

**MATH 56A SPRING 2008
STOCHASTIC PROCESSES**

KIYOSHI IGUSA

CONTENTS

Introduction	2
Example	2
Martingale	3
Markov process	3
Goal of the course	3
0. Differential and difference equations	4
0.1. Linear differential equations in one variable	4
0.2. Kermack-McKendrick	10
0.3. systems of first order equations	14
0.4. Linear difference equations	18
Homework 0	20

INTRODUCTION

In the first lecture, I discussed the concept of a stochastic process and gave a very quick introduction to two of the main concepts in this course: martingales and Markov processes. I also tried to convey the flavor and philosophy of the course.

A *stochastic process* is defined to be a random process which evolves with time. For example, if you toss two dice then you get the numbers 2 through 12 with a certain fixed probability distribution. This is standard probability theory. An example of a stochastic process might be: Toss two dice and get a total of X_1 . Then toss that many dice and get a total of X_2 and so on. As time goes on you will need a lot of dice!

Example. The next example I gave was the question: What is the probability that your family name will survive? The answer I got was 0. I.e., with probability 1, everyone on Earth will have the same last name. This is the male version. You get your last name from your father. But, you get your mitochondria from your mother. The female version is that everybody on Earth will eventually have the same mitochondria which is true!

The setup for the population extinction problem (which we will study more carefully later) is the following.

Start at time $t = 0$ with a male population of N_0 .

N_t is the male population after t generations.

X_1 is the number of male offspring from the first man,

X_2 is the number of male offspring from the second man, etc.

Then, the number of males in the next generation will be

$$N_1 = X_1 + X_2 + \cdots + X_{N_0}.$$

Assume that X_i are *independent identically distributed* (i.i.d.) random variables. In particular, they all have the same expected value:

$$\mathbb{E}(X_1) = \mathbb{E}(X_2) = \cdots = \mu.$$

So,

$$\mathbb{E}(N_1) = \mu N_0.$$

This repeats and we get

$$\mathbb{E}(N_t) = \mu^t N_0.$$

This is exponential growth.

Martingale. One thing that is good to do is to make a “martingale”:

$$M_t = \frac{N_t}{\mu^t}.$$

Then,

$$\mathbb{E}(M_t) = N_0.$$

This is constant. That makes M_t a martingale. (A *martingale* is a stochastic random variable which you expect to have the same value tomorrow as it has today.)

The “Martingale Convergence Theorem” now tells us that M_t converges to M_∞ . On page 119 of our book, it says that, if $\mu > 1$ then $\mathbb{E}(M_\infty) = \mathbb{E}(M_0)$ ($= N_0$ in this case). But, in class I argued that $\mu = 1$ and that the expected value of M_∞ will be zero!

Markov process. For this I converted the problem into a *Markov process*. This is defined to be a system in which there is a fixed set of states and each state there is a fixed probability of going to each other state. For example, in the *random walk* the states are the integer points on the real line. If you are at any point, the probability of going to the left one space is $1/2$ and the probability of going to the right one space is $1/2$.

For the population problem, the states are: $0, 1, 2, 3, 4, \dots$ and you are in state N_t at time t . Given N_t men in generation t , there is a certain probability of every possible number of males in the next generation. So, we have a Markov process. We will learn that, in a Markov system, there are only two types of states: “recurrent” and “transient”. A *recurrent* state is one that you keep coming back to with probability one. The only recurrent state is 0 (extinction). All other states must be *transient* which means you only go there a finite number of times. I will explain this later in the course.

Since all finite states except 0 are transient, the Markov process will “almost surely” (a.s.) go to 0 or ∞ . *Almost surely* means “with probability one.” I did not explain why infinity is not possible. In any case, the answer I got was $\mathbb{P}(M_\infty) = 0, a.s.$

Goal of the course. We will look more carefully at this and other example. But the main example that I am interested in is the *Black-Scholes equation*. Some of you already know this equation from economics where it is usually derived using a binomial distribution. We will do a more serious analysis of this equation using *stochastic integration*. Since this is the last topic in the book, we need to cover *the entire book!* We will go very fast, skipping some of the things at the beginning so that we can get to the end.

0. DIFFERENTIAL AND DIFFERENCE EQUATIONS

We have two days to go over the basics of linear differential equations. Differential equations is a one semester course and we don't have time to cover it in detail. However, to do Markov chains you just need to understand how first order linear differential equations work.

0.1. Linear differential equations in one variable. In the first lecture, I discussed linear differential equations (Diffeq's) in one variable of arbitrary order. I presented the problem and the complete solution but without proof. These missing proofs I appended at the end so that these notes will faithfully represent the style and content of the lectures.

