**MOL 437: A Brief Review of Linear Algebra**

An understanding of linear algebra is critical as a stepping-off point for understanding neural networks. This handout includes basic definitions, then quickly progresses to elementary but powerful techniques such as eigenbases. For your private edification, a few exercises are included, identified by bullets •; some exercises come with hints and answers. Take what you will from this handout, but be forewarned that future problem sets *will* require most of the concepts developed here, so it behooves you to be comfortable with them. Don't delay in asking a TA if you can't figure it out on your own.

# 1 STARTING DEFINITIONS

## 1.1 Matrix structure

Matrices are most often represented as rectangular arrays of scalars.[1] The $m \times n$ matrix $\mathbf{A}$ has $m$ rows and $n$ columns. The subscript notation $A_i$ is used to reference the $i$th row of the matrix, and $A_{ij}$ is used to reference the scalar in the $i$th row and $j$th column of $\mathbf{A}$.

For example, $\mathbf{A}$ is a 2×3 matrix:

$$\mathbf{A} = \begin{bmatrix} 4 & 2 & -5 \\ 1 & 0 & -8 \end{bmatrix}, \qquad A_{13} = -5$$

A *column vector* (quite often referred to simply as a vector) is an $n \times 1$ matrix, where $n$ is referred to as the *dimension* of the vector. The scalar $v_i$ is the $i$th element of vector $\mathbf{v}$.

A matrix with the same number of rows and columns is, not surprisingly, referred to as a *square matrix*. A commonly used notational convention is to use capital letters (i.e., $\mathbf{A}$) to denote matrices and lower case letters (i.e., $\mathbf{v}$) to denote vectors.

## 1.2 Matrix transpose

The transpose of $m \times n$ matrix $\mathbf{A}$ is denoted $\mathbf{A}^T$. $\mathbf{A}^T$ is an $n \times m$ matrix whose elements are:

$$(A^T)_{ij} = A_{ji}$$

The transpose of a column vector is called a *row vector*. An object twice transposed will produce the original object: $(\mathbf{A}^T)^T = \mathbf{A}$.

## 1.3 Addition and multiplication

Adding two matrices $\mathbf{A}$ and $\mathbf{B}$ results in a matrix whose elements are the sums of the corresponding elements from $\mathbf{A}$ and $\mathbf{B}$:

$$\text{If} \quad \mathbf{C} = \mathbf{A} + \mathbf{B}, \qquad \text{then} \quad C_{ij} = A_{ij} + B_{ij}$$

[1]In the examples presented, scalars will be real numbers, but in general they can be complex.

($\mathbf{A}$ and $\mathbf{B}$ must have the same dimensions to be able to add them together.) Addition is commutative and associative, just like regular addition.

A matrix $\mathbf{A}$ multiplied by a scalar $k$ produces a new $\mathbf{B} = k\mathbf{A}$ whose elements are the elements of $\mathbf{A}$ each multiplied by $k$. Multiplying two matrices together is more complicated: multiplying $m \times n$ matrix $\mathbf{A}$ by $n \times p$ matrix $\mathbf{B}$ produces an $m \times p$ matrix $\mathbf{C} = \mathbf{AB}$ whose elements are defined to be:

$$\text{If} \quad \mathbf{C} = \mathbf{AB}, \qquad \text{then} \quad C_{ik} = \sum_{j=1}^{n} A_{ij} B_{jk}$$

In this example, a $4 \times 2$ matrix is multiplied by a $2 \times 3$ matrix to produce a $4 \times 3$ matrix:

$$\begin{bmatrix} 2 & 3 \\ 4 & 0 \\ 3 & -2 \\ 5 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 \\ 5 & -1 & 3 \end{bmatrix} = \begin{bmatrix} 17 & 1 & 9 \\ 4 & 8 & 0 \\ -7 & 8 & -6 \\ 10 & 9 & 3 \end{bmatrix}$$

Note that matrix multiplication can only be performed between two matrices $\mathbf{A}$ and $\mathbf{B}$ if $\mathbf{A}$ has exactly as many columns as $\mathbf{B}$ has rows.

