2   Linear Algebra

Required Problems

10/19:   **1.** This question helps to complete the computations omitted in class. In deriving the Kepler orbits for the two-body problem, we have successfully reduced the differential equation satisfied by the curve $r = r(\varphi)$ to

$$\left(\frac{\mathrm{d}r}{\mathrm{d}\varphi}\right)^2 + r^2 = \frac{2GMr^3}{l_0^2} + \frac{2Er^4}{ml_0^2}$$

Show that the function $\mu = 1/r$ satisfies the differential equation

$$\left(\frac{\mathrm{d}\mu}{\mathrm{d}\varphi}\right)^2 + \mu^2 = \frac{2GM\mu}{l_0^2} + \frac{2E}{ml_0^2}$$

By differentiating with respect to φ again, this reduces to either $\mathrm{d}\mu/\mathrm{d}\varphi = 0$ or

$$\frac{\mathrm{d}^2\mu}{\mathrm{d}\varphi^2} + \mu - \frac{GM}{l_0^2} = 0$$

Find the general solution of the latter, hence conclude that $r = r(\varphi)$ represents a conic section. *Hint:* There is a very obvious particular solution.

Proof. We begin from the first differential equation and substitute $\mu = 1/r$ in the last step to yield the desired result.

$$\left(\frac{\mathrm{d}r}{\mathrm{d}\varphi}\right)^2 + r^2 = \frac{2GMr^3}{l_0^2} + \frac{2Er^4}{ml_0^2}$$

$$\left(-\frac{1}{r^2}\frac{\mathrm{d}r}{\mathrm{d}\varphi}\right)^2 + \frac{1}{r^2} = \frac{2GM}{l_0^2}\frac{1}{r} + \frac{2E}{ml_0^2}$$

$$\left[\frac{\mathrm{d}}{\mathrm{d}\varphi}\left(\frac{1}{r}\right)\right]^2 + \left(\frac{1}{r}\right)^2 = \frac{2GM}{l_0^2}\frac{1}{r} + \frac{2E}{ml_0^2}$$

$$\left(\frac{\mathrm{d}\mu}{\mathrm{d}\varphi}\right)^2 + \mu^2 = \frac{2GM\mu}{l_0^2} + \frac{2E}{ml_0^2}$$

The homogeneous version of the final differential equation is entirely analogous to the harmonic oscillator problem and thus has general (real) solution

$$\mu(\varphi) = \epsilon \cos(\varphi - \varphi_0)$$

for $\epsilon, \varphi_0 \in \mathbb{R}$. By inspection, we can take as our particular solution to the inhomogeneous system

$$\mu(\varphi) = \frac{GM}{l_0^2}$$

since it's second derivative (as a constant) is zero and it is the opposite of the inhomogeneous term. Thus, the general solution to the original inhomogeneous system is

$$\mu(\varphi) = \frac{GM}{l_0^2} + \epsilon \cos(\varphi - \varphi_0)$$

$$r(\varphi) = \frac{1}{GM/l_0^2 + \epsilon \cos(\varphi - \varphi_0)}$$

$$= \frac{\epsilon(l_0^2/GM\epsilon)}{1 + \epsilon \cos(\varphi - \varphi_0)}$$

which is exactly the polar form of the conic section with eccentricity ϵ and directrix $l_0^2/GM\epsilon$. $\square$

2. The general formula for the inverse of an $n \times n$ invertible matrix is very lengthy. However, for a 2×2 matrix
$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$
satisfying $ad - bc \neq 0$, there is a very simple formula. Try to find it; this could be very helpful if you can remember it.

Proof. Let A be the matrix given in the problem statement. We can determine A^{-1} by inspection as follows.

Let's focus on the right column of A^{-1} first, which we can denote $(x, y)^T$. We want $ax + by = 0$. One nice solution to this equation is $x = -b$ and $y = a$. Similarly, we can take the left column of A^{-1} to be $(d, -c)^T$. This choice of entries for A^{-1} yield the 0s in the right places, but the elements that should be 1 are instead $\det A = ad - bc$. Thus, we divide A^{-1} by $\det A$. This yields the following final formula

$$\boxed{A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}}$$

As a quick check, we have that

$$AA^{-1} = \frac{1}{ad - bc} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \qquad\qquad A^{-1}A = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$
$$= \frac{1}{ad - bc} \begin{pmatrix} ad - bc & 0 \\ 0 & ad - bc \end{pmatrix} \qquad\qquad = \frac{1}{ad - bc} \begin{pmatrix} ad - bc & 0 \\ 0 & ad - bc \end{pmatrix}$$
$$= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad\qquad\qquad\qquad\qquad = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

as expected.   □

3. Compute the determinant of the following matrices. Determine whether they are invertible or not.

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \qquad B = \begin{pmatrix} 2 & 2 & 3 & 6 \\ 1 & 3 & 4 & 2 \\ 0 & 0 & -1 & 2 \\ 0 & 0 & 1 & 2 \end{pmatrix} \qquad C = \begin{pmatrix} -1 & 2 & 1 \\ 3 & -1 & 2 \\ 2 & 1 & 3 \end{pmatrix}$$

