From Table Arrangements to Powerful Tools

CMEE Maths Week

Michael Mustri

January 19, 2023

Table of Contents

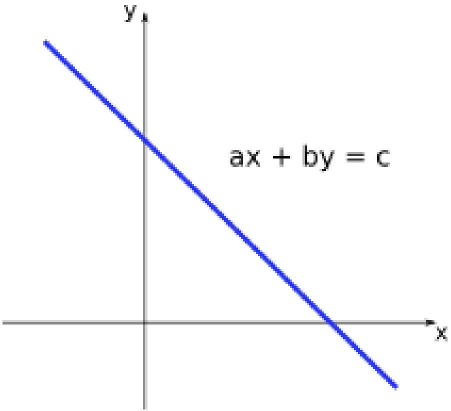
1 Systems of Linear Equations

The Line Revisited

Let's take another look at the equation for a line and arrange it in a slightly different fashion:

$$ax + by = c$$

Note how every point on the line satisfies the equality.



Degree one polynomials represent an example of a **linear transformation** and, along with degree zero polynomials, are the only example of a linear map in Euclidean spaces. Formally, a linear function must satisfy:

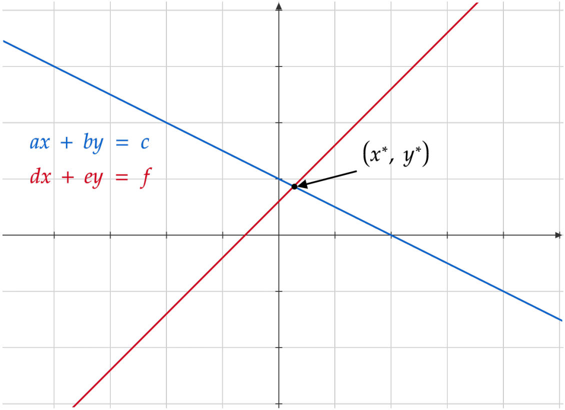
- $f(x+y) = f(x) + f(y)$ - Preserves vector addition.
- $f(ax) = af(x)$ - Preserves scalar multiplication.

These properties make linear functions well behaved and the field of linear algebra offers a number tools for dealing with them.

So let's imagine the simplest system of linear equations:

$$\begin{aligned} ax + by &= c \\ dx + ey &= f \end{aligned}$$

Solving it is equivalent to finding where the lines intersect.



Solving by Substitution

The most straightforward way of finding solutions to systems of linear equations is through **substitution**:

$$y = \frac{c}{b} - \frac{a}{b}x \quad \Rightarrow \quad dx + e\left(\frac{c}{b} - \frac{a}{b}x\right) = f$$

You isolate a variable in one of the equations, and substitute its value in the other equation. Then you simply solve for the remaining variables:

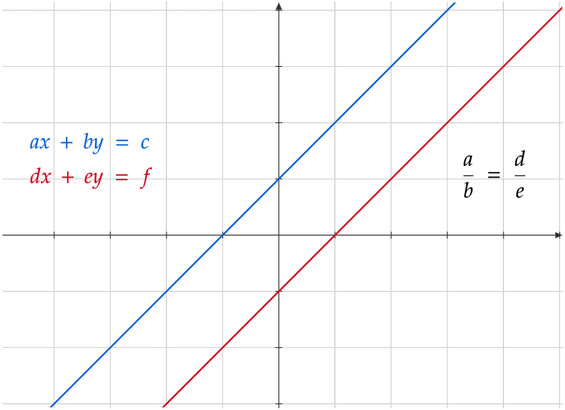
$$x\left(d - e\frac{a}{b}\right) = f - e\frac{c}{b} \quad \Rightarrow \quad x = \frac{fb - ec}{db - ea}$$

Undefined Solutions

However, solutions aren't always possible, and we can see this issue when:

$$\frac{a}{b} = \frac{d}{e}$$

That is, when both equations have the same slope, we call these systems **inconsistent**.

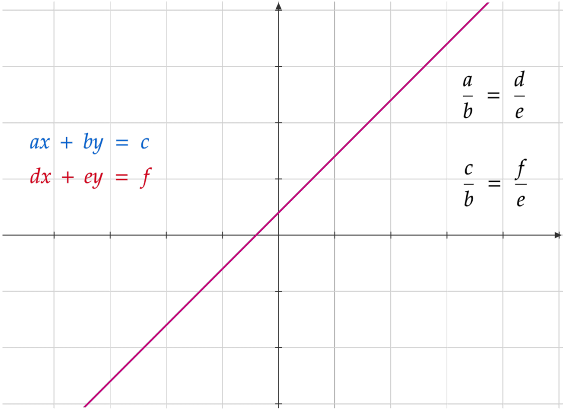


Infinite Solutions

But there is a scenario where parallel lines do intersect:

$$\frac{a}{b} = \frac{d}{e} \quad \text{and} \quad \frac{c}{b} = \frac{f}{e}$$

They are the same line, every point on the line is a solution.