Like ordinary multiplication, matrix multiplication is associative and distributive, but unlike ordinary multiplication, *it is not commutative*:

- $\mathbf{AB} \neq \mathbf{BA}$, in general

From the definitions of multiplication and transpose, we derive the following identity:

$$(\mathbf{AB})^T = \mathbf{B}^T \mathbf{A}^T$$

## 1.4 Inner product

The *inner product* (also known as the *dot product*) of $n$-dimensional vectors $\mathbf{x}$ and $\mathbf{y}$ is defined as $\mathbf{x}^T\mathbf{y}$ which is a scalar[2]. By our definitions of matrix transpose and matrix multiplication, this means that the inner product is the sum of the products of corresponding elements from the two vectors:

$$\mathbf{x}^T\mathbf{y} = \sum_{i=1}^{n} x_i y_i$$

If the inner product of two vectors is zero, they are said to be *orthogonal*, which has the usual geometric connotation of perpendicularity.

## 1.5 Square matrices

The *diagonal* of an $n \times n$ square matrix $\mathbf{A}$ are the elements $\mathbf{A}_{ii}$ running diagonally from the top left corner to the bottom right. A *diagonal matrix* is a matrix which has zeroes everywhere off the diagonal.

[2]When working with complex vectors, we use the inner product $\mathbf{x}^*\mathbf{y}$, which returns a real value when $\mathbf{y} = \mathbf{x}$. $\mathbf{x}^*$ is the complex conjugate of the transpose of $\mathbf{x}$. The complex conjugate $\overline{\mathbf{x}}$ has $Re[\overline{x}_i] = Re[x_i]$ and $Im[\overline{x}_i] = -Im[x_i]$.

The symbol $\mathbf{I}$ is reserved for a particular diagonal matrix known as the *identity matrix*, which has ones along its diagonal and zeroes elsewhere. It is the multiplicative identity for matrix multiplication of square matrices. In other words, given any $n \times n$ square matrix $\mathbf{A}$, it has the following property: $\mathbf{AI} = \mathbf{IA} = \mathbf{A}$.

The $n \times n$ square matrix $\mathbf{A}$ is called *invertible* if there exists a matrix denoted $\mathbf{A}^{-1}$ which statisfies:

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$$

If $\mathbf{A}^{-1}$ exists, it is called the *inverse matrix*. If $\mathbf{A}^{-1}$ does not exist, $\mathbf{A}$ is called a *singular* matrix. The inverse of the inverse matrix is simply the original matrix.

- Show that $(\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T$
- Show that $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$

All square matrices have a particular scalar value associated with them, known as the *determinant*, which is written as

$$\det\ \mathbf{A} = |\mathbf{A}|$$

For two dimensions, the formula for calculating the determinant is simple:

$$\begin{vmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{vmatrix} = A_{11}A_{22} - A_{12}A_{21}$$

In general, the algebraic formula for determinants is more complicated, but there is a simple and very useful recursive definition (which you can look up in any linear algebra book). For your amusement, the algebraic formula for the determinant of an $n \times n$ matrix can be summarized as

$$|\mathbf{A}| = \sum_{\sigma} sign(\sigma) A_{1\sigma_1} A_{2\sigma_2} \dots A_{n\sigma_n}$$

where the sum is over all $n!$ permutations of $(1 \dots n)$, and $sign(\sigma)$ is $+1$ if $\sigma$ is an *even* permutation, else $-1$ if $\sigma$ is an *odd* permutation[3]. Happily, determinants can be quite useful even without calculating them. A few facts about determinants include $|\mathbf{AB}| = |\mathbf{A}|\ |\mathbf{B}|$, and that if the determinant of a matrix is zero, the matrix is singular, which we'll use below.

## 2 EIGENVALUES AND EIGENVECTORS

A simple way to find if a matrix $\mathbf{A}$ is invertible or not is to find its determinant, since

$$\det \mathbf{A} = \mathbf{0} \qquad \text{if and only if} \qquad \mathbf{A} \text{ is not invertible}$$

If $\mathbf{A}$ is invertible, the only $\mathbf{x}$ satisfying $\mathbf{Ax} = \mathbf{0}$ is $\mathbf{x} = \mathbf{0}$ (why?). On the other hand, if $\mathbf{A}$ is *not* invertible, there can be interesting non-zero solutions for $\mathbf{x}$.

[3] *Even* and *odd* refer to how many swaps of adjacent elements are required to transform $(1 \dots n)$ into $\sigma$. For $n = 3$, the even permutations are $\{(1\ 2\ 3), (2\ 3\ 1), (3\ 1\ 2)\}$.

