Companion matrices

For any scalars $a_0, a_1, \ldots, a_{n-1}$ (with $n > 1$), the char polyn of the matrix

$$\begin{bmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \ldots & 0 & 1 \\ a_0 & a_1 & a_2 & \ldots & a_{n-1} \end{bmatrix}$$

is $(-1)^{n-1}[a_0 + a_1\lambda + a_2\lambda^2 + \cdots + a_{n-1}\lambda^{n-1} - \lambda^n]$.

Proof by induction: Easy to check for $n = 2$. Assume true for some n, and evaluate the determinant by cofactors down the first column. (See next slide.) The result will follow.

$$
\begin{vmatrix}
-\lambda & 1 & 0 & \dots & 0 \\
0 & -\lambda & 1 & \dots & 0 \\
\vdots & \vdots & \ddots & \ddots & \vdots \\
0 & 0 & \dots & -\lambda & 1 \\
a_0 & a_1 & a_2 & \dots & a_n - \lambda
\end{vmatrix}
$$

$$
\begin{aligned}
&= (-\lambda)(-1)^{1+1}((-1)^{n-1}[a_1 + a_2\lambda + a_3\lambda^2 + \cdots + a_n\lambda^{n-1} - \lambda^n]) \\
&\qquad + a_0(-1)^{n+1+1}(1) \\
&= (-1)^{n+1-1}[a_1\lambda + a_2\lambda^2 + a_3\lambda^3 + \cdots + a_n\lambda^n - \lambda^{n+1}] \\
&\qquad + (-1)^{n+1-1}a_0 \qquad \text{(1st term: Don't forget the } - \text{ on } \lambda) \\
&= (-1)^{n+1-1}[a_0 + a_1\lambda + a_2\lambda^2 + a_3\lambda^3 + \cdots + a_n\lambda^n - \lambda^{n+1}]\ ,
\end{aligned}
$$

as required.

Example

A certain gene has two forms, dominant A and recessive b, so that the organism in which this gene lives can have one of three genotypes, pure dominant AA, pure recessive bb or hybrid Ab. Suppose we always breed the females of our flock with hybrid males. What is the long-term distribution in our flock after many breeding seasons?

Let d_n be the fraction of pure dominants in the flock after n breeding seasons, h_n the fraction of hybrids and r_n the fraction of pure recessives.

Half the offspring of the pure dominants are pure dominants, and half are hybrids, so the offspring of the pure dominants will contribute to the fraction of the next generation $(1/2)d_n$ pure dominants and $(1/2)d_n$ hybrids (assuming all the females have the same number of offspring), Similarly for the other two.

So if we set $\mathbf{f}_n = (d_n, h_n, r_n)$,

$$\text{then } \mathbf{f}_{n+1} = A\mathbf{f}_n \text{ where } A = \begin{bmatrix} 1/2 & 1/2 & 0 \\ 1/4 & 1/2 & 1/4 \\ 0 & 1/2 & 1/2 \end{bmatrix}.$$

Note $d_n + h_n + r_n = 1$, i.e., $\begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \mathbf{f}_n = 1$.

Using `R`, I checked that

$$A^{1024} = \begin{bmatrix} 1/4 & 1/4 & 1/4 \\ 1/2 & 1/2 & 1/2 \\ 1/4 & 1/4 & 1/4 \end{bmatrix} .$$

So no matter what $\mathbf{f}_0$ was, $\mathbf{f}_{1024} = (1/4, 1/2, 1/4)$.

How come this works? Well, what are the eigvals of A?

$$\begin{vmatrix} 1/2-\lambda & 1/4 & 0 \\ 1/4 & 1/2-\lambda & 1/4 \\ 0 & 1/2 & 1/2-\lambda \end{vmatrix} = (1/2-\lambda)^3 - (1/4)(1/2-\lambda)$$

$$= (1/2-\lambda)(1/4-\lambda+\lambda^2-1/4) = (1/2-\lambda)(-\lambda)(1-\lambda) .$$

So if $\mathbf{u}, \mathbf{v}, \mathbf{w}$ are the eigvecs of $1/2, 0, 1$, they form a basis for $\mathbb{R}^3$.
(Continued)

And if we write $\mathbf{f}_0 = k\mathbf{u} + \ell\mathbf{v} + m\mathbf{w}$, then

$$\mathbf{f}_{1024} = A^{1024}\mathbf{f}_0 = k(1/2)^{1024}\mathbf{u} + \ell(0)^{1024}\mathbf{v} + m(1)^{1024}\mathbf{w} \approx m\mathbf{w}$$

And because all the $\mathbf{f}_n$'s add up to 1, the limiting distribution of percentages $\mathbf{f}_\infty$ is the eigvec corr to 1 that adds up to 1.