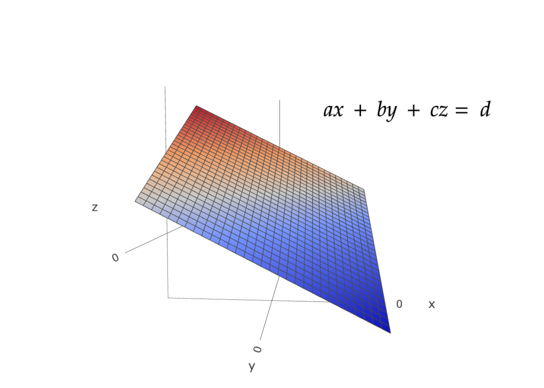


Moving Up a Dimension

Linear equations can be extended into higher dimensions, for example:

$$ax + by + cz = d$$

Is the equation for a plane in a 3-dimensional space.

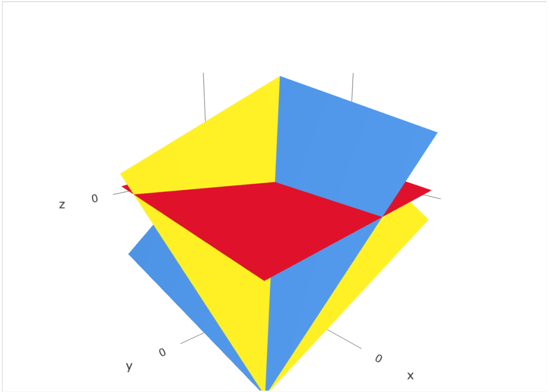


Systems of Three Variables

Much like with lines, solving systems with three variables means finding where the planes intersect.

$$
\begin{aligned}
ax + by + cz &= d \\
ex + fy + gz &= h \\
ix + jy + kz &= l
\end{aligned}
$$

Naturally, you need three equations to find a definite solution.



Row Operations

As systems get larger they become less amenable to solve using substitution. To sidestep this inconvenience we can build simpler equivalent systems using three types of **row operations**.

- Multiplication of a row by a non-zero scalar.
- Adding or subtracting one row from another.
- Rearranging the order of rows.

A row in this scenario refers to a single linear equation. Row operations work because they **preserve the linear relations** between columns.

Row Operations: An Example

We have a system of three linear equations and we order them by row:

$$\begin{aligned} -x + 2y + z &= 3 \ldots R_1^0 \\ x + y + 2z &= 6 \ldots R_2^0 \\ 3x + 2y - z &= 3 \ldots R_3^0 \end{aligned}$$

To get rid of the x variable we perform $R_2^1 = R_2^0 + R_1^0$ and $R_3^1 = R_3^0 + 3R_1^0$.

Now we have the equivalent system:

$$\begin{aligned} -x + 2y + z &= 3 \ldots R_1^1 \\ 0x + 3y + 3z &= 9 \ldots R_2^1 \\ 0x + 8y + 2z &= 12 \ldots R_3^1 \end{aligned}$$

We denote the new rows by the superscript R^1.

Row Operations: An Example

To get rid of the z variable, we perform $R_2^2 = 2R_2^1 - 3R_3^1$.

$$-x + 2y + z = 3 \ldots R_1^2$$
$$0x - 18y + 0z = -18 \ldots R_2^2$$
$$0x + 8y + 2z = 12 \ldots R_3^2$$

From R_2^2 it is obvious that $y = 1$.

Substituting $y = 1$ into R_3^1 we get:

$$8(1) + 2z = 12$$
$$z = \frac{4}{2} = 2$$

Finally, substituting $y = 1$ and $z = 2$ into R_1^0:

$$-x + 2(1) + (2) = 3$$
$$x = 1$$

So $x = 1$, $y = 1$, and $z = 2$.

Zero Rows

Analogous to lines, high dimensional systems of equations can be inconsistent, which can be easily assessed through row operations. Given an N dimensional system of equations, it is inconsistent if there exists a number $C \in \mathbb{R}$ such that:

$$R_i + CR_j = (0)x_1 + (0)x_2 + \cdots + (0)x_N$$

In other words, if a **linear composition** of any two rows results in a **zero row**, the system is inconsistent and has no defined solution.

Matrix Representation

Matrices present a convenient notation for systems of linear equations, for example a system of three equations can be expressed as:

$$\begin{pmatrix} a & b & c \\ e & f & g \\ i & j & k \end{pmatrix} \begin{pmatrix} x \\ y \\ x \end{pmatrix} = \begin{pmatrix} d \\ h \\ l \end{pmatrix}$$

Carrying out **matrix multiplication** on the left hand side of the expression leads back to the more familiar representation.