• Find a condition for which the equation

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$$

($\lambda$ a scalar) has interesting non-zero solutions for $\mathbf{x}$.

• Use this condition to write an equation that $\lambda$ must satisfy in order to get non-zero solutions to the following equation:

$$\begin{bmatrix} 5 & -1 \\ -2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{1}$$

Solve the equation you got for $\lambda$. You will get two possible values. Using one of them, find values for $x_1$ and $x_2$ that satisfy equation (1).

Note that if $\mathbf{x}$ satisfies $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$, so does $\alpha\mathbf{x}$. So you won't be able to solve for a *unique* $x_1$ and $x_2$, just for the direction that $\mathbf{x}$ should lie in. That direction is called the eigenvector direction, and all vectors parallel to it are eigenvectors with the same eigenvalue.

• Now use the other $\lambda$ to find the other eigenvector direction.

To restate, if $\mathbf{x} \neq \mathbf{0}$ and $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$ then $\lambda$ is an eigenvalue of $\mathbf{A}$ and $\mathbf{x}$ is an eigenvector of $\mathbf{A}$ with eigenvalue $\lambda$.

For larger square matrices, we can find eigenvalues and eigenvectors using the same approach you used for the $2 \times 2$ matrix. First, we look for values of $\lambda$ such that $\mathbf{A} - \mathbf{I}\lambda$ is singular, i.e., $|\mathbf{A} - \mathbf{I}\lambda| = 0$. Using the formula for determinants, this leads to a polynomial of degree $n$, which is called the *characteristic polynomial* of $\mathbf{A}$. You will recall from algebra that every polynomial of degree $n$ has exactly $n$ (not necessarily distinct) complex roots (some of which may be real, of course). Therefore, every matrix $\mathbf{A}$ has exactly $n$ (not necessarily distinct and possibly complex) eigenvalues. Once the eigenvalues are known, the eigenvectors can be determined.

• Find a matrix $\mathbf{A}$ for which 0 is an eigenvalue, and find all the eigenvectors.

A common convention is to chose eigenvectors to be unit vectors[4], i.e. $\mathbf{x}^T\mathbf{x} = 1$.

## 2.1 Some words about eigenvalues

If you write the eigenvector directions as column vectors and put them side by side, you get a new matrix— call it $\mathbf{E}$. Convince yourself that since the columns of $\mathbf{E}$ are eigenvectors, the following is true:

$$\mathbf{A}\mathbf{E} = \mathbf{E} \cdot \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

where $\lambda_1$ is the first column's eigenvalue, and $\lambda_2$ is the second column's eigenvalue We can multiply on the right by $\mathbf{E}^{-1}$ (assuming $\mathbf{E}$ is invertible) to get

$$\mathbf{A} = \mathbf{E} \cdot \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \cdot \mathbf{E}^{-1}$$

This expression helps us describe the true significance of the matrix $\mathbf{A}$, and why eigenvalues are so important.

[4]If the eigenvectors are complex, there is no obvious way to find a unique representation for the eigenvector by normalizing, since an eigenvector multiplied by any complex number is still an eigenvector, with the same eigenvalue. For convenience one can set $\mathbf{x}^*\mathbf{x} = 1$.

Multiplying on the left by $\mathbf{A}$ is the same as performing the following sequence of operations:

1. first multiplying by $\mathbf{E}^{-1}$. Think of this as doing a linear change of coordinates. That is, we change coordinates to some special coordinate system.
2. In that coordinate system, multiply by $\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$. But this is a particularly easy matrix to multiply with, since the coordinates don't mix! We can easily visualize what is going on: the first coordinate gets stretched (or squeezed) by $\lambda_1$, the second by $\lambda_2$.
3. Then go back to your original coordinates, by multiplying by the inverse of $\mathbf{E}^{-1}$, namely $\mathbf{E}$.

That is, there is some special coordinate system in which multiplying by $\mathbf{A}$ just stretches the two coordinates independently. Clearly this is the natural coordinate system for the problem, the one we want to be thinking in. Many times it is enough to know the eigenvalues: we know we could always transform the problem into the special system if we wanted to. We just pretend that we've already done the transformation.

This illustrates a very important concept that cannot be stressed enough: the real *guts* of a matrix, what it really does, <u>don't depend on what coordinate system we use to describe it</u>. Here, if $\mathbf{A}$ is a positive definite matrix, then $\mathbf{E}$ is an orthonormal matrix and represents simply a rotation.[5] Who cares if we rotate coordinates around? They're *our* coordinates, not the physical problem's. The eigenvalues are what really matter.

