

# Notes on Systems of Linear Difference Equations

Math 338, Eduardo D. Sontag

January 31, 2005

## 1 Systems of Difference Equations

Suppose we want to study the following second-order difference equation:

$$x(t+1) = x(t) + 4x(t-1).$$

In a manner totally analogous to what you may have done in a differential equations course when studying second order linear differential equations (like a harmonic oscillator), we may, instead, study a system of *two first order* equations. To do this, we write  $x(t)$  as  $x_1(t)$ , in order to think of it as the first coordinate of a vector, and introduce  $x_2(t) = x(t-1)$ , so that the above equation is equivalent to this system:

$$\begin{aligned}x_1(t+1) &= x_1(t) + 4x_2(t) \\x_2(t+1) &= x_1(t)\end{aligned}$$

or, in equivalent matrix form:

$$X(t+1) = AX(t) = \begin{pmatrix} 1 & 4 \\ 1 & 0 \end{pmatrix} X(t)$$

(or just “ $X^+ = AX$ ” using the “+” convention for shifting time by one unit), where  $X(t)$  is the vector

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

and  $A$  is the matrix

$$A = \begin{pmatrix} 1 & 4 \\ 1 & 0 \end{pmatrix}.$$

How do we solve  $X^+ = AX$ , where  $A$  is any given matrix? Well, if we iterate, we have

$$X(1) = AX(0),$$

$$X(2) = AX(1) = A(AX(0)) = A^2X(0),$$

$$X(3) = AX(2) = A(A^2X(0)) = A^3X(0),$$

and so on, so, for all  $t$ :

$$X(t) = A^tX(0).$$

We see that the problem is that of studying the powers  $A^t$ . For small  $t$ , we can compute these powers explicitly, and we can certainly use a computer to obtain the powers (see the link to the MATLAB-based notes, in the course website). But, especially for large matrices and large  $t$ , it makes more sense to use the following shortcut.

## 2 Computing Powers of a Matrix

We wish to calculate  $A^t$ . The key concept for simplifying the computation of matrix powers is that of *matrix similarity*. Suppose that we have found two matrices,  $\Lambda$  and  $S$ , where  $S$  is invertible, such that this formula holds:

$$A = S\Lambda S^{-1} \quad (1)$$

(one says that  $A$  and  $\Lambda$  are similar matrices). Then, we claim, it is true that also:

$$A^t = S\Lambda^t S^{-1} \quad (2)$$

for all  $t$ . Therefore, if the matrix  $\Lambda$  is one for which  $\Lambda^t$  is easy to find (for example, if it is a diagonal matrix), we can then multiply by  $S$  and  $S^{-1}$  to get  $A^t$ . To see why (2) is a consequence of (1), we just use the following “telescopic” property for powers:

$$A^t = \underbrace{[S\Lambda S^{-1}][S\Lambda S^{-1}] \dots [S\Lambda S^{-1}]}_{t \text{ times}} = S\Lambda^t S^{-1}$$

since all the in-between pairs  $S^{-1}S$  cancel out.

The basic theorem is this one:

**Theorem.** *For every  $n$  by  $n$  matrix  $A$ , one can find an invertible matrix  $S$  and an upper triangular matrix  $\Lambda$  such that (1) holds.*

Remember that an upper triangular matrix is one that has the following form:

$$\begin{pmatrix} \lambda_1 & * & * & \cdots & * & * \\ 0 & \lambda_2 & * & \cdots & * & * \\ 0 & 0 & \lambda_2 & \cdots & * & * \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_{n-1} & * \\ 0 & 0 & 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

where the stars are any numbers. The numbers  $\lambda_1, \dots, \lambda_n$  turn out to be the eigenvalues of  $A$ .

There are two reasons that this theorem is interesting. First, it provides a way to compute powers, because it is not difficult to find powers of upper triangular matrices, and second because it has important theoretical consequences regarding the behavior of powers as  $t \rightarrow \infty$ .

For purposes of this course, we will only use a special case: when  $A$  has  $n$  *distinct* eigenvalues, the matrix  $\Lambda$  can be picked to be *diagonal*. It is just a matrix which lists the eigenvalues  $\lambda_1, \dots, \lambda_n$  of  $A$ , in any order:

$$\begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \lambda_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_{n-1} & 0 \\ 0 & 0 & 0 & \cdots & 0 & \lambda_n \end{pmatrix}.$$

The matrix  $S$  is, in that case, the matrix that lists all *eigenvectors*  $v_1, \dots, v_n$  of  $A$  *in the same order*<sup>1</sup>, that is to say,  $v_1$  is so that  $Av_1 = \lambda_1 v_1$ ,  $v_2$  so that  $Av_2 = \lambda_2 v_2$ , and so on:

$$S = (v_1 \ v_2 \ \dots \ v_n).$$

<sup>1</sup>Eigenvectors are only defined up to nonzero multiples, but any other choice of any particular  $v_i$  will get “cancelled out” by the fact that  $S^{-1}$  appears in (1). In other words, we can pick any eigenvector and the end result will be the same.

