MTHE 237 — PROBLEM SET 11 SOLUTIONS

1. When a matrix A is diagonalizable, the corresponding diagonal matrix is the Jordan canonical form of A. Therefore, we already have a way of finding a basis that puts a matrix into its Jordan canonical form in the diagonalizable case — take the eigenvectors of the matrix as the basis.

The purpose of this problem is to develop a way of finding a basis of $\mathbb{R}^2$ in which a nondiagonalizable 2×2 matrix is in its Jordan canonical form.

Let A be a 2×2 matrix whose characteristic polynomial $p_A(z) = \det(A - zI) = (z - \lambda)^2$ has λ as a real double root.

Let

$$E_\lambda = \left\{ \mathbf{v} \in \mathbb{R}^2 : A\mathbf{v} = \lambda \mathbf{v} \right\}$$

be the eigenspace corresponding to λ.

Suppose that

$$\dim E_\lambda = 1.$$

Thus, λ is an eigenvalue of A with algebraic multiplicity 2 and geometric multiplicity 1.

The following theorem has many applications.

Thm (Cayley-Hamilton). *Let $p_A(z)$ be the characteristic polynomial of A. Then $p_A(A) = 0$.*

In our case, $p_A(z) = (z - \lambda)^2$, so the Cayley-Hamilton theorem says that $(A - \lambda I)^2 = 0$. This last fact is useful in finding the Jordan basis.

To be self-contained, let's begin by showing that $(A - \lambda I)^2 = 0$ with an elementary computation, without use of the Cayley-Hamilton theorem. Label the entries of A as follows:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

i) Show that

$$p_A(z) = z^2 - (a + d)z + (ad - bc).$$

The function $a + d$ of the entries of the matrix is called the *trace of* A and the function $ad - bc$ is called the *determinant of* A.

ii) Conclude from the hypothesis $p_A(z) = (z - \lambda)^2$ that

$$\begin{aligned} a + d &= 2\lambda \\ ad - bc &= \lambda^2. \end{aligned} \tag{1}$$

iii) Show that the two equalities of (1) imply that $\dfrac{(a - d)^2}{4} + bc = 0$.

iv) Show that $\lambda = (a + d)/2$ and, using the result of part iii), conclude that

$$(A - \lambda I)^2 = 0.$$

Now, let $\mathbf{w}$ be any nonzero vector in $\mathbb{R}^2$ but not in E_λ (such vectors exist, because $\mathbb{R}^2$ is two-dimensional and E_λ is one-dimensional). Let $\mathbf{v} = (A - \lambda I)\mathbf{w}$.

v) Show that $A\mathbf{w} = \mathbf{v} + \lambda\mathbf{w}$.

vi) Show that $(A - \lambda I)\mathbf{v} = 0$. Conclude that $A\mathbf{v} = \lambda\mathbf{v}$. Conclude that $\mathbf{v} \in E_\lambda$, and that $\mathbf{v}$ and $\mathbf{w}$ are linearly independent.

vii) Conclude that the matrix of A with respect to the ordered basis $\{\mathbf{v}, \mathbf{w}\}$ is

$$[A]_{\{\mathbf{v},\mathbf{w}\}} = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}.$$

Therefore, A is in Jordan canonical form in the basis $\{\mathbf{v}, \mathbf{w}\}$!

viii) Find a basis that puts the matrix

$$B = \begin{pmatrix} 10 & 4 \\ -9 & -2 \end{pmatrix}.$$

into Jordan canonical form. (Find the eigenvalue of B and its eigenspace, then choose a nonzero $\mathbf{w}$ outside of the eigenspace, and take $\mathbf{v} = (B - \lambda I)\mathbf{w}$.)

ix) Show that

$$\exp(Bt) = \begin{pmatrix} e^{4t} + 6t\,e^{4t} & 4t\,e^{4t} \\ -9t\,e^{4t} & e^{4t} - 6t\,e^{4t} \end{pmatrix}.$$

x) Sketch the flow lines of the following first-order system, and find a parametrization of the flow line through the given point:

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} 10 & 4 \\ -9 & -2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}, \qquad x(0) = 2,\ y(0) = -1.$$

The way one finds Jordan canonical forms of larger matrices has a similar flavour, building bases out of chains of vectors of the form $(A - \lambda I)^k\mathbf{w}$. A complete treatment, including a proof of existence of Jordan canonical form for an arbitrary matrix with complex entries, is done in Math 212. Alternatively, a good reference is the textbook Friedberg, Insel, Spence, *Linear Algebra*, §7.1–7.2.

Solution. i) We have

$$\begin{aligned} p_A(z) &= \det\begin{pmatrix} a - z & b \\ c & d - z \end{pmatrix} \\ &= (a - z)(d - z) - bc \\ &= ad - (a + d)z + z^2 - bc \\ &= z^2 - (a + d)z + (ad - bc). \end{aligned}$$

ii) Expanding $(z-\lambda)^2$, we have

$$(z-\lambda)^2 = z^2 - 2\lambda z + \lambda^2.$$

Because $\{1, z, z^2\}$ is a basis for the vector space of polynomials with real (or complex) coefficients of degree ≤ 2, we can compare coefficients of both sides of the equality

$$z^2 - (a+d)z + (ad-bc) = (z-\lambda)^2 = z^2 - 2\lambda z + \lambda^2$$

and conclude that

$$\begin{aligned} a+d &= 2\lambda \\ ad-bc &= \lambda^2. \end{aligned}$$

Alternatively, evaluate both polynomials at $z=0$ to conclude that $ad-bc=\lambda^2$. Then, differentiate both polynomials (getting $2z-(a+d) = 2z-2\lambda$) and evaluate the result at $z=0$ to conclude that $a+d=2\lambda$.

iii) From the equality $a+d=2\lambda$, we have $(a+d)/2=\lambda$. Therefore,

$$\frac{(a+d)^2}{4} = \lambda^2 = ad-bc.$$

Bringing all of the terms to the left side, we get

$$0 = \frac{(a^2+2ad+d^2)}{4} - ad + bc = \frac{(a^2-2ad+d^2)}{4} + bc = \frac{(a-d)^2}{4} + bc.$$

iv) The matrix of $(A-\lambda I)$ is

$$\begin{pmatrix} a-\frac{a+d}{2} & b \\ c & d-\frac{a+d}{2} \end{pmatrix} = \begin{pmatrix} \frac{a-d}{2} & b \\ c & \frac{d-a}{2} \end{pmatrix}.$$

Squaring, we get

$$\begin{aligned} (A-\lambda I)^2 = \begin{pmatrix} \frac{a-d}{2} & b \\ c & \frac{d-a}{2} \end{pmatrix}\begin{pmatrix} \frac{a-d}{2} & b \\ c & \frac{d-a}{2} \end{pmatrix} &= \begin{pmatrix} \frac{(a-d)^2}{4} + bc & \frac{a-d}{2}b + b\frac{d-a}{2} \\ c\frac{a-d}{2} + \frac{d-a}{2}c & bc + \frac{(d-a)^2}{4} \end{pmatrix} \\ &= \begin{pmatrix} \frac{(a-d)^2}{4} + bc & \frac{a-d}{2}b - b\frac{a-d}{2} \\ c\frac{a-d}{2} - \frac{a-d}{2}c & bc + \frac{(a-d)^2}{4} \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \end{aligned}$$

v) By construction, $\mathbf{v} = (A-\lambda I)\mathbf{w} = A\mathbf{w} - \lambda\mathbf{w}$, so that

$$A\mathbf{w} = \mathbf{v} + \lambda\mathbf{w}.$$

vi) From part iv), we know that $(A-\lambda I)^2 = 0$. Therefore,

$$(A-\lambda I)\mathbf{v} = (A-\lambda I)^2\mathbf{w} = 0.$$

It follows that $A\mathbf{v}-\lambda\mathbf{v} = 0$, so that $A\mathbf{v} = \lambda\mathbf{v}$. By definition of E_λ, we then have $\mathbf{v} \in E_\lambda$.