• Find the eigenvalues and eigenvectors of the following two matrices:

$$\begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix} \qquad \text{and} \qquad \begin{bmatrix} 10 & 2 & 2 \\ 2 & 22 & -5 \\ 2 & -5 & 22 \end{bmatrix}$$

The arguments we used above all rely on our casual assumption that $\mathbf{E}$ is invertible. This is usually an acceptable assumption to make, for the sorts of matrices commonly encountered in neural network theory. But it can be helpful to have some understanding of the other possibilities.[6]

## 2.2 Some more useful facts about eigenvectors

In the following, assume when necessary that $\mathbf{E}$ is invertible, with eigenvectors $\mathbf{e}_1 \dots \mathbf{e}_n$ and corresponding eigenvalues $\lambda_1 \dots \lambda_n$.

• If $\mathbf{C}$ has eigenvectors and eigenvalues $\{\mathbf{e}_i, \lambda_i\}$, then the matrix $\mathbf{B} = \mathbf{C} - \gamma\mathbf{I}$ has eigenvectors and eigenvalues $\{\mathbf{e}_i, \lambda_i - \gamma\}$.

• If $\mathbf{C}$ is a real symmetric matrix, (i.e. $C_{ij} = C_{ji}$), then all the eigenvalues of $\mathbf{C}$ are *real*. We can also choose all the eigenvectors to be real. [7]

[5] Positive definite matrix: a matrix $\mathbf{M}$ such that $\mathbf{x}^T\mathbf{M}\mathbf{x} > 0$ for all non-zero $\mathbf{x}$. Orthonormal matrix: One where $\mathbf{M}^T\mathbf{M} = \mathbf{I}$. The important point is that if these conditions are satisfied, the matrix $\mathbf{E}$ is just a coordinate rotation and/or a reflection. (Though reflections aren't properly rotations, we almost always include them when we say, abusing notation, "rotation matrix".

[6] A quick summary: Suppose the $n$ (possibly complex) eigenvalues of $\mathbf{M}$ are distinct. Then to each eigenvalue there is a unique (up to multiplication by a complex number) eigenvector, and all the eigenvectors are linearly independent (i.e. they span $\mathcal{C}^\backslash$). Now suppose there are $m$ eigenvalues with the same value $\lambda$. In this case, unfortunately, there might not be $m$ linearly independent eigenvectors all with the same eigenvalue $\lambda$, in which case $\mathbf{E}$ must be singular. In both these cases, there is nothing special about the eigenvalue 0 – the issue is only whether an eigenvalue is a multiple root of the characteristic polynomial.

[7] Hint: Start with a potentially complex eigenvalue $\lambda$ and its potentially complex eigenvector $\mathbf{x}$, satisfying $\mathbf{C}\mathbf{x} = \lambda\mathbf{x}$.

• If $\mathbf{C}$ is a real symmetric matrix, any two eigenvectors of $\mathbf{C}$ with different eigenvalues are **orthogonal**. [8]

• Let be $\mathbf{x}$ a random vector, and define the cross correlation matrix:

$$C_{ij} = \langle x_i x_j \rangle$$

that is, the $ij$th component of $\mathbf{C}$ is the expected value (i.e. mean value) of the product $x_i x_j$. $\mathbf{C}$ is symmetric, and all its eigenvalues are positive, i.e., $\lambda_i \geq 0$. [9]

Think about the vector defined by the row of intensities in one horizontal line on a television screen. What do you think the cross correlation matrix $\mathbf{C}$ of that vector looks like, averaged over many pictures? What do you think the principal eigenvector of $\mathbf{C}$ looks like? (The principal eigenvector is the one with largest eigenvalue.) What about the next few principal components? [10]

• We can express a vector $\mathbf{w}$ in terms of the complete orthonormal set of eigenvectors $\mathbf{e}_i$:

$$\mathbf{w} = \sum_i \omega_i \mathbf{e}_i$$

where $\omega_i = \mathbf{w}^T \mathbf{e}_i$. The $\omega_i$ are the components of $\mathbf{w}$ in the eigenvector basis.

