## Chapter 11

## Appendix C: A compendium of results from linear algebra

### 11.1 How to compute Jordan normal forms

In a linear algebra text, one expects the author to prove that an arbitrary square matrix is similar to a Jordan canonical form, and this proof is a messy affair ${ }^{1}$. We assume the reader has seen the definitions and the statement of the theorem but not really followed the proof. Here we accept that a Jordan normal form exists, and we ask, more simply, how to compute it. We break this problem into two sub-questions, focusing more on examples than theory: given a matrix $A$,

Q1: How can we decide what the normal form of $A$ is?
Q2: How can we find the similarity transformation that produces the normal form?
The first step in determining the normal form of $A$ is to find the eigenvalues of $A$. Of course finding eigenvalues analytically is an intractable problem in general. We work with hand-picked examples in which the eigenvalues are readily computed.

Example 1:

$$
A=\left[\begin{array}{cc}
5 & -2 \\
2 & 1
\end{array}\right]
$$

It is readily computed that $\operatorname{det}(A-\lambda I)=(\lambda-3)^{2}$. Thus, there are two possible Jordan forms for $A$,

$$
\mathbf{J}_{1}=\left[\begin{array}{ll}
3 & 0 \\
0 & 3
\end{array}\right] \quad \text { and } \quad \mathbf{J}_{2}=\left[\begin{array}{ll}
3 & 1 \\
0 & 3
\end{array}\right] .
$$

[^0]If $A$ were similar to $\mathbf{J}_{1}=3 I$, then every vector in $\mathbb{R}^{2}$ would be an eigenvector. However $\mathbf{v} \in \mathbb{R}^{2}$ is an eigenvector iff $(A-3 I) \mathbf{v}=0$, or writing this out

$$
\left[\begin{array}{ll}
2 & -2 \\
2 & -2
\end{array}\right] \mathbf{v}=0 .
$$

Obviously not every vector satisfies this equation, so $\mathbf{J}_{2}$ must be the normal form for $A$. Indeed, in hindsight we may see that if a $2 \times 2$ matrix has equal eigenvalues but is not equal to a multiple of the identity, then its Jordan normal form must be a $2 \times 2$ block.

Higher-dimensional examples in which there are double eigenvalues, but none of higher multiplicity, do not pose any additional difficulties, as we illustrate in some of the Exercises. Let us turn our attention to eigenvalues of multiplicity three.

Example 2: Consider

$$
A_{1}=\left[\begin{array}{ccc}
a & 1 & 1 \\
0 & a & 0 \\
0 & 0 & a
\end{array}\right] \quad A_{2}=\left[\begin{array}{ccc}
a & 1 & 1 \\
0 & a & 1 \\
0 & 0 & a
\end{array}\right] \quad A_{3}=\left[\begin{array}{ccc}
a & 0 & 1 \\
0 & a & 1 \\
0 & 0 & a
\end{array}\right]
$$

By inspection, $\lambda=a$ is the only eigenvalue of $A_{j}$. Thus the possible normal forms for $A_{j}$ are

$$
\mathbf{J}_{1}=\left[\begin{array}{ccc}
a & & \\
& a & \\
& & a
\end{array}\right] \quad \mathbf{J}_{2}=\left[\begin{array}{ccc}
a & 1 & \\
0 & a & \\
& & a
\end{array}\right] \quad \mathbf{J}_{3}=\left[\begin{array}{ccc}
a & 1 & 0 \\
0 & a & 1 \\
0 & 0 & a
\end{array}\right]
$$

where, to facilitate visualization, entries that are zero but lie outside of any Jordan block are left blank. We distinguish between cases by examining the dimension of the eigenspaces. These dimensions may be computed most easily by applying the "rank-plus-nullity" theorem (see Strang [?]), which gives us

$$
\operatorname{dim} \operatorname{ker}\left(\mathbf{J}_{j}-a I\right)=3-\operatorname{rank}\left(\mathbf{J}_{j}-a I\right)
$$

Thus $\mathbf{J}_{1}, \mathbf{J}_{2}, \mathbf{J}_{3}$ have eigenspaces of dimension $3,2,1$, respectively. Proceeding similarly, we find that $A_{1}, A_{2}, A_{3}$ have eigenspaces of dimension $2,1,2$, respectively. Since the dimension of eigenspaces is preserved under similarity transformations, we conclude that $A_{1}, A_{2}, A_{3}$ have Jordan forms $\mathbf{J}_{2}, \mathbf{J}_{3}, \mathbf{J}_{2}$, respectively.