In general a system of n variables $\{x_1, x_2, \ldots, x_n\}$ and m equations

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m \end{cases}$$

Can be represented in the form $A\mathbf{x} = \mathbf{b}$.

Where A is an $(m \times n)$ matrix, **x** and **b** are **vectors** of size $(n \times 1)$ and $(m \times 1)$ respectively.

$$A\mathbf{x} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{1m} & a_{2m} & \cdots & a_{nm} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} = \mathbf{b}$$

Note how multiplying a an $(m \times n)$ matrix by a $(n \times 1)$ vector gives us an $(m \times 1)$ vector.

Table of Contents

Matrix Operations

Let's go over some basic matrix operations. Given two matrices A and B, both with dimensions $(m \times n)$:

- $A = B$ if and only if $a_{ij} = b_{ij}$ for every $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, n\}$.
- $C = A + B$ is equivalent to $c_{ij} = a_{ij} + b_{ij}$, sums are performed element-wise.
- Given $k \in \mathbb{R}$, kA is an $(m \times n)$ matrix with elements ka_{ij}.

Properties of Matrix Addition

Matrix addition satisfies a few important properties:

- Matrix addition is commutative - $A + B = B + A$.
- Matrix addition is associative - $(A + B) + C = A + (B + C)$.
- The existence of the additive identity (null matrix) - $A + 0 = A$

The Transpose Matrix

Given an $m \times n$ matrix, A, with elements a_{ij}, we define the transpose of A (often denoted by A' or A^T) as the $(n \times m)$ matrix with elements $a'_{ij} = a_{ji}$.

$$A' = \begin{pmatrix} a'_{11} & a'_{12} & \cdots & a'_{1m} \\ a'_{21} & a'_{22} & \cdots & a'_{2m} \\ \vdots & & \ddots & \vdots \\ a'_{1n} & a'_{2n} & \cdots & a'_{mn} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix}$$

Note that the diagonal elements remain unchanged.

Consider two matrices A and B with sizes $(m \times l)$ and $(l \times n)$ respectively. We define the product of matrices A and B, $C = AB$, as the $(m \times n)$ matrix C such that:

$$c_{ij} = \sum_{k=1}^{l} a_{ik} b_{kj}$$

That is, the elements c_{ij} are composed by multiplying every element of the i^{th} row of A with the corresponding element of the j^{th} column of B. Note that for the product to be defined A must have the same number of rows as B has columns.

Properties of Matrix Multiplication

It can be shown that matrix multiplication satisfies the following:

- Is not commutative - $AB \neq BA$ (unless A and B are square, $n \times n$, matrices).
- Is associative - $(AB)C = A(BC)$.
- Is distributive on the left - $A(B + C) = AB + AC$.
- Is distributive on the right - $(B + C)A = BA + CA$.
- Existence of the multiplicative identity - $AI = A$ and $IA = A$.
- Multiplicative Property of Zero - $A0 = 0$ and $0A = 0$.

The Inverse Matrix

Focusing on the multiplicative identity of matrices we can define a multiplicative inverse: given an $(n \times n)$ matrix A, the **inverse of** A, denoted by A^{-1}, is the $(n \times n)$ matrix such that:

$$AA^{-1} = A^{-1}A = I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

That is, the product of a A with its inverse A^{-1}, is equal to the $(n \times n)$ identity matrix I_n. This is only holds for square matrices, however, it can be generalized.

Right and Left Inverse

Given an $(m \times n)$ matrix A, we define the **right inverse** of A as the $(n \times m)$ matrix B, such that:

$$AB = I_m$$

Similarly, the **left inverse** of A is the $(n \times m)$ matrix C, which satisfies:

$$CA = I_n$$

Finding inverse matrices is generally not trivial, and sometimes outright impossible. Fortunately, there are tools that can help us identify when a matrix is not invertible.

The **determinant**, denoted by $det(A)$, tells us how the **linear transformation** produced by its corresponding matrix changes distances and orientations in space. To illustrate this, let's look at an example.

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix} = \begin{pmatrix} x' \\ y' \end{pmatrix}$$

So here we have a 2×2 matrix multiplying a vector (x, y) and transforming it into another vector (x', y'). To visualize what this does to vectors, let's have a look at how it changes a unit square.

