A Brief Guide to H&S Chapter 5

(Notational note: H&S use “Det” for the determinant function, while most others, myself included, use “det”. The distinction is subtle and not worth worry. Also, “det” is easier to typeset.)

Our story thus far ...

Our current goal is to solve the system

\[ x' = A x, \]

where \( A \) is a constant \( n \times n \) matrix and \( x \in \mathbb{R}^n \). We will assume the results of the EUT (Chapter 8). For the purposes of Chapters 3, 4 and 5, we’ll consider \( n = 2 \) for specific examples, although it should be kept in mind that the results can be extended to higher \( n \).

Whatever technique we use to solve \( \clubsuit \), we will at some point have to find the eigenvalues and eigenvectors of \( A \), and keep track of them. [Disclaimer: It may often be possible to solve \( \clubsuit \) without explicit calculation of eigenvectors, but the resulting algebra is identical.] So, we’ll assume that we have obtained and solved the characteristic polynomial

\[ \det (A - \lambda I) = p(\lambda) = 0, \]

and identified as many eigenvectors as we can. For \( n = 2 \), solving \( p(\lambda) = 0 \) is high-school algebra, but we know the results of Appendix II, the Fundamental Theorem of Algebra, so we know that once we’ve used our high-school algebra, we’ve found all of the roots of the characteristic polynomial. Now, we use our solutions.

Chapter 3: \[ p(\lambda_1) = p(\lambda_2) = 0, \lambda_1, \lambda_2 \in \mathbb{R}, \lambda_1 \neq \lambda_2. \]

It’s easy to show that the corresponding eigenvectors \( f_1 \) and \( f_2 \) are linearly independent (Theorem 1, Pages 45-46), and so there exists a similarity transform

\[ B = Q A Q^{-1} \]
such that \( \mathbf{B} \) is diagonal, with the eigenvalues as the diagonal elements. **Theorem 2**
on Page 46 gives a constructive proof of this, namely showing that if the columns of \( \mathbf{P}^{-1} \) are the eigenvectors of \( \mathbf{A} \), then \( \mathbf{B} \) as given above has the desired form. Recall, however, that the proof of this theorem has a killer typo; see the notes on Diagonalization for Matrices with Real, Distinct Eigenvalues.

So, what’s the big whoop? Introduce a coordinate transformation \( \mathbf{y} = \mathbf{Q} \mathbf{x} \),

and note that

\[
\mathbf{x}' = \mathbf{A} \mathbf{Q}^{-1} \mathbf{Q} \mathbf{x}, \quad \mathbf{Q} \mathbf{x}' = \mathbf{Q} \mathbf{A} \mathbf{Q}^{-1} \mathbf{Q} \mathbf{x}, \quad \text{so that} \quad \mathbf{y}' = \mathbf{B} \mathbf{y}.
\]

But, since \( \mathbf{B} \) is diagonal, the last of the above string of equations can be solved for

\[
\mathbf{y} = \begin{bmatrix} e^{\lambda_1 t} & 0 \\ 0 & e^{\lambda_2 t} \end{bmatrix} \mathbf{y}_0 = \text{diag} \{ e^{\lambda_1 t}, e^{\lambda_2 t}, \ldots \} \mathbf{y}_0 = e^{\mathbf{B} t} \mathbf{y}_0.
\]

In the above, all off-diagonal elements in the displayed matrix are to be taken as zeros, and in the last step, the use of \( e^{\mathbf{B} t} \) is suggestive; the precise notation is a large part of what is done in Chapter 5. Our overall result is then

\[
\mathbf{x} = \mathbf{Q}^{-1} e^{\mathbf{B} t} \mathbf{Q} \mathbf{x}_0.
\]

**Chapter 4:** \( p(\lambda_1) = p(\lambda_2) = 0, \lambda_1, \lambda_2 \in \mathbb{C}, \lambda_1 \neq \lambda_2. \)

The fact is, the same techniques used in Chapter 3 for solving \( \mathbf{A} \) will work here, and much of the chapter is devoted to showing that everything we know about a real vector space can be extended to a complex vector space. However, there are other ways, and the text emphasizes the fact that we can solve real systems without resort to complex similar matrices. My opinion is that the crucial parts of the chapter are summarized in the figure on Page 56 (Chapter 3, of all places) and **Theorem 3** of Section 3.2. For detailed comments and an example regarding the latter, see the notes Similarity Transformations with Complex Eigenvalues. So, given a 2 \( \times \) 2 matrix \( \mathbf{A} \) with distinct complex eigenvalues, we can either:

1. Use a similarity transform as in Chapter 3, using the complex eigenvectors for the columns of \( \mathbf{Q}^{-1} \), resulting in a diagonal matrix, and recognizing that the matrix \( \mathbf{Q}^{-1} \text{diag} \{ \lambda_1, \lambda_2 \} \mathbf{Q} \) will be real. We did this explicitly only for the special case \( \mathbf{G} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \). For a more involved case, see the Addendum, the last page of these notes.
(2) Use a similarity transform as given in the Corollary on Page 68, and outlined in the notes Similarity Transformations with Complex Eigenvalues, resulting in a matrix of the form \[
\begin{bmatrix}
a & -b \\
b & a
\end{bmatrix}
\], where the eigenvalues are \(a \pm ib\), \(a\) and \(b\) real.

**Chapter 5:** Any possible solutions of \(p(\lambda) = 0\).

We have already seen that series formulations of the exponential function, \(e^x = \sum x^n/n!\), where the sum is from \(n = 0\) to \(\infty\), gives consistent results when the variable \(x\) is replaced by the purely imaginary \(i\theta\) or the matrix \(G t\), with \(G\) the “generator of rotations,” defined above (and elsewhere). The results were found to be

\[
e^{i\theta} = \cos \theta + i \sin \theta, \quad \exp (G t) = \begin{bmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{bmatrix} \equiv R(t).
\]

It has been hinted in passing, and could be shown with little difficulty, that for a diagonal matrix

\[
B = \begin{bmatrix}
\lambda_1 & 0 & 0 \\
0 & \lambda_2 & 0 \\
0 & 0 & \ddots
\end{bmatrix}, \quad \exp (B t) = \begin{bmatrix}
e^{\lambda_1 t} & 0 & 0 \\
0 & e^{\lambda_2 t} & 0 \\
0 & 0 & \ddots
\end{bmatrix}.
\]

We have also shown from the sum that for \(a\) and \(b\) real scalars,

\[
\exp \left( \begin{bmatrix} a & -b \\ b & a \end{bmatrix} t \right) = e^{at} R(bt).
\]

So Chapter 5 begins by formalizing this procedure when the argument of the exponential function is a matrix. To regard the infinite sum of matrices as anything other than a formal expression (B&C would call this a “formal sum”), we need to prove that the limit exists, and we only know how to find limits of scalars. So, we need to somehow decide on the “size” of a matrix.

We have already gone over Sections 1 & 2, where norms are “reviewed” and alternate norms exhibited for special cases. Section 2 is often not done, but it should be; any norm used for matrices might be suspect, but knowing that it doesn’t matter for the purpose of showing convergence should give some comfort. For all of the Propositions in this section, the propositions should be read and understood, and the proofs should be at least skimmed, but the details of the proofs need not be memorized.

**Section 5.3** should be read as carefully as possible. The *uniform norm*, defined at the bottom of Page 82, is used for many linear operators other than matrices.
There are some operators (but not on finite-dimensional spaces) for which this norm is not bounded, and such operators need to be watched carefully. I would say that this is one circumstance where the proofs to Lemma 1 should be read carefully and understood. Note that the previous work in Section 1 means that the proof of the Theorem at the bottom of Page 83 comes swiftly.

Lemma 2 is crucial, but the proof is a bit opaque. See the following page for a possibly helpful interpretation.

The Proposition on Pages 84-85 is crucial, but the proof of (d) is really wimpy. We’ve already done a better one in class, by explicit calculation. You should read and understand the remainder of the section.

Section 5.4 is more or less a recapitulation of everything done so far, with figures, and it should all be understood. If you see a similarity between what’s done in this section and what’s done in these notes, well, there’s a reason.

Due to our unorthodox skipping around (I’m looking at the “Warmup Assignment”), we’ve already covered Section 5.5, albeit with $n = 1$. Go over this again, and convince yourself that with $A$ a matrix instead of a scalar $A$, the derivation is unchanged, and the result (Equation (3) on Page 100) has the same form. Make sure you follow the example, recalling with pride that for this example $A = G$, $G$ being the generator of rotations, which should be begging for mercy by now.

