# GeoSci 236: A Linear Algebra Mini-Primer II

Gidon Eshel
491 Hinds
Dept. of the Geophysical Sciences,
5734 S. Ellis Ave., The Univ. of Chicago,
Chicago, IL 60637
(773) 702-0440, geshel@midway.uchicago.edu

August 1, 2006

Matrices and vectors are the higher-dimensional generalization of real numbers. As such, they are associated with both direction and magnitude. To be sure, real numbers have direction too; it's just always the *same* one: left or right along the real line. As soon as you throw in another dimension ($\Re^{\geq 2}$), both magnitude and direction can change, as Figure 1 shows. The thick gray horizontal line represents the number 3. Next, we take $3 \times 2$, shown by the slightly thiner black horizontal line. The magnitude changed (from 3 to 6), but the direction was conserved; both are along the real line. The other 2 lines of the Figure are the 2 $\Re^2$ vectors

$$\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \qquad \text{and} \qquad \begin{pmatrix} 1 & 2 \\ 3 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

(I hope it is obvious to you which is which...) Clearly, both magnitude and direction have changed.

Thus matrices *can* change magnitude and direction. (But they don't *have* to; just recall the identity matrix!) The direction is easy to understand – it stems from the multiplication rules, and depends on the elements of the matrices involved. But what about the magnitude? This is determined by one of the most fundamental properties of matrices - their *eigenvalues*. (The direction, too, can be cast in these terms, which we will get to a bit later.) Before we delve too deep, note that in this section, all matrices are square! Now you might ask, correctly, why we want to bother with square matrices when data matrices are almost always rectangular. This will become clear later.

To get a handle on eigenvalues, it's best to start with an example. To honor our biology-type classmates, consider the temporal evolution of an ecosystem with only 2 species, $x(t)$ and $y(t)$. Let's further envision that each species, if left alone, multiplies at a given rate. That is, in the absence of interaction, the system will satisfy

$$\frac{dx(t)}{dt} = ax(t)$$
$$\frac{dy(t)}{dt} = by(t).$$

Figure 1: Demonstration of direction and magnitude of vectors and matrices.

However, the species *do* interact; let's assume they compete over the same food source. Then, the less of $x$ the system has, the happier $y$ gets, and vice versa. Now, the above system of linear ODEs (ordinary differential equations) must be modified to take note of the competition,

$$\frac{dx(t)}{dt} = ax(t) - cy(t)$$

$$\frac{dy(t)}{dt} = by(t) - dx(t).$$

(You surely realize that this is a grotesque oversimplification of competition, I just need a system of linear ODEs...) Now this system can be written as the vector equation

$$\frac{d\mathbf{x}}{dt} \equiv \frac{d}{dt}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a & -c \\ -d & b \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} \equiv \mathbf{Ax}$$

(notice that the scalar $x$ is an element of the 2-vector $\mathbf{x}$).

Just like in the scalar case, we can always *try* to plug in a solution, and check whether it satisfies the equation. Let us choose, then

$$x(t) = \alpha e^{\lambda t} \qquad \text{and} \qquad y(t) = \beta e^{\lambda t},$$

with some amplitudes $\alpha$ and $\beta$. Note that we are trying to come up with *one* rate $\lambda$ that will describe the temporal evolution of both species (conceivably, we could have chosen 2 distinct rates, $\lambda_x$ and $\lambda_y$). The eigenvectors account for the internal distribution, as we will see shortly. Substituting the solutions into the 2 scalar equations, we get

$$\alpha\lambda e^{\lambda t} = a\alpha e^{\lambda t} - c\beta e^{\lambda t}$$

$$\beta\lambda e^{\lambda t} = -d\alpha e^{\lambda t} + b\beta e^{\lambda t}.$$

The exponent is common to all terms, and is nonzero. Hence the equations can be divided by it, yielding

$$\alpha\lambda = a\alpha - c\beta \qquad \text{and} \qquad \beta\lambda = -d\alpha + b\beta$$

or, in vector form,

$$\lambda\mathbf{s} \equiv \lambda \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} a & -c \\ -d & b \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \mathbf{A}\mathbf{s}.$$

Notice what this equation means. We are looking for a scalar–vector pair $(\lambda, \mathbf{s})$, or perhaps several $(\lambda_i, \mathbf{s}_i)$ pairs, such that when the vector in the pair is premultiplied by $\mathbf{A}$, its direction will not change, and the magnitude will change by a factor given by the scalar in the pair. Keep this in mind!

