Notes for Lecture 1

Mon, 1/13/2025

A crash course in linear algebra

Example 1. A typical 2×3 matrix is $\begin{bmatrix} 1 & 2 & 3\\ 4 & 5 & 6 \end{bmatrix}$.

It is composed of column vectors like $\begin{bmatrix} 2\\ 5 \end{bmatrix}$ and row vectors like $\begin{bmatrix} 1 & 2 & 3 \end{bmatrix}$.

Matrices (and vectors) of the same dimensions can be added and multiplied by a scalar:

For instance, $\begin{bmatrix} 1 & 2 & 3\\ 4 & 5 & 6 \end{bmatrix}+\begin{bmatrix} 1 & 0 & 2\\ 2 & 3 & -1 \end{bmatrix}=\begin{bmatrix} 2 & 2 & 5\\ 6 & 8 & 5 \end{bmatrix}$ or $3\cdot\begin{bmatrix} 1 & 2 & 3\\ 4 & 5 & 6 \end{bmatrix}=\begin{bmatrix} 3 & 6 & 9\\ 12 & 15 & 18 \end{bmatrix}$.

Remark. More generally, a **vector space** is an abstraction of a collection of objects that can be added and scaled: numbers, lists of numbers (like the above row and column vectors), arrays of numbers (like the above matrices), arrows, functions, polynomials, differential operators, solutions to homogeneous linear differential equations, ...

Example 2. The **transpose** A^T of A is obtained by interchanging roles of rows and columns.

For instance. $\begin{bmatrix} 1 & 2 & 3\\ 4 & 5 & 6 \end{bmatrix}^T=\begin{bmatrix} 1 & 4\\ 2 & 5\\ 3 & 6 \end{bmatrix}$

Example 3. Matrices of appropriate dimensions can also be **multiplied**.

This is based on the multiplication $\begin{bmatrix} a & b & c \end{bmatrix}\begin{bmatrix} x\\ y\\ z \end{bmatrix}=ax+by+cz$ of row and column vectors.

For instance. $\begin{bmatrix} 1 & -1 & 1\\ 2 & 1 & 3 \end{bmatrix}\begin{bmatrix} 1 & 0\\ -1 & 1\\ 2 & -2 \end{bmatrix}=\begin{bmatrix} 4 & -3\\ 7 & -5 \end{bmatrix}$

In general, we can multiply a $m\times n$ matrix A with a $n\times r$ matrix B to get a $m\times r$ matrix AB.

Its entry in row i and column j is defined to be $(AB)_{ij}=(\text{row } i \text{ of } A)\begin{bmatrix} \text{column}\\ j\\ \text{of } B \end{bmatrix}$.

Comment. One way to think about the multiplication $A\boldsymbol{x}$ is that the resulting vector is a linear combination of the columns of A with coefficients from $\boldsymbol{x}$. Similarly, we can think of $\boldsymbol{x}^TA$ as a combination of the rows of A.

Some nice properties of matrix multiplication are:

- There is an $n\times n$ identity matrix I (all entries are zero except the diagonal ones which are 1). It satisfies $AI=A$ and $IA=A$.
- The associative law $A(BC)=(AB)C$ holds. Hence, we can write ABC without ambiguity.
- The distributive laws including $A(B+C)=AB+AC$ hold.

Example 4. $\begin{bmatrix} 2 & 0\\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 2\\ 3 & 4 \end{bmatrix}\neq\begin{bmatrix} 1 & 2\\ 3 & 4 \end{bmatrix}\begin{bmatrix} 2 & 0\\ 0 & 1 \end{bmatrix}$, so we have no commutative law.

Example 5. $\begin{bmatrix} 3 & 1\\ 2 & 1 \end{bmatrix}\begin{bmatrix} 1 & -1\\ -2 & 3 \end{bmatrix}=\begin{bmatrix} 1 & 0\\ 0 & 1 \end{bmatrix}$

On the RHS we have the **identity matrix**, usually denoted I or I_2 (since it's the 2×2 identity matrix here).

Hence, the two matrices on the left are inverses of each other: $\begin{bmatrix} 3 & 1\\ 2 & 1 \end{bmatrix}^{-1}=\begin{bmatrix} 1 & -1\\ -2 & 3 \end{bmatrix}$, $\begin{bmatrix} 1 & -1\\ -2 & 3 \end{bmatrix}^{-1}=\begin{bmatrix} 3 & 1\\ 2 & 1 \end{bmatrix}$.

Armin Straub
straub@southalabama.edu

The **inverse** A^{-1} of a matrix A is characterized by $A^{-1}A = I$ and $AA^{-1} = I$.

Example 6. The following formula immediately gives us the inverse of a 2×2 matrix (if it exists). It is worth remembering!

$$\boxed{\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \qquad \text{provided that } ad - bc \neq 0}$$

Let's check that! $\frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \frac{1}{ad - bc} \begin{bmatrix} ad - bc & 0 \\ 0 & -cb + ad \end{bmatrix} = I_2$

In particular, a 2×2 matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is invertible $\Longleftrightarrow$ $ad - bc \neq 0$.

Recall that this is the **determinant**: $\det\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = ad - bc$.

$$\boxed{\det(A) = 0 \quad \Longleftrightarrow \quad A \text{ is not invertible}}$$

Example 7. The system $\begin{array}{rl} 7x_1 - 2x_2 &= 3 \\ 2x_1 + x_2 &= 5 \end{array}$ is equivalent to $\begin{bmatrix} 7 & -2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 5 \end{bmatrix}$. Solve it.

Solution. Multiplying (from the left!) by $\begin{bmatrix} 7 & -2 \\ 2 & 1 \end{bmatrix}^{-1} = \frac{1}{11} \begin{bmatrix} 1 & 2 \\ -2 & 7 \end{bmatrix}$ produces $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \frac{1}{11} \begin{bmatrix} 1 & 2 \\ -2 & 7 \end{bmatrix} \begin{bmatrix} 3 \\ 5 \end{bmatrix} = \frac{1}{11} \begin{bmatrix} 13 \\ 29 \end{bmatrix}$, which gives the solution of the original equations.

Example 8. (homework) Solve the system $\begin{array}{rl} x_1 + 2x_2 &= 1 \\ 3x_1 + 4x_2 &= -1 \end{array}$ (using a matrix inverse).

Solution. The equations are equivalent to $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

Multiplying by $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}^{-1} = -\frac{1}{2} \begin{bmatrix} 4 & -2 \\ -3 & 1 \end{bmatrix}$ produces $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = -\frac{1}{2} \begin{bmatrix} 4 & -2 \\ -3 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = -\frac{1}{2} \begin{bmatrix} 6 \\ -4 \end{bmatrix} = \begin{bmatrix} -3 \\ 2 \end{bmatrix}$.

Example 9. (homework) Solve the system $\begin{array}{rl} x_1 + 2x_2 &= 1 \\ 3x_1 + 4x_2 &= 2 \end{array}$ (using a matrix inverse).

Solution. The equations are equivalent to $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

Multiplying by $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}^{-1} = -\frac{1}{2} \begin{bmatrix} 4 & -2 \\ -3 & 1 \end{bmatrix}$ produces $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = -\frac{1}{2} \begin{bmatrix} 4 & -2 \\ -3 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = -\frac{1}{2} \begin{bmatrix} 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1/2 \end{bmatrix}$.

Comment. In hindsight, can you see this solution by staring at the equations?

Comment. Note how we can reuse the matrix inverse from the previous example.

The **determinant** of A, written as $\det(A)$ or $|A|$, is a number with the property that:

$$\begin{aligned} \det(A) \neq 0 \quad &\Longleftrightarrow \quad A \text{ is invertible} \\ &\Longleftrightarrow \quad A\boldsymbol{x} = \boldsymbol{b} \text{ has a (unique) solution } \boldsymbol{x} \text{ for all } \boldsymbol{b} \\ &\Longleftrightarrow \quad A\boldsymbol{x} = 0 \text{ is only solved by } \boldsymbol{x} = 0 \end{aligned}$$

Example 10. $\det\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = ad - bc$, which appeared in the formula for the inverse.

Armin Straub
straub@southalabama.edu

Example 11. (review) $\begin{bmatrix} 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 14 \end{bmatrix}$ whereas $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \end{bmatrix}$.

Review: Examples of differential equations we can solve

Let's start with one of the simplest (and most fundamental) differential equations (DE). It is **first-order** (only a first derivative) and **linear** with constant coefficients.

Example 12. Solve $y' = 3y$.

Solution. $y(x) = Ce^{3x}$

Check. Indeed, if $y(x) = Ce^{3x}$, then $y'(x) = 3Ce^{3x} = 3y(x)$.

Comment. Recall we can always easily check whether a function solves a differential equation. This means that (although you might be unfamiliar with the techniques for solving) you can use computer algebra systems like Sage to solve differential equations without trust issues.

To describe a unique solution, additional constraints need to be imposed.

Example 13. Solve the **initial value problem** (IVP) $y' = 3y$, $y(0) = 5$.

Solution. This has the unique solution $y(x) = 5e^{3x}$.

The following is a **nonlinear** differential equation. In general, such equations are much more complicated than linear ones. We can solve this particular one because it is **separable**.

Example 14. Solve $y' = xy^2$.

Solution. This DE is separable: $\frac{1}{y^2}\mathrm{d}y = x\,\mathrm{d}x$. Integrating, we find $-\frac{1}{y} = \frac{1}{2}x^2 + C$.

Hence, $y = -\frac{1}{\frac{1}{2}x^2 + C} = \frac{2}{D - x^2}$.

[Here, $D = -2C$ but that relationship doesn't matter; it only matters that the solution has a free parameter.]

Note. Note that we did not find the solution $y = 0$ (lost when dividing by y^2). It is called a **singular solution** because it is not part of the **general solution** (the one-parameter family found above). [Although, we can obtain it from the general solution by letting $D \to \infty$.]

Check. Compute y' and verify that the DE is indeed satisfied.

Armin Straub
straub@southalabama.edu

Review: Linear DEs

Linear DEs of order n are those that can be written in the form

$$y^{(n)} + P_{n-1}(x)\, y^{(n-1)} + \ldots + P_1(x) y' + P_0(x) y = f(x).$$

The corresponding **homogeneous linear DE** is the DE

$$y^{(n)} + P_{n-1}(x)\, y^{(n-1)} + \ldots + P_1(x) y' + P_0(x) y = 0,$$

and it plays an important role in solving the original linear DE.

Important. Note that a linear DE is **homogeneous** if and only if the zero function $y(x) = 0$ is a solution.

In terms of $D = \frac{\mathrm{d}}{\mathrm{d}x}$, the original DE becomes: $Ly = f(x)$ where L is the **differential operator**

$$L = D^n + P_{n-1}(x) D^{n-1} + \ldots + P_1(x) D + P_0(x).$$

The corresponding homogeneous linear DE is $Ly = 0$.

Linear DEs have a lot of structure that makes it possible to understand them more deeply. Most notably, their general solution always has the following structure:

(general solution of linear DEs) For a linear DE $Ly = f(x)$ of order n, the general solution always takes the form

$$y(x) = y_p(x) + C_1 y_1(x) + \ldots + C_n\, y_n(x),$$

where y_p is any single solution (called a **particular solution**) and $y_1, y_2, \ldots, y_n$ are solutions to the corresponding **homogeneous** linear DE $Ly = 0$.

Comment. If the linear DE is already homogeneous, then the zero function $y(x) = 0$ is a solution and we can use $y_p = 0$. In that case, the general solution is of the form $y(x) = C_1 y_1 + C_2 y_2 + \cdots + C_n y_n$.

Why? The key to this is that the differential operator L is **linear**, meaning that, for any functions $f_1(x), f_2(x)$ and any constants c_1, c_2, we have

$$L(c_1 f_1(x) + c_2 f_2(x)) = c_1 L(f_1(x)) + c_2 L(f_2(x)).$$

If this is not clear, consider first a case like $L = D^n$ or work through the next example for the order 2 case.

Example 15. (extra) Suppose that $L = D^2 + P(x) D + Q(x)$. Verify that the operator L is linear.

Solution. We need to show that the operator L satisfies

$$L(c_1 f_1(x) + c_2 f_2(x)) = c_1 L(f_1(x)) + c_2 L(f_2(x))$$

for any functions $f_1(x), f_2(x)$ and any constants c_1, c_2. Indeed:

$$\begin{aligned} L(c_1 f_1 + c_2 f_2) &= (c_1 f_1 + c_2 f_2)'' + P(x)(c_1 f_1 + c_2 f_2)' + Q(x)(c_1 f_1 + c_2 f_2) \\ &= c_1\{f_1'' + P(x) f_1' + Q(x) f_1\} + c_2\{f_2'' + P(x) f_2' + Q(x) f_2\} \\ &= c_1 \cdot L f_1 + c_2 \cdot L f_2 \end{aligned}$$

Armin Straub
straub@southalabama.edu

Example 16. Consider the following DEs. If linear, write them in operator form as $Ly = f(x)$.

(a) $y'' = xy$

(b) $x^2y'' + xy' = (x^2 + 4)y + x(x^2 + 3)$

(c) $y'' = y' + 2y + 2(1 - x - x^2)$

(d) $y'' = y' + 2y + 2(1 - x - y^2)$

Solution.

(a) This is a homogeneous linear DE: $\underbrace{(D^2 - x)}_{L}y = \underbrace{0}_{f(x)}$

Note. This is known as the Airy equation, which we will meet again later. The general solution is of the form $C_1y_1(x) + C_2y_2(x)$ for two special solutions y_1, y_2. [In the literature, one usually chooses functions called $\mathrm{Ai}(x)$ and $\mathrm{Bi}(x)$ as y_1 and y_2. See: https://en.wikipedia.org/wiki/Airy_function]

(b) This is an inhomogeneous linear DE: $\underbrace{(x^2D^2 + xD - (x^2 + 4))}_{L}y = \underbrace{x(x^2 + 3)}_{f(x)}$

Note. The corresponding homogeneous DE is an instance of the "modified Bessel equation" $x^2y'' + xy' - (x^2 + \alpha^2)y = 0$, namely the case $\alpha = 2$. Because they are important for applications (but cannot be written in terms of familiar functions), people have introduced names for two special solutions of this differential equation: $I_\alpha(x)$ and $K_\alpha(x)$ (called modified Bessel functions of the first and second kind).

It follows that the general solution of the modified Bessel equation is $C_1I_\alpha(x) + C_2K_\alpha(x)$.

In our case. The general solution of the homogeneous DE (which is the modified Bessel equation with $\alpha = 2$) is $C_1I_2(x) + C_2K_2(x)$. On the other hand, we can (do it!) easily check (this is coming from nowhere at this point!) that $y_p = -x$ is a particular solution to the original inhomogeneous DE.

It follows that the general solution to the original DE is $C_1I_2(x) + C_2K_2(x) - x$.

(c) This is an inhomogeneous linear DE: $\underbrace{(D^2 - D - 2)}_{L}y = \underbrace{2(1 - x - x^2)}_{f(x)}$

Note. We will recall in Example 17 that the corresponding homogeneous DE $(D^2 - D - 2)y = 0$ has general solution $C_1e^{2x} + C_2e^{-x}$. On the other hand, we can check that $y_p = x^2$ is a particular solution of the original inhomogeneous DE. (Do you recall from DE1 how to find this particular solution?)

It follows that the general solution to the original DE is $x^2 + C_1e^{2x} + C_2e^{-x}$.

(d) This is not a linear DE because of the term y^2. It cannot be written in the form $Ly = f(x)$.

Homogeneous linear DEs with constant coefficients

Example 17. Find the general solution to $y'' - y' - 2y = 0$.

Solution. We recall from *Differential Equations I* that e^{rx} solves this DE for the right choice of r.
Plugging e^{rx} into the DE, we get $r^2 e^{rx} - r e^{rx} - 2e^{rx} = 0$.
Equivalently, $r^2 - r - 2 = 0$. This is called the **characteristic equation**. Its solutions are $r = 2, -1$.
This means we found the two solutions $y_1 = e^{2x}$, $y_2 = e^{-x}$.
Since this a homogeneous linear DE, the general solution is $y = C_1 e^{2x} + C_2 e^{-x}$.

Solution. (operators) $y'' - y' - 2y = 0$ is equivalent to $(D^2 - D - 2)y = 0$.
Note that $D^2 - D - 2 = (D-2)(D+1)$ is the **characteristic polynomial**.
It follows that we get solutions to $(D-2)(D+1)y = 0$ from $(D-2)y = 0$ and $(D+1)y = 0$.
$(D-2)y = 0$ is solved by $y_1 = e^{2x}$, and $(D+1)y = 0$ is solved by $y_2 = e^{-x}$; as in the previous solution.

