# Introduction to Some Topics in Linear Algebra

Steve Kirkland

Department of Mathematics
University of Manitoba

Think of a population that consists of $n$ age groups. Let $u_i(m)$ denote the number of individuals in the $i$–th age group at time $m$. Let $f_i$, $i = 1, \ldots, n$, denote the fecundity of each individual in the $i$–th age group and let $p_i$, $i = 1, \ldots, n-1$ denote the proportion of individuals that survive from age $i$ to age $i+1$. Assume that both the fecundity and survival rates are independent of the time $m$. Then, as can be readily ascertained, the age vector at time $m+1$, $m \geq 0$, can be described by the matrix – vector relation:

$$u(m+1) \; = \; \begin{bmatrix} f_1 & f_2 & \ldots & \ldots & f_{n-1} & f_n \\ p_1 & 0 & \ldots & \ldots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & 0 & 0 \\ \vdots & \ldots & \ddots & \vdots & \vdots & \vdots \\ \vdots & \ldots & \ldots & \ddots & \vdots & \vdots \\ 0 & \ldots & \ldots & \ldots & p_{n-1} & 0 \end{bmatrix} u(m) \; =: Au(m).$$

Here $u(m) = \begin{bmatrix} u_1(m) & \dots & u_n(m) \end{bmatrix}^\top$ for all $m \geq 0$. Evidently we have $u(m) = A^m u(0)$, so that the sequence of population vectors is an example of the power method applied to $u(0)$.

"Pick a basis of eigenvectors and decompose $u(0)$ as a linear combination of eigenvectors, $\sum_j c_j v_j$, so that $u(m) = \sum_j c_j \lambda_j^m v_j$.

As $m \to \infty, u(m)/\mathbf{1}^\top u(m) \to v_1/\mathbf{1}^\top v_1$. Here $\lambda_1$ is the dominant eigenvalue. "

Can we justify this?

Imagine for now that we can justify those statements.

Denoting the all–ones vector by $\mathbf{1}$, the age distribution vector at time $m$ is given by $u(m)/\mathbf{1}^\top u(m)$, while the total size of the population at time $m$ is $\mathbf{1}^\top u(m)$.

As $m \to \infty$, we find that the population size is asymptotically growing like $\lambda_1^m$, the age distribution vector converges to an appropriately scaled eigenvector of $A$ corresponding to $\lambda_1$. In particular, the eigenvalue $\lambda_1$ has demographic significance, as it is interpreted as the asymptotic growth rate for the population under consideration.

Recall that for a square matrix $A$, a (complex) scalar $\lambda$ is an eigenvalue if there is a nonzero vector $v$, an eigenvector, such that $Av = \lambda v$.

Each eigenvalue is a root of the characteristic polynomial $\det(zI - A)$, and the algebraic multiplicity of $\lambda$ is its multiplicity as a root of the characteristic polynomial.

The geometric multiplicity is the dimension of the null space of $\lambda I - A$.

E.g. $A = \begin{bmatrix} 0 & \frac{3}{4} & \frac{1}{4} \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$, with characteristic polynomial $z^3 - \frac{3}{4}z - \frac{1}{4} = (z-1)(z+\frac{1}{2})^2$. Eigenvalue 1 with $\mathbf{1}$ as eigenvector; eigenvalue $-\frac{1}{2}$ with eigenspace spanned by $v = \begin{bmatrix} 1 \\ -2 \\ 4 \end{bmatrix}$.

So, no basis of eigenvectors, but observe that we "almost" have a third eigenvector: $\tilde{v} = \begin{bmatrix} -4 \\ 4 \\ 0 \end{bmatrix}$ satisfies $A\tilde{v} = -\frac{1}{2}\tilde{v} + v$.

For a matrix $A$ with eigenvalue $\lambda$, a Jordan chain associated with $\lambda$ is a list of nonzero vectors $v_1, \ldots, v_k$ such that $Av_1 = \lambda v_1, Av_2 = \lambda v_2 + v_1, Av_3 = \lambda v_3 + v_2, Av_k = \lambda v_k + v_{k-1}$.