Next we want to combine both terms that premultiply $\mathbf{s}$. However, recall that we cannot simply write

$$\left[\begin{pmatrix} a & -c \\ -d & b \end{pmatrix} - \lambda\right] \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

as this is dimensionally impossible (we are attempting to subtract a scalar from a matrix). What we can do, however, is

$$\left[\begin{pmatrix} a & -c \\ -d & b \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right] \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \qquad \text{or}$$

$$(\mathbf{A} - \lambda\mathbf{I})\,\mathbf{s} = \mathbf{0}.$$

This is the central equation of this discussion.

Let's write the equation explicitly,

$$\begin{pmatrix} a-\lambda & -c \\ -d & b-\lambda \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

i.e., the vector we are looking for is the nullspace of $\mathbf{B} \equiv (\mathbf{A} - \lambda\mathbf{I})$. For $\mathbf{B}$ to have a non-trivial nullspace, it must be singular. One way to check whether a matrix is singular or not, is to evaluate its *determinant*; if $det\,(\mathbf{A} - \lambda\mathbf{I}) = 0$, $\mathbf{B}$ has a nontrivial nullspace, as required. The determinant of a $2 \times 2$

$$\mathbf{D} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \qquad \text{is} \qquad det(\mathbf{D}) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$

Similarly, the determinant of a $3 \times 3$

$$\mathbf{D} = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \qquad \text{is} \qquad det(\mathbf{D}) = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = a(ci - fh) - b(di - fg) + c(dh - eg).$$

You can look up formulae for higher-dimensional determinants in linear algebra textbooks. However, note that you can figure out the rule from the above. The right-most expressions are called the *characteristic polynomial.* For an $N \times N$ matrix, the characteristic polynomial has $N$ roots, which are the eigenvalues.

Going back to the above 2-species competition scenario,

$$det(A) = \lambda^2 - (a+b)\lambda + ab - cd,$$

with roots

$$\lambda_{1,2} = \frac{(a+b) \pm \sqrt{(a+b)^2 - 4(ab-cd)}}{2}$$

, the eigenvalues.

For each eigenvalue, we solve the related equation for the corresponding eigenvector;

$$(\mathbf{A} - \lambda_i \mathbf{I})\, \mathbf{s}_i = \mathbf{0}, \qquad i = 1, 2, \cdots N.$$

Since we have already established that $(\mathbf{A} - \lambda_i \mathbf{I})$ has a nontrivial nullspace ($\lambda_i$ were chosen to ensure that), nontrivial $\mathbf{s}_i$ must exist.

Let's consider the numerical example

$$\mathbf{A} = \begin{pmatrix} 1 & -\frac{3}{2} \\ -\frac{1}{2} & 2 \end{pmatrix}.$$

Before we continue, let's spell out what this means. Element $(1,1)$ means that species $x$, when unmolested by $y$, grows exponentially with an e-folding timescale of 1 in whatever time-units we employ. For species $y$ the corresponding number is 2 [given by element $(2,2)$]. That is, $y$'s biology makes it able to exploit its available resources for expansion at twice the rate. The off-diagonal elements mean that the species are affected by the competition differently; species $x$ is rather sensitive to the fierce competition species $y$ puts up [element $(1,2)$], while species $y$ is less easily perturbed by the presence of species $x$. So much for population dynamics 101.