• If $\mathbf{Cw} = \mathbf{b}$, we can write an explicit solution for $\mathbf{w} = \mathbf{C}^{-1}\mathbf{b}$ in terms of $\mathbf{b}$ and $\{\mathbf{e}_i, \lambda_i\}$ (assuming all $\lambda_i > 0$):

$$\mathbf{C}^{-1} = \sum_i \frac{(\mathbf{e}_i^T \cdot \mathbf{b})}{\lambda_i} \mathbf{e}_i$$

# 3 LINEAR DIFFERENTIAL EQUATIONS

The simplest linear differential equation is the one variable equation

$$\frac{dx}{dt} = \lambda x$$

where $\lambda$ is a scalar.

• Confirm for yourself that its solution is an exponential,

$$x(t) = x(0)e^{\lambda t} \tag{2}$$

where $x(0)$ is the initial condition. Clearly, $x = 0$ is a fixed point of this equation since then $\dot{x} = 0$, that is, $x$ stays put.[11] We can ask about the *stability* of this fixed point: if we were to add a little

By convention, we've chosen a unit eigenvector (why can we always do this?) so $\mathbf{x}^*\mathbf{x} = 1$. Combine these two equations and $\mathbf{C}^* = \mathbf{C}$ to show that $\lambda^* = \lambda$. We now need to show that $\mathbf{x}$ is real; look at what $\mathbf{C}$ does to the real and imaginary components of $\mathbf{x}$ independently.

[8] Hint: if $\mathbf{x}_1$ and $\mathbf{x}_2$ are eigenvectors with eigenvalues $\lambda_1$ and $\lambda_2$, use the definition of an eigenvector to show that $\mathbf{x}_1^T\mathbf{x}_2 = \frac{\lambda_1}{\lambda_2}\mathbf{x}_1^T\mathbf{x}_2$, or something similar.

[9] Hint: to prove that all $\lambda_i \geq 0$, it is sufficient to show that $\mathbf{y}^T\mathbf{C}\mathbf{y} \geq 0$ for any vector $\mathbf{y}$.

[10] Hint 1: The cross correlation matrix has to be all positive, if the pixel intensities are positive. It is plausible that the correlation between two pixels should just be a function of their separation, so that the matrix should have a banded symmetrical appearance. (A matrix for which $C_{ij} = c(i-j)$ is called a Toeplitz matrix.) The correlation must be biggest on the diagonal because a pixel is most correlated with itself. Away from the diagonal, the correlations must get smaller, but not necessarily monotonically.

Hint 2: The principal eigenvector of an all–positive matrix must be an all–positive vector. Imagine multiplying a not quite all–positive vector repeatedly by $\mathbf{C}$, and think what happens to it; repeated multiplication yields a vector looking more and more like the principal eigenvector. Think of a monotonic Toeplitz $\mathbf{C}$ and you should confirm that this vector must end up looking like a symmetrical hump.

Hint 3: What is the next thing to having no changes of sign in a vector?

[11] $\dot{x}$ is a notational variant of $\frac{dx}{dt}$ and $\ddot{x}$ is a notational variant of $\frac{d^2x}{dt^2}$.

disturbance to $x$, would $x$ return to the fixed point, or would it shoot off in some direction? The stability of fixed points is of great practical importance in a world full of natural small random disturbances. For example, the bottom of a spherical bowl is a stable fixed point: fruit stays down there. But the top of a glass sphere is an unstable fixed point: we could very carefully balance an apple on top of it – but any small disturbance, and the apple will fall off.

- For $\frac{dx}{dt} = \lambda x$, convince yourself that $x = 0$ is a stable fixed point if $\lambda < 0$ and is unstable if $\lambda > 0$.

The phrasing and solution to the above problem assume $\lambda$ is real. What if $\lambda$ is complex? Convince yourself that equation 2 still holds. The imaginary part of $\lambda$ just represents an oscillation ($e^{iwt} = \cos wt + i \sin wt$). So the condition above, to be completely general, should really read "stable fixed point if the real part of $\lambda < 0$, unstable if the real part of $\lambda > 0$". What happens if the real part of $\lambda = 0$ exactly?

Now consider the following equation:

$$\ddot{x} = -\gamma\dot{x} - \omega x$$

One of the nice things about linear differential equations is that we can always take a single $n$-th order equation and turn it into $n$ coupled first-order equations by rewriting some of the variables. So we define $x_1 \equiv \dot{x}, \quad x_2 \equiv x$, to get the equivalent equations

$$\begin{bmatrix} \dot{x_1} \\ \dot{x_2} \end{bmatrix} = \begin{bmatrix} -\gamma & -\omega \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{3}$$

- Convince yourself that these equations indeed represent the same system.