The problem in degree $d = 2$ is to find a function $y = f(t)$ so that:

$$(0.1) \quad y'' + ay' + by + c = 0$$

where a, b, c are constants. (The *degree*, or *order* is the number of times that the variables are differentiated. In this case the degree is 2 since we have y'' .)

0.1.1. *particular solution.* A (one) solution $y = f_0(t)$ of this equation is called a *particular solution*. It is really easy to find:

$$y = f_0(t) = \frac{-c}{b}.$$

This is a constant function. It's derivative (and higher derivatives) are zero: $y' = y'' = 0$. So, when you plug it into Equation (0.1) you get

$$0 + 0 + by + c = 0 \Rightarrow y = -c/b.$$

If $b = 0$ then the answer is

$$y = f_0(t) = \frac{-c}{a}t.$$

This is also easy to see: $y' = -c/a$ and $y'' = 0$. So,

$$y'' + ay' + by + c = 0 + a(-c/a) + 0 + c = -c + c = 0.$$

Now, suppose you have another solution $y = f(t)$.

0.1.2. *homogeneous equation.*

Lemma 0.1. *If $f_0(t), f(t)$ are two solutions of the differential equation then the difference*

$$y = f(t) - f_0(t)$$

is a solution of the homogeneous equation

$$y'' + ay' + by = 0.$$

(This is the original equation minus the constant term c .)

This lemma implies that the solutions of the Diffeq are given by:

$$f(t) = f_0(t) + (\text{all solutions of the homogeneous equation})$$

The solutions of homogeneous equation have good theoretical properties:

Lemma 0.2. *If $f_1(t)$ is a solution of the homogeneous equation then so is $\alpha f_1(t)$ for any (constant) scalar α .*

Lemma 0.3. *If f_1, f_2 are two solutions of the homogeneous equation then so is $y = f_1 + f_2$.*

These two lemmas imply that we have a vector space:

Theorem 0.4. *The set of solutions of the homogeneous equation form a vector space. The dimension of this space is equal to the degree d of the differential equation.*

As you remember from linear algebra, every vector space has a basis: $f_1(t), f_2(t), \dots, f_d(t)$. So, every solution of the homogeneous equation is a linear combination of the basis elements:

$$a_1 f_1 + a_2 f_2 + \dots + a_d f_d.$$

This means that the solutions of the original Diffeq are given by:

$$f(t) = f_0(t) + a_1 f_1(t) + a_2 f_2(t) + \dots + a_d f_d(t)$$

where a_1, \dots, a_d are arbitrary scalars.

Thus, we need d linearly independent solutions of the homogeneous differential equation.

0.1.3. *solutions for homogeneous equation of degree one.* I first did the case of degree one. This is an equation of the form:

$$y' + ay = 0$$

or:

$$y' = -ay.$$

I.e., y is decreasing at a rate proportional to its size. This is the equation of exponential decay which you learn in calculus. The solution is

$$y = y_0 e^{-at}.$$

0.1.4. *solutions in higher order.* To explain the idea, I gave a specific example:

$$y'' - 5y' + 6y + 1 = 0.$$

The particular solution is $f_0(t) = -1/6$ and the homogeneous equation is:

$$y'' - 5y' + 6y = 0.$$

I rewrote this using the *differential operator* $\mathcal{D} = \frac{d}{dt}$:

$$\mathcal{D}^2 y - 5\mathcal{D}y + 6y = 0$$

or

$$(\mathcal{D}^2 - 5\mathcal{D} + 6)y = 0.$$

Now we can factor the operator (the thing that is operating on y):

$$(\mathcal{D} - 2) \underbrace{(\mathcal{D} - 3)}_{=0} y = 0.$$

To solve this, we look at the part $(\mathcal{D} - 3)y$. If this is zero then the whole thing is zero. But the equation

$$(\mathcal{D} - 3)y = 0$$

is the first order equation

$$y' - 3y = 0.$$

The solution is $y = C_1 e^{3t}$ for some constant C_1 . This gives one basis element $f_1(t) = e^{3t}$. To get the other one, we go back and rewrite the differential equation as:

$$(\mathcal{D} - 3)(\mathcal{D} - 2)y = 0.$$

Then we get the solution $y = C_2 e^{2t}$ and $f_2(t) = e^{2t}$. So, the general solution of the Diffeq is

$$f(t) = -\frac{1}{6} + C_1 e^{3t} + C_2 e^{2t}.$$

Theorem 0.5. *If we have a linear differential equation in one variable of order d given by a polynomial in \mathcal{D} of degree d with d distinct roots $\lambda_1, \lambda_2, \dots, \lambda_d$ then the functions $e^{\lambda_1 t}, e^{\lambda_2 t}, \dots, e^{\lambda_d t}$ form a basis for the vector space of all solutions of the associated homogeneous equation.*

In the example, $d = 2$, the polynomial is $\mathcal{D}^2 - 5\mathcal{D} + 6$ which has roots $\lambda_1 = 3, \lambda_2 = 2$. The analysis of the general case is very similar.