Related to that is a Jordan block:

$$J(\lambda)_k = \begin{bmatrix} \lambda & 1 & 0 & 0 & \ldots & 0 \\ 0 & \lambda & 1 & 0 & \ldots & 0 \\ \vdots & & \ddots & \ddots & & \vdots \\ \\ 0 & 0 & \ldots & 0 & \lambda & 1 \\ 0 & 0 & \ldots & 0 & 0 & \lambda \end{bmatrix}_{k \times k}$$

## Theorem

*Suppose that $A$ is an $n \times n$ real (or complex) matrix. There is a nonsingular complex matrix $S$, complex scalars $\lambda_1, \ldots, \lambda_q$, and $k_1, \ldots, k_q \in \mathbb{N}$ with $k_1 + \ldots + k_q = n$ such that*

$$A = S \begin{bmatrix} J(\lambda_1)_{k_1} & & \\ & \ddots & \\ & & J(\lambda_q)_{k_q} \end{bmatrix} S^{-1}.$$

Remarks: The matrix in the middle is known as the Jordan canonical form for $A$. The eigenvalues are $\lambda_1, \ldots, \lambda_q$, and the algebraic multiplicity of $\lambda_j$ is the number of times it appears on the diagonal, while the geometric multiplicity is the number of $J(\lambda_j)$ blocks that appear on the diagonal. Columns of $S$ are eigenvectors and generalised eigenvectors. Jordan matrix is unique up to reordering the blocks.

Proof outline:

a) $A = UTU^*$, $U$ unitary $T$ triangular (Schur's lemma, induction, start with an eigenvector as first column in $U$). Diagonal entries of $T$ can be taken in any prescribed order.

b) Take $T$ so that all the $\lambda_1$ come first on the diagonal, then the $\lambda_2$s etc. Show that $T$ is similar to a direct sum of upper triangular matrices each with common entry on the diagonal.

c) Show that any such diagonal block is similar to a direct sum of $J(\lambda)$ blocks.

To understand powers of $A$, we need to understand the powers of the Jordan canonical form.
For $m \in \mathbb{N}$,

$$J(\lambda)_k^m = \begin{bmatrix} \lambda^m & \binom{m}{1}\lambda^{m-1} & \binom{m}{2}\lambda^{m-2} & \binom{m}{3}\lambda^{m-3} & \dots & \binom{m}{k-1}\lambda^{m-k+1} \\ 0 & \lambda^m & \binom{m}{1}\lambda^{m-1} & \binom{m}{2}\lambda^{m-2} & \dots & \binom{m}{k-2}\lambda^{m-k+2} \\ \vdots & & \ddots & \ddots & & \vdots \\ 0 & 0 & \dots & 0 & \lambda^m & \binom{m}{1}\lambda^{m-1} \\ 0 & 0 & \dots & 0 & 0 & \lambda^m \end{bmatrix}$$

(Use induction on $m$.)

We have essentially justified the “pick a basis of eigenvectors” step in our Leslie model discussion.

"Dominant eigenvalue"

A square $n \times n$ matrix $A$ is nonnegative if $a_{jk} \geq 0, j, k = 1, \ldots, n$. Associated with $A$ is a directed graph on vertices labelled $1, \ldots, n$, with $j \to k$ iff $a_{jk} > 0$. We say that $A$ is irreducible if that directed graph is strongly connected – i.e. for any pair of vertices $j, k$ there is a suitable sequence of directed arcs by which we can walk from $j$ to $k$. $A$ is primitive if $A^m$ has all positive entries for some $m \in \mathbb{N}$.

Observe that notions of irreducibility/primitivity are combinatorial; they don't depend on the sizes of the positive entries, only on their placement relative to each other.

### Theorem

*Suppose that $A$ is a primitive nonnegative matrix. Then:*
*a) $A$ has a positive eigenvalue $r$;*
*b) there are left and right eigenvectors of $A$ corresponding to $r$ that have all positive entries;*
*c) if $\lambda$ is an eigenvalue of $A$ and $\lambda \neq r$, then $|\lambda| < r$;*
*d) $r$ is algebraically simple and geometrically simple;*
*e) the only left or right eigenvectors of $A$ having all positive entries are those corresponding to $r$.*

Remark: $r$ is called the Perron value, and positive left/right eigenvectors corresponding to $r$ are Perron vectors.

Proof's core idea:

Let $S = \{t \geq 0 | \exists x \in \mathbb{R}^n \ni x \geq 0, \mathbf{1}^\top x = 1, Ax \geq tx\}$. Then $r = \sup(S)$. The properties of the supremum, the fact that $A^m > 0$ for some $m \in \mathbb{N}$, and the triangle inequality, are then used to prove most of a)–e).