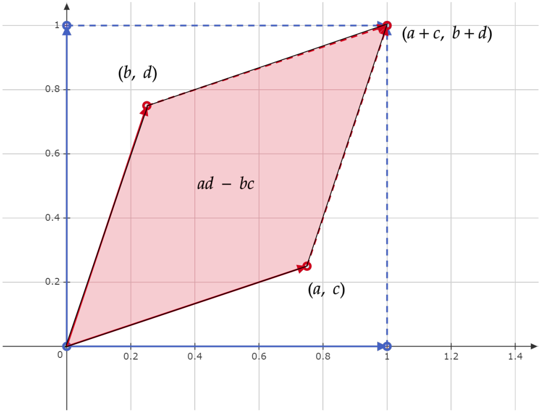
A unit square can be defined by four points on the plane: $(0,0)$, $(1,0)$, $(0,1)$, and $(1,1)$. If we multiply each of these points (as vectors) by our matrix, we get a new set of coordinates.

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} a \\ c \end{pmatrix} \qquad \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} b \\ d \end{pmatrix}$$

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} a+b \\ c+d \end{pmatrix} \qquad \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Plotting our new points on the plane we are able to see how the square has been transformed.

It can be shown that the **signed area** of the resulting parallelogram is $ad - bc$. Coincidentally, this is also the determinant of our transformation matrix.



Inversibility and the Determinant

From the previous example it should be clear that $det(A)$ functions as a **scaling factor** for the transformations produced by A. As such we may ask what it means for a determinant to be zero?

- Geometrically, the transformation collapses areas, volumes and hypervolumes to zero.
- Information is lost in the transformation.
- The components of the matrix are linearly dependent (essentially, they contain the same information).

A consequence of this is that matrices with $det(A) = 0$ are not invertible.

Earlier we saw that a system of equations can be represented in matrix notation:

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

Turns out, solving this system is equivalent to finding A^{-1}.

$$A^{-1}A\mathbf{x} = I_n\mathbf{x} = \mathbf{x} = A^{-1}\mathbf{b}$$

A natural conclusion from this is that, for the system to have a well defined solution, A needs to be a square, invertible matrix.

Table of Contents

Matrices and Structured Models

In a previous lecture we learned how to construct discrete time population models using sequences. With matrices we can take these models a step further to represent the interactions between various populations and subpopulations.

$$\begin{pmatrix} x_{t+1}^1 \\ x_{t+1}^2 \\ \vdots \\ x_{t+1}^m \end{pmatrix} = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix} \begin{pmatrix} x_t^1 \\ x_t^2 \\ \vdots \\ x_t^m \end{pmatrix}$$

Here we take the x_t^i to be the state of each subpopulation at a given time t, and the matrix A, to represent a particular transformation of the population for the next timestep. This matrix is often called the **projection matrix** or **Leslie matrix**.

Example: Life Stages of a Plant

Let's consider a plant whose individuals can be separated into three different life history stages: seedling, small plant, and large plant - (x^{sd}, x^{sp}, x^{lp}). We propose the following model:

$$\begin{pmatrix} x^{sd} \\ x^{sp} \\ x^{lp} \end{pmatrix}_{t+1} = \begin{pmatrix} 0 & f_{sp} & f_{lp} \\ g_{sd\to sp} & u_{sp} & r_{lp\to sp} \\ g_{sd\to lp} & g_{sp\to lp} & u_{lp} \end{pmatrix} \begin{pmatrix} x^{sd} \\ x^{sp} \\ x^{lp} \end{pmatrix}_{t}$$

Now, let's go through each row of the projection matrix and explain what it means.

Example: Seedling Recruitment

The first row of our projection matrix represents how the seedling population will change in the next generation.

$$\begin{pmatrix} 0 & f_{sp} & f_{lp} \end{pmatrix}$$

The zero in the first column signifies that seedlings will neither set seed, nor will they remain in the seedling stage for a whole generation. On the other hand, f_{sp} and f_{lp} are the **stage specific fecundity rates** for small plants and large plants respectively.

Example: Small Plant Development

The second row of A encodes the turnover and development of small plants.

$$\begin{pmatrix} g_{sd\to sp} & u_{sp} & r_{lp\to sp} \end{pmatrix}$$

Here we have three different things going on:

- $g_{sd\to sp}$ is the probability that a seedling will develop into a small plant.
- u_{sp} is the probability that a small plant will fail to develop into a large plant.
- $r_{lp\to sp}$ is the likelihood that a large plant will be reduced to a small plant.

Example: Large Plant Dynamics

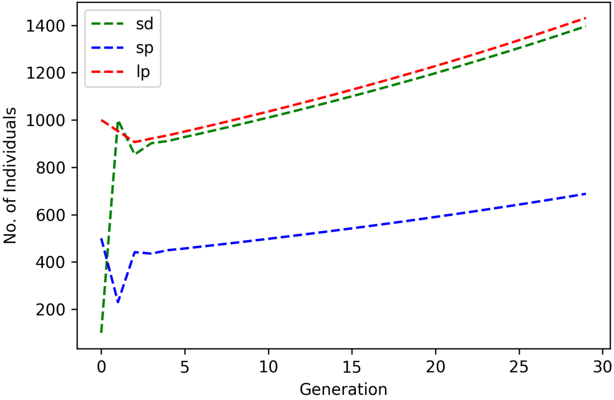
The last row of A expresses how smaller plants turn into large plants and the odds that a large plant will either be reduced to a small plant or survive.

$$\begin{pmatrix} g_{sd\to lp} & g_{sp\to lp} & u_{lp} \end{pmatrix}$$

- $g_{sd\to lp}$ - probability that a seedling will develop into a large plant.
- $g_{sp\to lp}$ - probability that a small plant will develop into a large plant.
- u_{lp} - probability that a large plant will continue to be a large plant.