There are two points which I expanded on in order to give a complete description of the answer:

- complex roots,
- multiple roots.

0.1.5. *complex roots.* What happens in the case when λ is a complex number? I gave an example to start:

$$y'' + 2y' + 5y = 0.$$

In terms of differential operators this is:

$$(\mathcal{D}^2 + 2\mathcal{D} + 5)y = 0.$$

The roots of this polynomial are:

$$\lambda_{\pm} = -1 \pm 2i, \quad i = \sqrt{-1}.$$

So, a basis for the solution space is $e^{\lambda_+ t}, e^{\lambda_- t}$. But what are these functions? Here I switched to letters: $\lambda_{\pm} = a \pm bi$ where $a = -1, b = 2$.

$$e^{\lambda_+ t} = e^{at} e^{bit} = e^{at} (\cos bt + i \sin bt)$$

$$e^{\lambda_- t} = e^{at} e^{-bit} = e^{at} (\cos bt - i \sin bt)$$

If we add these and divide by 2 or subtract and divide by $2i$ we get two other basic solutions of the homogeneous equation:

$$f_1(t) = e^{at} \cos bt$$

$$f_2(t) = e^{at} \sin bt$$

In our particular example, we have

$$f_1(t) = e^{-t} \cos 2t$$

$$f_2(t) = e^{-t} \sin 2t.$$

These two functions form the *real* basis for the 2-dimensional vector space of all solutions of the second order homogeneous diffeq.

0.1.6. *multiple roots.* Suppose that the polynomial has multiple roots. For example, suppose the equation is:

$$y'' + 4y' + 4y = 0$$

$$(\mathcal{D}^2 + 4\mathcal{D} + 4)y = 0.$$

This factors as:

$$(\mathcal{D} + 2)^2 y = 0.$$

The roots are $\lambda = -2, -2$. I.e., -2 is a *double root*.

We know that one of the solutions is $f_1(t) = e^{-2t}$. There must be one more. We cannot take the same function twice. The general answer is

$$f_2(t) = te^{\lambda t} = e^{-2t}.$$

If λ is a triple root we get $f_3 = t^2 e^{\lambda t}$ and so on.

The derivation, which I did not give in class, is easy:

$$\mathcal{D}(te^{\lambda t}) = e^{\lambda t} + t\lambda e^{\lambda t}.$$

So,

$$\begin{aligned}(\mathcal{D} - \lambda)te^{\lambda t} &= e^{\lambda t} \\ (\mathcal{D} - \lambda)(\mathcal{D} - \lambda)te^{\lambda t} &= (\mathcal{D} - \lambda)e^{\lambda t} = 0.\end{aligned}$$

This gives the complete description of the solution of a linear differential equation in one variable.

0.1.7. *proofs.* Students should feel free to skip this subsection and go on to the section on epidemic modeling. This is only for those who want to see all the details.

Going back to our original equation

$$y'' + ay' + by + c = 0$$

we rewrite this as

$$\begin{aligned}(\mathcal{D}^2 + a\mathcal{D} + b)y + c &= 0 \\ \varphi(y) &= -c\end{aligned}$$

where

$$\varphi = \mathcal{D}^2 + a\mathcal{D} + b.$$

The point is that φ is a *linear operator*, i.e.,

$$\varphi(\alpha f_1 + \beta f_2) = \alpha\varphi f_1 + \beta\varphi f_2$$

if α, β are constants. If $\varphi f_1 = \varphi f_2 = 0$ then this equation implies that $\varphi(\alpha f_1 + \beta f_2) = 0$. This proves Lemma 0.2, Lemma 0.3 and the first sentence in Theorem 0.4. It remains to prove Lemma 0.1 and the rest of Theorem 0.4.

Proof of Lemma 0.1. If f_0, f are solutions of the original Diffeq then $\varphi(f_0) = -c$ and $\varphi(f) = -c$. So,

$$\varphi(f - f_0) = \varphi f - \varphi f_0 = -c + c = 0.$$

I.e., $f - f_0$ is a solution of the homogeneous equation. \square

Proof of Theorem 0.4. The proof is by induction on d . If $d = 1$ then you learned, or should have learned, the following in calculus.

$$\begin{aligned}(\mathcal{D} - \lambda)y &= 0 \\ \mathcal{D}y &= \lambda y \\ \frac{dy}{dt} &= \lambda y.\end{aligned}$$

Separate variables:

$$\frac{dy}{y} = \lambda dt,$$

then integrate:

$$\int \frac{dy}{y} = \int \lambda dt.$$

The solution is:

$$\begin{aligned} \ln |y| &= \lambda t + C \\ y &= \pm e^C e^{\lambda t} = y_0 e^{\lambda t} \end{aligned}$$

where $y_0 = \pm e^C$ is constant, i.e., a scalar. So, $e^{\lambda t}$ is a basis. The vector space has dimension $d = 1$ as claimed. This case can be written as follows.