Remark: If $A$ is irreducible but not primitive, then a), b), d), e) still hold, and in c) $< \to \leq$.

If $A$ is reducible then some similar statements can be made, but further details depend on the specific combinatorial structure of the matrix and the sizes of the positive entries.

However, for any square nonnegative matrix, we can say that i) the spectral radius is an eigenvalue, and ii) associated with the spectral radius there are nonnegative left and right eigenvectors.

## Corollary

*Suppose that $A$ is an $n \times n$ irreducible nonnegative matrix with Perron value $r$. Let $B$ be another $n \times n$ nonnegative matrix such that $B \leq A$ (entrywise). Denoting the Perron value of $B$ by $\hat{r}$, we have $\hat{r} \leq r$. Further, if the entrywise inequality is strict in at least one position, then $\hat{r} < r$.*

Idea: Let $x$ be a right Perron vector for $A$ and $\hat{y}^\top$ be a left Perron vector for $B$. Note that $x > 0, \hat{y} \geq 0, \hat{y} \neq 0$. We have $\hat{y}^\top Bx \leq \hat{y}^\top Ax$, i.e., $\hat{r}\hat{y}^\top x = \hat{y}^\top Bx \leq \hat{y}^\top Ax = r\hat{y}^\top x$. (Observe that $\hat{y}^\top x > 0$.)

Suppose that the inequality is strict in the $(j, k)$ position. Let $C = A - \epsilon e_j e_k^\top$ where $\epsilon > 0$ is chosen so that $C$ is irreducible and $B \leq C$. Denote the left Perron vector for $C$ by $\tilde{y}^\top$ and Perron value by $\tilde{r}$. Then $\tilde{r}\tilde{y}^\top x = \tilde{y}^\top(A - \epsilon e_j e_k^\top)x = r\tilde{y}^\top x - \epsilon \tilde{y}_j x_k < r\tilde{y}^\top x$.

## Corollary

*Suppose that $A$ is a primitive nonnegative matrix with Perron value $r$ and left and right Perron vectors $y^\top, x$, normalized so that $y^\top x = 1$. Then as $m \to \infty$, $\frac{1}{r^m}A^m \to xy^\top$.*

Proof sketch:

From the Jordan Canonical Form: there is an invertible matrix $S$ so that $A^m = SJ^mS^{-1}$, where $J = [r] \oplus J(\lambda_2)_{k_2} \oplus \ldots \oplus J(\lambda_q)_{k_q}$ and $r > |\lambda_j|, j = 2, \ldots, q$.

Observe: the moduli of the entries in $\frac{1}{r^m}J(\lambda_j)^m_{k_j}$ are no larger than $m^{k_j-1}\left(\frac{|\lambda_j|}{r}\right)^m$. These last tend to 0, so that $\frac{1}{r^m}J^m \to e_1e_1^\top$ as $k \to \infty$. Note that $Se_1 = x, e_1^\top S^{-1} = y^\top$.

E.g. Matrix for the desert tortoise with the following stages: yearling, juvenile 1, juvenile 2, immature 1, immature 2, subadult, adult 1, adult 2.

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

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

The species is endangered – observe that in $x$, the sum of the entries corresponding to reproductive stages comprise only 11.44% of the total.

Suppose that we have a Leslie matrix

$$A = \begin{bmatrix} f_1 & f_2 & \dots & \dots & f_{n-1} & f_n \\ p_1 & 0 & \dots & \dots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & 0 & 0 \\ \vdots & \dots & \ddots & \vdots & \vdots & \vdots \\ \vdots & \dots & \dots & \ddots & \vdots & \vdots \\ 0 & \dots & \dots & \dots & p_{n-1} & 0 \end{bmatrix} \text{ with}$$

$p_j > 0, j = 1, \dots, n-1, f_n > 0$, and $\gcd\{j | f_j > 0\} = 1$. Denote the Perron value by $r$, and iterate $u(m) = Au(m-1), m \in \mathbb{N}$ with $u(0) \geq 0, u(0) \neq 0$. Then as $m \to \infty$,

$$u(m)/\mathbf{1}^\top u(m) \to \left( \frac{1}{1+\sum_{j=2}^{n} p_1 \cdots p_{j-1}/r^{j-1}} \right) \begin{bmatrix} 1 \\ \frac{p_1}{r} \\ \frac{p_1 p_2}{r^2} \\ \vdots \\ \frac{p_1 \dots p_{n-1}}{r^{n-1}} \end{bmatrix} \equiv x.$$

Our Perron vector $x$ is interpreted as the asymptotically stable age distribution.