If $a\mathbf{v} + b\mathbf{w} = 0$, then

$$a\mathbf{v} = -b\mathbf{w}.$$

Every element of the left side is in E_λ. Since $\mathbf{w}$ was chosen outside of E_λ, $-b\mathbf{w} \notin E_\lambda$ unless $b = 0$. Therefore, $b = 0$. Finally, since $\mathbf{v} \neq 0$ (otherwise we would have $(A-\lambda I)\mathbf{w} = 0$, so that $\mathbf{w} \in E_\lambda$), $a\mathbf{v} = 0$ only if $a = 0$. Therefore, $a = b = 0$, which shows that $\mathbf{v}$ and $\mathbf{w}$ are linearly independent.

vii) By definition, the matrix representing A with respect to the basis $\{\mathbf{v}, \mathbf{w}\}$ is

$$[A]_{\{\mathbf{v},\mathbf{w}\}} = \left([A\mathbf{v}]_{\{\mathbf{v},\mathbf{w}\}} \quad [A\mathbf{w}]_{\{\mathbf{v},\mathbf{w}\}}\right),$$

where $[A\mathbf{v}]_{\{\mathbf{v},\mathbf{w}\}}$ and $[A\mathbf{w}]_{\{\mathbf{v},\mathbf{w}\}}$ are the coordinate vectors of $A\mathbf{v}$ and $A\mathbf{w}$, respectively, with respect to the basis $\{\mathbf{v}, \mathbf{w}\}$.

By the previous parts,

$$A\mathbf{v} = \lambda\mathbf{v} + 0\mathbf{w},$$

so that

$$[A\mathbf{v}]_{\{\mathbf{v},\mathbf{w}\}} = \begin{pmatrix} \lambda \\ 0 \end{pmatrix}$$

and

$$A\mathbf{w} = \mathbf{v} + \lambda\mathbf{w},$$

so that

$$[A\mathbf{w}]_{\{\mathbf{v},\mathbf{w}\}} = \begin{pmatrix} 1 \\ \lambda \end{pmatrix}.$$

Therefore,

$$[A]_{\{\mathbf{v},\mathbf{w}\}} = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}.$$

viii) The characteristic polynomial of B is

$$\begin{aligned} p_B(z) = \det\begin{pmatrix} 10-z & 4 \\ -9 & -2-z \end{pmatrix} &= (10-z)(-2-z) + 36 \\ &= -20 + 8z + z^2 + 36 \\ &= (z-4)^2. \end{aligned}$$

Therefore, the only eigenvalue of B is $\lambda = 4$ (of algebraic multiplicity 2).

Looking for eigenvectors,

$$\ker(A-4I) = \ker\begin{pmatrix} 10-4 & 4 \\ -9 & -2-4 \end{pmatrix} = \ker\begin{pmatrix} 6 & 4 \\ -9 & -6 \end{pmatrix}$$

We see that $\dim\ker = \dim E_\lambda = 1$. (The two rows of $A-4I$ are linearly dependent, so $\dim\ker \ge 1$. However, the matrix is not the zero matrix, which is the only 2×2 matrix that has a two-dimensional kernel.)

The condition for a vector to be in the kernel of $A-4I$ is $6x+4y=0$, so the eigenspace E_λ is spanned by

$$\begin{pmatrix} 2 \\ -3 \end{pmatrix},$$

(there are many other choices of a spanning vector, of course).