Example 18. Solve $y'' - y' - 2y = 0$ with initial conditions $y(0) = 4$, $y'(0) = 5$.

Solution. From the previous example, we know that $y(x) = C_1 e^{2x} + C_2 e^{-x}$.
To match the initial conditions, we need to solve $C_1 + C_2 = 4$, $2C_1 - C_2 = 5$. We find $C_1 = 3$, $C_2 = 1$.
Hence the solution is $y(x) = 3e^{2x} + e^{-x}$.

Set $D = \frac{\mathrm{d}}{\mathrm{d}x}$. Every **homogeneous linear DE with constant coefficients** can be written as $p(D)y = 0$, where $p(D)$ is a polynomial in D, called the **characteristic polynomial**.

For instance. $y'' - y' - 2y = 0$ is equivalent to $Ly = 0$ with $L = D^2 - D - 2$.

Example 19. Find the general solution of $y''' + 7y'' + 14y' + 8y = 0$.

Solution. This DE is of the form $p(D)\,y = 0$ with characteristic polynomial $p(D) = D^3 + 7D^2 + 14D + 8$.
The characteristic polynomial factors as $p(D) = (D+1)(D+2)(D+4)$. (Don't worry! You won't be asked to factor cubic polynomials by hand.)
Hence, by the same argument as in Example 17, we find the solutions $y_1 = e^{-x}$, $y_2 = e^{-2x}$, $y_3 = e^{-4x}$. That's enough (independent!) solutions for a third-order DE.
The general solution therefore is $y(x) = C_1\, e^{-x} + C_2\, e^{-2x} + C_3\, e^{-4x}$.

This approach applies to any homogeneous linear DE with constant coefficients!

One issue is that roots might be repeated. In that case, we are currently missing solutions. The following result provides the missing solutions.

Theorem 20. Consider the homogeneous linear DE with constant coefficients $p(D)y = 0$.

- If r is a root of the characteristic polynomial and if k is its multiplicity, then k (independent) solutions of the DE are given by $x^j e^{rx}$ for $j = 0, 1, \ldots, k-1$.
- Combining these solutions for all roots, gives the general solution.

This is because the order of the DE equals the degree of $p(D)$, and a polynomial of degree n has (counting with multiplicity) exactly n (possibly **complex**) roots.

In the complex case. If $r = a \pm bi$ are roots of the characteristic polynomial and if k is its multiplicity, then $2k$ (independent) **real solutions** of the DE are given by $x^j e^{ax} \cos(bx)$ and $x^j e^{ax} \sin(bx)$ for $j = 0, 1, \ldots, k-1$.

Proof. Let r be a root of the characteristic polynomial of multiplicity k. Then $p(D) = q(D)\,(D - r)^k$.
We need to find k solutions to the simpler DE $(D - r)^k y = 0$.
It is natural to look for solutions of the form $y = c(x) e^{rx}$.
[We know that $c(x) = 1$ provides a solution. Note that this is the same idea as for variation of constants.]
Note that $(D - r)[c(x)e^{rx}] = (c'(x)e^{rx} + c(x) r e^{rx}) - r c(x)e^{rx} = c'(x)e^{rx}$.
Repeating, we get $(D - r)^2[c(x)e^{rx}] = (D - r)[c'(x)e^{rx}] = c''(x)e^{rx}$ and, eventually, $(D - r)^k[c(x)e^{rx}] = c^{(k)}(x)e^{rx}$.
In particular, $(D - r)^k y = 0$ is solved by $y = c(x)e^{rx}$ if and only if $c^{(k)}(x) = 0$.
The DE $c^{(k)}(x) = 0$ is clearly solved by x^j for $j = 0, 1, \ldots, k - 1$, and it follows that $x^j e^{rx}$ solves the original DE. □

Example 21. Find the general solution of $y''' = 0$.

Solution. We know from Calculus that the general solution is $y(x) = C_1 + C_2\,x + C_3\,x^2$.

Solution. The characteristic polynomial $p(D) = D^3$ has roots $0, 0, 0$. By Theorem 20, we have the solutions $y(x) = x^j\,e^{0x} = x^j$ for $j = 0, 1, 2$, so that the general solution is $y(x) = C_1 + C_2\,x + C_3\,x^2$.

Example 22. Find the general solution of $y''' - y'' - 5y' - 3y = 0$.

Solution. The characteristic polynomial $p(D) = D^3 - D^2 - 5D - 3 = (D - 3)(D + 1)^2$ has roots $3, -1, -1$.
By Theorem 20, the general solution is $y(x) = C_1 e^{3x} + (C_2 + C_3 x)e^{-x}$.

Example 23. Find the general solution of $y'' + y = 0$.

Solution. The characteristic polynomials is $p(D) = D^2 + 1 = 0$ which has no solutions over the reals.
Over the **complex numbers**, by definition, the roots are i and $-i$.
So the general solution is $y(x) = C_1\,e^{ix} + C_2\,e^{-ix}$.

Solution. On the other hand, we easily check that $y_1 = \cos(x)$ and $y_2 = \sin(x)$ are two solutions.
Hence, the general solution can also be written as $y(x) = D_1 \cos(x) + D_2 \sin(x)$.

Important comment. That we have these two different representations is a consequence of **Euler's identity**

$$e^{ix} = \cos(x) + i\sin(x).$$

Note that $e^{-ix} = \cos(x) - i\sin(x)$.
On the other hand, $\cos(x) = \frac{1}{2}(e^{ix} + e^{-ix})$ and $\sin(x) = \frac{1}{2i}(e^{ix} - e^{-ix})$.
[Recall that the first formula is an instance of $\operatorname{Re}(z) = \frac{1}{2}(z + \bar{z})$ and the second of $\operatorname{Im}(z) = \frac{1}{2i}(z - \bar{z})$.]

Example 24. Find the general solution of $y'' - 4y' + 13y = 0$.

Solution. The characteristic polynomial $p(D) = D^2 - 4D + 13$ has roots $2 + 3i, 2 - 3i$.
Hence, the general solution is $y(x) = C_1 e^{2x}\cos(3x) + C_2 e^{2x}\sin(3x)$.
Note. $e^{(2+3i)x} = e^{2x}e^{3ix} = e^{2x}(\cos(3x) + i\sin(3x))$

Example 25. **(review)** Find the general solution of $y''' - 3y' + 2y = 0$.

Solution. The characteristic polynomial $p(D) = D^3 - 3D + 2 = (D-1)^2(D+2)$ has roots $1, 1, -2$.
By Theorem 20, the general solution is $y(x) = (C_1 + C_2 x)e^x + C_2 e^{-2x}$.

Example 26. **(review)** Consider the function $y(x) = 7x - 5x^2 e^{4x}$. Find an operator $p(D)$ such that $p(D)y = 0$.

Comment. This is the same as determining a homogeneous linear DE with constant coefficients solved by $y(x)$.

Solution. In order for $y(x)$ to be a solution of $p(D)y = 0$, the characteristic roots must include $0, 0, 4, 4, 4$.
The simplest choice for $p(D)$ thus is $p(D) = D^2(D-4)^3$.

Inhomogeneous linear DEs: The method of undetermined coefficients

The **method of undetermined coefficients** allows us to solve certain inhomogeneous linear DEs $Ly = f(x)$ with constant coefficients..

It works if $f(x)$ is itself a solution of a homogeneous linear DE with constant coefficients (see previous example).

Example 27. Determine the general solution of $y'' + 4y = 12x$.

Solution. The DE is $p(D)y = 12x$ with $p(D) = D^2 + 4$, which has roots $\pm 2i$. Thus, the general solution is $y(x) = y_p(x) + C_1\cos(2x) + C_2\sin(2x)$. It remains to find a particular solution y_p.

Since $D^2 \cdot (12x) = 0$, we apply D^2 to both sides of the DE to get the **homogeneous** DE $D^2(D^2+4) \cdot y = 0$.

Its general solution is $C_1 + C_2 x + C_3\cos(2x) + C_4\sin(2x)$ and y_p must be of this form. Indeed, there must be a particular solution of the simpler form $y_p = C_1 + C_2 x$ (because $C_3\cos(2x) + C_4\sin(2x)$ can be added to any y_p).

It remains to find appropriate values C_1, C_2 such that $y_p'' + 4y_p = 12x$. Since $y_p'' + 4y_p = 4C_1 + 4C_2 x$, comparing coefficients yields $4C_1 = 0$ and $4C_2 = 12$, so that $C_1 = 0$ and $C_2 = 3$. In other words, $y_p = 3x$.

Therefore, the general solution to the original DE is $y(x) = 3x + C_1\cos(2x) + C_2\sin(2x)$.

Example 28. Determine the general solution of $y'' + 4y' + 4y = e^{3x}$.

Solution. The DE is $p(D)y = e^{3x}$ with $p(D) = D^2 + 4D + 4 = (D+2)^2$, which has roots $-2, -2$. Thus, the general solution is $y(x) = y_p(x) + (C_1 + C_2 x)e^{-2x}$. It remains to find a particular solution y_p.

Since $(D-3)e^{3x} = 0$, we apply $(D-3)$ to the DE to get the **homogeneous** DE $(D-3)(D+2)^2 y = 0$.

Its general solution is $(C_1 + C_2 x)e^{-2x} + C_3 e^{3x}$ and y_p must be of this form. Indeed, there must be a particular solution of the simpler form $y_p = Ae^{3x}$.

To determine the value of C, we plug into the original DE: $y_p'' + 4y_p' + 4y_p = (9 + 4 \cdot 3 + 4)Ae^{3x} \stackrel{!}{=} e^{3x}$. Hence, $A = 1/25$. Therefore, the general solution to the original DE is $y(x) = (C_1 + C_2 x)e^{-2x} + \frac{1}{25}e^{3x}$.

Solution. (same, just shortened) In schematic form:

	homogeneous DE	inhomogeneous part
characteristic roots	$-2, -2$	3
solutions	e^{-2x}, xe^{-2x}	e^{3x}

This tells us that there exists a particular solution of the form $y_p = Ae^{3x}$. Then the general solution is

$$y = y_p + C_1 e^{-2x} + C_2 x e^{-2x}.$$

So far, we didn't need to do any calculations (besides determining the roots)! However, we still need to determine the value of A (by plugging into the DE as above), namely $A = \frac{1}{25}$. For this reason, this approach is often called the **method of undetermined coefficients**.

Armin Straub
straub@southalabama.edu

We found the following recipe for solving nonhomogeneous linear DEs with constant coefficients:

That approach works for $p(D)y = f(x)$ whenever the right-hand side $f(x)$ is the solution of some homogeneous linear DE with constant coefficients: $q(D)f(x) = 0$

(method of undetermined coefficients) To find a particular solution y_p to an inhomogeneous linear DE with constant coefficients $p(D)y = f(x)$:

- Determine the characteristic roots of the homogeneous DE and corresponding solutions.
- Find the roots of $q(D)$ so that $q(D)f(x) = 0$. [This does not work for all $f(x)$.]
  Let $y_{p,1}, y_{p,2}, \ldots$ be the additional solutions (when the roots are added to those of the homogeneous DE).

Then there exist (unique) C_i so that

$$y_p = C_1 y_{p,1} + C_2 y_{p,2} + \ldots$$

To find the values C_i, we need to plug y_p into the original DE.

Why? To see that this approach works, note that applying $q(D)$ to both sides of the inhomogeneous DE $p(D)y = f(x)$ results in $q(D)p(D)y = 0$ which is homogeneous. We already know that the solutions to the homogeneous DE can be added to any particular solution y_p. Therefore, we can focus only on the additional solutions coming from the roots of $q(D)$.

For which $f(x)$ does this work? By Theorem 20, we know exactly which $f(x)$ are solutions to homogeneous linear DEs with constant coefficients: these are linear combinations of exponentials $x^j e^{rx}$ (which includes $x^j e^{ax}\cos(bx)$ and $x^j e^{ax}\sin(bx)$).

Example 29. Determine the general solution of $y'' + 4y' + 4y = 7e^{-2x}$.

Solution. The homogeneous DE is $y'' + 4y' + 4y = 0$ (note that $D^2 + 4D + 4 = (D+2)^2$) and the inhomogeneous part is $7e^{-2x}$.

	homogeneous DE	inhomogeneous part
characteristic roots	$-2, -2$	-2
solutions	e^{-2x}, xe^{-2x}	$x^2 e^{-2x}$

This tells us that there exists a particular solution of the form $y_p = Cx^2 e^{-2x}$. To find the value of C, we plug into the DE.

$y_p' = C(-2x^2 + 2x)e^{-2x}$

$y_p'' = C(4x^2 - 8x + 2)e^{-2x}$

$y_p'' + 4y_p' + 4y_p = 2Ce^{-2x} \overset{!}{=} 7e^{-2x}$

It follows that $C = \frac{7}{2}$, so that $y_p = \frac{7}{2}x^2 e^{-2x}$. Hence the general solution is

$$y(x) = \left(C_1 + C_2 x + \frac{7}{2}x^2\right)e^{-2x}.$$

Example 30. Consider the DE $y'' + 4y' + 4y = 2e^{3x} - 5e^{-2x}$.

(a) What is the simplest form (with undetermined coefficients) of a particular solution?

(b) Determine a particular solution using our results from Examples 28 and 29.

(c) Determine the general solution.

Solution.

(a) Note that $D^2+4D+4=(D+2)^2$.

	homogeneous DE	inhomogeneous part
characteristic roots	$-2,-2$	$3,-2$
solutions	e^{-2x}, xe^{-2x}	e^{3x}, x^2e^{-2x}

Hence, there has to be a particular solution of the form $y_p=Ae^{3x}+Bx^2e^{-2x}$.

To find the (unique) values of A and B, we can plug into the DE. Alternatively, we can break the problem into two pieces as illustrated in the next part.

(b) Write the DE as $Ly=2e^{3x}-5e^{-2x}$ where $L=D^2+4D+4$. In Example 28 we found that $y_1=\frac{1}{25}e^{3x}$ satisfies $Ly_1=e^{3x}$. Also, in Example 29 we found that $y_2=\frac{7}{2}x^2e^{-2x}$ satisfies $Ly_2=7e^{-2x}$.

By linearity, it follows that $L(Ay_1+By_2)=ALy_1+BLy_2=Ae^{3x}+7Be^{-2x}$.

To get a particular solution y_p of our DE, we need $A=2$ and $7B=-5$.

Hence, $y_p=2y_1-\frac{5}{7}y_2=\frac{2}{25}e^{3x}-\frac{5}{2}x^2e^{-2x}$.

Comment. Of course, if we hadn't previously solved Examples 28 and 29, we could have plugged the result from the first part into the DE to determine the coefficients A and B. On the other hand, breaking the inhomogeneous part $(2e^{3x}-5e^{-2x})$ up into pieces (here, e^{3x} and e^{-2x}) can help keep things organized, especially when working by hand.

(c) The general solution is $\frac{2}{25}e^{3x}-\frac{5}{2}x^2e^{-2x}+(C_1+C_2x)e^{2x}$.

Example 31. Consider the DE $y''-2y'+y=5\sin(3x)$.

(a) What is the simplest form (with undetermined coefficients) of a particular solution?

(b) Determine a particular solution.

(c) Determine the general solution.

Solution. Note that $D^2-2D+1=(D-1)^2$.

	homogeneous DE	inhomogeneous part
characteristic roots	$1,1$	$\pm 3i$
solutions	e^x, xe^x	$\cos(3x), \sin(3x)$

(a) This tells us that there exists a particular solution of the form $y_p=A\cos(3x)+B\sin(3x)$.

(b) To find the values of A and B, we plug into the DE.

$y_p'=-3A\sin(3x)+3B\cos(3x)$

$y_p''=-9A\cos(3x)-9B\sin(3x)$

$y_p''-2y_p'+y_p=(-8A-6B)\cos(3x)+(6A-8B)\sin(3x)\overset{!}{=}5\sin(3x)$

Equating the coefficients of $\cos(x)$, $\sin(x)$, we obtain the two equations $-8A-6B=0$ and $6A-8B=5$.

Solving these, we find $A=\frac{3}{10}$, $B=-\frac{2}{5}$. Accordingly, a particular solution is $y_p=\frac{3}{10}\cos(3x)-\frac{2}{5}\sin(3x)$.

(c) The general solution is $y(x)=\frac{3}{10}\cos(3x)-\frac{2}{5}\sin(3x)+(C_1+C_2x)e^x$.

Example 32. Consider the DE $y''-2y'+y=5e^{2x}\sin(3x)+7xe^x$. What is the simplest form (with undetermined coefficients) of a particular solution?

