Analysis of Linear Equations

Motivation – To determine the dynamics of a system of equations in more than one variable.

Consider the recursion equations for any model that describes the change in state of a population from one generation ($x_i[t]$) to the next ($x_i[t+1]$). To make it easier to write, we will use x_i to denote the variables in the current generation and x_i' to denote the variables in the next generation. In this handout, we will consider only LINEAR functions of the variables (e.g. $x_1' = j_{11}x_1 + j_{12}x_2$ but not $x_1' = j_{11}x_1x_2$). If there are n variables then there will be n functions describing the change in these variables:

$$\begin{aligned} x_1' &= j_{11}x_1 + j_{12}x_2 + ... j_{1n}x_n, \\ x_2' &= j_{21}x_1 + j_{22}x_2 + ... j_{2n}x_n, \\ &... \\ x_n' &= j_{n1}x_1 + j_{n2}x_2 + ... j_{nn}x_n \end{aligned} \tag{1}$$

(e.g. in a predator-prey model, $n = 2$ since we have to track both the number of predators and the number of prey). Since the equations are linear, we can also write these equations in matrix form:

$$\begin{pmatrix} x_1' \\ x_2' \\ ... \\ x_n' \end{pmatrix} = \begin{pmatrix} j_{11} & j_{12} & ... & j_{1n} \\ j_{21} & j_{22} & ... & j_{2n} \\ ... & ... & ... & ... \\ j_{n1} & j_{n2} & ... & j_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ ... \\ x_n \end{pmatrix} \tag{2}$$

Denoting the matrix by $\mathbf{J}$ and the vector of x_i by $\vec{x}$, we can then write equation (2) as

$$\vec{\mathbf{x}}' = \mathbf{J}\vec{\mathbf{x}}. \tag{3}$$

$\mathbf{J}$ is known as a *transition* matrix, since it describes how the population vector changes from one generation to the next. To find out where the population will be at some generation t (described by the vector $\vec{\mathbf{x}}[\mathbf{t}]$), we can use equation (3) over and over again: $\vec{\mathbf{x}}[\mathbf{t}] = \mathbf{J}\ \vec{\mathbf{x}}[\mathbf{t-1}] = \mathbf{J^2}\ \vec{\mathbf{x}}[\mathbf{t-2}] ... = \mathbf{J^t}\ \vec{\mathbf{x}}[\mathbf{0}]$. In most cases, it will be hard to find out what $\mathbf{J^t}$ equals directly, so we must digress for a moment to review some basic theorems from linear algebra that can help. These will be used to determine what happens to the vector, $\vec{\mathbf{x}}$ over time.

A Digression into Linear Algebra – A number λ is an *eigenvalue* of matrix $\mathbf{J}$ if there exists a non-zero vector, $\vec{\mathbf{v}}$, that satisfies the equation:

$$\mathbf{J}\vec{\mathbf{v}} = \vec{\mathbf{v}}\lambda. \tag{4}$$

Every vector satisfying this relation is an *eigenvector* of $\mathbf{J}$ belonging to the eigenvalue, λ. To find the eigenvalues of a matrix, note that we can rearrange[1] equation (4) as $\mathbf{J}\vec{\mathbf{v}} - \lambda\vec{\mathbf{v}} = (\mathbf{J} - \lambda\mathbf{I})\vec{\mathbf{v}} = \vec{\mathbf{0}}$, where $\mathbf{I}$ is the *identity matrix* (a diagonal matrix with ones along the diagonal), and $\vec{\mathbf{0}}$ is a vector of zeros. A matrix, like $(\mathbf{J} - \lambda\mathbf{I})$, which equals zero when multiplied by some non-zero vector $\vec{\mathbf{v}}$ is called *singular*. Singular matrices have the property that their determinant equals zero. This means that the determinant of $(\mathbf{J} - \lambda\mathbf{I})$ equals zero, which is written as $|\mathbf{J} - \lambda\mathbf{I}| = 0$. This determinant is an n^{th} degree polynomial in λ, the roots of which are the eigenvalues of the matrix $\mathbf{J}$: $\lambda_1, \lambda_2, ...\lambda_n$. For example, in the $n = 2$ case,