Moreover, because the columns of A also added up to 1, the columns of A^∞ are all that same eigvec.

Markov process

Genetic example from last unit was an example of a *discrete linear Markov process,* where each discrete step is a vector, and the transition from each to the next is multiplication by a *Markov matrix,* M:

(1) All entries in M are nonnegative, and

(2) Each column adds to 1 (i.e., $\begin{bmatrix} 1 & 1 & \dots & 1 \end{bmatrix} M = \begin{bmatrix} 1 & 1 & \dots & 1 \end{bmatrix}$).

Because columns add to 1, $\det(M - 1I) = 0$, so 1 is an eigval — a corr eigvec $\mathbf{v}$ is a *steady state,* because $M\mathbf{v} = \mathbf{v}$. If 1 is (strictly) largest eigval in abs val, no matter where the process starts, it approaches a steady state in the long run.

Bad example

$$M = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \text{has eigvals } 1, -1:$$

Steady state is (a scalar multiple of) (1,1). But multiplying (a, b) repeatedly by M alternates with (b, a), rather than converging to a steady state.
(Eigvec corr to -1 is $(1, -1)$, so 1 isn't strictly largest in abs val.)

Perron-Frobenius Theorem for $A > 0$

Theorem

Suppose A is an $n \times n$ matrix with all (strictly) positive entries. Then the largest eigval λ_{max} of A in absolute value is (real) positive and has an eigvec with strictly positive entries.

For a proof, see the Unit 11 notes.

Now because, if M is a Markov matrix, then the sums of $\mathbf{f}$ and of $M\mathbf{f}$ are equal, so M can't have an eigval greater than 1. Thus, if M is Markov and has <u>all</u> its entries positive, then its largest eigval is 1, and multiplying by it again and again will move any vector to a multiple of the corresponding eigvec.

Differential equations I

A "solution" to the diff eqn

$$\frac{dx}{dt} = 3x$$

is a function $x(t)$ that makes this equation true. It is easy to see that $x = e^{3t}$ works — in fact, $x = Ce^{3t}$ works for any scalar C. The family of solutions of this "homogeneous" diff eqn is a 1-dim subspace of the vector space of real-valued functions.

The solutions of $dy/dt = -4y$ are $y = De^{-4t}$. So, if we write $\mathbf{u} = (x, y)$ — a vector-valued function of t — the solutions to the *system of homogeneous linear diff eqns (with constant coefficients)*

$$\frac{d\mathbf{u}}{dt} = \begin{bmatrix} 3 & 0 \\ 0 & -4 \end{bmatrix} \mathbf{u} \qquad \text{are} \qquad \mathbf{u} = \begin{bmatrix} Ce^{3t} \\ De^{-4t} \end{bmatrix}.$$

Example:

$$\frac{d\mathbf{u}}{dt} = A\mathbf{u} \quad \text{where} \quad A = \begin{bmatrix} .8 & .3 \\ .2 & .7 \end{bmatrix}, \quad \mathbf{u} = \begin{bmatrix} x \\ y \end{bmatrix}$$

We can check that $A = S\Lambda S^{-1}$ where

$$S = \begin{bmatrix} .6 & 1 \\ .4 & -1 \end{bmatrix} \quad \text{and} \quad \Lambda = \begin{bmatrix} 1 & 0 \\ 0 & .5 \end{bmatrix}$$

and because the entries in S^{-1} are constants, we have $d(S^{-1}\mathbf{u})/dt) = S^{-1}(d\mathbf{u}/dt)$, so if we set $\mathbf{v} = S^{-1}\mathbf{u}$, then we want to solve

$$\frac{d\mathbf{v}}{dt} = \Lambda\mathbf{v} \quad \Longrightarrow \quad \mathbf{v} = \begin{bmatrix} Ce^{t} \\ De^{.5t} \end{bmatrix}.$$

(Continued)

So

$$\mathbf{u} = S\mathbf{v} = \begin{bmatrix} .6Ce^{t} + De^{.5t} \\ .4Ce^{t} - De^{.5t} \end{bmatrix} = Ce^{t}\begin{bmatrix} .6 \\ .4 \end{bmatrix} + De^{.5t}\begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

In general, for the system $d\mathbf{u}/dt = A\mathbf{u}$, if A has distinct eigvals $\lambda_1, \ldots, \lambda_n$ with corr eigvecs $\mathbf{x}_1, \ldots, \mathbf{x}_n$, then

$$e^{\lambda_1 t}\mathbf{x}_1, \ , \ldots , \ e^{\lambda_n t}\mathbf{x}_n$$

is a basis for the space of solutions.