Solution. Since $D^2-2D+1=(D-1)^2$, the characteristic roots are $1,1$. The roots for the inhomogeneous part are $2\pm 3i, 1, 1$. Hence, there has to be a particular solution of the form $y_p=Ae^{2x}\cos(3x)+Be^{2x}\sin(3x)+Cx^2e^x+Dx^3e^x$.

(We can then plug into the DE to determine the (unique) values of the coefficients A,B,C,D.)

Armin Straub
straub@southalabama.edu

Example 33. (homework) What is the shape of a particular solution of $y'' + 4y' + 4y = x\cos(x)$?

Solution. The characteristic roots are $-2, -2$. The roots for the inhomogeneous part are $\pm i, \pm i$. Hence, there has to be a particular solution of the form $y_p = (C_1 + C_2 x)\cos(x) + (C_3 + C_4 x)\sin(x)$.

Continuing to find a particular solution. To find the value of the C_j's, we plug into the DE.

$$y_p' = (C_2 + C_3 + C_4 x)\cos(x) + (C_4 - C_1 - C_2 x)\sin(x)$$

$$y_p'' = (2C_4 - C_1 - C_2 x)\cos(x) + (-2C_2 - C_3 - C_4 x)\sin(x)$$

$$\begin{aligned} y_p'' + 4y_p' + 4y_p = {} & (3C_1 + 4C_2 + 4C_3 + 2C_4 + (3C_2 + 4C_4)x)\cos(x) \\ & + (-4C_1 - 2C_2 + 3C_3 + 4C_4 + (-4C_2 + 3C_4)x)\sin(x) \overset{!}{=} x\cos(x). \end{aligned}$$

Equating the coefficients of $\cos(x)$, $x\cos(x)$, $\sin(x)$, $x\sin(x)$, we get the equations $3C_1 + 4C_2 + 4C_3 + 2C_4 = 0$, $3C_2 + 4C_4 = 1$, $-4C_1 - 2C_2 + 3C_3 + 4C_4 = 0$, $-4C_2 + 3C_4 = 0$.

Solving (this is tedious!), we find $C_1 = -\frac{4}{125}$, $C_2 = \frac{3}{25}$, $C_3 = -\frac{22}{125}$, $C_4 = \frac{4}{25}$.

Hence, $y_p = \left(-\frac{4}{125} + \frac{3}{25}x\right)\cos(x) + \left(-\frac{22}{125} + \frac{4}{25}x\right)\sin(x)$.

Armin Straub
straub@southalabama.edu

Example 34. (review) What is the shape of a particular solution of $y''+4y'+4y=4e^{3x}\sin(2x)-x\sin(x)$.

Solution. The characteristic roots are $-2,-2$. The roots for the inhomogeneous part roots are $3\pm2i,\pm i,\pm i$. Hence, there has to be a particular solution of the form
$y_p=C_1e^{3x}\cos(2x)+C_2e^{3x}\sin(2x)+(C_3+C_4x)\cos(x)+(C_5+C_6x)\sin(x)$.

Continuing to find a particular solution. To find the values of $C_1,\ldots,C_6$, we plug into the DE. But this final step is so boring that we don't go through it here. Computers (currently?) cannot afford to be as selective; mine obediently calculated: $y_p=-\frac{4}{841}e^{3x}(20\cos(2x)-21\sin(2x))+\frac{1}{125}((-22+20x)\cos(x)+(4-15x)\sin(x))$

Sage

In practice, we are happy to let a machine do tedious computations. Let us see how to use the open-source computer algebra system **Sage** to do basic computations for us.

Sage is freely available at `sagemath.org`. Instead of installing it locally (it's huge!) we can conveniently use it in the cloud at `cocalc.com` from any browser.

[For basic computations, you can also simply use the textbox on our course website.]

Sage is built as a **Python** library, so any Python code is valid. For starters, we will use it as a fancy calculator.

Example 35. To solve the differential equation $y''+4y'+4y=7e^{-2x}$, as we did in Example 29, we can use the following:

```
>>> x = var('x')
>>> y = function('y')(x)
>>> desolve(diff(y,x,2) + 4*diff(y,x) + 4*y == 7*exp(-2*x), y)
```

$$\frac{7}{2}\,x^2\,e^{(-2\,x)}+(K_2\,x+K_1)\,e^{(-2\,x)}$$

This confirms, as we had found, that the general solution is $y(x)=\left(C_1+C_2x+\frac{7}{2}x^2\right)e^{-2x}$.

Example 36. Similarly, Sage can solve initial value problems such as $y''-y'-2y=0$ with initial conditions $y(0)=4$, $y'(0)=5$.

```
>>> x = var('x')
>>> y = function('y')(x)
>>> desolve(diff(y,x,2) - diff(y,x) - 2*y == 0, y, ics=[0,4,5])
```

$$3\,e^{(2\,x)}+e^{(-x)}$$

This matches the (unique) solution $y(x)=3e^{2x}+e^{-x}$ that we derived in Example 18.

Higher order. Unfortunately, the command `desolve` currently only works like this for differential equations of first and second order. To likewise solve a third-order differential equation, we can use the function `desolve_laplace` instead. For instance, to solve the IVP $y'''=3y''-4y$ with $y(0)=1$, $y'(0)=-2$, $y''(0)=3$, use

```
>>> desolve_laplace(diff(y,x,3) == 3*diff(y,x,2) - 4*y, y, ics=[0,1,-2,3])
```

$$x\,e^{(2\,x)}-\frac{2}{3}\,e^{(2\,x)}+\frac{5}{3}\,e^{(-x)}$$

to find that the unique solution is $y(x)=\frac{1}{3}(3x-2)e^{2x}+\frac{5}{3}e^{-x}$.

Armin Straub
straub@southalabama.edu

More on differential operators

Example 37. We have been factoring differential operators like $D^2+4D+4=(D+2)^2$.

Things become much more complicated when the coefficients are not constant!

For instance, the linear DE $y''+4y'+4xy=0$ can be written as $Ly=0$ with $L=D^2+4D+4x$. However, in general, such operators cannot be factored (unless we allow as coefficients functions in x that we are not familiar with). [On the other hand, any ordinary polynomial can be factored over the complex numbers.]

One indication that things become much more complicated is that x and D do not commute: $xD\neq Dx$!!

Indeed, $(xD)f(x)=xf'(x)$ while $(Dx)f(x)=\frac{\mathrm{d}}{\mathrm{d}x}[xf(x)]=f(x)+xf'(x)=(1+xD)f(x)$.

This computation shows that, in fact, $Dx=xD+1$.

Review. Linear DEs are those that can be written as $Ly=f(x)$ where L is a linear differential operator: namely,

$$L=p_n(x)D^n+p_{n-1}(x)D^{n-1}+\ldots+p_1(x)D+p_0(x). \tag{1}$$

Recall that the operators xD and Dx are not the same: instead, $Dx=xD+1$.

We say that an operator of the form (1) is in **normal form**.

For instance. xD is in normal form, whereas Dx is not in normal form. It follows from the previous example that the normal form of Dx is $xD+1$.

Example 38. Let $a=a(x)$ be some function.

(a) Write the operator Da in normal form [normal form means as in (1)].

(b) Write the operator D^2a in normal form.

Solution.

(a) $(Da)f(x)=\frac{\mathrm{d}}{\mathrm{d}x}[a(x)\,f(x)]=a'(x)f(x)+a(x)f'(x)=(a'+aD)f(x)$

Hence, $Da=aD+a'$.

(b) $(D^2a)f(x)=\frac{\mathrm{d}^2}{\mathrm{d}x^2}[a(x)\,f(x)]=\frac{\mathrm{d}}{\mathrm{d}x}[a'(x)f(x)+a(x)f'(x)]=a''(x)f(x)+2a'(x)f'(x)+a(x)f''(x)$
$=(a''+2a'D+aD^2)f(x)$

Hence, $D^2a=aD^2+2a'D+a''$.

Armin Straub
straub@southalabama.edu

Example 39. (review) Let $a = a(x)$ be some function.

(a) Write the operator Da in normal form [normal form means as in (1)].

(b) Write the operator $D^2 a$ in normal form.

Solution.

(a) $(Da)f(x) = \frac{\mathrm{d}}{\mathrm{d}x}[a(x)\,f(x)] = a'(x)f(x) + a(x)f'(x) = (a' + aD)f(x)$
Hence, $Da = aD + a'$.

(b) $(D^2 a)f(x) = \frac{\mathrm{d}^2}{\mathrm{d}x^2}[a(x)\,f(x)] = \frac{\mathrm{d}}{\mathrm{d}x}[a'(x)f(x) + a(x)f'(x)] = a''(x)f(x) + 2a'(x)f'(x) + a(x)f''(x)$
$= (a'' + 2a'D + aD^2)f(x)$
Hence, $D^2 a = aD^2 + 2a'D + a''$.
Alternatively. We can also use $Da = aD + a'$ from the previous part and work with the operators directly:
$D^2 a = D(Da) = D(aD + a') = DaD + Da' = (aD + a')D + a'D + a'' = aD^2 + 2a'D + a''$.

Example 40. Suppose that a and b depend on x. Expand $(D + a)(D + b)$ in normal form.

Solution. $(D + a)(D + b) = D^2 + Db + aD + ab = D^2 + (bD + b') + aD + ab = D^2 + (a + b)D + ab + b'$

Comment. Of course, if b is a constant, then $b' = 0$ and we just get the familiar expansion.
Comment. At this point, it is not surprising that, in general, $(D + a)(D + b) \neq (D + b)(D + a)$.

Example 41. Suppose we want to factor $D^2 + pD + q$ as $(D + a)(D + b)$. [p, q, a, b depend on x]

(a) Spell out equations to find a and b.

(b) Find all factorizations of D^2. [An obvious one is $D^2 = D \cdot D$ but there are others!]

Solution.

(a) Matching coefficients with $(D + a)(D + b) = D^2 + (a + b)D + ab + b'$ (we expanded this in the previous example), we find that we need

$$p = a + b, \quad q = ab + b'.$$

Equivalently, $a = p - b$ and $q = (p - b)b + b'$. The latter is a nonlinear (!) DE for b. Once solved for b, we obtain a as $a = p - b$.

(b) This is the case $p = q = 0$. The DE for b becomes $b' = b^2$.
Because it is separable (show all details!), we find that $b(x) = \frac{1}{C - x}$ or $b(x) = 0$.
Since $a = -b$, we obtain the factorizations $D^2 = \left(D - \frac{1}{C - x}\right)\left(D + \frac{1}{C - x}\right)$ and $D^2 = D \cdot D$.
Our computations show that there are no further factorizations.

Comment. Note that this example illustrates that factorization of differential operators is not unique!
For instance, $D^2 = D \cdot D$ and $D^2 = \left(D + \frac{1}{x}\right) \cdot \left(D - \frac{1}{x}\right)$ (the case $C = 0$ above).
Comment. In general, the nonlinear DE for b does not have any polynomial or rational solution (or, in fact, any solution that can be expressed in terms of functions that we are familiar with).

Solving linear recurrences with constant coefficients

Motivation: Fibonacci numbers

The numbers $0, 1, 1, 2, 3, 5, 8, 13, 21, 34, \ldots$ are called **Fibonacci numbers**.

They are defined by the recursion $F_{n+1} = F_n + F_{n-1}$ and $F_0 = 0$, $F_1 = 1$.

How fast are they growing?

Have a look at ratios of Fibonacci numbers: $\frac{2}{1} = 2$, $\frac{3}{2} = 1.5$, $\frac{5}{3} \approx 1.667$, $\frac{8}{5} = 1.6$, $\frac{13}{8} = 1.625$, $\frac{21}{13} = 1.615$, $\frac{34}{21} = 1.619$, …

These ratios approach the **golden ratio** $\varphi = \frac{1+\sqrt{5}}{2} = 1.618\ldots$

In other words, it appears that $\lim_{n\to\infty} \frac{F_{n+1}}{F_n} = \frac{1+\sqrt{5}}{2}$.

We will soon understand where this is coming from.

We can derive all of that using the same ideas as in the case of linear differential equations. The crucial observation that we can write the recursion in operator form:

> $F_{n+1} = F_n + F_{n-1}$ is equivalent to $(N^2 - N - 1)F_n = 0$.
>
> Here, N is the shift operator: $Na_n = a_{n+1}$.

Comment. Recurrence equations are discrete analogs of differential equations.

For instance, recall that $f'(x) = \lim_{h\to 0} \frac{1}{h}[f(x+h) - f(x)]$.

Setting $h = 1$, we get the rough estimate $f'(x) \approx f(x+1) - f(x)$ so that D is (roughly) approximated by $N - 1$.

Example 42. Find the general solution to the recursion $a_{n+1} = 7a_n$.

Solution. Note that $a_n = 7a_{n-1} = 7 \cdot 7a_{n-2} = \cdots = 7^n a_0$.

Hence, the general solution is $a_n = C \cdot 7^n$.

Comment. This is analogous to $y' = 7y$ having the general solution $y(x) = Ce^{7x}$.

Solving recurrence equations

Example 43. ("warmup") Let the sequence a_n be defined by the recursion $a_{n+2} = a_{n+1} + 6a_n$ and the initial values $a_0 = 1$, $a_1 = 8$. Determine the first few terms of the sequence.

Solution. $a_2 = a_1 + 6a_0 = 14$, $a_3 = a_2 + 6a_1 = 62$, $a_4 = 146$, …

Comment. In the next example, we get ready to solve this recursion and to find an explicit formula for the sequence a_n.

Example 44. ("warmup") Find the general solution to the recursion $a_{n+2} = a_{n+1} + 6a_n$.

Solution. The recursion can be written as $p(N)a_n = 0$ where $p(N) = N^2 - N - 6 = (N-3)(N+2)$.

Since $(N-3)a_n = 0$ has solution $a_n = C \cdot 3^n$, and since $(N+2)a_n = 0$ has solution $a_n = C \cdot (-2)^n$ (compare previous example), we conclude that the general solution is $a_n = C_1 \cdot 3^n + C_2 \cdot (-2)^n$.

Comment. This must indeed be the general solution, because the two degrees of freedom C_1, C_2 allow us to match any initial conditions $a_0 = A$, $a_1 = B$: the two equations $C_1 + C_2 = A$ and $3C_1 - 2C_2 = B$ in matrix form are $\begin{bmatrix} 1 & 1 \\ 3 & -2 \end{bmatrix}\begin{bmatrix} C_1 \\ C_2 \end{bmatrix} = \begin{bmatrix} A \\ B \end{bmatrix}$, which always has a (unique) solution because $\det\left(\begin{bmatrix} 1 & 1 \\ 3 & -2 \end{bmatrix}\right) = -5 \neq 0$.

Armin Straub
straub@southalabama.edu

Review. The recurrence $a_{n+1} = 5a_n$ has general solution $a_n = C \cdot 5^n$.

In operator form, the recurrence is $(N-5)a_n = 0$, where $p(N) = N - 5$ is the characteristic polynomial. The characteristic root 5 corresponds to the solution 5^n.

This is analogous to the case of DEs $p(D)y = 0$ where a root r of $p(D)$ corresponds to the solution e^{rx}.

Example 45. (cont'd) Let the sequence a_n be defined by $a_{n+2} = a_{n+1} + 6a_n$ and $a_0 = 1$, $a_1 = 8$.

(a) Determine the first few terms of the sequence.

(b) Find a formula for a_n.

(c) Determine $\lim_{n \to \infty} \frac{a_{n+1}}{a_n}$.

Solution.

(a) $a_2 = a_1 + 6a_0 = 14$, $a_3 = a_2 + 6a_1 = 62$, $a_4 = 146$, ...

(b) The recursion can be written as $p(N)a_n = 0$ where $p(N) = N^2 - N - 6$ has roots $3, -2$.

Hence, $a_n = C_1\, 3^n + C_2\, (-2)^n$ and we only need to figure out the two unknowns C_1, C_2. We can do that using the two initial conditions: $a_0 = C_1 + C_2 = 1$, $a_1 = 3C_1 - 2C_2 = 8$.

Solving, we find $C_1 = 2$ and $C_2 = -1$ so that, in conclusion, $a_n = 2 \cdot 3^n - (-2)^n$.

Comment. Such a formula is sometimes called a **Binet-like formula** (because it is of the same kind as the Binet formula for the Fibonacci numbers that we can derive in the same manner).

(c) It follows from our formula that $\lim_{n \to \infty} \frac{a_{n+1}}{a_n} = 3$ (because $|3| > |-2|$ so that 3^n dominates $(-2)^n$).

To see this, we need to realize that, for large n, 3^n is much larger than $(-2)^n$ so that we have $a_n \approx 2 \cdot 3^n$ when n is large. Hence, $\frac{a_{n+1}}{a_n} \approx \frac{2 \cdot 3^{n+1}}{2 \cdot 3^n} = 3$.