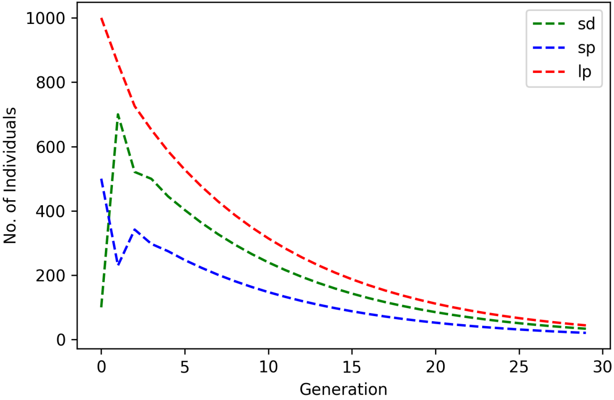
Example: Population Dynamics - Growth

Now let’s have a look at what this model looks like when the population experiences sustained growth.



Example: Population Dynamics - Extinction

But under certain circumstances, the population dwindles towards extinction



Predicting Trends

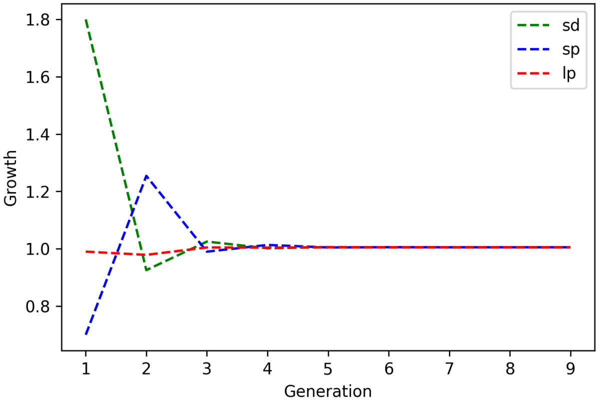
The fact that our projection matrix A is **linear** allows us to make some powerful inferences:

- Regardless of the initial population distributions, every subpopulation will eventually increase by the same **growth rate**.
- As the population continues to evolve each subpopulation will achieve a constant proportion of the total population called the **stable age distribution**.

This independence from initial starting conditions is called **ergodicity**, and the resulting fixed values can be calculated directly from the projection matrix.

The Global Growth Rate

Independent of the initial population distribution, every subpopulation will achieve the same growth rate:



Growth Rate and the Dominant Eigenvalue

Saying that every population will eventually grow at the same rate is equivalent to saying that, for a large enough t:

$$\mathbf{x}_{t+1} = \lambda \mathbf{x}_t$$

Here λ represents the population's long term growth rate and it just so happens that λ is the projection matrix's **dominant eigenvalue**.

The Stable Age Distribution

Independent of the initial population distribution, every subpopulation will achieve a constant proportion of the total population.

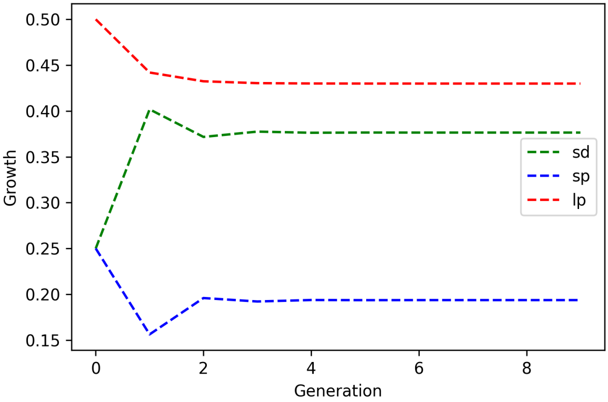


Table of Contents

4 Linear Transformations

Let's go back a little and elaborate upon the concept of **linear transformations**. Earlier, we saw that a function $f(x)$ is linear if and only if:

- $f(x+y) = f(x) + f(y)$
- $f(ax) = af(x)$

However, now that we know what vectors and matrices are we can make a more general and precise statement.