You will have noticed that we already wrote this down in matrix form. We will now get a chance to use what we saw in section 2. Call the vector on the left $\dot{\mathbf{x}}$, the matrix on the right hand side $\mathbf{A}$, and the vector on the right hand side $\mathbf{x}$, so the equation is $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$.

We said that we can find special coordinates where our matrix doesn't mix coordinates (that is, it is a diagonal matrix). Suppose that we find matrices $\mathbf{E}$ and $\Lambda$, where $\Lambda$ is a diagonal matrix that holds the eigenvalues, such that $\mathbf{A} = \mathbf{E}\Lambda\mathbf{E}^{-1}$, as in section 2.1. Then

$$\dot{\mathbf{x}} = \mathbf{E}\Lambda\mathbf{E}^{-1}\mathbf{x}$$

Multiplying on the left by $\mathbf{E}^{-1}$, and remembering that as a linear operation in commutes with differentiation by time, we get

$$\frac{d}{dt}(\mathbf{E}^{-1}\mathbf{x}) = \Lambda(\mathbf{E}^{-1}\mathbf{x})$$

Let's just say that we define new coordinates $\mathbf{x}' = \mathbf{E}^{-1}\mathbf{x}$. Then we get an equation that looks like

$$\begin{bmatrix} \dot{x}'_1 \\ \dot{x}'_2 \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} x'_1 \\ x'_2 \end{bmatrix}$$

But this is just two completely sparate equations, each one in the simple single-variable form we saw at the beginning of this section! We know how to solve that, and how to know whether their fixed point is stable; and since these equations are the same as our original ones (simply represented in different coordinates), if these are stable so are the original ones, and vice-versa.

- In the following system,

$$\begin{bmatrix} \dot{x_1} \\ \dot{x_2} \end{bmatrix} = \begin{bmatrix} -2 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

is the fixed-point (0,0) stable or unstable? Why? You need to consider both equations at once.

- In equation 3, if $\gamma = 3$ and $\omega = 1$, is (0,0) stable or unstable?
- How about if $\gamma = 2$ and $\omega = 2$?

Note that the fixed point doesn't always has to be at (0,0). We just put it there in these examples for simplicity. The eigenvalue analysis still holds, however.

# 4 THE TRACE IS THE SUM OF THE EIGENVALUES

Take an $n$ by $n$ matrix $\mathbf{A}$. Then

$$\text{Tr } \mathbf{A} \equiv \sum_i A_{ii}$$

is called the Trace of $\mathbf{A}$. Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of $\mathbf{A}$, with corresponding eigenvectors $\mathbf{e}_1, \ldots, \mathbf{e}_n$. Then Tr $\mathbf{A} = \sum_i \lambda_i$. We will show this below in two ways.

First note what this means for dynamical systems: if the matrix that describes the linearization about a given fixed point of the dynamics has Trace equal to zero, then either (1) all its eigenvalues have zero real part; or (2) some have a negative real part and some have a positive real part. In the second (more usual) case, therefore, the fixed point is a saddle and is unstable.

## Method 1 (easy but not beautiful)

$$\text{Tr } (\mathbf{AB}) = \sum_i \left( \sum_j A_{ij} B_{ji} \right)$$

simply from the definition of matrix multiplication. The order in which we do the sums doesn't matter, however, so we can quickly see that

$$\text{Tr } (\mathbf{AB}) = \sum_j \sum_i A_{ij} B_{ji} = \sum_j \sum_i B_{ji} A_{ij} = \text{Tr } (\mathbf{BA})$$

that is, Trace is commutative.

Now recall (if necessary from the basic math class) that $\mathbf{A}$ can be written as

$$\mathbf{A} = \mathbf{E} \mathbf{\Lambda} \mathbf{E}^{-1}$$

where the columns of $\mathbf{E}$ are the eigenvectors of $\mathbf{A}$ and $\mathbf{\Lambda}$ is a diagonal matrix with the eigenvalues of $\mathbf{A}$ as its diagonal elements. (We have assumed $\mathbf{E}$ is invertible.) Then

$$\text{Tr } \mathbf{A} = \text{Tr } (\mathbf{E\Lambda E}^{-1}) = \text{Tr } ((\mathbf{E\Lambda})\mathbf{E}^{-1}) = \text{Tr } (\mathbf{E}^{-1}(\mathbf{E\Lambda})) = \text{Tr } ((\mathbf{E}^{-1}\mathbf{E})\mathbf{\Lambda}) = \text{Tr } \mathbf{\Lambda}$$

This last is just the sum of the eigenvalues.