Example 3:

$$
A_{1}=\left[\begin{array}{cccc}
a & 0 & 0 & 1 \\
0 & a & 0 & 1 \\
0 & 0 & a & 0 \\
0 & 0 & 0 & a
\end{array}\right] \quad A_{2}=\left[\begin{array}{cccc}
a & 0 & 1 & 0 \\
0 & a & 0 & 1 \\
0 & 0 & a & 0 \\
0 & 0 & 0 & a
\end{array}\right] \quad A_{3}=\left[\begin{array}{cccc}
a & 1 & 0 & 0 \\
0 & a & 0 & 1 \\
0 & 0 & a & 0 \\
0 & 0 & 0 & a
\end{array}\right]
$$

The possible Jordan forms are

$$
\begin{gathered}
\mathbf{J}_{1}=\left[\begin{array}{llll}
a & & & \\
& a & & \\
& & a & \\
& & & a
\end{array}\right] \quad \mathbf{J}_{2}=\left[\begin{array}{llll}
a & 1 & & \\
0 & a & & \\
& & a & \\
& & & a
\end{array}\right] \quad \mathbf{J}_{3}=\left[\begin{array}{llll}
a & 1 & 0 & \\
0 & a & 1 & \\
0 & 0 & a & \\
& & & a
\end{array}\right] \\
\mathbf{J}_{4}=\left[\begin{array}{lllll}
a & 1 & 0 & 0 \\
0 & a & 1 & 0 \\
0 & 0 & a & 1 \\
0 & 0 & 0 & a
\end{array}\right] \quad \mathbf{J}_{5}=\left[\begin{array}{llll}
a & 1 & & \\
0 & a & & \\
& & a & 1 \\
& 0 & a
\end{array}\right]
\end{gathered}
$$

Proceeding as above, we compute that $\mathbf{J}_{1}, \mathbf{J}_{2}, \mathbf{J}_{3}, \mathbf{J}_{4}, \mathbf{J}_{5}$ have eigenspaces of dimension $4,3,2,1,2$, respectively. We can see potential trouble here in that $\mathbf{J}_{3}$ and $\mathbf{J}_{5}$ both have two-dimensional eigenspaces. Now $A_{1}, A_{2}, A_{3}$ have eigenspaces of dimension $3,2,2$ respectively. Thus we may conclude that $A_{1}$ has $\mathbf{J}_{2}$ as its normal form, but the dimension of the eigenspace does not distinguish between $\mathbf{J}_{3}$ and $\mathbf{J}_{5}$ for $A_{2}$ and $A_{3}$. For this task we turn to generalized eigenvectors: a vector $\mathbf{v} \in \mathbb{R}^{d}$ is called a generalized eigenvector of a matrix $A$ with eigenvalue $\lambda$ if for some power $p$

$$
(A-\lambda I)^{p} \mathbf{v}=0
$$

Choosing $p=2$, we compute that $\left(\mathbf{J}_{j}-a I\right)^{2}$ has a three-dimensional null space for $j=3$ and a four-dimensional null space for $j=5$. On the other hand, $\left(A_{j}-a I\right)^{2}$ has a four-dimensional null space if $j=2$ and a three-dimensional null space if $j=3$. Thus the normal forms for $A_{2}, A_{3}$ are $\mathbf{J}_{5}, \mathbf{J}_{3}$, respectively.

Now we turn to the second question above, finding the similarity matrix $S$ such that $S^{-1} A S$ produces the Jordan form of $A$. As we shall see, the columns of $S$ are generalized eigenvectors of $A$ (cf. Proposition ??).