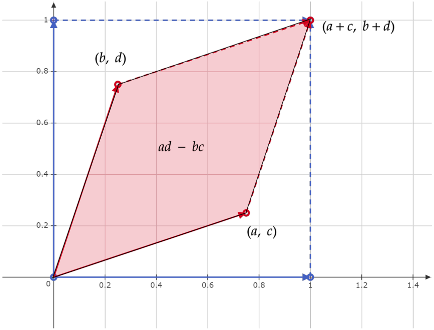
We say a linear transformation is a function $A : V \to W$, that maps one **vector space** (V) to another (W) while satisfying the following:

$$A(a\mathbf{v} + b\mathbf{w}) = aA(\mathbf{v}) + bA(\mathbf{w})$$

Where **v** and **w** are vectors in V and $a, b \in \mathbb{R}$. In short, $A(\mathbf{v})$ preserves both vector addition as well as scalar multiplication. Additionally, let us note that all matrices represent some form of linear transformation.

A Word on Vector Spaces

Formally, vector spaces are sets whose elements satisfy a number of conditions (vector addition, distributive, scalar multiplication, etc.). For now we'll stick to Euclidean vectors, which are the traditional components coordinate spaces, but keep in mind that this is only a small subset of possible vector spaces.



Other Types of Vector Spaces

A few examples of vector spaces:

- Matrix vector spaces
- Polynomial vector spaces
- Function spaces
- ...

Table of Contents

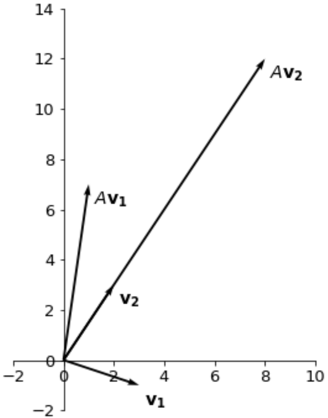
5 Eigenvalues and Eigenvectors

Eigenvalues and Eigenvectors

For a linear transformation given by the $n \times n$ matrix A, there exists a vector **v** such that:

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

Where $\lambda \in \mathbb{C}$ is some constant. We call λ an **eigenvalue** of A and **v** is its corresponding **eigenvector**.



The Characteristic Polynomial

We can manipulate some of the terms in our last equation to reach:

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

Where I is the identity matrix. If this is true, then the determinant of $(A - \lambda I)$ is necessarily zero.

$$det(A - \lambda I) = \begin{vmatrix} a_{11} - \lambda & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} - \lambda & \cdots & a_{n2} \\ \vdots & & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{vmatrix} = 0$$

Solving for λ, we are left with an n^{th} degree polynomial, this is known as the **characteristic polynomial** of A.

Solving for λ

All possible solutions to the characteristic polynomial are **eigenvalues** of A, and let us further note that eigenvalues can be complex numbers. Let's look at a simple example: the 2×2 matrix:

$$\begin{pmatrix} a-\lambda & b \\ c & d-\lambda \end{pmatrix} \mathbf{v} = 0$$

Hence:

$$\begin{vmatrix} a-\lambda & b \\ c & d-\lambda \end{vmatrix} = (a-\lambda)(d-\lambda) - cb = \lambda^2 - (a+d)\lambda + (ad-cb) = 0$$

Solving for λ

Solving this quadratic equation renders:

$$\lambda_{1,2} = \frac{(a+d) \pm \sqrt{(a+d)^2 - 4(ad-cb)}}{2}$$

Making a few substitutions for $\tau = (a+d)$ and noting that $det(A) = (ad-cb) = \Delta$, we have:

$$\lambda_{1,2} = \frac{\tau \pm \sqrt{\tau^2 - 4\Delta}}{2}$$

We call τ the **trace** of A, the sum of the diagonal elements. Note that for $\tau^2 > 4\Delta$, both of our eigenvalues are real numbers, for $\tau^2 < 4\Delta$ we have complex solutions, and when $\tau^2 = 4\Delta$ we only have one solution, $\frac{\tau}{2}$.

Finding the Eigenvectors

We've now know how to find λ, but we still need to find $\mathbf{v}$, the **eigenvectors**. Assuming λ_i is an eigenvalue of A, we can make the following statement:

$$A\mathbf{v} = \lambda_i \mathbf{v} \quad \Rightarrow \quad (A - \lambda_i I)\mathbf{v} = 0$$

And this is equivalent to finding a vector $\mathbf{v}$ such that:

$$\begin{pmatrix} a_{11} - \lambda & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} - \lambda & \cdots & a_{n2} \\ \vdots & & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Finding the Eigenvectors

The previous relationship leads to a system of equations which is undefined (because $det(A - \lambda_i I) = 0$), however, it is undefined in the sense that it has an infinite number of solutions. Let's retake our previous example using a 2×2 matrix:

$$\begin{pmatrix} a - \lambda_i & b \\ c & d - \lambda_i \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Rendering the following system of equations:

$$(a - \lambda_i)x + by = 0$$
$$cx + (d - \lambda_i)y = 0$$

Since this system of equations has an infinite number of solutions, we can safely assume that for one of them $y = 1$.