Lemma 0.6. *The kernel $\ker(\mathcal{D} - \lambda)$ of the linear operator $\mathcal{D} - \lambda$ is one dimensional and is spanned by $e^{\lambda t}$.*

Now suppose the theorem holds for degree $d - 1$. Since any polynomial in \mathcal{D} can be factor as a product of linear factors $\mathcal{D} - \lambda$ for complex numbers λ , we can write:

$$\varphi = \psi(\mathcal{D} - \lambda)$$

where ψ is a polynomial of degree $d - 1$ in \mathcal{D} . By induction on d we know that the solution space of the homogeneous Diffeq $\psi y = 0$ is a vector space of dimension $d - 1$, i.e.,

$$\dim \ker \psi = d - 1.$$

But,

$$f(t) \in \ker \varphi \Leftrightarrow \psi(\mathcal{D} - \lambda)f(t) = 0 \Leftrightarrow (\mathcal{D} - \lambda)f(t) \in \ker \psi.$$

So, $\mathcal{D} - \lambda$ is a linear mapping

$$\mathcal{D} - \lambda = L : \ker \varphi \rightarrow \ker \psi.$$

We just showed that the kernel of $\mathcal{D} - \lambda$ is one-dimensional. So,

$$\begin{aligned} \dim(\ker \varphi) &= \dim(\ker L) + \dim(\text{im } L) \\ &\leq 1 + \dim \ker \psi = 1 + (d - 1) = d. \end{aligned}$$

This means that there are *at most* d linearly independent solutions of a homogeneous linear Diffeq of order d . But, in the discussion above we found d linearly independent solutions. So, we know that

$$\dim \ker \varphi \geq d.$$

So, we must have

$$\dim \ker \varphi = d$$

as claimed. □

0.2. Kermack-McKendrick. This is from the book *Epidemic Modelling, An Introduction*, D.J. Daley & J.Gani, Cambridge University Press. Kermack-McKendrick is the most common model for the general epidemic. It is usually more realistic with many subpopulations with different characteristics. But we are only interested in the concept, not in an accurate model. So, I made two simplifying assumptions:

- The population is homogeneous
- No births or deaths by other means

Since there are no births, the size of the population N is constant.

This model is similar to the Markov processes which we will study starting next week: There are “states” and people move from one state to another according to certain rules. In a Markov process, the movement is random. Here it is deterministic.

In this model there are three states:

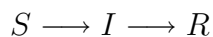
- S:** = susceptible
- I:** = infected
- R:** = removed (immune)

Let $x = \#S, y = \#I, z = \#R$. So

$$N = x + y + z.$$

I assume that $z_0 = 0$ (If there are any “removed” people at $t = 0$ we ignore them.)

As time passes, susceptible people become infected and infected “re-cover” and become immune. So the size of S decreases and the size of R increases. People move as shown by the arrows:



0.2.1. recovery rate. The infected recover at an exponential rate. We assume that the infection has a half-life, say one week, and half of the infected will recover in that time and $3/4$ will recover in two weeks, etc. So, the number of infected tends to decrease at the rate proportional to its size. However, there are also newly infected which keep appearing. So, this recovery process only describes the flow $I \rightarrow R$. The equation is:

$$\boxed{\frac{dz}{dt} = \gamma y, \quad \gamma > 0.}$$

The rate of change of y is equal to the rate of infection minus the rate of recovery.

0.2.2. *infection rate.* The infection rate is given by the *Law of mass action* which says:

The rate of interaction between two different subsets of the population is proportional to the product of the number of elements in each subset.

So,

$$\boxed{\frac{dx}{dt} = -\beta xy, \quad \beta > 0.}$$

To solve these equations, we divide them:

$$\frac{dx}{dz} = \frac{dx/dt}{dz/dt} = \frac{-\beta xy}{\gamma y} = \frac{-\beta x}{\gamma}$$

This is a linear differential equation with solution

$$x = x_0 \exp\left(\frac{-\beta}{\gamma} z\right) = x_0 e^{-z/\rho}$$

where $\rho := \gamma/\beta$ is called the *threshold* population size. This is an exponential decay equation. It says that the size of the susceptible population is decreasing at an alarming rate. Bad news!