We have to choose a vector $\mathbf{w}$ outside of E_λ. There are many possibilities. A simple one is $\mathbf{w} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

Then, by the method developed above, take

$$\mathbf{v} = (A-4I)\mathbf{w} = \begin{pmatrix} 6 & 4 \\ -9 & -6 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 6 \\ -9 \end{pmatrix}.$$

The (ordered) basis

$$\left\{ \mathbf{v} = \begin{pmatrix} 6 \\ -9 \end{pmatrix},\ \mathbf{w} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\}$$

puts B into its Jordan canonical form.

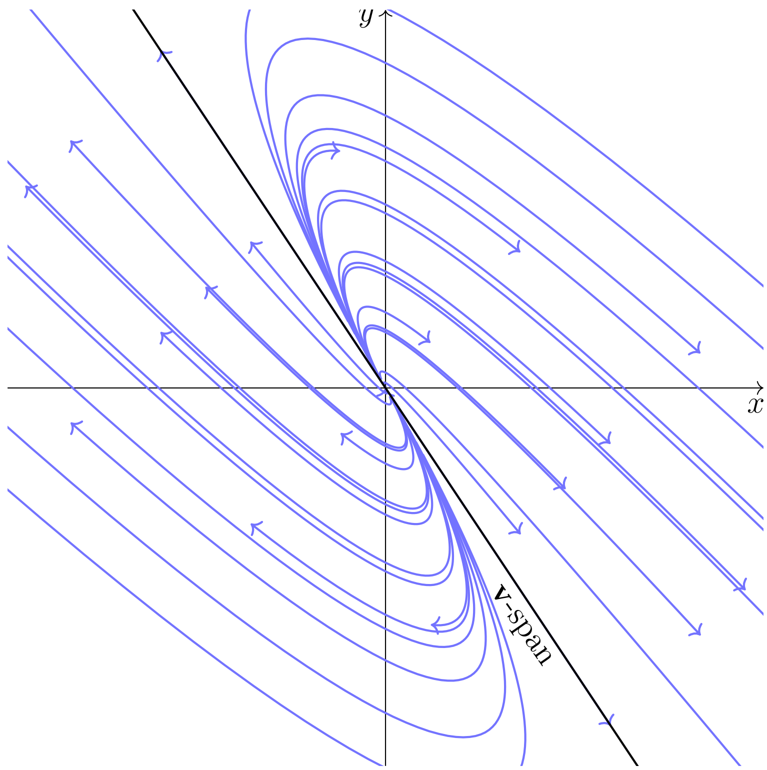
We have

$$B = \begin{pmatrix} 6 & 1 \\ -9 & 0 \end{pmatrix}\begin{pmatrix} 4 & 1 \\ 0 & 4 \end{pmatrix}\frac{1}{9}\begin{pmatrix} 0 & -1 \\ 9 & 6 \end{pmatrix}.$$

ix) Continuing from the previous part,

$$\begin{aligned}
\exp(Bt) &= \begin{pmatrix} 6 & 1 \\ -9 & 0 \end{pmatrix}\exp\left(\begin{pmatrix} 4t & t \\ 0 & 4t \end{pmatrix}\right)\frac{1}{9}\begin{pmatrix} 0 & -1 \\ 9 & 6 \end{pmatrix} \\
&= \begin{pmatrix} 6 & 1 \\ -9 & 0 \end{pmatrix}\begin{pmatrix} e^{4t} & t\,e^{4t} \\ 0 & e^{4t} \end{pmatrix}\frac{1}{9}\begin{pmatrix} 0 & -1 \\ 9 & 6 \end{pmatrix} \\
&= \frac{1}{9}\begin{pmatrix} 6 & 1 \\ -9 & 0 \end{pmatrix}\begin{pmatrix} 9t\,e^{4t} & -e^{4t}+6t\,e^{4t} \\ 9e^{4t} & 6e^{4t} \end{pmatrix} \\
&= \frac{1}{9}\begin{pmatrix} 54t\,e^{4t}+9e^{4t} & -6e^{4t}+36t\,e^{4t}+6e^{4t} \\ -81t\,e^{4t} & 9e^{4t}-54t\,e^{4t} \end{pmatrix} \\
&= \begin{pmatrix} 6t\,e^{4t}+e^{4t} & 4t\,e^{4t} \\ -9t\,e^{4t} & e^{4t}-6t\,e^{4t} \end{pmatrix}.
\end{aligned}$$