Recall Example 1, where

$$
A=\left[\begin{array}{cc}
5 & -2 \\
2 & 1
\end{array}\right], \quad \text { with } J=\left[\begin{array}{ll}
3 & 1 \\
0 & 3
\end{array}\right]
$$

its Jordan form. Observe that, with respect to the standard basis $\mathbf{e}_{1}, \mathbf{e}_{2}$ for $\mathbb{R}^{2}$, the matrix $J$ satisfies

$$
(J-3 I) \mathbf{e}_{1}=0 \quad(J-3 I) \mathbf{e}_{2}=\mathbf{e}_{1} .
$$

To match this behavior for $A$, we need to find vectors $\mathbf{v}_{1}, \mathbf{v}_{2}$ such that

$$
(A-3 I) \mathbf{v}_{1}=0 \quad(A-3 I) \mathbf{v}_{2}=\mathbf{v}_{1}
$$

and then the matrix $S=\operatorname{Col}\left(\mathbf{v}_{1}, \mathbf{v}_{2}\right)$ will achieve the required transformation. (Note that $(A-3 I)^{2} \mathbf{v}_{2}=0$, so $\mathbf{v}_{2}$ is a generalized eigenvector.) One possible choice is

$$
S=\left[\begin{array}{cc}
1 & 1 / 2 \\
1 & 0
\end{array}\right]
$$

In the Exercises we ask the reader to check that this matrix performs the desired task. Incidentally, there is great latitude in the choice of $S$, more so than in the case of distinct eigenvalues.

More subtle issues may arise in cases of higher multiplicity. Let $A$ be the first of the three matrices considered in that Example 2, and let $J$ be its Jordan form. Observe that $J$ satisfies

$$
(J-a I) \mathbf{e}_{1}=0, \quad(J-a I) \mathbf{e}_{2}=\mathbf{e}_{1}, \quad(J-a I) \mathbf{e}_{3}=0
$$

Thus we need to find vectors $\mathbf{v}_{1}, \mathbf{v}_{2}, \mathbf{v}_{3}$ such that

$$
\begin{equation*}
(A-a I) \mathbf{v}_{1}=0, \quad(A-a I) \mathbf{v}_{2}=\mathbf{v}_{1}, \quad(A-a I) \mathbf{v}_{3}=0 \tag{11.1}
\end{equation*}
$$

and let $S=\operatorname{Col}\left(\mathbf{v}_{1}, \mathbf{v}_{2}, \mathbf{v}_{3}\right)$. Note that $\mathbf{v}_{1}$ and $\mathbf{v}_{3}$ are eigenvectors of $A$, but $\mathbf{v}_{1}$ must be chosen with care in order that the middle equation in (11.1), which is inhomogeneous, has a solution. Now the eigenspace of $A$ is spanned by

$$
\left[\begin{array}{l}
1 \\
0 \\
0
\end{array}\right] \quad\left[\begin{array}{c}
0 \\
1 \\
-1
\end{array}\right] .
$$

Suppose $\mathbf{v}_{1}$ is a linear combination of these vectors with coefficients $\alpha, \beta$. Writing out the middle equation in (11.1), we have

$$
\left[\begin{array}{lll}
0 & 1 & 1 \\
0 & 0 & 0 \\
0 & 0 & 0
\end{array}\right]\left[\begin{array}{l}
x \\
y \\
z
\end{array}\right]=\left[\begin{array}{c}
\alpha \\
\beta \\
-\beta
\end{array}\right] .
$$

To have a solution we need $\beta=0$; to avoid trivialities we need $\alpha \neq 0$. Thus

$$
S=\left[\begin{array}{ccc}
1 & 0 & 0 \\
0 & 1 & 1 \\
0 & 0 & -1
\end{array}\right]
$$

where we have chosen $\alpha=1$, is one of the possible similarity matrices that transforms $A$ to its Jordan form.

In the Exercises we ask the reader to carry out this procedure and check that it works for several of the matrices considered above.