Alternatively, to be very precise, we can observe that (by dividing each term by 3^n)

$$\frac{a_{n+1}}{a_n} = \frac{2 \cdot 3^{n+1} - (-2)^{n+1}}{2 \cdot 3^n - (-2)^n} = \frac{2 \cdot 3 + 2\left(-\frac{2}{3}\right)^n}{2 \cdot 1 - \left(-\frac{2}{3}\right)^n} \quad \overset{\text{as } n \to \infty}{\longrightarrow} \quad \frac{2 \cdot 3 + 0}{2 \cdot 1 - 0} = 3.$$

Example 46. Consider the sequence a_n defined by $a_{n+2} = a_{n+1} + 2a_n$ and $a_0 = 1$, $a_1 = 8$.

(a) Determine the first few terms of the sequence.

(b) Find a formula for a_n.

(c) Determine $\lim_{n \to \infty} \frac{a_{n+1}}{a_n}$.

Solution.

(a) $a_2 = 10$, $a_3 = 26$

(b) The recursion can be written as $p(N)a_n = 0$ where $p(N) = N^2 - N - 2$ has roots $2, -1$.

Hence, $a_n = C_1\, 2^n + C_2\, (-1)^n$ and we only need to figure out the two unknowns C_1, C_2. We can do that using the two initial conditions: $a_0 = C_1 + C_2 = 1$, $a_1 = 2C_1 - C_2 = 8$.

Solving, we find $C_1 = 3$ and $C_2 = -2$ so that, in conclusion, $a_n = 3 \cdot 2^n - 2(-1)^n$.

(c) It follows from the formula $a_n = 3 \cdot 2^n - 2(-1)^n$ that $\lim_{n \to \infty} \frac{a_{n+1}}{a_n} = 2$.

Comment. In fact, this already follows from $a_n = C_1\, 2^n + C_2\, (-1)^n$ provided that $C_1 \neq 0$. Since $a_n = C_2\, (-1)^n$ (the case $C_1 = 0$) is not compatible with $a_0 = 1$, $a_1 = 8$, we can conclude $\lim_{n \to \infty} \frac{a_{n+1}}{a_n} = 2$ without computing the actual values of C_1 and C_2.

Example 47. ("warmup") Find the general solution to the recursion $a_{n+2} = 4a_{n+1} - 4a_n$.

Solution. The recursion can be written as $p(N)a_n = 0$ where $p(N) = N^2 - 4N + 4$ has roots $2, 2$.
So a solution is 2^n and, from our discussion of DEs, it is probably not surprising that a second solution is $n \cdot 2^n$.
Hence, the general solution is $a_n = C_1 \cdot 2^n + C_2 \cdot n \cdot 2^n = (C_1 + C_2 n) \cdot 2^n$.

Comment. This is analogous to $(D-2)^2 y' = 0$ having the general solution $y(x) = (C_1 + C_2 x)e^{2x}$.

Check! Let's check that $a_n = n \cdot 2^n$ indeed satisfies the recursion $(N-2)^2 a_n = 0$.
$(N-2)n \cdot 2^n = (n+1)2^{n+1} - 2n \cdot 2^n = 2^{n+1}$, so that $(N-2)^2 n \cdot 2^n = (N-2)2^{n+1} = 0$.

Combined, we obtain the following analog of Theorem 20 for recurrence equations (RE):

Comment. Sequences that are solutions to such recurrences are called **constant recursive** or **C-finite**.

Theorem 48. Consider the homogeneous linear RE with constant coefficients $p(N)a_n = 0$.

- If r is a root of the characteristic polynomial and if k is its multiplicity, then k (independent) solutions of the RE are given by $n^j r^n$ for $j = 0, 1, \ldots, k-1$.
- Combining these solutions for all roots, gives the general solution.

Moreover. If r is the sole largest root by absolute value among the roots contributing to a_n, then $a_n \approx C r^n$ (if r is not repeated—what if it is?) for large n. In particular, it follows that

$$\lim_{n \to \infty} \frac{a_{n+1}}{a_n} = r.$$

Advanced comment. Things can get weird if there are several roots of the same absolute value. Consider, for instance, the case $a_n = 2^n + (-2)^n$. Can you see that, in this case, the limit $\lim_{n \to \infty} \frac{a_{n+1}}{a_n}$ doesn't even exist?

Example 49. Find the general solution to the recursion $a_{n+3} = 2a_{n+2} + a_{n+1} - 2a_n$.

Solution. The recursion can be written as $p(N)a_n = 0$ where $p(N) = N^3 - 2N^2 - N + 2$ has roots $2, 1, -1$.
(Here, we may use some help from a computer algebra system to find the roots.)
Hence, the general solution is $a_n = C_1 \cdot 2^n + C_2 + C_3 \cdot (-1)^n$.

Example 50. Find the general solution to the recursion $a_{n+3} = 3a_{n+2} - 4a_n$.

Solution. The recursion can be written as $p(N)a_n = 0$ where $p(N) = N^3 - 3N^2 + 4$ has roots $2, 2, -1$. (Again, we may use some help from a computer algebra system to find the roots.)
Hence, the general solution is $a_n = (C_1 + C_2 n) \cdot 2^n + C_3 \cdot (-1)^n$.

Armin Straub
straub@southalabama.edu

Theorem 51. (Binet's formula) $F_n = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right]$

Proof. The recursion $F_{n+1} = F_n + F_{n-1}$ can be written as $p(N)a_n = 0$ where $p(N) = N^2 - N - 1$ has roots

$$\lambda_1 = \frac{1+\sqrt{5}}{2} \approx 1.618, \quad \lambda_2 = \frac{1-\sqrt{5}}{2} \approx -0.618.$$

Hence, $F_n = C_1 \cdot \lambda_1^n + C_2 \cdot \lambda_2^n$ and we only need to figure out the two unknowns C_1, C_2. We can do that using the two initial conditions: $F_0 = C_1 + C_2 \overset{!}{=} 0$, $F_1 = C_1 \cdot \frac{1+\sqrt{5}}{2} + C_2 \cdot \frac{1-\sqrt{5}}{2} \overset{!}{=} 1$.

Solving, we find $C_1 = \frac{1}{\sqrt{5}}$ and $C_2 = -\frac{1}{\sqrt{5}}$ so that, in conclusion, $F_n = \frac{1}{\sqrt{5}}(\lambda_1^n - \lambda_2^n)$, as claimed. □

Comment. For large n, $F_n \approx \frac{1}{\sqrt{5}}\lambda_1^n$ (because λ_2^n becomes very small). In fact, $F_n = \operatorname{round}\left(\frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n\right)$.

Back to the quotient of Fibonacci numbers. In particular, because λ_1^n dominates λ_2^n, it is now transparent that the ratios $\frac{F_{n+1}}{F_n}$ approach $\lambda_1 = \frac{1+\sqrt{5}}{2} \approx 1.618$. To be precise, note that

$$\frac{F_{n+1}}{F_n} = \frac{\frac{1}{\sqrt{5}}(\lambda_1^{n+1} - \lambda_2^{n+1})}{\frac{1}{\sqrt{5}}(\lambda_1^n - \lambda_2^n)} = \frac{\lambda_1^{n+1} - \lambda_2^{n+1}}{\lambda_1^n - \lambda_2^n} = \frac{\lambda_1 - \lambda_2\left(\frac{\lambda_2}{\lambda_1}\right)^n}{1 - \left(\frac{\lambda_2}{\lambda_1}\right)^n} \xrightarrow{n\to\infty} \frac{\lambda_1 - 0}{1 - 0} = \lambda_1.$$

In fact, it follows from $\lambda_2 < 0$ that the ratios $\frac{F_{n+1}}{F_n}$ approach λ_1 in the alternating fashion that we observed numerically earlier. Can you see that?

Example 52. Consider the sequence a_n defined by $a_{n+2} = 4a_{n+1} + 9a_n$ and $a_0 = 1$, $a_1 = 2$. Determine $\lim_{n\to\infty} \frac{a_{n+1}}{a_n}$.

Solution. The recursion can be written as $p(N)a_n = 0$ where $p(N) = N^2 - 4N - 9$ has roots $\frac{4 \pm \sqrt{52}}{2} \approx 5.6056$, -1.6056. Both roots have to be involved in the solution in order to get integer values.
We conclude that $\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = 2 + \sqrt{13} \approx 5.6056$ (because $|5.6056| > |-1.6056|$).

Example 53. (extra) Consider the sequence a_n defined by $a_{n+2} = 2a_{n+1} + 4a_n$ and $a_0 = 0$, $a_1 = 1$. Determine $\lim_{n\to\infty} \frac{a_{n+1}}{a_n}$.

First few terms of sequence. $0, 1, 2, 8, 24, 80, 256, 832, \ldots$
These are actually related to Fibonacci numbers. Indeed, $a_n = 2^{n-1}F_n$. Can you prove this directly from the recursions? Alternatively, this follows from comparing the Binet-like formulas.

Solution. Proceeding as in the previous example, we find $\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = 1 + \sqrt{5} \approx 3.23607$.

Comment. With just a little more work, we find the Binet-like formula $a_n = \frac{(1+\sqrt{5})^n - (1-\sqrt{5})^n}{2\sqrt{5}}$.

Crash course: Eigenvalues and eigenvectors

If $A\boldsymbol{x} = \lambda\boldsymbol{x}$ (and $\boldsymbol{x} \neq \boldsymbol{0}$), then $\boldsymbol{x}$ is an **eigenvector** of A with **eigenvalue** λ (just a number).

Note that, for the equation $A\boldsymbol{x} = \lambda\boldsymbol{x}$ to make sense, A needs to be a square matrix (i.e. $n \times n$).

Key observation:

$$\begin{aligned} & A\boldsymbol{x} = \lambda\boldsymbol{x} \\ \iff & A\boldsymbol{x} - \lambda\boldsymbol{x} = \boldsymbol{0} \\ \iff & (A - \lambda I)\boldsymbol{x} = \boldsymbol{0} \end{aligned}$$

This homogeneous system has a nontrivial solution $\boldsymbol{x}$ if and only if $\det(A - \lambda I) = 0$.

To find eigenvectors and eigenvalues of A:

(a) First, find the eigenvalues λ by solving $\det(A - \lambda I) = 0$.

$\det(A - \lambda I)$ is a polynomial in λ, called the **characteristic polynomial** of A.

(b) Then, for each eigenvalue λ, find corresponding eigenvectors by solving $(A - \lambda I)\boldsymbol{x} = \boldsymbol{0}$.

Example 54. Determine the eigenvalues and eigenvectors of $A = \begin{bmatrix} 8 & -10 \\ 5 & -7 \end{bmatrix}$.

Solution. The characteristic polynomial is:

$\det(A - \lambda I) = \det\left(\begin{bmatrix} 8-\lambda & -10 \\ 5 & -7-\lambda \end{bmatrix}\right) = (8-\lambda)(-7-\lambda) + 50 = \lambda^2 - \lambda - 6 = (\lambda - 3)(\lambda + 2)$

Hence, the eigenvalues are $\lambda = 3$ and $\lambda = -2$.

- To find an eigenvector for $\lambda = 3$, we need to solve $\begin{bmatrix} 5 & -10 \\ 5 & -10 \end{bmatrix}\boldsymbol{x} = \boldsymbol{0}$.
  Hence, $\boldsymbol{x} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ is an eigenvector for $\lambda = 3$.
- To find an eigenvector for $\lambda = -2$, we need to solve $\begin{bmatrix} 10 & -10 \\ 5 & -5 \end{bmatrix}\boldsymbol{x} = \boldsymbol{0}$.
  Hence, $\boldsymbol{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is an eigenvector for $\lambda = -2$.

Check! $\begin{bmatrix} 8 & -10 \\ 5 & -7 \end{bmatrix}\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 3 \end{bmatrix} = 3 \cdot \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 8 & -10 \\ 5 & -7 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ -2 \end{bmatrix} = -2 \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

On the other hand, a random other vector like $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ is not an eigenvector: $\begin{bmatrix} 8 & -10 \\ 5 & -7 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -12 \\ -9 \end{bmatrix} \neq \lambda\begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

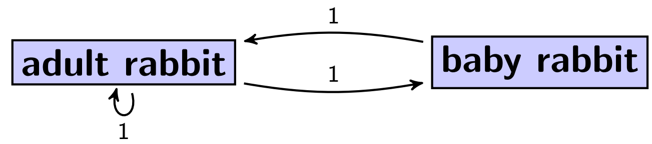
Example 55. (homework) Determine the eigenvalues and eigenvectors of $A = \begin{bmatrix} 1 & -6 \\ 1 & -4 \end{bmatrix}$.

Solution. (final answer only) $\boldsymbol{x} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ is an eigenvector for $\lambda = -2$, and $\boldsymbol{x} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ is an eigenvector for $\lambda = -1$.

Preview: A system of recurrence equations equivalent to the Fibonacci recurrence

Example 56. We model rabbit reproduction as follows.

Each month, every pair of adult rabbits produces one pair of baby rabbit as offspring. Meanwhile, it takes baby rabbits one month to mature to adults.



Comment. In this simplified model, rabbits always come in male/female pairs and no rabbits die. Though these features might make it sound fairly useless, the model may have some merit when describing populations under ideal conditions (unlimited resources) and over short time (no deaths).

Historical comment. The question how many rabbits there are after one year, when starting out with a pair of baby rabbits is famously included in the 1202 textbook of the Italian mathematician Leonardo of Pisa, known as Fibonacci.

If we start with one baby rabbit pair, how many adult rabbit pairs are there after n months?

Solution. Let a_n be the number of adult rabbit pairs after n months. Likewise, b_n is the number of baby rabbit pairs. The transition from one month to the next is given by $a_{n+1} = a_n + b_n$ and $b_{n+1} = a_n$. Using $b_n = a_{n-1}$ (from the second equation) in the first equation, we obtain $a_{n+1} = a_n + a_{n-1}$.

The initial conditions are $a_0 = 0$ and $a_1 = 1$ (the latter follows from $b_0 = 1$).

It follows that the number b_n of adult rabbit pairs are precisely the Fibonacci numbers F_n.

Comment. Note that the transition from one month to the next is described by in matrix-vector form as

$$\begin{bmatrix} a_{n+1} \\ b_{n+1} \end{bmatrix} = \begin{bmatrix} a_n + b_n \\ a_n \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a_n \\ b_n \end{bmatrix}.$$

Writing $\boldsymbol{a}_n = \begin{bmatrix} a_n \\ b_n \end{bmatrix}$, this becomes $\boldsymbol{a}_{n+1} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \boldsymbol{a}_n$ with $\boldsymbol{a}_0 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

Consequently, $\boldsymbol{a}_n = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^n \boldsymbol{a}_0 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^n \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

Looking ahead. Can you see how, starting with the Fibonacci recurrence $F_{n+2} = F_{n+1} + F_n$, we can arrive at this same system?

Solution. Set $\boldsymbol{a}_n = \begin{bmatrix} F_{n+1} \\ F_n \end{bmatrix}$. Then $\boldsymbol{a}_{n+1} = \begin{bmatrix} F_{n+2} \\ F_{n+1} \end{bmatrix} = \begin{bmatrix} F_{n+1} + F_n \\ F_{n+1} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} F_{n+1} \\ F_n \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \boldsymbol{a}_n$.

Systems of recurrence equations

Example 57. (review) Consider the sequence a_n defined by $a_{n+2} = 4a_n - 3a_{n+1}$ and $a_0 = 1$, $a_1 = 2$. Determine $\lim_{n\to\infty} \frac{a_{n+1}}{a_n}$.

Solution. The recursion can be written as $p(N)a_n = 0$ where $p(N) = N^2 + 3N - 4$ has roots $1, -4$. Hence, the general solution is $a_n = C_1 + C_2 \cdot (-4)^n$. We can see that both roots have to be involved in the solution (in other words, $C_1 \neq 0$ and $C_2 \neq 0$) because $a_n = C_1$ and $a_n = C_2 \cdot (-4)^n$ are not consistent with the initial conditions. We conclude that $\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = -4$ (because $|-4| > |1|$).

Example 58. Write the (second-order) RE $a_{n+2} = 4a_n - 3a_{n+1}$, with $a_0 = 1$, $a_1 = 2$, as a system of (first-order) recurrences.

Solution. Write $b_n = a_{n+1}$.

Then, $a_{n+2} = 4a_n - 3a_{n+1}$ translates into the first-order system $\begin{cases} a_{n+1} = b_n \\ b_{n+1} = 4a_n - 3b_n \end{cases}$.