Note that  $S$  has  $n$  columns; its  $i$ th column is the  $n$ -vector  $v_i$ .

(Two stronger theorems are possible. One is the “Jordan canonical form” theorem, which provides a matrix  $\Lambda$  that is not only upper triangular but which has an even more special structure. Jordan canonical forms are not very useful from a computational point of view, because they are what is known in numerical analysis as “numerically unstable”, meaning that small perturbations of  $A$  can give one totally different Jordan forms. A second strengthening is the “Schur unitary triangularization theorem” which says that one can pick the matrix  $S$  to be unitary. This last theorem is extremely useful in practice, and is implemented in many numerical algorithms.)

We do not prove the theorem here in general, but only show it for  $n = 2$ ; the general case can be proved in much the same way, by means of a recursive process.

We start the proof by remembering that every matrix has at least one eigenvalue, let us call it  $\lambda$ , and an associated eigenvector,  $v$ . That is to say,  $v$  is a vector different from zero, and

$$Av = \lambda v. \tag{3}$$

To find  $\lambda$ , we find a root of the characteristic equation

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

which, for two-dimensional systems is the same as the equation

$$\lambda^2 - \text{trace}(A)\lambda + \det(A) = 0$$

and, recall,

$$\begin{aligned} \text{trace} \begin{pmatrix} a & b \\ c & d \end{pmatrix} &= a + d \\ \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} &= ad - bc. \end{aligned}$$

(There are, for 2 by 2 matrices, either two real eigenvalues, one real eigenvalue with multiplicity two, or two complex eigenvalues. In the last case, the two complex eigenvalues must be conjugates of each other.) An eigenvector associated to an eigenvalue  $\lambda$  is then found by solving the linear equation

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

(there are an infinite number of solutions; we pick any nonzero one).

With an eigenvalue  $\lambda$  and eigenvector  $v$  found, we next pick *any* vector  $w$  with the property that the two vectors  $v$  and  $w$  are linearly independent. For example, if

$$v = \begin{pmatrix} a \\ b \end{pmatrix}$$

and  $a$  is not zero, we can take

$$w = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

(what would you pick for  $w$  if  $a$  were zero?). Now, since the set  $\{v, w\}$  forms a basis of two-dimensional space, we can find coefficients  $c$  and  $d$  so that

$$Aw = cv + dw. \tag{4}$$

We can summarize both (3) and (4) in one matrix equation:

$$A(v w) = (v w) \begin{pmatrix} \lambda & c \\ 0 & d \end{pmatrix}.$$

We let  $S = (v \ w)$  and

$$\Lambda = \begin{pmatrix} \lambda & c \\ 0 & d \end{pmatrix}.$$

Then,

$$AS = S\Lambda$$

which is the same as what we wanted to prove, namely  $A = SAS^{-1}$ . Actually, we can even say more. It is a fundamental fact in linear algebra that, if two matrices are similar, then their eigenvalues must be the same. Now, the eigenvalues of  $\Lambda$  are  $\lambda$  and  $d$ , because the eigenvalues of any triangular matrix are its diagonal elements. Therefore, since  $A$  and  $\Lambda$  are similar,  $d$  is also an eigenvalue of  $A$ .

## 2.1 The Three Cases for $n = 2$

The following special cases are worth discussing in detail:

1.  $A$  has two different real eigenvalues.
2.  $A$  has two complex conjugate eigenvalues.
3.  $A$  has a repeated real eigenvalue.

In cases 1 and 2, one can always find a *diagonal* matrix  $\Lambda$ . To see why this is true, let us go back to the proof, but now, instead of taking just any linearly independent vector  $w$ , let us pick a special one, namely an eigenvector corresponding to the other eigenvalue of  $A$ :

$$Aw = \mu w.$$

This vector is always linearly independent of  $v$ , so the proof can be completed as before. Notice that  $\Lambda$  is now diagonal, because  $d = \mu$  and  $c = 0$ .