## Method 2 (much more interesting concepts here)

Recall that for any square matrix the eigenvalues are found by obtaining the solutions to the characteristic polynomial:

$$\det(\mathbf{A} - \lambda \mathbf{I}) = 0$$

Suppose, for example, that $\mathbf{A}$ is a 2-by-2 matrix. Then we end up with a quadratic, something like

$$\lambda^2 + \alpha\lambda + \beta = 0$$

Now we make an interesting statement: *the matrix* $\mathbf{A}$ *satisfies the same characteristic polynomial as its eigenvalues.* In the example above, this means that

$$\mathbf{A}^2 + \alpha\mathbf{A} + \beta = 0 \tag{4}$$

To see this, suppose the eigenvectors of $\mathbf{A}$ span the entire space $\Re^n$. (That is, there are $n$ of them and they are linearly independent– equivalent to the condition that $\mathbf{E}$ be invertible in method 1 above.) Then *any* vector $\mathbf{v}$ can be represented as a sum of the eigenvectors times some linear coefficients: $\mathbf{v} = \sum_i c_i \mathbf{e}_i$.

Multiply the left hand side of equation (4) above on the right by $\mathbf{v}$ to get

$$\begin{aligned}
(\mathbf{A}^2 + \alpha\mathbf{A} + \beta)\left(\sum_i c_i \mathbf{e}_i\right) &= \sum_i c_i \left(\mathbf{A}^2 + \alpha\mathbf{A} + \beta\right)\mathbf{e}_i \\
&= \sum_i c_i \mathbf{e}_i \left(\lambda_i^2 + \alpha\lambda_i + \beta\right) \\
&= 0
\end{aligned} \tag{5}$$

since $\lambda_i^2 + \alpha\lambda_i + \beta = 0$ for every $\lambda_i$. (Remember, the $\mathbf{e}_i$ are eigenvectors.) Hence the left hand side of equation 4, multiplied by *any* vector, always gives zero. This necessarily means that (4) is true.

Ok, now we know that (4) is true. What does this tell us about $\mathbf{A}$? Well, it describes $\mathbf{A}$ in terms of what it *does*; that is, in terms of how it operates. $\mathbf{A}$ is a linear operator (it operates on vectors, transforming them to new vectors)— and when you apply it twice, add $\alpha$ times applying it once, and add $\beta$ times the vector you started from, you get zero.

This tells us about $\mathbf{A}$ *without* making any reference to the matrix representation of $\mathbf{A}$, or to the coordinates $\mathbf{A}$ might be described in. So if we were to rotate, or change coordinates, so that the matrix representation of $\mathbf{A}$ were to change, this wouldn't change (4). *The characteristic polynomial is invariant to coordinate transformations of the form* $\mathbf{TAT}^{-1}$. That is, the coefficients in the polynomial will not change.

Since $\mathbf{TAT}^{-1}$ is precisely the form of the transformation used to diagonalize $\mathbf{A}$, we know that the characteristic polynomial is invariant under diagonalization.

Thinking of matrices as representation-independent operators is a very powerful concept. Let's use it now, and move in for the kill with respect to Trace being sum of eigenvalues.

<u>Define</u> Trace as (minus the second coefficient) of the characteristic polynomial.[12] In the example above, say, this would be $-\alpha$. Trace is then clearly the sum of diagonal terms (because of how $\det(\mathbf{A} - \lambda\mathbf{I})$ is calculated), *whatever representation* $\mathbf{A}$ *is in.* And if we choose the representation that diagonalizes $\mathbf{A}$, this is the sum of the eigenvalues. QED.

Another fun one is that the product of the diagonal entries in a square matrix is always just the product of the eigenvalues.

[12]For $n$ by $n$ matrices I mean by second coefficient the constant that multiplies $\lambda^{n-1}$; and we choose as sign convention that the $\lambda^n$ term in the polynomial be positive.