$$(\mathbf{J} - \lambda\mathbf{I}) = \begin{pmatrix} j_{11} - \lambda & j_{12} \\ j_{21} & j_{22} - \lambda \end{pmatrix} \tag{5}$$

so that

$$|\mathbf{J} - \lambda\mathbf{I}| = (j_{11} - \lambda)(j_{22} - \lambda) - j_{21}j_{12} = \lambda^2 - \lambda(j_{11} + j_{22}) + j_{11}j_{22} - j_{21}j_{12} = 0. \tag{6}$$

The two roots[2] to this equation are the two eigenvalues.

The analysis of the transition matrix $\mathbf{J}$ can be simplified by changing the coordinate system (or *basis*[3]). That is, we can look at the recursions from a different vantage point and they'll look simpler, but all we've done is change our viewpoint and not the dynamical behavior of the system. If $\mathbf{J}$ has n linearly independent eigenvectors, then $\mathbf{J}$ can be transformed into another coordinate system in which the transition matrix is a diagonal matrix, $\mathbf{D}$, which is much easier to analyze and whose diagonal elements are the corresponding n

[1] Using the distributive law for matrix multiplication.

[2] The two roots can be found using the quadratic formula: $\lambda_1 = \frac{-b+\sqrt{b^2-4ac}}{2a}$ and $\lambda_2 = \frac{-b-\sqrt{b^2-4ac}}{2a}$

[3] The *basis* is the co-ordinate system in which the vectors are measured. For instance, in a regular two dimensional plot, the x-axis and the y-axis provide the co-ordinate system in which everything is measured. The basis in which measurements are taken can be changed or *transformed*. This basically superimposes a different grid onto the system of equations but doesn't change their behavior. For instance, you could transform x-y coordinates into polar coordinates, but that wouldn't change what was happening over time.

eigenvalues:

$$\mathbf{D} = \begin{pmatrix} \lambda_1 & 0 & ... & 0 \\ 0 & \lambda_2 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & \lambda_n \end{pmatrix} \tag{7}$$

(In linear algebra terms, matrix $\mathbf{J}$ is *similar* to matrix $\mathbf{D}$.) The advantage of performing this transformation is that while $\mathbf{J^t}$ is hard to compute, $\mathbf{D^t}$ is easy to compute.

To change coordinate systems so that the transition matrix is diagonal we do the following. Let $\mathbf{A}$ equal the matrix whose columns are the n eigenvectors ($\vec{\mathbf{v}}$) that satisfy equation (4), then $\mathbf{A}$ corresponds to a *transformation* matrix from the original co-ordinate system (which represented the number or frequency of each type separately), into the co-ordinate system based on the n eigenvectors, $\{\vec{\mathbf{v}}_\mathbf{1}, \vec{\mathbf{v}}_\mathbf{2} ... \vec{\mathbf{v}}_\mathbf{n}\}$. To transform a matrix from one basis into another, the following operation is performed:

$$\mathbf{A^{-1}JA}, \tag{8}$$

where $\mathbf{A^{-1}}$ is the *inverse* of $\mathbf{A}$, which means that $\mathbf{A^{-1}A} = \mathbf{AA^{-1}} = \mathbf{I}$. Since each of the columns in matrix $\mathbf{A}$ are eigenvectors, we can use equation (4) to show that $\mathbf{JA} = \mathbf{AD}$. For instance, when $n = 2$,

$$\begin{pmatrix} j_{11} & j_{12} \\ j_{21} & j_{22} \end{pmatrix} \begin{pmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{pmatrix} = \begin{pmatrix} \lambda_1 v_{11} & \lambda_2 v_{12} \\ \lambda_1 v_{21} & \lambda_2 v_{22} \end{pmatrix} = \begin{pmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \tag{9}$$

where

$$\mathbf{J} = \begin{pmatrix} j_{11} & j_{12} \\ j_{21} & j_{22} \end{pmatrix}, \quad \mathbf{A} = \begin{pmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{pmatrix}, \quad \text{and} \quad \mathbf{D} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \tag{10}$$

This works because each of the columns in the $\mathbf{A}$ matrix is an eigenvector. Since $\mathbf{JA} = \mathbf{AD}$, we can multiply both sides of the equation on the left by $\mathbf{A^{-1}}$ to get $\mathbf{A^{-1}JA} = \mathbf{A^{-1}AD} = \mathbf{D}$. In other words, in the new basis composed of eigenvectors, the transition matrix is $\mathbf{D}$.