Let's eigen-analyze $\mathbf{A}$. First,

$$\mathbf{A} - \lambda \mathbf{I} = \begin{pmatrix} 1-\lambda & -\frac{3}{2} \\ -\frac{1}{2} & 2-\lambda \end{pmatrix}$$

yields the characteristic equation $\lambda^2 - 3\lambda + 5/4 = 0$, with roots $(1/2,\ 5/2)$. The first root yields

$$\frac{1}{2}\begin{pmatrix} 1 & -3 \\ -1 & 3 \end{pmatrix}\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \Longrightarrow \quad \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \beta \begin{pmatrix} 3 \\ 1 \end{pmatrix} \quad \Longrightarrow \quad \mathbf{s}_1 = \begin{pmatrix} 3 \\ 1 \end{pmatrix},$$

with unconstrained $\beta$. We must check this result, and convince ourselves that the problem is in fact solved with $\lambda = 1/2$ and $\mathbf{s} = \begin{pmatrix} 3 & 1 \end{pmatrix}^T$;

$$\begin{pmatrix} 1 & -\frac{3}{2} \\ -\frac{1}{2} & 2 \end{pmatrix}\begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 - \frac{3}{2} \\ -\frac{3}{2} + 2 \end{pmatrix} = \begin{pmatrix} \frac{3}{2} \\ \frac{1}{2} \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 3 \\ 1 \end{pmatrix},$$

as required.

The second root yields

$$\frac{1}{2}\begin{pmatrix} -3 & -3 \\ -1 & -1 \end{pmatrix}\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \Longrightarrow \quad \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \beta \begin{pmatrix} -1 \\ 1 \end{pmatrix} \quad \Longrightarrow \quad \mathbf{s}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

The test

$$\begin{pmatrix} 1 & -\frac{3}{2} \\ -\frac{1}{2} & 2 \end{pmatrix}\begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 - \frac{3}{2} \\ \frac{1}{2} + 2 \end{pmatrix} = \frac{5}{2}\begin{pmatrix} -1 \\ 1 \end{pmatrix},$$

Figure 2: Time evolution of 2 examples of the 2-species eigensystem. The top panels show the total number of individuals (species 1 plus species 2) as a function of time. The lower panels show the individual species evolution.

is also satisfied, and our problem is solved. We can write down victoriously the complete solution to

$$\frac{d}{dt}\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 2 & -3 \\ -1 & 4 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

as the superposition of the 2 pure exponentials governing the evolution of the 2 modes (the 2 eigenvalue/eigenvector pairs)

$$\mathbf{x}(t) = A_1\, exp\left(\frac{1}{2}t\right)\mathbf{s}_1 + A_2\, exp\left(\frac{5}{2}t\right)\mathbf{s}_2,$$

where the amplitudes $A_1$ and $A_2$ are determined by the initial conditions. Let's pick $x(0) = y(0) = 1000$, in which case

$$\begin{pmatrix} 1000 \\ 1000 \end{pmatrix} = A_1\begin{pmatrix} 3 \\ 1 \end{pmatrix} + A_2\begin{pmatrix} -1 \\ 1 \end{pmatrix} \qquad \Longrightarrow \qquad A_1 = A_2 = 500,$$

yielding finally

$$\mathbf{x}(t) = 500\left[exp\left(\frac{1}{2}t\right)\begin{pmatrix} 3 \\ 1 \end{pmatrix} + exp\left(\frac{5}{2}t\right)\begin{pmatrix} -1 \\ 1 \end{pmatrix}\right].$$

The left panels of Fig. 2 shows the results of the system over one half time unit.

The right panels correspond to the different system whose governing matrix is

$$\mathbf{A} = \begin{pmatrix} 1 & -2 \\ -2 & 1 \end{pmatrix}.$$

The characteristic polynomial is

$$(1-\lambda)^2 - 4 = 0,$$

with roots $(3, -1)$. The corresponding eigenvectors are

$$\lambda_1 = 3,\ \mathbf{s}_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad \text{and} \quad \lambda_2 = 1,\ \mathbf{s}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

The usual tests yield

$$\begin{pmatrix} 1 & -2 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = 3 \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$
$$\begin{pmatrix} 1 & -2 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = -1 \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

as required. We choose to have the same initial total number of individuals (2000), which dictates the complete solution

$$\mathbf{x}(t) = 1000 \left[ e^{3t} \begin{pmatrix} 1 \\ -1 \end{pmatrix} + e^{-t} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right].$$

As the right panels of Fig. 2 show, the small difference in the growth rate of the fastest growing mode (the largest eigenvalue) is sufficient to give very different time-behavior from that of the system in the previous example. Note that it is often useful (and customary) to normalize the eigenvectors (which I did not do above).