Asymptotically, we have $\mathbf{1}^\top u(m) = \text{constant} \times r^m + e(m)$, where $\frac{e(m)}{r^m} \to 0$ as $m \to \infty$.

So, asymptotically, the size of the population is growing geometrically at rate $r$.

Suppose that we have our primitive Leslie matrix $A$, and we want to perturb it to $A + \epsilon E$ (also irreducible & nonnegative matrix). How is the Perron vector (stable age distribution) $x$ affected?

$Ax = rx$, so differentiating wrt $\epsilon$, $A'x + Ax' = r'x + rx'$. There's good information in $x'$ and we would like to find it. Observe that $(rI - A)x' = A'x - r'x = Ex - r'x$. It's inconvenient that our coefficient matrix $rI - A$ is a singular, but it has some extra structure – i.e. all offdiagonal entries are nonpositive, and there is a positive eigenvector (a null vector in this case).

A matrix $M$ is called an M–matrix if it can be written as $sI - A$ where $A$ is nonnegative and $s \geq$ spectral radius of A.

The 'M' is apparently in honour of Minkowski.

Closely related is the notion of a Z–matrix: a square matrix $B$ is a Z–matrix if all of its off–diagonal entries are nonpositive.

Properties:

If $M$ is a nonsingular M–matrix, then $M^{-1} \geq 0$. Idea is: $M = sI - A$, where $s$ is larger than the spectral radius of $A$. Observe that $\sum_{j=0}^{\infty} \frac{1}{s^j} A^j$ converges (JCF), and hence $M^{-1} = \sum_{j=0}^{\infty} \frac{1}{s^{j+1}} A^j$. If $A$ happens to be irreducible, then in fact $M^{-1}$ is positive.

Suppose that we have a Z–matrix $B$, and wlog suppose that $b_{11} = \max b_{jj}$. Write $B$ as $B = b_{11} I - A$, where $a_{jj} = b_{11} - b_{jj}, a_{jk} = |b_{jk}|, j \neq k$. Evidently $B$ is an M–matrix iff $b_{11}$ is at least as large as the spectral radius of $A$, say $r$.

Suppose that there is a positive vector $v$ such that $Bv \geq 0$, i.e. $b_{11} v \geq Av$. Let $y^\top$ be a nonnegative left eigenvector of $A$ corresponding to $r$. Then $b_{11} y^\top v \geq r y^\top v$ so that $b_{11} \geq r$.

A Z–matrix $B$ is an M–matrix if all of its eigenvalues lie in the right half–plane.

$B = b_{11}I - A$, and $A$ has spectral radius $r$ as an eigenvalue, so $B$ has $b_{11} - r$ as an eigenvalue. If all eigenvalues of $B$ have nonnegative real part, then necessarily $b_{11} \geq r$.

There are plenty of equivalent characterizations of M–matrices, e.g. in terms of principal minors, inverse positivity, Lyapunov stability, etc.

As suggested already, M–matrices arise naturally in perturbation theory for Perron eigenvectors/eigenvalues. I will discuss this more in Lecture 9.

In certain SIR epidemic models on patch networks, an M–matrix $M$ associated with the network arises naturally. The spectral radius of $M^{-1}$ then measures the invasibility of the epidemic.

Other applications: convergence of iterative methods for large sparse linear systems, electrical networks, discretization of the Laplacian operator, Markov chains.

References:

Horn & Johnson, *Matrix Analysis,* $2^{nd}$ *edition* , Cambridge, 2013.

Seneta, *Non-negative Matrices and Markov Chains,* $2^{nd}$ *edition*, Springer, 1981.

Berman & Plemmons, *Nonnegative Matrices in the Mathematical Sciences,* $2^{nd}$ *edition*, SIAM, 1994.

Doak, Kareiva and Klepetka, Modeling population viability for the desert tortoise in the western Mojave Desert, *Ecological Applications* 4, 1994.