To transform a vector (e.g. $\vec{x}$) from the old basis into the new basis, the following operation is performed:

$$\vec{\mathbf{x}}_{\mathbf{new}} = \mathbf{A^{-1}}\vec{\mathbf{x}}. \tag{11}$$

Back to Analysing a System of Linear Equations – Take the recursions described by equations (3) and multiply both sides by $\mathbf{A^{-1}}$ on the left. We then get:

$$\begin{aligned} \mathbf{A^{-1}}\vec{\mathbf{x}}\,' &= \mathbf{A^{-1}J}\,\vec{\mathbf{x}} = \mathbf{A^{-1}J\;I}\,\vec{\mathbf{x}} \\ \rightarrow \vec{\mathbf{x}}\,'_{\mathbf{new}} &= \mathbf{A^{-1}J\;AA^{-1}}\,\vec{\mathbf{x}} \\ \rightarrow \vec{\mathbf{x}}\,'_{\mathbf{new}} &= \mathbf{D}\,\vec{\mathbf{x}}_{\mathbf{new}}. \end{aligned} \tag{12}$$

Equation (12) gives the recursions viewed from the new basis. The wonderous trick of all this is that equation (12) is easy to iterate:

$$\vec{\mathbf{x}}_{\mathbf{new}}[\mathbf{t}] = \mathbf{D}\,\vec{\mathbf{x}}_{\mathbf{new}}[\mathbf{t-1}] = \mathbf{D^2}\,\vec{\mathbf{x}}_{\mathbf{new}}[\mathbf{t-2}] = ... = \mathbf{D^t}\,\vec{\mathbf{x}}_{\mathbf{new}}[\mathbf{0}], \tag{13}$$

where

$$\mathbf{D^t} = \begin{pmatrix} \lambda_1^t & 0 & ... & 0 \\ 0 & \lambda_2^t & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & \lambda_n^t \end{pmatrix}. \tag{14}$$

It would not have been so easy to find $\mathbf{J^t}$!

We can see what is happening in the original basis by noting that the equation $\mathbf{D} = \mathbf{A^{-1}JA}$ may be multiplied by matrix $\mathbf{A}$ on the left and matrix $\mathbf{A^{-1}}$ on the right to give $\mathbf{ADA^{-1}} = \mathbf{AA^{-1}\;J\;AA^{-1}} = \mathbf{J}$. This is the transformation to go from the new basis back to the old one. For instance, after an amount of time t, we can find $\mathbf{J^t}$ by writing it as $(\mathbf{ADA^{-1}})^{\mathbf{t}}$. This can be written as the product of $(\mathbf{ADA^{-1}})$ times itself t times, but $(\mathbf{ADA^{-1}})(\mathbf{ADA^{-1}}) = \mathbf{AD(A^{-1}A)DA^{-1}} = \mathbf{AD^2A^{-1}}$, etc. Therefore, the transition matrix for the population over a period of time, t, is $\mathbf{J^t} = (\mathbf{ADA^{-1}})^{\mathbf{t}} = \mathbf{AD^tA^{-1}}$. This provides the *general solution* to the recursion equations.

Summary – Although a transition matrix may be difficult to iterate to determine how a linear system of equations changes over time, we can transform the recursions into a new basis (specified by the eigenvectors) in which the transition matrix is a diagonal matrix. It is then easy to iterate the diagonal matrix to find out where the population will be any time in the future. Since a change in basis is simply a change in 'vantage point', this transformation doesn't change the behavior of the dynamics. In fact, we can back-transform to get the general solution in the original basis from the general solution in the new basis.