Finding the Eigenvectors

By arbitrarily assigning $y = 1$ we've reduced the problem to finding x, which follows from:

$$(a - \lambda_i)x + b(1) = 0 \quad \Rightarrow \quad x = \frac{b}{\lambda_i - a}$$

Therefore, the eigenvector corresponding to λ_i is:

$$\mathbf{v}_i = \begin{pmatrix} \frac{b}{\lambda_i - a} \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{\lambda_i - d}{c} \\ 1 \end{pmatrix}$$

Note that it doesn't matter which equation you choose to solve, they will all lead to the same result. Moreover, **any scalar multiple of an eigenvector is still an eigenvector**.

Eigenvalues and Eigenvectors: A Generalization

The previous example using a 2×2 matrix is the really the best case scenario, finding eigenvectors and eigenvalues becomes increasingly more difficult for higher dimensions. In essence, however, the idea is no more complicated:

- Eigenvalues will always be roots of the characteristic polynomial of A.
- Eigenvectors will always correspond to directions which do not suffer rotation under transformations by A.

Table of Contents

6 Vector Decomposition

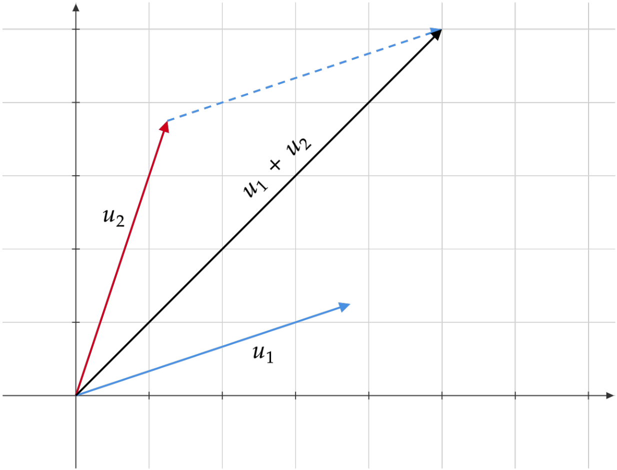
Let's return to our 2×2 matrix A. Assuming A has eigenvalues λ_1 and λ_2 with corresponding eigenvectors $\mathbf{u}_1$ and $\mathbf{u}_2$, if $\lambda_1 \neq \lambda_2$, we say that $\mathbf{u}_1$ and $\mathbf{u}_2$ are **linearly independent**, meaning:

$$\mathbf{u}_1 + c\mathbf{u}_2 \neq 0, \quad \forall c \in \mathbb{R}$$

In short, the directions corresponding to $\mathbf{u}_1$ and $\mathbf{u}_2$ are not on the same line. When this occurs we say that $\mathbf{u}_1$ and $\mathbf{u}_2$ are a **basis** of this vector space and any vector $\mathbf{v}$ in this space can be written as a **linear composition** of $\mathbf{u}_1$ and $\mathbf{u}_2$.

$$\mathbf{v} = a_1\mathbf{u}_1 + a_2\mathbf{u}_2$$

Vector Decomposition



Transforming the Eigenspace

This property of linearly independent eigenvectors has an interesting consequence. Given a vector $\mathbf{v}$ we can see that transforming it by A renders:

$$A\mathbf{v} = A(a_1\mathbf{u}_1 + a_2\mathbf{u}_2) = \lambda_1 a_1 \mathbf{u}_1 + \lambda_2 a_2 \mathbf{u}_2$$

Subsequent transformations will produce an analogous result:

$$A(A\mathbf{v}) = A(\lambda_1 a_1\mathbf{u}_1 + \lambda_2 a_2\mathbf{u}_2) = \lambda_1^2 a_1 \mathbf{u}_1 + \lambda_2^2 a_2 \mathbf{u}_2$$

And more generally:

$$A^n\mathbf{v} = \lambda_1^n a_1 \mathbf{u}_1 + \lambda_2^n a_2 \mathbf{u}_2$$

Transforming the Eigenspace

What the expression - $A^n\mathbf{v} = \lambda_1^n a_1 \mathbf{u}_1 + \lambda_2^n a_2 \mathbf{u}_2$ - is telling us, is that any linear transformation acting upon a vector **v**, will stretch (or shrink) that vector by a factor λ_i, along the corresponding **eigendirections**.