I’ve read Section 5.6, but not critically. When I do, I’ll update these notes. For now, and to explain why I’m looking at this section carefully, go back to the beginning of the chapter, and read the last paragraph of the introduction, Pages 74-75. I have great reservations about the remarks regarding the dearth of encounters with higher-order systems. The authors qualify their remarks with “other kinds of applied mathematics” and “seldom”, but I’m not convinced and I’m not sure they’re convinced either. Stay tuned.
An Interpretation of the Proof of Lemma 2

The use of the symbols $\Sigma'$ and $\Sigma''$, with the associated subscripts, are better interpreted by arranging the terms $A_jB_k$ in a table, as suggested below; please note that although the presentation is similar, this is a table, not a matrix.

\[
\begin{array}{cccccccc}
A_0B_0 & A_0B_1 & \cdots & A_0B_{n-1} & A_0B_n & : & A_0B_{n+1} & \cdots & A_0B_{2n-1} & A_0B_{2n} \\
A_1B_0 & A_1B_1 & \cdots & A_1B_{n-1} & A_1B_n & : & A_1B_{n+1} & \cdots & A_1B_{2n-1} & \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \\
A_{n-1}B_0 & A_{n-1}B_1 & \cdots & A_{n-1}B_{n-1} & A_{n-1}B_n & : & A_{n-1}B_{n+1} & \\
A_{n}B_0 & A_{n}B_1 & \cdots & A_{n}B_{n-1} & A_{n}B_n & : & \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \\
A_{n+1}B_0 & A_{n+1}B_1 & \cdots & \\
\vdots & \vdots & \ddots & \\
A_{2n-1}B_0 & A_{2n-1}B_1 & \\
A_{2n} & \\
\end{array}
\]

The column of vertical dots and the row of horizontal dots is meant to suggest a division of the terms into a square and two triangles. The sum of the terms in the square (upper left terms) corresponds to the text’s product $\alpha_n\beta_n$, the sum of the terms in the upper right triangle is the sum $\Sigma' A_jB_k$ and the sum of the terms in the lower left triangle is $\Sigma'' A_jB_k$. Note that only terms with $j+k \leq 2n$ are included, as this is what is used in the definition of

\[
C = \lim_{n \to \infty} \gamma_{2n}.
\]
Addendum - Nontrivial Example of Diagonalization Using Complex Eigenvalues

A close look at the proof of Theorem 2 on Page 46 shows that the reality of the eigenvalues and eigenvectors was not needed. If the reality condition is not imposed, the result still holds, but of course the diagonal matrix \( \text{diag} \{ \lambda_1, \lambda_2, \ldots \} \) will not be real.

A specific example was done in class, with \( A = G = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \), the generator of rotations, and indeed we found that the solution to

\[
x' = Ax, \quad x(0) = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}
\]

was

\[
x(t) = \begin{bmatrix} c_1 \cos t & -c_2 \sin t \\ c_1 \sin t & c_2 \cos t \end{bmatrix}.
\]

For another example, consider again the matrix \( \begin{bmatrix} -1 & -1 \\ 1 & 0 \end{bmatrix} \), with eigenvalues \( \mu_{1,2} = (-1 \pm i\sqrt{3})/2 \) and corresponding eigenvectors

\[
f_1 = \begin{bmatrix} \mu_1 \\ 1 \end{bmatrix}, \quad f_2 = \begin{bmatrix} \mu_2 \\ 1 \end{bmatrix}
\]

(note that \( f_1 = f_2 \)). Then, proceeding as in Chapter 3, we have

\[
Q^{-1} = \begin{bmatrix} \mu_1 & \mu_2 \\ 1 & 1 \end{bmatrix}, \quad Q = \frac{1}{\mu_1 - \mu_2} \begin{bmatrix} 1 & -\mu_2 \\ 1 & \mu_1 \end{bmatrix}.
\]

A basic calculation shows that

\[
Q A Q^{-1} = Q \begin{bmatrix} \mu_1 f_1 & \mu_2 f_2 \end{bmatrix} = Q \begin{bmatrix} \mu_1^2 & \mu_2^2 \\ \mu_1 & \mu_2 \end{bmatrix} = \frac{1}{\mu_1 - \mu_2} \begin{bmatrix} 1 & -\mu_2 \\ 1 & \mu_1 \end{bmatrix} \begin{bmatrix} \mu_1^2 & \mu_2^2 \\ \mu_1 & \mu_2 \end{bmatrix} = \begin{bmatrix} \mu_1 & 0 \\ 0 & \mu_2 \end{bmatrix}.
\]

At this point, the calculation of the matrix

\[
Q^{-1} \begin{bmatrix} e^{\mu_1 t} & 0 \\ 0 & e^{\mu_2 t} \end{bmatrix} Q
\]

is more than a little bit tedious; see the notes More on Complex Eigenvalues, linked from the 18.03(4)-ESG page, but be aware that those notes use a variation on the E&P notation, with \( \lambda_\alpha, \lambda_\beta \) instead of \( \mu_1, \mu_2 \).