(Proof that  $v$  and  $w$  are linearly independent: if  $\alpha v + \beta w = 0$ , then  $\alpha\lambda v + \beta\mu w = A(\alpha v + \beta w) = 0$ . On the other hand, multiplying  $\alpha v + \beta w = 0$  by  $\lambda$  we would have  $\alpha\lambda v + \beta\lambda w = 0$ . Subtracting we would obtain  $\beta(\lambda - \mu)w = 0$ , and as  $\lambda - \mu \neq 0$  we would arrive at the conclusion that  $\beta w = 0$ . But  $w$ , being an eigenvector, is required to be nonzero, so we would have to have  $\beta = 0$ . Plugging this back into our linear dependence would give  $\alpha v = 0$ , which would require  $\alpha = 0$  as well. This shows us that there are no nonzero coefficients  $\alpha$  and  $\beta$  for which  $\alpha v + \beta w = 0$ , which means that the eigenvectors  $v$  and  $w$  are linearly independent.)

So, in the case that  $A$  has two distinct eigenvalues  $\lambda, \mu$ , we can write

$$\Lambda = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}$$

and therefore, for every  $t$ :

$$\Lambda^t = \begin{pmatrix} \lambda^t & 0 \\ 0 & \mu^t \end{pmatrix}.$$

Notice that in cases 1 and 3, the matrices  $\Lambda$  and  $S$  are both real. In case 2, the matrices  $\Lambda$  and  $S$  are, in general, not real. In that case, suppose that  $Av = \lambda v$ ; then, taking complex conjugates gives:

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

and we note that:

$$\bar{\bar{\lambda}} \neq \lambda$$

because  $\lambda$  is not real. So, we can always pick  $w =$  the conjugate of  $v$ . It turns out that solutions can be re-expressed in terms of trigonometric functions. You have seen (or will see) this in a differential equations course.

Next, let's consider Case 3 (the repeated real eigenvalue case). We have that

$$\Lambda = \begin{pmatrix} \lambda & c \\ 0 & \lambda \end{pmatrix}$$

so we can also write  $\Lambda = \lambda I + cN$ , where  $N$  is the following matrix:

$$N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Observe that:

$$(\lambda I + cN)^2 = (\lambda I)^2 + c^2 N^2 + 2\lambda cN = \lambda^2 I + 2\lambda cN$$

(because  $N^2 = 0$ ) and, for the general power  $t$ , recursively:

$$\begin{aligned} A^t = (\lambda I + cN)^t &= (\lambda^{t-1} I + (t-1)\lambda^{t-2} cN) (\lambda I + cN) \\ &= \lambda^t I + (t-1)\lambda^{t-1} cN + \lambda^{t-1} cN + (t-1)\lambda^{t-2} c^2 N^2 \\ &= \lambda^t I + t\lambda^{t-1} cN. \end{aligned}$$

This means that

$$\Lambda^t = \begin{pmatrix} \lambda^t & ct\lambda^{t-1} \\ 0 & \lambda^t \end{pmatrix}$$

for every  $t$ .

### 3 A Shortcut

If we just want to find the form of the general solution of  $X^+ = AX$ , we do not need to actually calculate the powers of  $A$  and the inverse of the matrix  $S$ .

Let us first take the cases of different eigenvalues (real or complex, that is, cases 1 or 2; it doesn't matter which one at this point). As we saw,  $\Lambda$  can be taken to be the diagonal matrix consisting of these eigenvalues (which we call here  $\lambda$  and  $\mu$ ), and  $S = (vw)$  just lists the two eigenvectors as its columns, in the respective order. We then know that the solution of every initial value problem  $X^+ = AX$ ,  $X(0) = X_0$  will be of the following form:

$$X(t) = A^t X_0 = S \Lambda^t S^{-1} X_0 = (vw) \begin{pmatrix} \lambda^t & 0 \\ 0 & \mu^t \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = a\lambda^t v + b\mu^t w$$

where we just wrote  $S^{-1}X_0$  as a column vector of general coefficients  $a$  and  $b$ . In conclusion:

*The general solution of  $X^+ = AX$ , when  $A$  has two eigenvalues  $\lambda$  and  $\mu$  with respective eigenvectors  $v$  and  $w$ , is of the form*

$$a\lambda^t v + b\mu^t w \tag{5}$$

*for some constants  $a$  and  $b$ .*

So, one approach to solving such linear difference equations is to first find eigenvalues and eigenvectors, write the solution in the above general form, and then plug-in the initial condition in order to figure out what are the right constants. We did an example of this in class, and more examples are given in the MATLAB handout.