However, something happens before we all die: the number of infected y goes to zero and the infection stops!

Since $N = x + y + z$ is fixed we can find y as a function of z :

$$y = N - x - z = N - x_0 e^{-z/\rho} - z$$

Differentiating gives:

$$\frac{dy}{dz} = \frac{x_0}{\rho} e^{-z/\rho} - 1$$

$$\frac{d^2y}{dz^2} = -\frac{x_0}{\rho^2} e^{-z/\rho} < 0$$

So, the function is concave down with initial slope

$$\boxed{\frac{dy}{dz} = \frac{x_0}{\rho} - 1.}$$

I graphed these functions for different values of the parameters to show you what this means.

0.2.3. *Case 1: $x_0 > \rho$.* When the initial susceptible population size is greater than the threshold ρ , the infected population increases at the beginning. This is because

$$x_0 > \rho \quad \Rightarrow \quad \frac{x_0}{\rho} > 1 \quad \Rightarrow \quad \frac{dy}{dz} = \frac{x_0}{\rho} - 1 > 0.$$

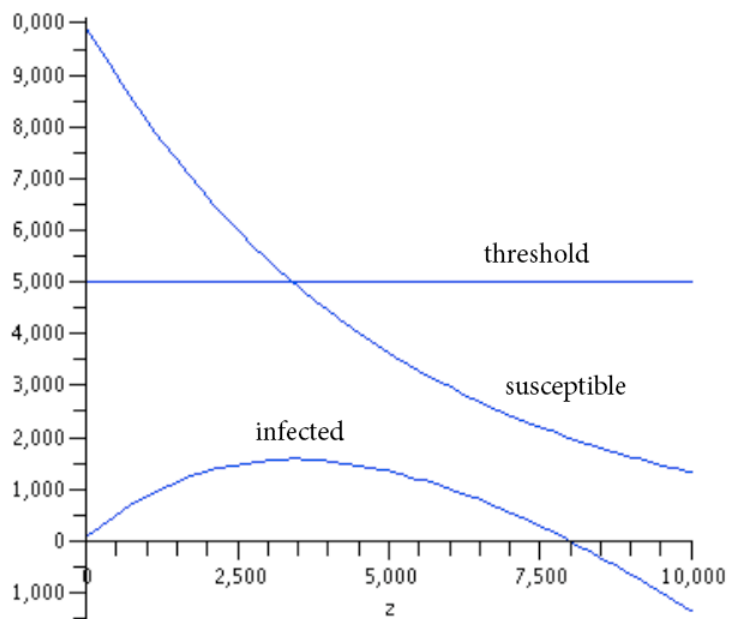
However, it eventually comes back down, although this may not be obvious. Here is the plot in the case when

$$N = 10,000$$

$$x_0 = 9,900$$

$$y_0 = 100$$

$$\rho = 5,000.$$



The plot shows x, y, ρ as a function of z . (ρ is constant.)

Note that there are approximately 2,000 uninfected at the end of the epidemic. (The infected line crosses the z axis at $z = 8,000$ and $x = N - z = 2,000$ is the final value.)

In fact, this model predicts that there will always be survivors of any epidemic. I.e., there will always be people who never get infected.

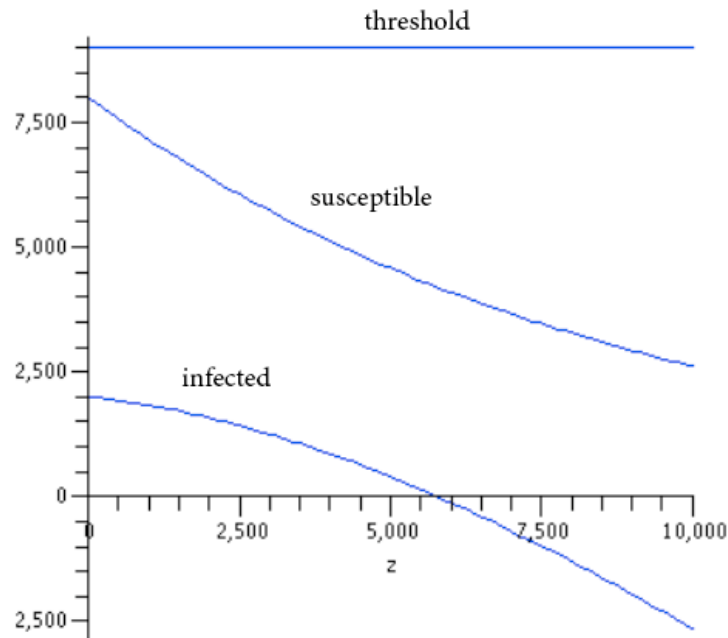
0.2.4. *case 2: $x_0 < \rho$.* If the initial susceptible population size is less than the threshold ρ , the infected population is decreasing at the beginning. Since the infected curve is concave down, it decreases even faster as z increases. Here is the plot in the case

$$N = 10,000$$

$$x_0 = 8,000$$

$$y_0 = 2,000$$

$$\rho = 9,000.$$



Here, almost half of the population survives the epidemic.