x) The flow lines look as follows



The flow line with $\begin{pmatrix} x(0) \\ y(0) \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$ is

$$\exp(Bt)\begin{pmatrix} 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 6t\,e^{4t} + e^{4t} & 4t\,e^{4t} \\ -9t\,e^{4t} & e^{4t} - 6t\,e^{4t} \end{pmatrix}\begin{pmatrix} 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 8t\,e^{4t} + 2e^{4t} \\ -12t\,e^{4t} - e^{4t} \end{pmatrix}.$$

2. For each of the following two systems, find the solution using the Method of Variation of Parameters.

i) $\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 0 \\ t \end{pmatrix}, \qquad \begin{pmatrix} x(0) \\ y(0) \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$

ii) $\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} 2 & 3 \\ 1 & 4 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \qquad \begin{pmatrix} x(0) \\ y(0) \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$

Solution. i) The matrix $C = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ is already of the form $\begin{pmatrix} \sigma & \omega \\ -\omega & \sigma \end{pmatrix}$, so we have

$$\exp(Ct) = \begin{pmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{pmatrix}.$$

By Variation of Parameters, the solution of the system

$$\dot{\mathbf{x}} = C\mathbf{x} + \begin{pmatrix} 0 \\ t \end{pmatrix}, \qquad \mathbf{x}(0) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

is

$$\mathbf{x}(t) = \exp(Ct)\mathbf{x}_0 + \int_0^t \exp(C(t-\tau)) \begin{pmatrix} 0 \\ \tau \end{pmatrix} d\tau.$$

Let's first compute the integrals —

$$\begin{aligned} \exp(C(t-\tau)) \begin{pmatrix} 0 \\ \tau \end{pmatrix} &= \begin{pmatrix} \cos(t-\tau) & \sin(t-\tau) \\ -\sin(t-\tau) & \cos(t-\tau) \end{pmatrix} \begin{pmatrix} 0 \\ \tau \end{pmatrix} \\ &= \begin{pmatrix} \tau \sin(t-\tau) \\ \tau \cos(t-\tau) \end{pmatrix}. \end{aligned}$$

Integrating by parts,

$$\begin{aligned} \int_0^t \tau \sin(t-\tau)\, d\tau &= \tau \cos(t-\tau)\big|_{\tau=0}^{\tau=t} - \int_0^t \cos(t-\tau)\, d\tau \\ &= t\cos(0) - [-\sin(t-\tau)]_{\tau=0}^{\tau=t} \\ &= t - (-\sin(0) + \sin(t)) \\ &= t - \sin(t). \end{aligned}$$

Similarly,

$$\begin{aligned} \int_0^t \tau \cos(t-\tau)\, d\tau &= -\tau \sin(t-\tau)\big|_{\tau=0}^{\tau=t} - \int_0^t -\sin(t-\tau)\, d\tau \\ &= 0 - [-\cos(t-\tau)]_{\tau=0}^{\tau=t} \\ &= \cos(0) - \cos(t) \\ &= 1 - \cos(t). \end{aligned}$$