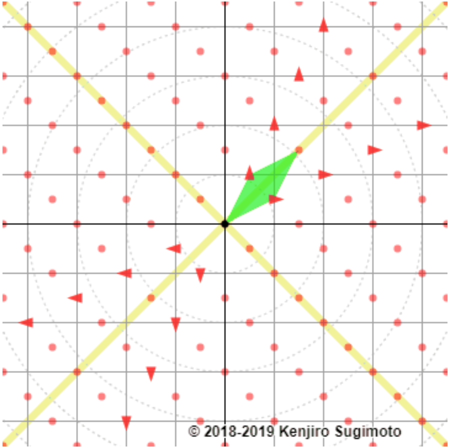


Table of Contents

Back when we saw structured models I mentioned in passing that, given a linear projection matrix, a population's growth rate would eventually converge on the **dominant eigenvalue** for every subpopultion.

$$\mathbf{x}_{t+1} = \lambda^* \mathbf{x}_t$$

The **dominant eigenvalue** of the projection matrix is the eigenvalue with the largest real part, if all eigenvalues are real, it is simply the largest eigenvalue. This should be evident from our previous section, but let's motivate it further.

The Fast Eigendirection

We've seen that a linear transformation stretches and shrinks vectors along its eigendirections, and this holds for systems of any size:

$$A^n\mathbf{v} = \lambda_1^n a_1 \mathbf{u}_1 + \lambda_2^n a_2 \mathbf{u}_2 + \cdots + \lambda_m^n a_m \mathbf{u}_m$$

Now, let's assume that one of the eigenvalues of A is larger than all others and denote it by λ_i. Defining a new transformation $A^* = \frac{1}{\lambda_i}A$, it follows from the definition of linear transformations that A^*'s eigenvalues are scaled by a factor of $\frac{1}{\lambda_i}$ and that the eigenvectors remain unchanged:

$$(A^*)^n\mathbf{v} = \left(\frac{\lambda_1}{\lambda_i}\right)^n a_1\mathbf{u}_1 + \left(\frac{\lambda_2}{\lambda_i}\right)^n a_2\mathbf{u}_2 + \cdots + \left(\frac{\lambda_i}{\lambda_i}\right)^n a_i\mathbf{u}_i + \cdots + \left(\frac{\lambda_m}{\lambda_i}\right)^n a_m\mathbf{u}_m$$

The Fast Eigendirection

$$(A^*)^n \mathbf{v} = \left(\frac{\lambda_1}{\lambda_i}\right)^n a_1\mathbf{u}_1 + \left(\frac{\lambda_2}{\lambda_i}\right)^n a_2\mathbf{u}_2 + \cdots + \left(\frac{\lambda_i}{\lambda_i}\right)^n a_i\mathbf{u}_i + \cdots + \left(\frac{\lambda_m}{\lambda_i}\right)^n a_m\mathbf{u}_m$$

From the above equation we note a few things: first $\left(\frac{\lambda_i}{\lambda_i}\right) = 1$, and since λ_i is the largest eigenvalue, $\left(\frac{\lambda_j}{\lambda_i}\right) < 1$ for any $j \neq i$. Thus, for large enough n, $\left(\frac{\lambda_j}{\lambda_i}\right)^n \to 0$ and:

$$(A^*)^n \mathbf{v} \to a_i\mathbf{u}_i \quad \Rightarrow \quad A^n\mathbf{v} \to \lambda_i^n a_i\mathbf{u}_i$$

In other words, $A^n\mathbf{v}$ approaches the eigendirection corresponding to the dominant eigenvalue, this is known as the **fast eigendirection**

Implications for Structured Models

To conclude, a few things worth noting about dominant eigenvalues:

- If $\lambda^* < 1$ the entire population will tend towards extinction.
- Every subpopulation will approach the same growth rate determined by λ^*.
- The **stable age distribution** is defined by the **dominant eigenvector**.

Structured Models and the Dominant Eigenvalue

As mentioned before, the dominant eigenvalue determines the long-term growth rate of structured models defined by linear transformations.

$$X_t = A^t \mathbf{v}_0 \quad \Rightarrow \quad \lim_{t\to\infty} X_t \to \lambda_i^t a_i \mathbf{u}_i$$

Notice how eventually the entire population will assume exponential growth for $\lambda_i > 1$ or decay for $\lambda_i < 1$.

The Dominant Eigenvector and Stable Age Distribution

The stable age distribution can be inferred by observing that as:

$$\lim_{t \to \infty} X_t \to \lambda_i^t a_i \mathbf{u}_i$$

each subpopulation assumes a proportion of the whole population determined by the eigenvector $\mathbf{u}_i = (u_i^{(1)}, u_i^{(2)}, \ldots, u_i^{(n)})$. So the relative of the abundance of subpopultaion $X_n^{(j)}$ converges to a limit defined by the dominant eigenvector:

$$\lim_{t \to \infty} \frac{X_t^{(j)}}{\sum_{k=1}^{n} X_t^{(k)}} \to \frac{u_i^{(j)}}{\sum_{k=1}^{n} u_i^{(k)}}$$