Another way to look at it is that this is the tail end of the epidemic. The worst is over. In case 1 we saw the beginning of the epidemic.

Exercise 0.7. Prove that the highest point in the infected curve occurs when the susceptible curve crosses the threshold.

0.3. systems of first order equations.

- (1) Linear differential equations can be reduced to first order.
- (2) First order equations are matrix equations.
- (3) Exponential of a matrix.
- (4) Eigenvalues and eigenvectors.
- (5) Diagonalization of a matrix.
- (6) Solution.

0.3.1. reduction to first order.

Theorem 0.8. *Any n th order linear differential equation in m variables can be reduced to a first order linear diffeq in nm variables.*

Here is an example with $n = m = 2$.

$$\begin{aligned}x'' &= 2x' + 3y' + 6x + 7 \\y'' &= x' + 9y - 5.\end{aligned}$$

(Differentiation is with respect to time t .) The theorem says you need $nm = 4$ variables, i.e., 2 more. Call them v, w and let

$$v = x', \quad w = y'$$

Then we have four 1st order linear diffeq's in 4 variables:

$$\begin{aligned}v' &= 2v + 3w + 6x + 7 \\w' &= v + 9y - 5 \\x' &= v \\y' &= w.\end{aligned}$$

You always get derivatives of the variable on the left and linear combinations of the variables on the right.

This example is too complicated. So, I started over with a simpler example.

Example 0.9. Solve the following 2nd order differential equation:

$$y'' + y' - 6y = 0$$

with initial conditions:

$$y_0 = 0, \quad y'_0 = 5.$$

(Initial position at the origin with initial velocity 5.)

Following the procedure, you introduce another variable $z := y'$. Then the 1st order equations are:

$$\begin{aligned}y' &= z \\z' &= 6y - z\end{aligned}$$

0.3.2. *matrix form.* In matrix form this is:

$$\frac{d}{dt} \begin{pmatrix} y \\ z \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 6 & -1 \end{pmatrix} \begin{pmatrix} y \\ z \end{pmatrix}$$

Which we can write as

$$\boxed{\frac{d}{dt} Y = AY}$$

where

$$Y = \begin{pmatrix} y \\ z \end{pmatrix}, \quad A = \begin{pmatrix} 0 & 1 \\ 6 & -1 \end{pmatrix}.$$

0.3.3. *exponential of a matrix.* The solution of this equations is

$$\boxed{Y = e^{tA} Y_0}$$

where $Y_0 = \begin{pmatrix} y_0 \\ z_0 \end{pmatrix}$ and the matrix exponential e^{tA} is defined by:

$$e^{tA} := I_2 + tA + \frac{t^2 A^2}{2} + \frac{t^3 A^3}{3!} + \dots = \sum_{k=0}^{\infty} \frac{t^k A^k}{k!}$$

To prove that this is the solution, you differentiate the sum term by term:

$$\frac{d}{dt} e^{tA} = \frac{d}{dt} \sum \frac{t^k A^k}{k!} = \sum \frac{d}{dt} \frac{t^k A^k}{k!} = \sum \frac{kt^{k-1} A^k}{k!}$$

which simplifies to

$$\sum \frac{t^{k-1} A^k}{(k-1)!} = A \sum \frac{t^{k-1} A^{k-1}}{(k-1)!} = A e^{tA}.$$

So,

$$\frac{d}{dt} e^{tA} = A e^{tA}$$

0.3.4. *eigenvalues and eigenvectors.* In order to compute e^{tA} we need the eigenvalues and eigenvectors of the matrix A . These are given by the equation

$$\boxed{AX = \lambda X = X(\lambda)}.$$

In this example we have:

$$\underbrace{\begin{pmatrix} 0 & 1 \\ 6 & -1 \end{pmatrix}}_A \underbrace{\begin{pmatrix} 1 \\ 2 \end{pmatrix}}_X = \begin{pmatrix} 2 \\ 4 \end{pmatrix} = \underbrace{2}_{\lambda} \underbrace{\begin{pmatrix} 1 \\ 2 \end{pmatrix}}_X = \begin{pmatrix} 1 \\ 2 \end{pmatrix} (2)$$

(If λ is a scalar, it belongs on the left. If (λ) is a 1×1 matrix, it must be on the right by the rules of matrix multiplication.)