Therefore,

$$\begin{aligned} \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} &= \begin{pmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \begin{pmatrix} t - \sin(t) \\ 1 - \cos(t) \end{pmatrix} \\ &= \begin{pmatrix} \cos(t) + \sin(t) + t - \sin(t) \\ -\sin(t) + \cos(t) + 1 - \cos(t) \end{pmatrix} \\ &= \begin{pmatrix} \cos(t) + t \\ 1 - \sin(t) \end{pmatrix}. \end{aligned}$$

ii) We computed the exponential of the matrix $D = \begin{pmatrix} 2 & 3 \\ 1 & 4 \end{pmatrix}$ in Problem Set 09. We have

$$\exp(Dt) = \frac{1}{4} \begin{pmatrix} 3e^t + e^{5t} & -3e^t + 3e^{5t} \\ -e^t + e^{5t} & e^t + 3e^{5t} \end{pmatrix}.$$

Then,

$$\exp(D(t-\tau)) \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{1}{4} \begin{pmatrix} -3e^{t-\tau} + 3e^{5(t-\tau)} \\ e^{t-\tau} + 3e^{5(t-\tau)} \end{pmatrix}.$$

Integrating, we have

$$\begin{aligned}
\int_0^t \frac{1}{4}\left(-3e^{t-\tau}+3e^{5(t-\tau)}\right)\,d\tau &= \left[\frac{3}{4}e^{t-\tau}\right]_{\tau=0}^{\tau=t}+\left[-\frac{3}{20}e^{5(t-\tau)}\right]_{\tau=0}^{\tau=t}\\
&= \frac{3}{4}-\frac{3}{4}e^t+\left(-\frac{3}{20}+\frac{3}{20}e^{5t}\right)\\
&= \frac{3}{5}-\frac{3}{4}e^t+\frac{3}{20}e^{5t}.
\end{aligned}$$

Similarly,

$$\begin{aligned}
\int_0^t \frac{1}{4}\left(e^{t-\tau}+3e^{5(t-\tau)}\right)\,d\tau &= \left[-\frac{1}{4}e^{t-\tau}\right]_{\tau=0}^{\tau=t}+\left[-\frac{3}{20}e^{5(t-\tau)}\right]_{\tau=0}^{\tau=t}\\
&= -\frac{1}{4}+\frac{1}{4}e^t-\frac{3}{20}+\frac{3}{20}e^{5t}\\
&= -\frac{2}{5}+\frac{1}{4}e^t+\frac{3}{20}e^{5t}.
\end{aligned}$$

The solution is

$$\begin{aligned}
\begin{pmatrix}x(t)\\y(t)\end{pmatrix} &= \exp(Dt)\begin{pmatrix}1\\0\end{pmatrix}+\int_0^t \exp(D(t-\tau))\begin{pmatrix}0\\1\end{pmatrix}\,d\tau\\
&= \frac{1}{4}\begin{pmatrix}3e^t+e^{5t} & -3e^t+3e^{5t}\\ -e^t+e^{5t} & e^t+3e^{5t}\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}+\begin{pmatrix}\frac{3}{5}-\frac{3}{4}e^t+\frac{3}{20}e^{5t}\\ -\frac{2}{5}+\frac{1}{4}e^t+\frac{3}{20}e^{5t}\end{pmatrix}\\
&= \frac{1}{4}\begin{pmatrix}3e^t+e^{5t}\\ -e^t+e^{5t}\end{pmatrix}+\begin{pmatrix}\frac{3}{5}-\frac{3}{4}e^t+\frac{3}{20}e^{5t}\\ -\frac{2}{5}+\frac{1}{4}e^t+\frac{3}{20}e^{5t}\end{pmatrix}\\
&= \frac{1}{20}\begin{pmatrix}15e^t+5e^{5t}\\ -5e^t+5e^{5t}\end{pmatrix}+\frac{1}{20}\begin{pmatrix}12-15e^t+3e^{5t}\\ -8+5e^t+3e^{5t}\end{pmatrix}\\
&= \frac{1}{20}\begin{pmatrix}12+8e^{5t}\\ -8+8e^{5t}\end{pmatrix}\\
&= \frac{1}{5}\begin{pmatrix}3+2e^{5t}\\ -2+2e^{5t}\end{pmatrix}.
\end{aligned}$$