Now let's take another look at the last example,

$$\mathbf{A} = \begin{pmatrix} 1 & -2 \\ -2 & 1 \end{pmatrix},$$

$$\lambda_1 = 3,\ \mathbf{s}_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad \text{and} \quad \lambda_2 = 1,\ \mathbf{s}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Let's first normalize the eigenvectors to unit norm, and form an eigenvector matrix **E** whose columns are the renormalized eigenvectors

$$\mathbf{E} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}.$$

Next, we obtain **E**'s inverse

$$\mathbf{E}^{-1} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}.$$

With these matrices,

$$\mathbf{E}^{-1}\mathbf{A}\mathbf{E} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ -2 & 1 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} =$$
$$= \frac{1}{2} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix} \overset{\text{def}}{=} \mathbf{\Lambda}.$$

Notice that $\mathbf{\Lambda}$ is the eigenvalue matrix, with the eigenvalues along the diagonal. This is one of the many important aspects of eigenanalysis:

$$\mathbf{E}^{-1}\mathbf{A}\mathbf{E} = \mathbf{\Lambda} \quad \text{or} \quad \mathbf{A} = \mathbf{E}\mathbf{\Lambda}\mathbf{E}^{-1}.$$

If the eigenvectors are mutually orthogonal, this can be simplified even further, to

$$\mathbf{E}^T\mathbf{A}\mathbf{E} = \mathbf{\Lambda} \qquad \text{or} \qquad \mathbf{A} = \mathbf{E}\mathbf{\Lambda}\mathbf{E}^T.$$

Let's look at some interesting and revealing examples.

- Failure to diagonalize: With

  $$\mathbf{A} = \begin{pmatrix} 0 & 3 \\ 0 & 0 \end{pmatrix},$$

  the characteristic polynomial is $(-\lambda)^2 = 0$, with $\lambda_{1,2} = 0$. The corresponding eigenvectors satisfay

  $$\begin{pmatrix} 0 & 3 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

  which yields $\mathbf{e}_1 = \begin{pmatrix} 1 & 0 \end{pmatrix}^T$ only. In this case there are not enough eigenvectors to form $\mathbf{E}$ of the necessary dimension $(2 \times 2)$, and diagonalization fails. Note that this is *not* because of $\lambda_{1,2} = 0$, or even because $\lambda_1 = \lambda_2$; it is the twice-repeated eigenvalue, whose *algebraic multiplicity* is 2, but whose *geometric muliplicity* fails to achieve the required 2 ($\mathbf{A}$ has only a one dimensional nullspace).

- Things are great when $\mathbf{A}$ yields an orthonormal $\mathbf{E}$, as with, e.g.,

  $$\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

  The characteristic polynomial is $(1-\lambda)^2 - 1 = 0$, with $\lambda_1 = 0$ and $\lambda_2 = 2$. The eigenvector equations give

  $$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \Longrightarrow \quad \mathbf{e}_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

  $$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = 2\begin{pmatrix} \alpha \\ \beta \end{pmatrix} \quad \Longrightarrow \quad \mathbf{e}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

  which we normalize and use to construct

  $$\mathbf{E} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}.$$

  Now notice that

  $$\mathbf{E}^T\mathbf{E} = \left(\frac{1}{\sqrt{2}}\right)^2 \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = \mathbf{I},$$

  i.e., $\mathbf{E}$'s transpose is also its inverse!! this is a general property og orthonormal matrices, which makes them very convenient.

- Powers of a matrix. Consider the square of the above

$$\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix},$$

$$\mathbf{A}^2 = \mathbf{A}\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}.$$

The characteristic equation, $(2-\lambda)^2 - 4 = 0$, yields $\lambda_1 = 0$ and $\lambda_2 = 4$, the square of **A**'s eigenvalues. The eigenvectors are

$$\begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \Longrightarrow \quad \mathbf{e}_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

$$\begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = 4\begin{pmatrix} \alpha \\ \beta \end{pmatrix} \quad \Longrightarrow \quad \mathbf{e}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

just like **A**'s. This too is a general property of diagonalizable matrices; their eigenspace is invariant under raising to an arbitrary power, while the eigenvalues are raised to the required power.