Let $\boldsymbol{a}_n = \begin{bmatrix} a_n \\ b_n \end{bmatrix}$. Then, in matrix form, the RE is $\boldsymbol{a}_{n+1} = \begin{bmatrix} 0 & 1 \\ 4 & -3 \end{bmatrix} \boldsymbol{a}_n$, with $\boldsymbol{a}_0 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

Equivalently. Write $\boldsymbol{a}_n = \begin{bmatrix} a_n \\ a_{n+1} \end{bmatrix}$. Then we obtain the above system as

$$\boldsymbol{a}_{n+1} = \begin{bmatrix} a_{n+1} \\ a_{n+2} \end{bmatrix} = \begin{bmatrix} a_{n+1} \\ 4a_n - 3a_{n+1} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 4 & -3 \end{bmatrix} \begin{bmatrix} a_n \\ a_{n+1} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 4 & -3 \end{bmatrix} \boldsymbol{a}_n, \quad \boldsymbol{a}_0 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

Comment. It follows that $\boldsymbol{a}_n = \begin{bmatrix} 0 & 1 \\ 4 & -3 \end{bmatrix}^n \boldsymbol{a}_0 = \begin{bmatrix} 0 & 1 \\ 4 & -3 \end{bmatrix}^n \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Solving (systems of) REs is equivalent to computing powers of matrices!

Comment. We could also write $\boldsymbol{a}_n = \begin{bmatrix} a_{n+1} \\ a_n \end{bmatrix}$ (with the order of the entries reversed). In that case, the system is

$$\boldsymbol{a}_{n+1} = \begin{bmatrix} a_{n+2} \\ a_{n+1} \end{bmatrix} = \begin{bmatrix} 4a_n - 3a_{n+1} \\ a_{n+1} \end{bmatrix} = \begin{bmatrix} -3 & 4 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a_{n+1} \\ a_n \end{bmatrix} = \begin{bmatrix} -3 & 4 \\ 1 & 0 \end{bmatrix} \boldsymbol{a}_n, \quad \boldsymbol{a}_0 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

Comment. Recall that the **characteristic polynomial** of a matrix M is $\det(M - \lambda I)$. Compute the characteristic polynomial of both $M = \begin{bmatrix} 0 & 1 \\ 4 & -3 \end{bmatrix}$ and $M = \begin{bmatrix} -3 & 4 \\ 1 & 0 \end{bmatrix}$. In both cases, we get $\lambda^2 + 3\lambda - 4$, which matches the polynomial $p(N)$ (also called characteristic polynomial!) in the previous example. This will always happen and explains why both are referred to as the characteristic polynomial.

Example 59. Write $a_{n+3} - 4a_{n+2} + a_{n+1} + 6a_n = 0$ as a system of (first-order) recurrences.

Solution. Write $\boldsymbol{a}_n = \begin{bmatrix} a_n \\ a_{n+1} \\ a_{n+2} \end{bmatrix}$. Then we obtain the system

$$\boldsymbol{a}_{n+1} = \begin{bmatrix} a_{n+1} \\ a_{n+2} \\ a_{n+3} \end{bmatrix} = \begin{bmatrix} a_{n+1} \\ a_{n+2} \\ 4a_{n+2} - a_{n+1} - 6a_n \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -6 & -1 & 4 \end{bmatrix} \begin{bmatrix} a_n \\ a_{n+1} \\ a_{n+2} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -6 & -1 & 4 \end{bmatrix} \boldsymbol{a}_n.$$

In summary, the RE in matrix form is $\boldsymbol{a}_{n+1} = M\boldsymbol{a}_n$ with M the matrix above.

Important comment. Given a first-order system $\boldsymbol{a}_{n+1} = M\boldsymbol{a}_n$, it is clear that the solution satisfies $\boldsymbol{a}_n = M^n \boldsymbol{a}_0$. If you know how to compute matrix powers M^n, this means you can solve recurrences! On the other hand, we will proceed the other way around. We solve the recurrence and then use that to determine M^n.

Solving systems of recurrence equations

The following summarizes how we can solve systems of recurrence equations using eigenvectors. As a bonus, we obtain a way to compute matrix powers.

Each step is spelled out in Example 60 below.

(solving systems of REs) To solve $\boldsymbol{a}_{n+1} = M\boldsymbol{a}_n$, determine the eigenvectors of M.

- Each λ-eigenvector $\boldsymbol{v}$ provides a solution: $\boldsymbol{a}_n = \boldsymbol{v}\lambda^n$ [assuming that $\lambda \neq 0$]
- If there are enough eigenvectors, these combine to the general solution.
  In that case, we get a **fundamental matrix (solution)** Φ_n by placing each solution vector into one column of Φ_n.
- If desired, we can compute the **matrix powers** M^n using any fundamental matrix Φ_n as

$$M^n = \Phi_n \Phi_0^{-1}.$$

Note that M^n is the unique matrix solution to $\boldsymbol{a}_{n+1} = M\boldsymbol{a}_n$ with $\boldsymbol{a}_0 = I$ (the identity matrix).

Application: the unique solution to $\boldsymbol{a}_{n+1} = M\boldsymbol{a}_n$, $\boldsymbol{a}_0 = \boldsymbol{c}$ is given by $\boldsymbol{a}_n = M^n\boldsymbol{c}$.

Why? If $\boldsymbol{a}_n = \boldsymbol{v}\lambda^n$ for a λ-eigenvector $\boldsymbol{v}$, then $\boldsymbol{a}_{n+1} = \boldsymbol{v}\lambda^{n+1}$ and $M\boldsymbol{a}_n = M\boldsymbol{v}\lambda^n = \lambda\boldsymbol{v}\cdot\lambda^n = \boldsymbol{v}\lambda^{n+1}$.

Where is this coming from? When solving single linear recurrences, we found that the basic solutions are of the form cr^n where $r \neq 0$ is a root of the characteristic polynomials. To solve $\boldsymbol{a}_{n+1} = M\boldsymbol{a}_n$, it is therefore natural to look for solutions of the form $\boldsymbol{a}_n = \boldsymbol{c}r^n$ (where $\boldsymbol{c} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$). Note that $\boldsymbol{a}_{n+1} = \boldsymbol{c}r^{n+1} = r\boldsymbol{a}_n$.

Plugging into $\boldsymbol{a}_{n+1} = M\boldsymbol{a}_n$ we find $\boldsymbol{c}r^{n+1} = M\boldsymbol{c}r^n$.

Cancelling r^n (just a nonzero number!), this simplifies to $r\boldsymbol{c} = M\boldsymbol{c}$.

In other words, $\boldsymbol{a}_n = \boldsymbol{c}r^n$ is a solution if and only if $\boldsymbol{c}$ is an r-eigenvector of M.

Not enough eigenvectors? In that case, we know what to do as well (at least in principle): instead of looking only for solutions of the type $\boldsymbol{a}_n = \boldsymbol{v}\lambda^n$, we also need to look for solutions of the type $\boldsymbol{a}_n = (\boldsymbol{v}n + \boldsymbol{w})\lambda^n$. Note that this can only happen if an eigenvalue is a repeated root of the characteristic polynomial.

Matrix solutions. A matrix Φ_n is a **matrix solution** to $\boldsymbol{a}_{n+1} = M\boldsymbol{a}_n$ if $\Phi_{n+1} = M\Phi_n$. Φ_n being a matrix solution is equivalent to each column of Φ_n being a normal (vector) solution. If the general solution of $\boldsymbol{a}_{n+1} = M\boldsymbol{a}_n$ can be obtained as the linear combination of the columns of Φ_n, then Φ_n is a **fundamental matrix solution**.

Why can we compute matrix powers this way? Recall that, given a first-order system $\boldsymbol{a}_{n+1} = M\boldsymbol{a}_n$, it is clear that the solution satisfies $\boldsymbol{a}_n = M^n\boldsymbol{a}_0$. Likewise, a fundamental matrix solution Φ_n to the same recurrence satisfies $\Phi_n = M^n\Phi_0$. Multiplying both sides by Φ_0^{-1} (on the right!) we conclude that $\Phi_n\Phi_0^{-1} = M^n$.

Already know how to compute matrix powers? If you have taken linear algebra classes, you may have learned that matrix powers M^n can be computed by diagonalizing the matrix M. The latter hinges on computing eigenvalues and eigenvectors of M as well. Compare the two approaches!

Example 60. Let $M = \begin{bmatrix} 8 & -10 \\ 5 & -7 \end{bmatrix}$.

(a) Determine the general solution to $\boldsymbol{a}_{n+1} = M\boldsymbol{a}_n$.

(b) Determine a **fundamental matrix solution** to $\boldsymbol{a}_{n+1} = M\boldsymbol{a}_n$.

(c) Compute M^n.

(d) Solve $\boldsymbol{a}_{n+1} = M\boldsymbol{a}_n$, $\boldsymbol{a}_0 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

Armin Straub
straub@southalabama.edu

Solution.

(a) Recall that each λ-eigenvector $\boldsymbol{v}$ of M provides us with a solution: $\boldsymbol{a}_n = \boldsymbol{v}\lambda^n$

We computed in Example 54 that $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ is an eigenvector for $\lambda = 3$, and $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is an eigenvector for $\lambda = -2$.

Hence, the general solution is $C_1\begin{bmatrix} 2 \\ 1 \end{bmatrix}3^n + C_2\begin{bmatrix} 1 \\ 1 \end{bmatrix}(-2)^n$.

(b) Note that we can write the general solution as

$\boldsymbol{a}_n = C_1\begin{bmatrix} 2 \\ 1 \end{bmatrix}3^n + C_2\begin{bmatrix} 1 \\ 1 \end{bmatrix}(-2)^n = \begin{bmatrix} 2\cdot 3^n & (-2)^n \\ 3^n & (-2)^n \end{bmatrix}\begin{bmatrix} C_1 \\ C_2 \end{bmatrix}$.

We call $\Phi_n = \begin{bmatrix} 2\cdot 3^n & (-2)^n \\ 3^n & (-2)^n \end{bmatrix}$ the corresponding **fundamental matrix (solution)**.

Note that our general solution is precisely $\Phi_n \boldsymbol{c}$ with $\boldsymbol{c} = \begin{bmatrix} C_1 \\ C_2 \end{bmatrix}$.

Observations.

(a) The columns of Φ_n are (independent) solutions of the system.

(b) Φ_n solves the RE itself: $\Phi_{n+1} = M\Phi_n$.
[Spell this out in this example! That Φ_n solves the RE follows from the definition of matrix multiplication.]

(c) It follows that $\Phi_n = M^n \Phi_0$. Equivalently, $\Phi_n \Phi_0^{-1} = M^n$. (See next part!)

(c) Note that $\Phi_0 = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$, so that $\Phi_0^{-1} = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}$. It follows that

$$M^n = \Phi_n \Phi_0^{-1} = \begin{bmatrix} 2\cdot 3^n & (-2)^n \\ 3^n & (-2)^n \end{bmatrix}\begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 2\cdot 3^n - (-2)^n & -2\cdot 3^n + 2(-2)^n \\ 3^n - (-2)^n & -3^n + 2(-2)^n \end{bmatrix}.$$

Check. Let us verify the formula for M^n in the cases $n = 0$ and $n = 1$:

$M^0 = \begin{bmatrix} 2-1 & -2+2 \\ 1-1 & -1+2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

$M^1 = \begin{bmatrix} 2\cdot 3 - (-2) & -2\cdot 3 + 2(-2) \\ 3 - (-2) & -3 + 2(-2) \end{bmatrix} = \begin{bmatrix} 8 & -10 \\ 5 & -7 \end{bmatrix}$

(d) $\boldsymbol{a}_n = M^n \boldsymbol{a}_0 = \begin{bmatrix} 2\cdot 3^n - (-2)^n & -2\cdot 3^n + 2(-2)^n \\ 3^n - (-2)^n & -3^n + 2(-2)^n \end{bmatrix}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -2\cdot 3^n + 3(-2)^n \\ -3^n + 3(-2)^n \end{bmatrix}$

Sage. Once we are comfortable with these computations, we can let Sage do them for us.

```
>>> M = matrix([[8,-10],[5,-7]])
```

```
>>> M^2
```

$$\begin{pmatrix} 14 & -10 \\ 5 & -1 \end{pmatrix}$$

Verify that this matrix matches what our formula for M^n produces for $n = 2$. In order to reproduce the general formula for M^n, we need to first define n as a symbolic variable:

```
>>> n = var('n')
```

```
>>> M^n
```

$$\begin{pmatrix} 2\cdot 3^n - (-2)^n & -2\cdot 3^n + 2\,(-2)^n \\ 3^n - (-2)^n & -3^n + 2\,(-2)^n \end{pmatrix}$$

Note that this indeed matches our earlier formula. Can you see how we can read off the eigenvalues and eigenvectors of M from this formula for M^n? Of course, Sage can readily compute these for us directly using, for instance, `M.eigenvectors_right()`. Try it! Can you interpret the output?

Example 61. (review) Write the (second-order) RE $a_{n+2} = a_{n+1} + 2a_n$, with $a_0 = 0$, $a_1 = 1$, as a system of (first-order) recurrences.

Solution. If $\boldsymbol{a}_n = \begin{bmatrix} a_n \\ a_{n+1} \end{bmatrix}$, then $\boldsymbol{a}_{n+1} = \begin{bmatrix} a_{n+1} \\ a_{n+2} \end{bmatrix} = \begin{bmatrix} a_{n+1} \\ a_{n+1} + 2a_n \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix} \boldsymbol{a}_n$ with $\boldsymbol{a}_0 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

Example 62. Let $M = \begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix}$.

(a) Determine the general solution to $\boldsymbol{a}_{n+1} = M\boldsymbol{a}_n$.

(b) Determine a fundamental matrix solution to $\boldsymbol{a}_{n+1} = M\boldsymbol{a}_n$.

(c) Compute M^n.

(d) Solve $\boldsymbol{a}_{n+1} = M\boldsymbol{a}_n$, $\boldsymbol{a}_0 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

Solution.

(a) Recall that each λ-eigenvector $\boldsymbol{v}$ of M provides us with a solution: namely, $\boldsymbol{a}_n = \boldsymbol{v}\lambda^n$.

The characteristic polynomial is: $\det(A - \lambda I) = \det\left(\begin{bmatrix} -\lambda & 1 \\ 2 & 1-\lambda \end{bmatrix}\right) = \lambda^2 - \lambda - 2 = (\lambda - 2)(\lambda + 1)$.

Hence, the eigenvalues are $\lambda = 2$ and $\lambda = -1$.

- $\lambda = 2$: Solving $\begin{bmatrix} -2 & 1 \\ 2 & -1 \end{bmatrix} \boldsymbol{v} = \boldsymbol{0}$, we find that $\boldsymbol{v} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ is an eigenvector for $\lambda = 2$.
- $\lambda = -1$: Solving $\begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix} \boldsymbol{v} = \boldsymbol{0}$, we find that $\boldsymbol{v} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ is an eigenvector for $\lambda = -1$.

Hence, the general solution is $C_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} 2^n + C_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix} (-1)^n$.

(b) Note that $C_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} 2^n + C_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix} (-1)^n = \begin{bmatrix} 2^n & -(-1)^n \\ 2 \cdot 2^n & (-1)^n \end{bmatrix} \begin{bmatrix} C_1 \\ C_2 \end{bmatrix}$.

Hence, a fundamental matrix solution is $\Phi_n = \begin{bmatrix} 2^n & -(-1)^n \\ 2 \cdot 2^n & (-1)^n \end{bmatrix}$.

Comment. Other choices are possible and natural. For instance, the order of the two columns is based on our choice of starting with $\lambda = 2$. Also, the columns can be scaled by any constant (for instance, using $-\boldsymbol{v}$ instead of $\boldsymbol{v}$ for $\lambda = -1$ above, we end up with the same Φ_n but with the second column scaled by -1). In general, if Φ_n is a fundamental matrix solution, then so is $\Phi_n C$ where C is an invertible 2×2 matrix.