### 11.2 The Routh-Hurwitz criterion

It is astonishingly easy to determine whether a polynomial with real coefficients has all its zeros in the left-half-plane. For example, for the two polynomials

$$
\begin{array}{llc}
Q_{1}(\lambda) & = & \lambda^{4}+2 \lambda^{3}+3 \lambda^{2}+2 \lambda+1 \\
Q_{2}(\lambda) & = & \lambda^{5}+2 \lambda^{4}+3 \lambda^{3}+3 \lambda^{2}+2 \lambda+1,
\end{array}
$$

the calculations in Table 11.2 show that the first has all its zeros in $\{\Re \lambda<0\}$ while the second has at least one zero in $\{\Re \lambda \geq 0\}$, respectively. Let us explain these calculations in the context of a general polynomial

$$
P(\lambda)=\lambda^{n}+c_{1} \lambda^{n-1}+c_{2} \lambda^{n-2}+\ldots+c_{n-1} \lambda+c_{n} .
$$

The algorithm is slightly different, depending on whether $n$ is even or odd. Reflecting this difference we define $\nu=[n / 2]$ where [•] is the greatest-integer function: thus $n=2 \nu$ if $n$ is even and $n=2 \nu+1$ if $n$ is odd. The algorithm forms an $(n+1) \times(\nu+1)$ matrix $A$ as follows. The first two rows of $A$ contain the coefficients of even and odd powers of $\lambda$ :

$$
\begin{array}{rrrrl}
a_{1 l}: & 1 & c_{2} & c_{4} & \ldots \\
a_{2 l}: & c_{1} & c_{3} & c_{5} & \ldots
\end{array} .
$$

(If $n$ is even, then 0 is inserted as the last entry of the second row, as in the table on the left.) Subsequent rows, $3,4, \ldots, n+1$, are calculated inductively from products that resemble $2 \times 2$ determinants

$$
\begin{equation*}
a_{k+1, l}=a_{k, l} a_{k-1, l+1}-a_{k, l+1} a_{k-1, l} . \tag{11.2}
\end{equation*}
$$

In words, computation of $a_{k+1, l}$ involves selecting entries from the two preceding rows and from the same column as $a_{k+1, l}$ and the one to the right. In calculating the last column $(l=\nu+1)$, entries $a_{k, \nu+2}$ or $a_{k-1, \nu+2}$ outside the appropriate range are assumed to be zero, as has been done in Table 11.2. Then we have:

Theorem 11.2.1. All the zeros of $P$ lie in the open left half plane iff all entries $a_{k 1}, k=1, \ldots, n+1$, in the first column of the above matrix are positive.

If the calculation produces a zero row, as in the table on the right, then the calculation is stopped and there is at least one zero in closed right half plane. Indeed, note that $Q_{2}( \pm i)=0$. A root of $P$ on the imaginary axis will cause a zero row, but a zero row may arise under other circumstances, also.

This theorem is proved in Section 4.2 of Engelberg's book. Although the proof requires careful reading, it is not terribly difficult, just clever. In cases where some of the zeros of $P(\lambda)$ lie in the right half plane, it is usually possible to deduce how many zeros lie there.

|  | 1 | 2 | 3 |
| :--- | :--- | :--- | :--- |
| 1 | 1 | 3 | 1 |
| 2 | 2 | 2 | 0 |
| 3 | 4 | 2 | 0 |
| 4 | 4 | 0 | 0 |
| 5 | 8 | 0 | 0 |


|  | 1 | 2 | 3 |
| :--- | :---: | :---: | :---: |
| 1 | 1 | 3 | 2 |
| 2 | 2 | 3 | 1 |
| 3 | 3 | 3 | 0 |
| 4 | 3 | 3 | 0 |
| 5 | 0 | 0 | 0 |
| 6 | - | - | - |

Table 11.1: The matrices $\left\{a_{k l}\right\}$ in the Routh-Hurwitz calculations for $Q_{1}(\lambda)=\lambda^{4}+2 \lambda^{3}+3 \lambda^{2}+2 \lambda+1$ (left table) and $Q_{2}(\lambda)=\lambda^{5}+2 \lambda^{4}+3 \lambda^{3}+3 \lambda^{2}+2 \lambda+1$ (right table). Values of $k$ from 1 to $n+1$ appear in the first column of each table; values for $l$ from 1 to $\nu+1$ appear in the top row. The two rows $\left\{a_{k l}: l=1,2\right\}$, which come directly from the coefficients of the polynomial, are separated from later rows that come from the calculation indicated in (11.2).