$\lambda_1 = 2$ is an eigenvalue of our matrix A with eigenvector $X_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$.

The other eigenvalue is $\lambda_2 = -3$ with eigenvector $X_2 = \begin{pmatrix} 1 \\ -3 \end{pmatrix}$:

$$\begin{pmatrix} 0 & 1 \\ 6 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ -3 \end{pmatrix} = \begin{pmatrix} -3 \\ 9 \end{pmatrix} = \begin{pmatrix} 1 \\ -3 \end{pmatrix} (-3)$$

0.3.5. *diagonalization.* Put these together to get:

$$\underbrace{\begin{pmatrix} 0 & 1 \\ 6 & -1 \end{pmatrix}}_A \underbrace{\begin{pmatrix} 1 & 1 \\ 2 & -3 \end{pmatrix}}_{Q=(X_1 X_2)} = \begin{pmatrix} 2 & -3 \\ 4 & 9 \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 1 \\ 2 & -3 \end{pmatrix}}_{Q=(X_1 X_2)} \underbrace{\begin{pmatrix} 2 & 0 \\ 0 & -3 \end{pmatrix}}_D.$$

Or:

$$AQ = QD$$

where $Q = (X_1 X_2)$ is the matrix whose columns are the eigenvectors of A and $D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$ is the diagonal matrix whose diagonal entries are the eigenvalues. The equation $AQ = QD$ should be rewritten as:

$$\boxed{A = QDQ^{-1}.}$$

This is called the *diagonalization* of A .

0.3.6. *powers of A .* After we successfully diagonalized A , we can raise it to a power: First,

$$A^2 = QD \underbrace{Q^{-1}Q}_{I_2} DQ^{-1} = QDI_2DQ^{-1} = QD^2Q^{-1}.$$

By induction we get:

$$A^k = QD^kQ^{-1}.$$

Divide by $k!$ and sum over k to get:

$$e^{tA} = Qe^{tD}Q^{-1} = Q \begin{pmatrix} e^{t\lambda_1} & 0 \\ 0 & e^{t\lambda_2} \end{pmatrix} Q^{-1}$$

0.3.7. *solution.* The inverse of

$$Q = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 2 & -3 \end{pmatrix}$$

is given by

$$Q^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \frac{1}{-5} \begin{pmatrix} -3 & -1 \\ -2 & 1 \end{pmatrix} = \begin{pmatrix} 3/5 & 1/5 \\ 2/5 & -1/5 \end{pmatrix}.$$

So,

$$Y = e^{tA}Y_0 = Qe^{tD}Q^{-1}Y_0.$$

Putting in all the numbers, including $y_0 = 0, z_0 = y'_0 = 5$, we get:

$$\begin{aligned} \begin{pmatrix} y \\ z \end{pmatrix} &= \begin{pmatrix} 1 & 1 \\ 2 & -3 \end{pmatrix} \begin{pmatrix} e^{2t} & 0 \\ 0 & e^{-3t} \end{pmatrix} \begin{pmatrix} 3/5 & 1/5 \\ 2/5 & -1/5 \end{pmatrix} \begin{pmatrix} 0 \\ 5 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 1 \\ 2 & -3 \end{pmatrix} \begin{pmatrix} e^{2t} & 0 \\ 0 & e^{-3t} \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 1 \\ 2 & -3 \end{pmatrix} \begin{pmatrix} e^{2t} \\ -e^{-3t} \end{pmatrix} = \begin{pmatrix} e^{2t} - e^{-3t} \\ 2e^{2t} + 3e^{-3t} \end{pmatrix}. \end{aligned}$$

In other words,

$$\begin{aligned} y &= e^{2t} - e^{-3t} \\ z &= y' = 2e^{2t} + 3e^{-3t}. \end{aligned}$$

This algorithm works if the eigenvalues of A are distinct.

0.3.8. *review of linear algebra.* The eigenvalues of a square matrix are defined to be the solutions of the equation

$$\det(A - \lambda I) = 0$$

but there is a trick to use for 2×2 matrices. The determinant of a matrix is always the product of its eigenvalues:

$$\det A = \lambda_1 \lambda_2 = -6$$

The trace (sum of diagonal entries) is equal to the sum of the eigenvalues:

$$\text{tr} A = \lambda_1 + \lambda_2 = -1$$

So, $\lambda_1 = 2, \lambda_2 = -3$. The eigenvectors are the solutions of the linear equations

$$AX = \lambda X$$

or

$$(A - \lambda I_2)X = 0.$$

0.4. Linear difference equations.

0.4.1. *discrete time.* To make the transition

differential equations \Rightarrow Markov chains

we convert to *discrete time* and *row vectors*¹.

$$t \in \mathbb{R}_+(\text{continuous time}) \Rightarrow n \in \mathbb{Z}_+(\text{discrete time})$$

$$Y : \text{column vector} \Rightarrow X : \text{row vector.}$$

The solution we get is

$$X = X_0 e^{tA} = X_0 P^n$$

where $P = e^A$ and $n = t$ (because it is an integer). The exponential matrix function is replaced with positive integer powers of a matrix P .