To find a particular solution to an "initial value problem" — a linear system of differential equations as above together with some initial values, of the form $\mathbf{u}(0) = (b_1, \ldots, b_n) = \mathbf{b}$ — we would use the fact that $e^{\lambda 0} = 1$ for every λ, and solve the system

$$\begin{bmatrix} \mathbf{x}_1 & \ldots & \mathbf{x}_n \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = \mathbf{b}$$

to find the coefficients c_i of the basis above.

Example (ctnd): For the initial value problem

$$\frac{d\mathbf{u}}{dt} = \begin{bmatrix} .8 & .3 \\ .2 & .7 \end{bmatrix} \mathbf{u} \ , \qquad \mathbf{u}(0) = \begin{bmatrix} 5 \\ 0 \end{bmatrix} \ :$$

All solutions to the diff eqn are

$$\mathbf{u} = Ce^{t} \begin{bmatrix} .6 \\ .4 \end{bmatrix} + De^{.5t} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \ ,$$

and substituting $t = 0$ gives

$$C \begin{bmatrix} .6 \\ .4 \end{bmatrix} + D \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 5 \\ 0 \end{bmatrix} \ , \qquad \text{so}$$

$$\begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} .6 & 1 \\ .4 & -1 \end{bmatrix}^{-1} \begin{bmatrix} 5 \\ 0 \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \end{bmatrix} \ .$$

Higher order diff eqns I

Given the diff eqn

$$x''' - 3x'' - x' + 3x = 0 \quad :$$

DE class: Suppose $x = e^{kt}$ is a soln. What's k?

$$k^3 e^{kt} - 3k^2 e^{kt} - 3ke^{kt} + 3e^{kt} = 0$$

But e^{kt} is not zero, so $k^3 - 3k^2 - 3k + 3 = 0$.

Higher order diff eqns II

Given the diff eqn

$$x''' - 3x'' - x' + 3x = 0 \ :$$

LA class: Let $y = x'$, $z = y'$, and $\mathbf{u} = (x, y, z)$. Then the diff eqn is equivalent to the system

$$x' = y \qquad y' = z \qquad z' = -3x + y + 3z, \qquad \text{or}$$

$$\frac{d\mathbf{u}}{dt} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -3 & 1 & 3 \end{bmatrix} \mathbf{u} \ .$$

To solve the system, diagonalize the matrix — but it's a companion matrix: Char polyn is $(-1)^2[-3 + \lambda + 3\lambda^2 - \lambda^3]$. Look familiar?

Continuing the LA solution: Roots of $-3+\lambda+3\lambda^2-\lambda^3$ are $3, 1, -1$, and corresponding eigvecs are

$$\begin{bmatrix} 1 \\ 3 \\ 9 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

respectively. So general solution is

$$Ce^{3t}\begin{bmatrix} 1 \\ 3 \\ 9 \end{bmatrix} + De^{t}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + Ee^{-t}\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}.$$

We only really want $x(t)$, so we can read off first coordinate. And if we add an initial condition, like $x(0)=p$, $x'(0)=q$ $x''(0)=r$, then to find C, D, E we need to solve

$$\begin{bmatrix} 1 & 1 & 1 \\ 3 & 1 & -1 \\ 9 & 1 & 1 \end{bmatrix}\begin{bmatrix} C \\ D \\ E \end{bmatrix} = \begin{bmatrix} p \\ q \\ r \end{bmatrix}.$$

Just from the eigvals?

Suppose the solution to $\frac{d\mathbf{u}}{dt} = A\mathbf{u}$ is

$$\mathbf{u}(t) = Ce^{\lambda t}\mathbf{x} + De^{\mu t}\mathbf{y} \ .$$

Case 1: λ, μ real, $\lambda < \mu$. Then as $t \to \infty$, $e^{\mu t}$ grows faster than $e^{\lambda t}$, so $\mathbf{u}(t)$ gets closer to a multiple of $\mathbf{y}$.