Reference: Shlomo Engelberg, A Mathematical Introduction to Control Theory, Series in Electrical and Computer Engineering, Vol. 2, Imperial College Press, London, 2005. Duke Catalogue QA402.3.E527.

If $A$ is a $d \times d$ matrix with real entries, then in principle one could calculate the characteristic polynomial of $A$ and apply the Routh-Hurwitz criterion to it to determine whether the eigenvalues of $A$ lie in the left half plane. However, calculating the characteristic polynomial of a moderately large matrix by hand is not a pleasant task. (One could of course resort to symbolic computations to obtain the characteristic polynomial, but if the computer is involved, one might as well compute eigenvalues directly.) Thus we refrain even from formulating an analogue of Theorem ?? for $4 \times 4$ matrices. However, let us use the Routh-Hurwitz criterion to handle the $3 \times 3$ case.

Proof of Proposition ??. Let $A$ be a $3 \times 3$ matrix with characteristic polynomial

$$
\operatorname{det}(A-\lambda I)=-\left[\lambda^{3}+c_{1} \lambda^{2}+c_{2} \lambda+c_{3}\right]
$$

The table applying the Routh-Hurwitz criterion to this polynomial is shown in Table 11.2. Thus the roots of this polynomial are all in the left half plane iff

$$
\begin{equation*}
\text { (a) } c_{1}>0, \quad \text { (b) } c_{1} c_{2}-c_{3}>0, \quad \text { (c) } c_{3}>0 \tag{11.3}
\end{equation*}
$$

|  | 1 | 2 |
| :---: | :---: | :---: |
| 1 | 1 | $c_{2}$ |
| 2 | $c_{1}$ | $c_{3}$ |
| 3 | $c_{1} c_{2}-c_{3}$ | 0 |
| 4 | $c_{3}\left(c_{1} c_{2}-c_{3}\right)$ | 0 |

Table 11.2: The matrix $\left\{a_{k l}\right\}$ in the Routh-Hurwitz calculations for the general cubic $\lambda^{3}+c_{1} \lambda^{2}+c_{2} \lambda+c_{3}$.

The coefficients $c_{j}$ are related to the eigenvalues of $A$ through

$$
\begin{array}{ccc}
c_{1} & = & -\left(\lambda_{1}+\lambda_{2}+\lambda_{3}\right) \\
c_{2} & = & \lambda_{1} \lambda_{2}+\lambda_{2} \lambda_{3}+\lambda_{3} \lambda_{1} \\
c_{3} & = & -\lambda_{1} \lambda_{2} \lambda_{3} .
\end{array}
$$

Thus it is apparent that (11.3a) and (c) are equivalent to Conditions (i) and (iii) of Theorem ??, and the equivalence of (11.3b) with Condition (ii) follows on observing that

$$
c_{2}=\frac{1}{2}\left[(\operatorname{tr} A)^{2}-\operatorname{tr}\left(A^{2}\right)\right] .
$$

### 11.3 Continuity of eigenvalues of a matrix with respect to its entries

Near simple eigenvalues, the dependence of eigenvalues of a matrix on its entries is as nice as one could wish. Specifically, we have the following

Proposition 11.3.1. Let $\lambda_{1}$ be a simple eigenvalue of a $d \times d$ matrix $A_{0}$. There is an $\varepsilon>0$ and a neighborhood $\mathcal{U}$ of $A_{0}$ in $\mathbb{R}^{d^{2}}$ such any matrix $A \in \mathcal{U}$ has exactly one eigenvalue in the disk $\left\{\left|z-\lambda_{1}\right|<\varepsilon\right\}$, and moreover this eigenvalue is a differentiable function on $\mathcal{U}$.