0.4.2. *one variable, second order.* Discrete time equations have a one variable higher order form and a multivariable first order matrix form. We will start with a one variable second order equation in discrete time.

The problem is to find a sequence of numbers $f(n)$ where n ranges over all the integers from K to N ($K \leq n \leq N$) so that

$$(0.2) \quad f(n) = af(n-1) + bf(n+1)$$

The theory is the same as for Diffeq's. This is homogeneous, linear, second order. So, it has two linearly independent solutions.

The continuous solutions were $e^{t\lambda}$. In discrete time this is

$$e^{t\lambda} = c^n$$

where $c = e^\lambda$ and $n = t$ is an integer. So, $f(n) = c^n$ where you have to solve for c :

$$c^n = ac^{n-1} + bc^{n+1}$$

$$(0.3) \quad bc^2 - c + a = 0$$

$$c = \frac{1 \pm \sqrt{1 - 4ab}}{2b}$$

There were two cases.

Case 1: ($4ab \neq 1$) When the quadratic equation (0.3) has two roots c_1, c_2 then the linear combinations of c_1^n and c_2^n give all the solutions of the homogeneous linear recursion (0.2).

Case 2: ($4ab = 1$) In this case there is only one root $c = \frac{1}{2b}$ and the two independent solutions are $f(n) = c^n$ and nc^n . The second solution is the discrete form of:

$$te^{t\lambda} = nc^n$$

¹I will show you later the rigorous proof that transposing the matrix changes deterministic differential equations into random stochastic processes!!

where $t = n, c = e^\lambda$.

0.4.3. *examples.* An example of Case 1:

Example 0.10. (Fibonacci numbers) These are given by $f(0) = 1, f(1) = 1$ and $f(n+1) = f(n) + f(n-1)$ or:

$$f(n) = f(n+1) - f(n-1)$$

This is $a = -1, b = 1$. The roots of the quadratic equation are $c = \frac{1 \pm \sqrt{5}}{2}$. So,

$$f(n) = \frac{1}{\sqrt{5}} \left(\frac{1 + \sqrt{5}}{2} \right)^n - \frac{1}{\sqrt{5}} \left(\frac{1 - \sqrt{5}}{2} \right)^n$$

This is a rational number since it is Galois invariant (does not change if you switch the sign of $\sqrt{5}$). However, it is not clear from the formula why it is an integer.

Here is an example of Case 2:

Example 0.11. Solve the linear difference equation

$$f(n) = \frac{f(n-1) + f(n+1)}{2}$$

with initial conditions $f(0) = 1, f(1) = 2$. In this case $a = b = \frac{1}{2}$. So, $4ab = 1$ and $c = \frac{1}{2b} = 1$. This means the two linearly independent solutions are

$$\begin{aligned} f_1(n) &= c^n = 1 \\ f_2(n) &= nc^n = n. \end{aligned}$$

So,

$$f(n) = xf_1(n) + yf_2(n) = x + ny.$$

Plugging in $n = 0, 1$ we get:

$$\begin{aligned} f(0) &= 1 = x, \\ f(1) &= 2 = x + y. \end{aligned}$$

So, $x = y = 1$ and

$$f(n) = 1 + n.$$

HOMEWORK 0

Linear diffeq's and difference equations

Four problems due next Monday, Jan 28:

0.1. In the Kermack-McKendrick model, prove that the number of infected reaches its highest point when the number of susceptibles is equal to the threshold (or at $t = 0$ at the beginning of the recorded/modeled time period).

0.2.

Find all functions $x(t), y(t)$ so that

$$x'(t) = 5x - y$$

$$y'(t) = 3x + y$$

Find the particular solution with initial position $(x_0, y_0) = (1, 3)$.

0.3. Find all functions f from integers to complex numbers so that

$$f(n+1) = 4f(n) - 5f(n-1).$$

Now find the solution when $f(0) = f(1) = 2$ and explain why it is real.

0.4. (From the book)

Find the function $f(n), n = 0, 1, 2, 3, \dots$ so that $f(0) = 0$

$$f(n) = \frac{1}{3} [f(n-1) + f(n+1) + f(n+2)], \quad n \geq 1$$

$$\lim_{n \rightarrow \infty} f(n) = 1.$$