(c) We compute $M^n = \Phi_n \Phi_0^{-1}$ using $\Phi_n = \begin{bmatrix} 2^n & -(-1)^n \\ 2 \cdot 2^n & (-1)^n \end{bmatrix}$. Since $\Phi_0^{-1} = \begin{bmatrix} 1 & -1 \\ 2 & 1 \end{bmatrix}^{-1} = \frac{1}{3} \begin{bmatrix} 1 & 1 \\ -2 & 1 \end{bmatrix}$, we have

$$M^n = \Phi_n \Phi_0^{-1} = \begin{bmatrix} 2^n & -(-1)^n \\ 2 \cdot 2^n & (-1)^n \end{bmatrix} \frac{1}{3} \begin{bmatrix} 1 & 1 \\ -2 & 1 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 2^n + 2(-1)^n & 2^n - (-1)^n \\ 2 \cdot 2^n - 2(-1)^n & 2 \cdot 2^n + (-1)^n \end{bmatrix}.$$

(d) $\boldsymbol{a}_n = M^n \boldsymbol{a}_0 = \frac{1}{3} \begin{bmatrix} 2^n + 2(-1)^n & 2^n - (-1)^n \\ 2 \cdot 2^n - 2(-1)^n & 2 \cdot 2^n + (-1)^n \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 2^n - (-1)^n \\ 2 \cdot 2^n + (-1)^n \end{bmatrix}$

Alternative solution of the first part. We saw in Example 61 that this system can be obtained from $a_{n+2} = a_{n+1} + 2a_n$ if we set $\boldsymbol{a} = \begin{bmatrix} a_n \\ a_{n+1} \end{bmatrix}$. In Example 46, we found that this RE has solutions $a_n = 2^n$ and $a_n = (-1)^n$. Correspondingly, $\boldsymbol{a}_{n+1} = \begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix} \boldsymbol{a}_n$ has solutions $\boldsymbol{a}_n = \begin{bmatrix} 2^n \\ 2^{n+1} \end{bmatrix}$ and $\boldsymbol{a}_n = \begin{bmatrix} (-1)^n \\ (-1)^{n+1} \end{bmatrix}$.

These combine to the general solution $C_1 \begin{bmatrix} 2^n \\ 2^{n+1} \end{bmatrix} + C_2 \begin{bmatrix} (-1)^n \\ (-1)^{n+1} \end{bmatrix}$ (equivalent to our solution above).

Alternative for last part. Solve the RE from Example 61 to find $a_n = \frac{1}{3}(2^n - (-1)^n)$. The above is $\boldsymbol{a}_n = \begin{bmatrix} a_n \\ a_{n+1} \end{bmatrix}$.

We have learned how to compute M^n for a matrix M using its eigenvalues and eigenvectors, as well as solve the system $\boldsymbol{a}_{n+1} = M\boldsymbol{a}_n$. For diagonal matrices, all this is much simpler:

Example 63. If $M = \begin{bmatrix} 3 & & & \\ & -2 & & \\ & & 5 & \\ & & & 1 \end{bmatrix}$, what is M^n?

Also: what is the solution to $\boldsymbol{a}_{n+1} = M\boldsymbol{a}_n$?

Comment. Entries that are not printed are meant to be zero (to make the structure of the 4×4 matrix more visibly transparent).

Solution. $M^n = \begin{bmatrix} 3^n & & & \\ & (-2)^n & & \\ & & 5^n & \\ & & & 1 \end{bmatrix}$

If this isn't clear to you, multiply out M^2. What happens?

Also: $\boldsymbol{a}_{n+1} = M\boldsymbol{a}_n$ with $\boldsymbol{a}_n = \begin{bmatrix} a_n \\ b_n \\ c_n \\ d_n \end{bmatrix}$ decouples into $\begin{array}{c} a_{n+1} = 3a_n \\ b_{n+1} = -2b_n \\ c_{n+1} = 5c_n \\ d_{n+1} = d_n \end{array}$ which is solved by $\boldsymbol{a}_n = \begin{bmatrix} a_n \\ b_n \\ c_n \\ d_n \end{bmatrix} = \begin{bmatrix} 3^n a_0 \\ (-2)^n b_0 \\ 5^n c_0 \\ d_0 \end{bmatrix}$.

Example 64. (extra practice)

(a) Write the recurrence $a_{n+3} - 4a_{n+2} + a_{n+1} + 6a_n = 0$ as a system $\boldsymbol{a}_{n+1} = M\boldsymbol{a}_n$ of (first-order) recurrences.

(b) Determine a fundamental matrix solution to $\boldsymbol{a}_{n+1} = M\boldsymbol{a}_n$.

(c) Compute M^n.

Solution.

(a) If $\boldsymbol{a}_n = \begin{bmatrix} a_n \\ a_{n+1} \\ a_{n+2} \end{bmatrix}$, then the RE becomes $\boldsymbol{a}_{n+1} = M\boldsymbol{a}_n$ with $M = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -6 & -1 & 4 \end{bmatrix}$.

(b) Because we started with a single (third-order) equation, we can avoid computing eigenvectors and eigenvalues (indeed, we will find these as a byproduct).

By factoring the characteristic equation $N^3 - 4N^2 + N + 6 = (N-3)(N-2)(N+1)$, we find that the characteristic roots are $3, 2, -1$ (these are also precisely the eigenvalues of M).

Hence, $a_n = C_1 \cdot 3^n + C_2 \cdot 2^n + C_3 \cdot (-1)^n$ is the general solution to the initial RE.

Correspondingly, a fundamental matrix solution of the system is $\Phi_n = \begin{bmatrix} 3^n & 2^n & (-1)^n \\ 3 \cdot 3^n & 2 \cdot 2^n & -(-1)^n \\ 9 \cdot 3^n & 4 \cdot 2^n & (-1)^n \end{bmatrix}$.

Note. This tells us that $\begin{bmatrix} 1 \\ 3 \\ 9 \end{bmatrix}$ is a 3-eigenvector, $\begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}$ a 2-eigenvector, and $\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$ a -1-eigenvector of M.

(c) Since $\Phi_{n+1} = M\Phi_n$, we have $\Phi_n = M^n \Phi_0$ so that $M^n = \Phi_n \Phi_0^{-1}$. This allows us to compute that:

$$M^n = \frac{1}{12} \begin{bmatrix} -6 \cdot 3^n + 12 \cdot 2^n + 6(-1)^n & -3 \cdot 3^n + 8 \cdot 2^n - 5(-1)^n & 3 \cdot 3^n - 4 \cdot 2^n + (-1)^n \\ -18 \cdot 3^n + 24 \cdot 2^n - 6(-1)^n & \ldots & \ldots \\ -54 \cdot 3^n + 48 \cdot 2^n + 6(-1)^n & \ldots & \ldots \end{bmatrix}$$

Systems of differential equations

Review. Check out Examples 61 and 62 again. Below we will repeat the same steps, replacing recurrences with differential equations as well as λ^n with $e^{\lambda x}$.

Example 65. Write the (second-order) initial value problem $y'' = y' + 2y$, $y(0) = 0$, $y'(0) = 1$ as a first-order system.

Solution. If $\boldsymbol{y} = \begin{bmatrix} y \\ y' \end{bmatrix}$, then $\boldsymbol{y}' = \begin{bmatrix} y' \\ y'' \end{bmatrix} = \begin{bmatrix} y' \\ y' + 2y \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} y \\ y' \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix}\boldsymbol{y}$ with $\boldsymbol{y}(0) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

This is exactly how we proceeded in Example 61.

Homework. Solve this IVP to find $y(x) = \frac{1}{3}(e^{2x} - e^{-x})$. Then compare with the next example.

Example 66. (preview) Let $M = \begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix}$.

(a) Determine the general solution to $\boldsymbol{y}' = M\boldsymbol{y}$.

(b) Determine a fundamental matrix solution to $\boldsymbol{y}' = M\boldsymbol{y}$.

(c) Solve $\boldsymbol{y}' = M\boldsymbol{y}$, $\boldsymbol{y}(0) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

Solution. In Example 62, we only need to replace 2^n by e^{2x} (root 2) and $(-1)^n$ by e^{-x} (root -1)!

(a) The general solution is $C_1\begin{bmatrix} 1 \\ 2 \end{bmatrix}e^{2x} + C_2\begin{bmatrix} -1 \\ 1 \end{bmatrix}e^{-x}$.

(b) A fundamental matrix solution is $\Phi(x) = \begin{bmatrix} e^{2x} & -e^{-x} \\ 2 \cdot e^{2x} & e^{-x} \end{bmatrix}$.

(c) $\boldsymbol{y}(x) = \frac{1}{3}\begin{bmatrix} e^{2x} - e^{-x} \\ 2 \cdot e^{2x} + e^{-x} \end{bmatrix}$

Preview. The special fundamental matrix M^n will be replaced by e^{Mx}, the **matrix exponential**.

Example 67. Write the (third-order) differential equation $y''' = 3y'' - 2y' + y$ as a system of (first-order) differential equations.

Solution. If $\boldsymbol{y} = \begin{bmatrix} y \\ y' \\ y'' \end{bmatrix}$, then $\boldsymbol{y}' = \begin{bmatrix} y' \\ y'' \\ y''' \end{bmatrix} = \begin{bmatrix} y' \\ y'' \\ 3y'' - 2y' + y \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -2 & 3 \end{bmatrix}\begin{bmatrix} y \\ y' \\ y'' \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -2 & 3 \end{bmatrix}\boldsymbol{y}$.

For short, $\boldsymbol{y}' = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -2 & 3 \end{bmatrix}\boldsymbol{y}$

Comment. This is one reason why we care about systems of DEs, even if we work with just one function.

Example 68. Consider the following system of (second-order) initial value problems:

$$\begin{aligned} y_1'' &= 2y_1' - 3y_2' + 7y_2 \\ y_2'' &= 4y_1' + y_2' - 5y_1 \end{aligned} \qquad y_1(0) = 2,\ y_1'(0) = 3,\ y_2(0) = -1,\ y_2'(0) = 1$$

Write it as a first-order initial value problem in the form $\boldsymbol{y}' = M\boldsymbol{y}$, $\boldsymbol{y}(0) = \boldsymbol{y}_0$.

Armin Straub
straub@southalabama.edu

Solution. If $\boldsymbol{y}=\begin{bmatrix} y_1 \\ y_2 \\ y_1' \\ y_2' \end{bmatrix}$, then $\boldsymbol{y}'=\begin{bmatrix} y_1' \\ y_2' \\ y_1'' \\ y_2'' \end{bmatrix}=\begin{bmatrix} y_1' \\ y_2' \\ 2y_1'-3y_2'+7y_2 \\ 4y_1'+y_2'-5y_1 \end{bmatrix}=\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 7 & 2 & -3 \\ -5 & 0 & 4 & 1 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \\ y_1' \\ y_2' \end{bmatrix}=\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 7 & 2 & -3 \\ -5 & 0 & 4 & 1 \end{bmatrix}\boldsymbol{y}$.

For short, the system translates into $\boldsymbol{y}'=\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 7 & 2 & -3 \\ -5 & 0 & 4 & 1 \end{bmatrix}\boldsymbol{y}$ with $\boldsymbol{y}(0)=\begin{bmatrix} 2 \\ -1 \\ 3 \\ 1 \end{bmatrix}$.

Solving systems of differential equations

We can solve the system $\boldsymbol{y}'=M\boldsymbol{y}$ exactly as we solved $\boldsymbol{a}_{n+1}=M\boldsymbol{a}_n$.

The only difference is that we replace each λ^n (for characteristic root / eigenvalue λ) with $e^{\lambda x}$. In fact, as shown in the examples below, we can translate back and forth at any stage.

(solving systems of DEs) To solve $\boldsymbol{y}'=M\boldsymbol{y}$, determine the eigenvectors of M.

- Each λ-eigenvector $\boldsymbol{v}$ provides a solution: $\boldsymbol{y}(x)=\boldsymbol{v}e^{\lambda x}$
- If there are enough eigenvectors, these combine to the general solution.
  In that case, we get a **fundamental matrix (solution)** $\Phi(x)$ by placing each solution vector into one column of $\Phi(x)$.
- If desired, we can find the **matrix exponential** e^{Mx} using any fundamental matrix $\Phi(x)$:

$$e^{Mx}=\Phi(x)\Phi(0)^{-1}.$$

Note that e^{Mx} is the unique matrix solution to $\boldsymbol{y}'=M\boldsymbol{y}$, $\boldsymbol{y}(0)=I$ (the identity matrix).

Application: the unique solution to $\boldsymbol{y}'=M\boldsymbol{y}$, $\boldsymbol{y}(0)=\boldsymbol{c}$ is given by $\boldsymbol{y}(x)=e^{Mx}\boldsymbol{c}$.

Note. Unlike with M^n, it might not be clear what the **matrix exponential** e^{Mx} really is. One way to think about it is that we are defining e^{Mx} as the solution to the IVP $\boldsymbol{y}'=M\boldsymbol{y}$, $\boldsymbol{y}(0)=I$. This is equivalent to how one can define the ordinary exponential e^x as the solution to $y'=y$, $y(0)=1$.

[In a little bit, we will also discuss how to think about the matrix exponential e^{Mx} using power series.]

Comment. If there are not enough eigenvectors, then we know what to do (at least in principle): instead of looking only for solutions of the type $\boldsymbol{y}(x)=\boldsymbol{v}e^{\lambda x}$, we also need to look for solutions of the type $\boldsymbol{y}(x)=(\boldsymbol{v}x+\boldsymbol{w})e^{\lambda x}$. Note that this can only happen if an eigenvalue is a repeated root of the characteristic polynomial.

Why does this work? Compare this to our method of solving systems of REs and for computing matrix powers M^n. The above conclusion about systems of DEs can be deduced along the same lines as what we did for REs:

- For instance, for the first part, let us look for solutions of $\boldsymbol{y}'=M\boldsymbol{y}$ of the form $\boldsymbol{y}(x)=\boldsymbol{v}e^{\lambda x}$.
  Note that $\boldsymbol{y}'=\lambda\boldsymbol{v}e^{\lambda x}=\lambda\boldsymbol{y}$. Plugging into $\boldsymbol{y}'=M\boldsymbol{y}$, we find $\lambda\boldsymbol{y}=M\boldsymbol{y}$.
  In other words, $\boldsymbol{y}(x)=\boldsymbol{v}e^{\lambda x}$ is a solution if and only if $\boldsymbol{v}$ is a λ-eigenvector of M.
- If $\Phi(x)$ is a fundamental matrix solution, then so is $\Psi(x)=\Phi(x)C$ for every constant matrix C. (Why?!)
  Therefore, $\Psi(x)=\Phi(x)\Phi(0)^{-1}$ is a fundamental matrix solution with $\Psi(0)=\Phi(0)\Phi(0)^{-1}=I$.
  But e^{Mx} is defined to be the unique such solution, so that $\Psi(x)=e^{Mx}$.

Example 69. (homework) Let $M = \begin{bmatrix} -1 & 6 \\ -1 & 4 \end{bmatrix}$.

(a) Determine the general solution to $\boldsymbol{y}' = M\boldsymbol{y}$.

(b) Determine a fundamental matrix solution to $\boldsymbol{y}' = M\boldsymbol{y}$.

(c) Compute e^{Mx}.

(d) Solve the initial value problem $\boldsymbol{y}' = M\boldsymbol{y}$ with $\boldsymbol{y}(0) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

(e) Compute M^n.

(f) Solve $\boldsymbol{a}_{n+1} = M\boldsymbol{a}_n$ with $\boldsymbol{a}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

Solution.

(a) We determine the eigenvectors of M. The characteristic polynomial is:
$\det(M - \lambda I) = \det\left(\begin{bmatrix} -1-\lambda & 6 \\ -1 & 4-\lambda \end{bmatrix}\right) = (-1-\lambda)(4-\lambda) + 6 = \lambda^2 - 3\lambda + 2 = (\lambda - 1)(\lambda - 2)$
Hence, the eigenvalues are $\lambda = 1$ and $\lambda = 2$.

- $\lambda = 1$: Solving $\begin{bmatrix} -2 & 6 \\ -1 & 3 \end{bmatrix}\boldsymbol{v} = \boldsymbol{0}$, we find that $\boldsymbol{v} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ is an eigenvector for $\lambda = 1$.
- $\lambda = 2$: Solving $\begin{bmatrix} -3 & 6 \\ -1 & 2 \end{bmatrix}\boldsymbol{v} = \boldsymbol{0}$, we find that $\boldsymbol{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ is an eigenvector for $\lambda = 2$.

Hence, the general solution is $C_1\begin{bmatrix} 3 \\ 1 \end{bmatrix}e^x + C_2\begin{bmatrix} 2 \\ 1 \end{bmatrix}e^{2x}$.

(b) The corresponding fundamental matrix solution is $\Phi = \begin{bmatrix} 3e^x & 2e^{2x} \\ e^x & e^{2x} \end{bmatrix}$.