Proof. We prove this result by applying the implicit-function theorem to solve for $\lambda$ in the equation for eigenvalues,

$$
f(\lambda, A)=\operatorname{det}(A-\lambda I)=0
$$

Now $f\left(\lambda, A_{0}\right)$ is a product of eigenvalues $\left(\lambda_{1}-\lambda\right) \ldots\left(\lambda_{d}-\lambda\right)$. Differentiation of this
product with respect to $\lambda$ gives $d$ terms, but only one of them is nonzero at $\lambda=\lambda_{1}$ :

$$
\frac{\partial f}{\partial \lambda}\left(\lambda_{1}, A_{0}\right)=-\left(\lambda_{2}-\lambda_{1}\right) \ldots\left(\lambda_{d}-\lambda_{1}\right)
$$

Since $\lambda_{1}$ is a simple eigenvalue, this product is nonzero, which completes the proof.

Suppose that $A_{0}$ has a simple eigenvalue $\lambda_{1}$, and consider a one-parameter family of perturbations, $A_{0}+\varepsilon B$. It follows from the proposition that there is a smoothly varying eigenvalue $\lambda_{1}\left(A_{0}+\varepsilon B\right)$ that equals $\lambda_{1}$ when $\varepsilon=0$. In Exercise ?? we give a formula for calculating the derivative of this eigenvalue with respect to $\varepsilon$ at $\varepsilon=0$.

Near multiple eigenvalues, the dependence of eigenvalues of a matrix on its entries is complicated by a difficulty familiar from complex function theory. For example, consider the matrix function

$$
A(\alpha, \beta)=\left[\begin{array}{cc}
0 & 1 \\
\alpha+i \beta & 0
\end{array}\right]
$$

which has eigenvalues $\lambda_{j}(\alpha, \beta)= \pm \sqrt{\alpha+i \beta}$. We claim it is impossible to define these square roots as continuous functions of $\alpha, \beta$ in a neighborhood of zero in $\mathbb{R}^{2}$. To see this, suppose otherwise that there are continuous eigenvalues $\lambda_{j}(\alpha, \beta)$. On the positive $\alpha$-axis, the eigenvalues are $\pm \sqrt{\alpha}$. Index the eigenvalues so that $\lambda_{1}$ is positive on the positive $\alpha$-axis, and let us restrict $\lambda_{1}$ to a small circle that encloses the origin: i.e., let

$$
\Lambda(\phi)=\lambda_{1}(\varepsilon \cos \phi, \varepsilon \sin \phi), \quad \text { where } 0 \leq \phi<2 \pi .
$$

Calcluation then shows that

$$
\begin{equation*}
\Lambda(\phi)=\sqrt{\varepsilon} e^{i \phi / 2} \tag{11.4}
\end{equation*}
$$

If $\lambda_{1}$ were continuous we would have $\Lambda(2 \pi)=\Lambda(0)$, but in fact (11.4) implies that $\Lambda(2 \pi)=-\Lambda(0)$. This contradiction proves the claim.

The reader may protest that the matrix in this example has complex entries, but here is a $4 \times 4$ matrix with real entries that exhibits the same difficulty:

$$
A(\alpha, \beta)=\left[\begin{array}{cc}
0 & I \\
B(\alpha, \beta) & 0
\end{array}\right]
$$

where 0 is the $2 \times 2$ zero matrix, $I$ is the $2 \times 2$ identity matrix, and

$$
B(\alpha, \beta)=\left[\begin{array}{cc}
\alpha & -\beta \\
\beta & \alpha
\end{array}\right]
$$

Although at a multiple eigenvalue, one cannot define individual eigenvalues con-
tinuously, nonetheless the group of eigenvalues does vary continuously, in the sense of the following

Proposition 11.3.2. Let $\lambda_{1}$ be an eigenvalue of a $d \times d$ matrix $A_{0}$ of multiplicity $k$. There is an $\varepsilon_{0}>0$ with the property that for all $\varepsilon<\varepsilon_{0}$ there is a neighborhood $\mathcal{U}$ of $A_{0}$ in $\mathbb{R}^{d^{2}}$ such any matrix $A \in \mathcal{U}$ has exactly $k$ eigenvalue in the disk $\left\{\left|z-\lambda_{1}\right|<\varepsilon\right\}$.