In the case of non-real eigenvalues, recall that we showed that the two eigenvalues must be conjugates of each other, and the two eigenvectors may be picked to be conjugates of each other. Let us show now

that we can write (5) in a form which does not involve any complex numbers. In order to do so, we start by decomposing the first vector function which appears in (5) into its real and imaginary parts:

$$\lambda^t v = X_1(t) + iX_2(t) \quad (6)$$

(let us not worry for now about what the two functions  $X_1$  and  $X_2$  look like). Since  $\mu$  is the conjugate of  $\lambda$  and  $w$  is the conjugate of  $v$ , the second term is:

$$\mu^t w = X_1(t) - iX_2(t). \quad (7)$$

So we can write the general solution shown in (5) also like this:

$$a(X_1 + iX_2) + b(X_1 - iX_2) = (a + b)X_1 + i(a - b)X_2. \quad (8)$$

Now, it is easy to see that  $a$  and  $b$  must be conjugates of each other. (Do this as an optional homework problem. Use the fact that these two coefficients are the components of  $S^{-1}X_0$ , and the fact that  $X_0$  is real and that the two columns of  $S$  are conjugates of each other.) This means that *both coefficients  $a + b$  and  $i(a - b)$  are real numbers*. Calling these coefficients “ $k_1$ ” and “ $k_2$ ”, we can summarize the complex case like this:

*The general solution of  $X^+ = AX$ , when  $A$  has a non-real eigenvalue  $\lambda$  with respective eigenvector  $v$ , is of the form*

$$k_1 X_1(t) + k_2 X_2(t) \quad (9)$$

*for some real constants  $k_1$  and  $k_2$ . The functions  $X_1$  and  $X_2$  are found by the following procedure: calculate the product  $\lambda^t v$  and separate it into real and imaginary parts as in Equation (6).*

What do  $X_1$  and  $X_2$  really look like? One way to compute the power  $\lambda^t$  is as follows. We first write  $\lambda$  in polar form,  $\lambda = re^{i\beta}$ , where  $r > 0$  (why is  $r \neq 0$ ?) and  $\beta$  is some real number (the argument of  $\lambda$ ). Next, we let  $\alpha = \ln r$ , so that  $r = e^\alpha$ . In summary,  $\lambda = e^\alpha e^{i\beta}$ , and therefore Euler’s formula gives:

$$\lambda^t = e^{\alpha t} e^{i\beta t} = e^{\alpha t} (\cos \beta t + i \sin \beta t) = e^{\alpha t} \cos \beta t + i e^{\alpha t} \sin \beta t = r^t \cos \beta t + i r^t \sin \beta t$$

and we can then find  $X_1$  and  $X_2$  by using a little complex algebra.

Finally, in case 3 (repeated eigenvalues) we can write, instead:

$$\begin{aligned} X(t) &= A^t X_0 = S \Lambda^t S^{-1} X_0 = (v \ w) \begin{pmatrix} \lambda^t & ct\lambda^t \\ 0 & \mu^t \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} \\ &= a \lambda^t v + b \lambda^t (ctv + w). \end{aligned}$$

When  $c = 0$  we have from  $A = S\Lambda S^{-1}$  that  $A$  must have been the diagonal matrix

$$\begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}$$

to start with (because  $S$  and  $\Lambda$  commute). When  $c \neq 0$ , we can write  $k_2 = bc$  and redefine  $w$  as  $\frac{1}{c}w$ . Note that then (4) becomes  $Aw = v + \lambda w$ , that is,  $(A - \lambda I)w = v$ . Any vector  $w$  with this property is linearly independent from  $v$  (why?).

So we conclude, for the case of repeated eigenvalues:

*The general solution of  $X^+ = AX$ , when  $A$  has a repeated (real) eigenvalue  $\lambda$  is either of the form  $\lambda^t X_0$  (if  $A$  is a diagonal matrix) or, otherwise, is of the form*

$$k_1 \lambda^t v + k_2 \lambda^t (tv + w) \quad (10)$$

*for some real constants  $k_1$  and  $k_2$ , where  $v$  is an eigenvector corresponding to  $\lambda$  and  $w$  is any vector which satisfies  $(A - \lambda I)w = v$ .*

Observe that  $(A - \lambda I)^2 w = (A - \lambda I)v = 0$ . general, one calls any nonzero vector such that  $(A - \lambda I)^k w = 0$  a *generalized eigenvector* (of order  $k$ ) of the matrix  $A$  (since, when  $k = 1$ , we have eigenvectors).

## 4 Exercises

1. In each of the following, factor the matrix  $A$  into a product  $S\Lambda S^{-1}$ , with  $\Lambda$  diagonal:

a.  $A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$

b.  $A = \begin{pmatrix} 5 & 6 \\ -1 & -2 \end{pmatrix}$

c.  $A = \begin{pmatrix} 2 & -8 \\ 1 & -4 \end{pmatrix}$

d.  $A = \begin{pmatrix} 2 & 2 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & -1 \end{pmatrix}$

2. For each of the matrices in Exercise 1, use the  $S\Lambda S^{-1}$  factorization to calculate  $A^6$  (do *not* just multiply  $A$  by itself).
3. Solve  $X' = AX$  with initial condition  $X(0) = (-2, 9)$ , where  $A$  is the matrix in part (b). (You can try the others too; I just wrote answers for this part.)