- $\mu > 0$: $\mathbf{u}(t) \to \infty$ (or $-\infty$ if $D < 0$).
- $\mu = 0$: $\mathbf{u}(t) \to D\mathbf{y}$. (So in diff eqns, an eigval of 0, not 1, gives a steady state.)
- $\mu < 0$: $\mathbf{u}(t) \to 0$.

Case 2: $\lambda, \mu = a \pm bi$ (a, b real):

$$e^{(a \pm bi)t} = e^{at}(\cos(bt) \pm i \sin(bt)) \; .$$

Because λ, μ are complex conjugates, so are eigvecs $\mathbf{x}, \mathbf{y}$ and coefficients C, D; so we can rewrite

$$\mathbf{u}(t) = e^{at} \cos(bt)\mathbf{x}' + e^{at} \sin(bt)\mathbf{y}'$$

(where the primes just mean new values, not derivatives).

- $a > 0$: $\mathbf{u}(t)$ spirals outward.
- $a = 0$: $\mathbf{u}(t)$ rotates around a point.
- $a < 0$: $\mathbf{u}(t)$ spirals inward toward a fixed point (the origin, in this case).

Example

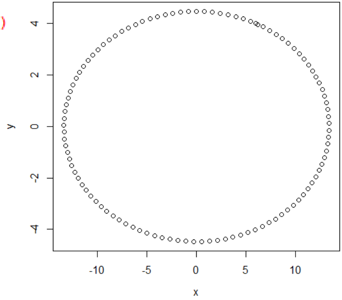
$$\frac{d\mathbf{u}}{dt} = A\mathbf{u} \text{ where } A = \begin{bmatrix} 0 & .9 \\ -.1 & 0 \end{bmatrix} \text{ and } \mathbf{u}(0) = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$$

$$\begin{vmatrix} -\lambda & .9 \\ -.1 & -\lambda \end{vmatrix} = \lambda^2 + .09 \;, \quad \text{so eigvals are } .3i, -.3i.$$

You can check the eigvecs are for $(-3i, 1), (3i, 1)$ respectively, and $C = 2 + i$, $D = 2 - i$.

$$\mathbf{u}(t) = (2+i)e^{.3it}\begin{bmatrix} -3i \\ 1 \end{bmatrix} + (2-i)e^{-.3it}\begin{bmatrix} 3i \\ 1 \end{bmatrix}$$
$$= (2+i)(\cos(.3t) + i\sin(.3t)\begin{bmatrix} -.3i \\ 1 \end{bmatrix}$$
$$+ (2-i)(\cos(.3t) - i\sin(.3t))\begin{bmatrix} .3i \\ 1 \end{bmatrix}$$
$$= \begin{bmatrix} 6\cos(.3t) + 12\sin(.3t) \\ 4\cos(.3t) - 2\sin(.3t) \end{bmatrix} = \cos(.3t)\begin{bmatrix} 6 \\ 4 \end{bmatrix} + \sin(.3t)\begin{bmatrix} 12 \\ -2 \end{bmatrix}$$

```
> t<-seq(0,21, by=.2)
> x<-6*cos(.3*t)+12*sin(.3*t)
> y<-4*cos(.3*t)-2*sin(.3*t)
> plot(x,y)
```



Matrix exponential

Just as $e^x = 1 + x + x^2/2! + x^3/3! + \ldots$, it is sometimes useful to use

$$e^A = I + A + (1/2!)A^2 + (1/3!)A^3 + \cdots = \sum_{n=0}^{\infty}(1/n!)A^n \ .$$

We get some natural results, like $d(e^{At})/dt = Ae^{At}$ if A is constant. But some things fail: $e^{A+B} \neq e^A e^B$ in general. Because some matrices are "nilpotent", some exponentials are easy to compute:

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix} \quad : \quad A^3 = O \text{ , so}$$

$$e^A = I + A + (1/2)A^2 = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} .$$

Also, if $A\mathbf{x} = \lambda\mathbf{x}$, then

$$e^A\mathbf{x} = \sum(1/n!)A^n\mathbf{x} = \sum(1/n!)\lambda^n\mathbf{x} = e^\lambda\mathbf{x} .$$