The conclusion of this result is a standard epsilon-delta characterization of continuity, with three differences: (i) a set of eigenvalues, rather than a single eigenvalue, is being bounded, (ii) an upper bound is needed on epsilon, and (iii) delta is relaced by the neighborhood $\mathcal{U}$. Two remarks: (i) For $\varepsilon_{0}$ one may use any number less than the minimum separation between $\lambda_{1}$ and the other eigenvalue of $A_{0}$. (ii) The maximum diameter of $\mathcal{U}$ scales like $\varepsilon^{1 / k}$ as $\varepsilon \rightarrow 0$.

The proposition is easily proved with complex-function theory, but, since this subject is not a pre-requisite for this text, we do not give the proof here.

Corollary 11.3.3. Suppose all the eigenvalues of $A$ lie in the left half plane $\{\Re \lambda<$ $0\}$. For any perturbation matrix B, for sufficiently small $\varepsilon$ all the eigenvalues of $A+\varepsilon B$ also lie in the left half plane.

The reader is asked to derive this corollary in the Exercises.
If one restricts attention to symmetric matrices, then all eigenvalues are real, and one may define individual eigenvalues continuously by ordering them. For example, we may define $\lambda_{1}(A)$ to be the smallest eigenvalue of $A ; \lambda_{2}(A)$ to be the next smallest eigenvalue; etc. (Ties do not matter for these definitions.) However, even though with this convention the eigenvalues are continuous, they are not differentiable: this is demonstrated by the matrix

$$
A(\alpha, \beta)=\left[\begin{array}{cc}
\alpha & \beta \\
\beta & -\alpha
\end{array}\right]
$$

which has eigenvalues $\pm \sqrt{\alpha^{2}+\beta^{2}}$.

### 11.4 Fast-slow systems

Recall problem from Chapter 1. More general system

$$
\begin{aligned}
\varepsilon x^{\prime} & =-x+\mathbf{b}^{T} \mathbf{y} \\
\mathbf{y}^{\prime} & =x \mathbf{c}+B \mathbf{y}
\end{aligned}
$$

Coefficient matrix

$$
A=\left[\begin{array}{cc}
-\varepsilon^{-1} & \varepsilon^{-1} \mathbf{b}^{T} \\
\mathbf{c} & B
\end{array}\right]
$$

Compare two approaches:

1. Solve for $x=\mathbf{b}^{T} \mathbf{y}$. Substitute into $\mathbf{y}$-eqn, get reduced system

$$
\mathbf{y}^{\prime}=\left(B+\mathbf{c b}^{T}\right) \mathbf{y} .
$$

2. Full system.

Full system has one e-value approx equal to $-\varepsilon^{-1}$. Show other eigenvalues are same, modulo $\varepsilon$ to some fractional power. In particular, reduced system is stable iff full system is.

### 11.5 Exercises

Prove Corollary 11.3.3.

Suppose that $A_{0}$ has a simple eigenvalue $\lambda_{1}$, and let $\Lambda(\varepsilon)$ be the eigenvalue of $A_{0}+\varepsilon B$ that equals $\lambda_{1}$ when $\varepsilon=0$. In Exercise ?? we give a formula for calculating $d \Lambda / d \varepsilon(0)$.

1. In case $A_{0}$ has block-diagonal form, show derivative is 1,1 -entry of $B$.
2. Suppose $S^{-1} A_{0} S$ is block diagonal. Note that first col of $S$ is eigenvector, say $\mathbf{v}$, of $A_{0}$, first row of $S^{-1}$ is eigenvector, say $\mathbf{w}$, of $A^{T}$. Argue that

$$
d \Lambda / d \varepsilon(0)=\langle\mathbf{w}, B \mathbf{v}\rangle
$$


[^0]:    ${ }^{1}$ One of the best treatments of Jordan forms of which we are aware is in Appendix B of Strang [?]

