal{R}_0 \approx \sum_{k=1}^{n} u_k R_0^{(k)}$, where $u^T = \left[\begin{array}{cccc} u_1 & u_2 & \dots & u_n \end{array} \right]$ is the right null vector of the Laplacian matrix L, normalised so that $u^T \mathbf{1} = 1$.

Notice that u depends only on the network, not on the disease dynamics. Hence, we can study the influence of that network structure on $\mathcal{R}_0$.

Suppose that we perturb the movement matrix so that $m_{ij} \to m_{ij} + \epsilon$. This yields $\tilde{L} = L + \epsilon(e_j - e_i)e_j^T \equiv L + E$. Similar to the Markov chain perturbation setting, $\tilde{u} = (I + L^{\#}E)^{-1}u$. Since $E = \epsilon(e_j - e_i)e_j^T$, $(I + L^{\#}E)^{-1} = I - \frac{\epsilon}{1+\epsilon(L^{\#}_{jj} - L^{\#}_{ji})} L^{\#}(e_j - e_i)e_j^{\top}$.

Upshot:

$$\tilde{\mathcal{R}}_0 = \mathcal{R}_0 - \frac{\epsilon u_j}{1 + \epsilon(L^{\#}_{jj} - L^{\#}_{ji})} \sum_{k=1}^{n} (L^{\#}_{kj} - L^{\#}_{ki}) \mathcal{R}_0^{(k)}.$$

Knowledge of $L^{\#}$ gives insight into how changes in the network affect $\tilde{\mathcal{R}}_0$.

References:

Kirkland and Neumann, *Group Inverses for M–matrices and their Applications*, CRC, 2013.

Myer, The role of the group generalized inverse in the theory of finite Markov chains, *SIAM Review* 17, 1975.

Kirkland, Shuai, van den Driessche and Wang, Impact of varying community networks on disease invasion, *SIAM Journal on Applied Mathematics* 81, 2021.