(c) Note that $\Phi(0) = \begin{bmatrix} 3 & 2 \\ 1 & 1 \end{bmatrix}$, so that $\Phi(0)^{-1} = \begin{bmatrix} 1 & -2 \\ -1 & 3 \end{bmatrix}$. It follows that

$$e^{Mx} = \Phi(x)\Phi(0)^{-1} = \begin{bmatrix} 3e^x & 2e^{2x} \\ e^x & e^{2x} \end{bmatrix}\begin{bmatrix} 1 & -2 \\ -1 & 3 \end{bmatrix} = \begin{bmatrix} 3e^x - 2e^{2x} & -6e^x + 6e^{2x} \\ e^x - e^{2x} & -2e^x + 3e^{2x} \end{bmatrix}.$$

(d) The solution to the IVP is $\boldsymbol{y}(x) = e^{Mx}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3e^x - 2e^{2x} & -6e^x + 6e^{2x} \\ e^x - e^{2x} & -2e^x + 3e^{2x} \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -3e^x + 4e^{2x} \\ -e^x + 2e^{2x} \end{bmatrix}$.

Note. If we hadn't already computed e^{Mx}, we would use the general solution and solve for the appropriate values of C_1 and C_2. Do it that way as well!

(e) From the first part, it follows that $\boldsymbol{a}_{n+1} = M\boldsymbol{a}_n$ has general solution $C_1\begin{bmatrix} 3 \\ 1 \end{bmatrix} + C_2\begin{bmatrix} 2 \\ 1 \end{bmatrix}2^n$.
(Note that $1^n = 1$.)
The corresponding fundamental matrix solution is $\Phi_n = \begin{bmatrix} 3 & 2\cdot 2^n \\ 1 & 2^n \end{bmatrix}$.
As above, $\Phi_0 = \begin{bmatrix} 3 & 2 \\ 1 & 1 \end{bmatrix}$, so that $\Phi(0)^{-1} = \begin{bmatrix} 1 & -2 \\ -1 & 3 \end{bmatrix}$ and

$$M^n = \Phi_n\Phi_0^{-1} = \begin{bmatrix} 3 & 2\cdot 2^n \\ 1 & 2^n \end{bmatrix}\begin{bmatrix} 1 & -2 \\ -1 & 3 \end{bmatrix} = \begin{bmatrix} 3 - 2\cdot 2^n & -6 + 6\cdot 2^n \\ 1 - 2^n & -2 + 3\cdot 2^n \end{bmatrix}.$$

Important. Compare with our computation for e^{Mx}. Can you see how this was basically the same computation? Write down M^n directly from e^{Mx}.

(f) The (unique) solution is $\boldsymbol{a}_n = M^n\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 - 2\cdot 2^n & -6 + 6\cdot 2^n \\ 1 - 2^n & -2 + 3\cdot 2^n \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -3 + 4\cdot 2^n \\ -1 + 2\cdot 2^n \end{bmatrix}$.

Important. Again, compare with the earlier IVP! Without work, we can write down one from the other.

We purposefully omit details of some computations in the next example to highlight how it proceeds along the same lines as Example 60.

Important. In fact, we can translate back and forth (without additional computations) by simply replacing 3^n and $(-2)^n$ by e^{3x} and e^{-2x}.

Example 70. (extra practice) Let $M = \begin{bmatrix} 8 & -10 \\ 5 & -7 \end{bmatrix}$.

(a) Determine the general solution to $\boldsymbol{y}' = M\boldsymbol{y}$.

(b) Determine a fundamental matrix solution to $\boldsymbol{y}' = M\boldsymbol{y}$.

(c) Compute e^{Mx}.

(d) Solve the initial value problem $\boldsymbol{y}' = M\boldsymbol{y}$ with $\boldsymbol{y}(0) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

Solution. (See Example 60 for more details on the analogous computations.)

(a) Recall that each λ-eigenvector $\boldsymbol{v}$ of M provides us with a solution: namely, $\boldsymbol{y}(x) = \boldsymbol{v}e^{\lambda x}$.
We computed earlier that $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ is an eigenvector for $\lambda = 3$, and $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is an eigenvector for $\lambda = -2$.
Hence, the general solution is $C_1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} e^{3x} + C_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} e^{-2x}$.

(b) The corresponding fundamental matrix solution is $\Phi(x) = \begin{bmatrix} 2 \cdot e^{3x} & e^{-2x} \\ e^{3x} & e^{-2x} \end{bmatrix}$.
[Note that our general solution is precisely $\Phi(x) \begin{bmatrix} C_1 \\ C_2 \end{bmatrix}$.]

(c) Since $\Phi(0) = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$, we have $\Phi(0)^{-1} = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}$. It follows that

$$e^{Mx} = \Phi(x)\Phi(0)^{-1} = \begin{bmatrix} 2 \cdot e^{3x} & e^{-2x} \\ e^{3x} & e^{-2x} \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 2 \cdot e^{3x} - e^{-2x} & -2 \cdot e^{3x} + 2e^{-2x} \\ e^{3x} - e^{-2x} & -e^{3x} + 2e^{-2x} \end{bmatrix}.$$

Check. Let us verify the formula for e^{Mx} in the simple case $x = 0$: $e^{M0} = \begin{bmatrix} 2-1 & -2+2 \\ 1-1 & -1+2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

(d) The solution to the IVP is $\boldsymbol{y}(x) = e^{Mx} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \cdot e^{3x} + 2e^{-2x} \\ -e^{3x} + 2e^{-2x} \end{bmatrix}$ (the second column of e^{Mx}).

Sage. We can compute the matrix exponential in Sage as follows:

```
>>> M = matrix([[8,-10],[5,-7]])
>>> exp(M*x)
```

$$\begin{pmatrix} \left(2\, e^{(5\,x)} - 1\right) e^{(-2\,x)} & -2\,\left(e^{(5\,x)} - 1\right) e^{(-2\,x)} \\ \left(e^{(5\,x)} - 1\right) e^{(-2\,x)} & -\left(e^{(5\,x)} - 2\right) e^{(-2\,x)} \end{pmatrix}$$

Note that this indeed matches the result of our computation.

[By the way, the variable x is pre-defined as a symbolic variable in Sage. That's why, unlike for n in the computation of M^n, we did not need to use `x = var('x')` first.]

Example 71. Suppose that $e^{Mx} = \frac{1}{10}\begin{bmatrix} e^x + 9e^{2x} & 3e^x - 3e^{2x} \\ 3e^x - 3e^{2x} & 9e^x + e^{2x} \end{bmatrix}$.

(a) Without doing any computations, determine M^n.

(b) What is M?

(c) Without doing any computations, determine the eigenvalues and eigenvectors of M.

(d) From those, write down a simple fundamental matrix solution to $\boldsymbol{y}' = M\boldsymbol{y}$.

(e) From that fundamental matrix solution, how can we compute e^{Mx}? (If we didn't know it already...)

(f) Having computed e^{Mx}, what is a simple check that we can (should!) make?

Solution.

(a) Since e^x and e^{2x} correspond to eigenvalues 1 and 2, we just need to replace these by $1^n = 1$ and 2^n:

$$M^n = \frac{1}{10}\begin{bmatrix} 1 + 9 \cdot 2^n & 3 - 3 \cdot 2^n \\ 3 - 3 \cdot 2^n & 9 + 2^n \end{bmatrix}$$

(b) We can simply set $n = 1$ in our formula for M^n, to get $M = \frac{1}{10}\begin{bmatrix} 19 & -3 \\ -3 & 11 \end{bmatrix}$.

(c) The eigenvalues are 1 and 2 (because e^{Mx} contains the exponentials e^x and e^{2x}).

Looking at the coefficients of e^x in the first column of e^{Mx}, we see that $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$ is a 1-eigenvector.

[We can also look the second column of e^{Mx}, to obtain $\begin{bmatrix} 3 \\ 9 \end{bmatrix}$ which is a multiple and thus equivalent.]

Likewise, by looking at the coefficients of e^{2x}, we see that $\begin{bmatrix} 9 \\ -3 \end{bmatrix}$ or, equivalently, $\begin{bmatrix} -3 \\ 1 \end{bmatrix}$ is a 2-eigenvector.

Comment. To see where this is coming from, keep in mind that, associated to a λ-eigenvector $\boldsymbol{v}$, we have the corresponding solution $\boldsymbol{y}(x) = \boldsymbol{v}e^{\lambda x}$ of the DE $\boldsymbol{y}' = M\boldsymbol{y}$. On the other hand, the columns of e^{Mx} are solutions to that DE and, therefore, must be linear combinations of these $\boldsymbol{v}e^{\lambda x}$.

(d) From the eigenvalues and eigenvectors, we know that $\begin{bmatrix} 1 \\ 3 \end{bmatrix}e^x$ and $\begin{bmatrix} -3 \\ 1 \end{bmatrix}e^{2x}$ are solutions (and that the general solutions consists of the linear combinations of these two).

Selecting these as the columns, we obtain the fundamental matrix solution $\Phi(x) = \begin{bmatrix} e^x & -3e^{2x} \\ 3e^x & e^{2x} \end{bmatrix}$.

Comment. The *fundamental* refers to the fact that the columns combine to the general solution. The *matrix solution* means that $\Phi(x)$ itself satisfies the DE: namely, we have $\Phi' = M\Phi$. That this is the case is a consequence of matrix multiplication (namely, say, the second column of $M\Phi$ is defined to be M times the second column of Φ; but that column is a vector solution and therefore solves the DE).

(e) We can compute e^{Mx} as $e^{Mx} = \Phi(x)\Phi(0)^{-1}$.

If $\Phi(x) = \begin{bmatrix} e^x & -3e^{2x} \\ 3e^x & e^{2x} \end{bmatrix}$, then $\Phi(0) = \begin{bmatrix} 1 & -3 \\ 3 & 1 \end{bmatrix}$ and, hence, $\Phi(0)^{-1} = \frac{1}{10}\begin{bmatrix} 1 & 3 \\ -3 & 1 \end{bmatrix}$. It follows that

$$e^{Mx} = \Phi(x)\Phi(0)^{-1} = \begin{bmatrix} e^x & -3e^{2x} \\ 3e^x & e^{2x} \end{bmatrix}\frac{1}{10}\begin{bmatrix} 1 & 3 \\ -3 & 1 \end{bmatrix} = \frac{1}{10}\begin{bmatrix} e^x + 9e^{2x} & 3e^x - 3e^{2x} \\ 3e^x - 3e^{2x} & 9e^x + e^{2x} \end{bmatrix}.$$

(f) We can check that e^{Mx} equals the identity matrix if we set $x = 0$:

$$\frac{1}{10}\begin{bmatrix} e^x + 9e^{2x} & 3e^x - 3e^{2x} \\ 3e^x - 3e^{2x} & 9e^x + e^{2x} \end{bmatrix} \overset{x=0}{\rightsquigarrow} \frac{1}{10}\begin{bmatrix} 1+9 & 3-3 \\ 3-3 & 9+1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

This check does not require much effort and can even be done in our head while writing down e^{Mx}. There is really no excuse for not doing it!

Armin Straub
straub@southalabama.edu

Another perspective on the matrix exponential

Review. We achieved the milestone to introduce a **matrix exponential** in such a way that we can treat a system of DEs, say $\boldsymbol{y}' = M\boldsymbol{y}$ with $\boldsymbol{y}(0) = \boldsymbol{c}$, just as if the matrix M was a number: namely, the unique solution is simply $\boldsymbol{y} = e^{Mx}\boldsymbol{c}$.

The price to pay is that the matrix e^{Mx} requires some work to actually compute (and proceeds by first determining a different matrix solution $\Phi(x)$ using eigenvectors and eigenvalues). We offer below another way to think about e^{Mx} (using Taylor series).

(exponential function) e^x is the unique solution to $y' = y$, $y(0) = 1$.

From here, it follows that $e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots$.

The latter is the Taylor series for e^x at $x = 0$ that we have seen in Calculus II.

Important note. We can actually construct this infinite sum directly from $y' = y$ and $y(0) = 1$.

Indeed, observe how each term, when differentiated, produces the term before it. For instance, $\frac{\mathrm{d}}{\mathrm{d}x}\frac{x^3}{3!} = \frac{x^2}{2!}$.

Review. We defined the **matrix exponential** e^{Mx} as the unique matrix solution to the IVP

$$\boldsymbol{y}' = M\boldsymbol{y}, \quad \boldsymbol{y}(0) = I.$$

We next observe that we can also make sense of the matrix exponential e^{Mx} as a power series.

Theorem 72. Let M be $n \times n$. Then the **matrix exponential** satisfies

$$e^M = I + M + \frac{1}{2!}M^2 + \frac{1}{3!}M^3 + \ldots$$

Proof. Define $\Phi(x) = I + Mx + \frac{1}{2!}M^2x^2 + \frac{1}{3!}M^3x^3 + \ldots$

$$\begin{aligned}\Phi'(x) &= \frac{\mathrm{d}}{\mathrm{d}x}\left[I + Mx + \frac{1}{2!}M^2x^2 + \frac{1}{3!}M^3x^3 + \ldots\right] \\ &= 0 + M + M^2x + \frac{1}{2!}M^3x^2 + \ldots = M\Phi(x).\end{aligned}$$

Clearly, $\Phi(0) = I$. Therefore, $\Phi(x)$ is the fundamental matrix solution to $\boldsymbol{y}' = M\boldsymbol{y}$, $\boldsymbol{y}(0) = I$.

But that's precisely how we defined e^{Mx} earlier. It follows that $\Phi(x) = e^{Mx}$. Now set $x = 1$. □

Example 73. If $A = \begin{bmatrix} 2 & 0 \\ 0 & 5 \end{bmatrix}$, then $A^{100} = \begin{bmatrix} 2^{100} & 0 \\ 0 & 5^{100} \end{bmatrix}$.

Example 74. If $A = \begin{bmatrix} 2 & 0 \\ 0 & 5 \end{bmatrix}$, then $e^A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 2 & 0 \\ 0 & 5 \end{bmatrix} + \frac{1}{2!}\begin{bmatrix} 2^2 & 0 \\ 0 & 5^2 \end{bmatrix} + \cdots = \begin{bmatrix} e^2 & 0 \\ 0 & e^5 \end{bmatrix}$.

Clearly, this works to obtain e^D for every diagonal matrix D.

In particular, for $Ax = \begin{bmatrix} 2x & 0 \\ 0 & 5x \end{bmatrix}$, $e^{Ax} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 2x & 0 \\ 0 & 5x \end{bmatrix} + \frac{1}{2!}\begin{bmatrix} (2x)^2 & 0 \\ 0 & (5x)^2 \end{bmatrix} + \cdots = \begin{bmatrix} e^{2x} & 0 \\ 0 & e^{5x} \end{bmatrix}$.

The following is a preview of how the matrix exponential deals with repeated characteristic roots.

Example 75. Determine e^{Ax} for $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$.

Solution. If we compute eigenvalues, we find that we get $\lambda = 0, 0$ (multiplicity 2) but there is only one 0-eigenvector (up to multiples). This means we are stuck with this approach—however, see next extra section how we could still proceed.

The key here is to observe that $A^2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$. It follows that $e^{Ax} = I + Ax = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & x \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & x \\ 0 & 1 \end{bmatrix}$.

Extra: The case of repeated eigenvalues with too few eigenvectors

Review. To construct a fundamental matrix solution $\Phi(x)$ to $\boldsymbol{y}' = M\boldsymbol{y}$, we compute eigenvectors: Given a λ-eigenvector $\boldsymbol{v}$, we have the corresponding solution $\boldsymbol{y}(x) = \boldsymbol{v}e^{\lambda x}$.

If there are enough eigenvectors, we can collect these as columns to obtain $\Phi(x)$.

The next example illustrates how to proceed if there are not enough eigenvectors.

In that case, instead of looking only for solutions of the type $\boldsymbol{y}(x) = \boldsymbol{v}e^{\lambda x}$, we also need to look for solutions of the type $\boldsymbol{y}(x) = (\boldsymbol{v}x + \boldsymbol{w})e^{\lambda x}$. This can only happen if an eigenvalue is a repeated root of the characteristic polynomial.

Example 76. Let $M = \begin{bmatrix} 8 & 4 \\ -1 & 4 \end{bmatrix}$.

(a) Determine the general solution to $\boldsymbol{y}' = M\boldsymbol{y}$.

(b) Determine a fundamental matrix solution to $\boldsymbol{y}' = M\boldsymbol{y}$.

(c) Compute e^{Mx}.

(d) Solve the initial value problem $\boldsymbol{y}' = M\boldsymbol{y}$ with $\boldsymbol{y}(0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

Solution.

(a) We determine the eigenvectors of M. The characteristic polynomial is:

$\det(M - \lambda I) = \det\left(\begin{bmatrix} 8-\lambda & 4 \\ -1 & 4-\lambda \end{bmatrix}\right) = (8-\lambda)(4-\lambda) + 4 = \lambda^2 - 12\lambda + 36 = (\lambda - 6)(\lambda - 6)$

Hence, the eigenvalues are $\lambda = 6, 6$ (meaning that 6 has multiplicity 2).

- To find eigenvectors $\boldsymbol{v}$ for $\lambda = 6$, we need to solve $\begin{bmatrix} 2 & 4 \\ -1 & -2 \end{bmatrix}\boldsymbol{v} = \boldsymbol{0}$.
  Hence, $\boldsymbol{v} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$ is an eigenvector for $\lambda = 6$. There is no independent second eigenvector.
- We therefore search for a solution of the form $\boldsymbol{y}(x) = (\boldsymbol{v}x + \boldsymbol{w})e^{\lambda x}$ with $\lambda = 6$.
  $\boldsymbol{y}'(x) = (\lambda\boldsymbol{v}x + \lambda\boldsymbol{w} + \boldsymbol{v})e^{\lambda x} \overset{!}{=} M\boldsymbol{y} = (M\boldsymbol{v}x + M\boldsymbol{w})e^{\lambda x}$
  Equating coefficients of x, we need $\lambda\boldsymbol{v} = M\boldsymbol{v}$ and $\lambda\boldsymbol{w} + \boldsymbol{v} = M\boldsymbol{w}$.
  Hence, $\boldsymbol{v}$ must be an eigenvector (which we already computed); we choose $\boldsymbol{v} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$.
  [Note that any multiple of $\boldsymbol{y}(x)$ will be another solution, so it doesn't matter which multiple of $\begin{bmatrix} -2 \\ 1 \end{bmatrix}$ we choose.]
  $\lambda\boldsymbol{w} + \boldsymbol{v} = M\boldsymbol{w}$ or $(M - \lambda)\boldsymbol{w} = \boldsymbol{v}$ then becomes $\begin{bmatrix} 2 & 4 \\ -1 & -2 \end{bmatrix}\boldsymbol{w} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$.
  One solution is $\boldsymbol{w} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$. [We only need one.]

Hence, the general solution is $C_1\begin{bmatrix} -2 \\ 1 \end{bmatrix}e^{6x} + C_2\left(\begin{bmatrix} -2 \\ 1 \end{bmatrix}x + \begin{bmatrix} -1 \\ 0 \end{bmatrix}\right)e^{6x}$.

(b) The corresponding fundamental matrix solution is $\Phi = \begin{bmatrix} -2e^{6x} & -(2x+1)e^{6x} \\ e^{6x} & xe^{6x} \end{bmatrix}$.

(c) Note that $\Phi(0) = \begin{bmatrix} -2 & -1 \\ 1 & 0 \end{bmatrix}$, so that $\Phi(0)^{-1} = \begin{bmatrix} 0 & 1 \\ -1 & -2 \end{bmatrix}$. It follows that

$$e^{Mx} = \Phi(x)\Phi(0)^{-1} = \begin{bmatrix} -2e^{6x} & -(2x+1)e^{6x} \\ e^{6x} & xe^{6x} \end{bmatrix}\begin{bmatrix} 0 & 1 \\ -1 & -2 \end{bmatrix} = \begin{bmatrix} (2x+1)e^{6x} & 4xe^{6x} \\ -xe^{6x} & -(2x-1)\,e^{6x} \end{bmatrix}.$$

(d) The solution to the IVP is $\boldsymbol{y}(x) = e^{Mx}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} (2x+1)e^{6x} & 4xe^{6x} \\ -xe^{6x} & -(2x-1)\,e^{6x} \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} (2x+1)e^{6x} \\ -xe^{6x} \end{bmatrix}$.

Armin Straub
straub@southalabama.edu

Phase portraits and phase plane analysis

Our goal is to visualize the solutions to systems of equations. This works particularly well in the case of systems of two differential equations. A system that can be written as

$$\begin{aligned}\frac{\mathrm{d}x}{\mathrm{d}t} &= f(x,y)\\ \frac{\mathrm{d}y}{\mathrm{d}t} &= g(x,y)\end{aligned}$$

is called **autonomous** because it doesn't depend on the independent variable t.

Comment. Can you show that if $x(t)$ and $y(t)$ are a pair of solutions, then so is the pair $x(t+t_0)$ and $y(t+t_0)$?

We can visualize solutions to such a system by plotting the points $(x(t), y(t))$ for increasing values of t so that we get a curve (and we can attach an arrow to indicate the direction we're flowing along that curve). Each such curve is called the **trajectory** of a solution.

Even better, we can do such a **phase portrait** without solving to get a formula for $(x(t), y(t))$! That's because we can combine the two equations to get $\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{g(x,y)}{f(x,y)}$, which allows us to make a slope field! If a trajectory passes through a point (x,y), then we know that the slope at that point must be $\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{g(x,y)}{f(x,y)}$.

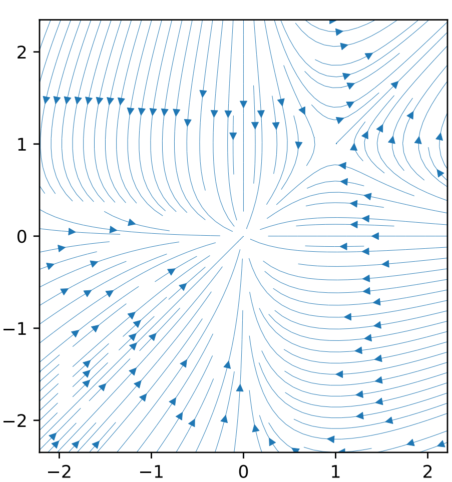
This allows us to sketch trajectories. However, it does not tell us everything about the corresponding solution $(x(t), y(t))$ because we don't know at which times t the solution passes through the points on the curve.

However, we can visualize the speed with which a solution passes through the trajectory by attaching to a point (x,y) not only the slope $\frac{g(x,y)}{f(x,y)}$ but the vector $\begin{bmatrix} f(x,y)\\ g(x,y)\end{bmatrix}$. That vector has the same direction as the slope but it also tells us in which direction we are moving and how fast (by its magnitude).

Example 77. Sketch some trajectories for the system $\frac{\mathrm{d}x}{\mathrm{d}t} = x\cdot(y-1)$, $\frac{\mathrm{d}y}{\mathrm{d}t} = y\cdot(x-1)$.

Solution. Let's look at the point $(x,y) = (2,-1)$, for instance. Then the DEs tell us that $\frac{\mathrm{d}x}{\mathrm{d}t} = x\cdot(y-1) = -4$ and $\frac{\mathrm{d}y}{\mathrm{d}t} = y\cdot(x-1) = -1$. We therefore attach the vector $\left(\frac{\mathrm{d}x}{\mathrm{d}t}, \frac{\mathrm{d}y}{\mathrm{d}t}\right) = (-4,-1)$ to $(x,y) = (2,-1)$.

Note that if we use $\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{y\cdot(x-1)}{x\cdot(y-1)}$ directly, we find the slope $\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{-1}{-4} = \frac{1}{4}$. This is slightly less information because it doesn't tell us that we are moving "left and down" as the arrows in the following plot indicate:



Armin Straub
straub@southalabama.edu

Comment. In this example, we can solve the slope-field equation $\frac{dy}{dx} = \frac{y(x-1)}{x(y-1)}$ using separation of variables.

Do it! We end up with the implicit solutions $y - \ln|y| = x - \ln|x| + C$.

If we plot these curves for various values of C, we get trajectories in the plot above. However, note that none of this solving is needed for plotting by itself.

Sage. We can make Sage create such phase portraits for us!

```
>>> x,y = var('x y')
```

```
>>> streamline_plot((x*(y-1),y*(x-1)), (x,-3,3), (y,-3,3))
```

Equilibrium solutions

(x_0, y_0) is an **equilibrium point** of the system $\frac{dx}{dt} = f(x,y)$, $\frac{dy}{dt} = g(x,y)$ if

$$f(x_0, y_0) = 0 \quad \text{and} \quad g(x_0, y_0) = 0.$$

In that case, we have the **constant (equilibrium) solution** $x(t) = x_0$, $y(t) = y_0$.

Comment. Equilibrium points are also called **critical points** (or stationary points or rest points).

In a phase portrait, the equilibrium solutions are just a single point.

Recall that every other solution $(x(t), y(t))$ corresponds to a curve (parametrized by t), called the **trajectory** of the solution (and we can adorn it with an arrow that indicates the direction of the "flow" of the solution).

We can learn a lot from how solutions behave near equilibrium points.

An equilibrium point is called:

- **stable** if all nearby solutions remain close to the equilibrium point;
- **asymptotically stable** if all nearby solutions remain close and "flow into" the equilibrium;
- **unstable** if it is not stable (some nearby solutions "flow away" from the equilibrium).

Comment. Note that asymptotically stable is a stronger condition than stable. A typical example of a stable, but not asymptotically stable, equilibrium point is one where nearby solutions loop around the equilibrium point without coming closer to it.

Advanced comment. For asymptotically stable, we kept the condition that nearby solutions remain close because there are "weird" instances where trajectories come arbitrarily close to the equilibrium, then "flow away" but eventually "flow into" (this would constitute an unstable equilibrium point).

Example 78. (cont'd) Consider again the system $\frac{dx}{dt} = x \cdot (y-1)$, $\frac{dy}{dt} = y \cdot (x-1)$.

(a) Determine the equilibrium points.

(b) Using the phase portrait from Example 77, classify the stability of each equilibrium point.

Solution.

(a) We solve $x(y-1) = 0$ (that is, $x = 0$ or $y = 1$) and $y(x-1) = 0$ (that is, $x = 1$ or $y = 0$).
We conclude that the equilibrium points are $(0,0)$ and $(1,1)$.

(b) $(0,0)$ is asymptotically stable (because all nearby solutions "flow into" $(0,0)$).
$(1,1)$ is unstable (because some nearby solutions "flow away" from $(1,1)$).

Comment. We will soon learn how to determine stability without the need for a plot.

Comment. If you look carefully at the phase portrait near $(1,1)$, you can see that certain solutions get attracted at first to $(1,1)$ and then "flow away" at the last moment. This suggests that there is a single trajectory which actually "flows into" $(1,1)$. This constellation is typical and is called a **saddle point**.

Phase portraits of autonomous linear differential equations

Example 79. Consider the system $\frac{dx}{dt} = y - 5x$, $\frac{dy}{dt} = 4x - 2y$.

(a) Determine the general solution.

(b) Make a phase portrait. Can you connect it with the general solution?

(c) Determine all equilibrium points and their stability.

Solution.

(a) Note that we can write this is in matrix form as $\begin{bmatrix} x \\ y \end{bmatrix}' = M \begin{bmatrix} x \\ y \end{bmatrix}$ with $M = \begin{bmatrix} -5 & 1 \\ 4 & -2 \end{bmatrix}$.

M has -1-eigenvector $\begin{bmatrix} 1 \\ 4 \end{bmatrix}$ as well as -6-eigenvector $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$.

Hence, the general solution is $\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = C_1 \begin{bmatrix} 1 \\ 4 \end{bmatrix} e^{-t} + C_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix} e^{-6t}$.

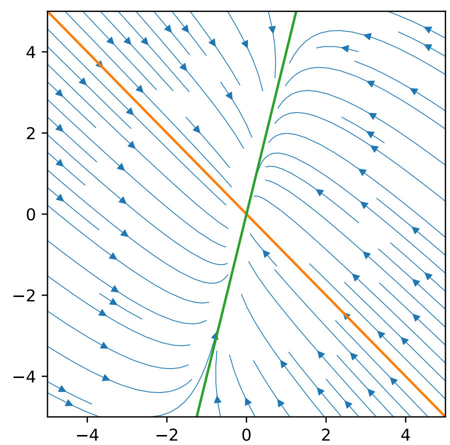
(b) We can have Sage make such a plot for us:

```
>>> x,y = var('x y')
    streamline_plot((-5*x+y,4*x-2*y), (x,-4,4), (y,-4,4))
```

Question. In our plot, we also highlighted two lines through the origin. Can you explain their significance?

Explanation. The lines correspond to the special solutions $C_1 \begin{bmatrix} 1 \\ 4 \end{bmatrix} e^{-t}$ (green) and $C_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix} e^{-6t}$ (orange). For each, the trajectories consist of points that are multiples of the vectors $\begin{bmatrix} 1 \\ 4 \end{bmatrix}$ and $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$, respectively.

Note that each such solution starts at a point on one of the lines and then "flows" into the origin. (Because e^{-t} and e^{-6t} approach zero for large t.)



Question. Consider a point like $(4,4)$. Can you explain why the trajectory through that point doesn't go somewhat straight to $(0,0)$ but rather flows nearly parallel the orange line towards the green line?

Explanation. A solution through $(4,4)$ is of the form $\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = C_1 \begin{bmatrix} 1 \\ 4 \end{bmatrix} e^{-t} + C_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix} e^{-6t}$ (like any other solution). Note that, if we increase t, then e^{-6t} becomes small much faster than e^{-t}.

As a consequence, we quickly get $\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} \approx C_1 \begin{bmatrix} 1 \\ 4 \end{bmatrix} e^{-t}$, where the right-hand side is on the green line.

(c) The only equilibrium point is $(0,0)$ and it is asymptotically stable.

We can see this from the phase portrait but we can also determine it from the DE and our solution: first, solving $y - 5x = 0$ and $4x - 2y = 0$ we only get the unique solution $x = 0, y = 0$, which means that only $(0,0)$ is an equilibrium point. On the other hand, the general solution shows that every solution approaches $(0,0)$ as $t \to \infty$ because both e^{-t} and e^{-6t} approach 0.

In general. This is typical: if both eigenvalues are negative, then the equilibrium is asymptotically stable. If at least one eigenvalue is positive, then the equilibrium is unstable.

Armin Straub
straub@southalabama.edu

Example 80. Consider the system $\frac{dx}{dt} = 5x - y$, $\frac{dy}{dt} = 2y - 4x$.

(a) Determine the general solution.

(b) Make a phase portrait.

(c) Determine all equilibrium points and their stability.

Solution.

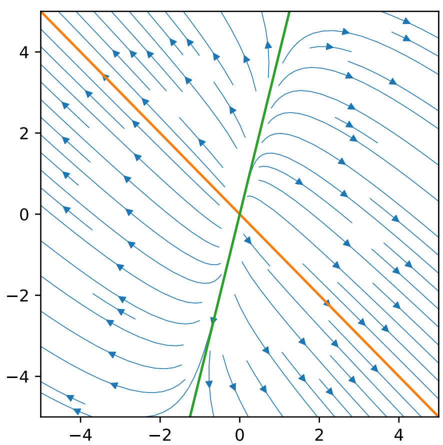
(a) Note that we can write this is in matrix form as $\begin{bmatrix} x \\ y \end{bmatrix}' = M \begin{bmatrix} x \\ y \end{bmatrix}$ with $M = -\begin{bmatrix} -5 & 1 \\ 4 & -2 \end{bmatrix}$, where the matrix is exactly -1 times what it was in Example 79.

Consequently, M has 1-eigenvector $\begin{bmatrix} 1 \\ 4 \end{bmatrix}$ as well as 6-eigenvector $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$. (Can you explain why the eigenvectors are the same and the eigenvalues changed sign?)

Thus, the general solution is $\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = C_1 \begin{bmatrix} 1 \\ 4 \end{bmatrix} e^t + C_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix} e^{6t}$.

(b) We again have Sage make the plot for us:

```
>>> x,y = var('x y')
    streamline_plot((5*x-y,-4*x+2*y), (x,-4,4), (y,-4,4))
```



Note that the phase portrait is identical to the one in Example 79, except that the arrows are reversed.

(c) The only equilibrium point is $(0,0)$ and it is unstable.

We can see this from the phase portrait but we can also see it readily from our general solution $\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = C_1 \begin{bmatrix} 1 \\ 4 \end{bmatrix} e^t + C_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix} e^{6t}$ because e^t and e^{6t} go to ∞ as $t \to \infty$.

In general. If at least one eigenvalue is positive, then the equilibrium is unstable.

Example 81. Suppose the system $\frac{dx}{dt} = f(x, y)$, $\frac{dy}{dt} = g(x, y)$ has general solution $\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = C_1 \begin{bmatrix} 1 \\ 4 \end{bmatrix} e^{-t} + C_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix} e^{6t}$. Determine all equilibrium points and their stability.

Solution. Recall that equilibrium points correspond to constant solutions. Clearly, the only constant solution is the zero solution $\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. Equivalently, the only equilibrium point is $(0,0)$.

Since $e^{6t} \to \infty$ as $t \to \infty$, we conclude that the equilibrium is unstable. (Note that the solution $C_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix} e^{6t}$ starts arbitrarily near to $(0,0)$ but always "flows away").

Armin Straub
straub@southalabama.edu