Proof. We have that

$$\det A = 1[5 \cdot 9 - 6 \cdot 8] - 2[4 \cdot 9 - 6 \cdot 7] + 3[4 \cdot 8 - 5 \cdot 7]$$
$$\boxed{\det A = 0}$$

so $\boxed{A \text{ is not invertible.}}$

Since B is block upper triangular, we know that

$$\det B = \det B_1 \cdot \det B_2$$
$$= [2 \cdot 3 - 2 \cdot 1] \cdot [-1 \cdot 2 - 2 \cdot 1]$$
$$\boxed{\det B = -16}$$

so $\boxed{B \text{ is invertible.}}$

We have that

$$\det C = -1[(-1)(3) - (2)(1)] - 2[(3)(3) - (2)(2)] + 1[(3)(1) - (-1)(2)]$$
$$\boxed{\det C = 0}$$

so $\boxed{C \text{ is not invertible.}}$ □

4. Determine whether the following linear systems admit solution(s); if they do, write down the solution (or the formula for the general solution).

(1)

$$\begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

Proof. By inspection, A is a dimension 2 matrix of rank 2, so it $\boxed{\text{admits a unique solution.}}$ We now row-reducing the augmented matrix.

$$\left(\begin{array}{cc|c} 1 & 2 & -1 \\ 2 & -1 & 1 \end{array} \right) \cong \left(\begin{array}{cc|c} 1 & 0 & \frac{1}{5} \\ 0 & 1 & -\frac{3}{5} \end{array} \right)$$

Therefore, the solution is

$$\boxed{x = \begin{pmatrix} \frac{1}{5} \\ -\frac{3}{5} \end{pmatrix}}$$

$\square$

(2)

$$\begin{pmatrix} -1 & 2 & 1 \\ 3 & -1 & 2 \\ 2 & 1 & 3 \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

Proof. By inspection, A is a dimension 3 matrix of rank 2 and the b vector is in the column space of A, so it $\boxed{\text{admits a family of solutions.}}$ We now row-reducing the augmented matrix.

$$\left(\begin{array}{ccc|c} -1 & 2 & 1 & 1 \\ 3 & -1 & 2 & 2 \\ 2 & 1 & 3 & 3 \end{array} \right) \cong \left(\begin{array}{ccc|c} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{array} \right)$$

Therefore, the family of solutions is given by

$$\boxed{x = \begin{pmatrix} 1 - x^3 \\ 1 - x^3 \\ x^3 \end{pmatrix}}$$

for $x^3 \in \mathbb{R}$. $\square$

(3)

$$\begin{pmatrix} -1 & 2 & 1 \\ 3 & -1 & 2 \\ 2 & 1 & 3 \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

Proof. No promising solution immediately appears by inspection, so we row reduce and evaluate the results.

$$\left(\begin{array}{ccc|c} -1 & 2 & 1 & 1 \\ 3 & -1 & 2 & 0 \\ 2 & 1 & 3 & 1 \end{array} \right) \cong \left(\begin{array}{ccc|c} 1 & 0 & 1 & \frac{1}{5} \\ 0 & 1 & 1 & \frac{3}{5} \\ 0 & 0 & 0 & 0 \end{array} \right)$$

It follows that A $\boxed{\text{admits a family of solutions.}}$ In particular, these are given by

$$\boxed{x = \begin{pmatrix} \frac{1}{5} - x^3 \\ \frac{3}{5} - x^3 \\ x^3 \end{pmatrix}}$$

for $x^3 \in \mathbb{R}$. $\square$

5. Find the connecting matrix from the basis $\begin{pmatrix} p_1 & p_2 & p_3 \end{pmatrix}$ to the new basis $\begin{pmatrix} q_1 & q_2 & q_3 \end{pmatrix}$, where

$$\begin{pmatrix} p_1 & p_2 & p_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & -1 \\ 1 & 2 & 0 \\ 0 & -1 & 2 \end{pmatrix} \qquad \begin{pmatrix} q_1 & q_2 & q_3 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & -1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

That is, represent q_1, q_2, q_3 as linear combinations of p_1, p_2, p_3.

Proof. P is the connecting matrix from the standard basis (e_1, e_2, e_3) to (p_1, p_2, p_3). Likewise, Q is the connecting matrix from (e_1, e_2, e_3) to (q_1, q_2, q_3). It follows that if we want A to be the connecting matrix from (p_1, p_2, p_3) to (q_1, q_2, q_3), then we can do the transformation stepwise, i.e., take a vector represented in (p_1, p_2, p_3) to its representation in (e_1, e_2, e_3) using P^{-1} and then to its representation in (q_1, q_2, q_3) using Q. Indeed, the desired connecting matrix is

$$A = QP^{-1}$$

$$\boxed{A = \frac{1}{5} \begin{pmatrix} -2 & 2 & -1 \\ 5 & 0 & 5 \\ -1 & 1 & 2 \end{pmatrix}}$$

Direct computation can confirm that $Ap_i = q_i$ for $i = 1, 2, 3$.

With respect to representing q_1, q_2, q_3 as linear combinations of p_1, p_2, p_3, we can solve the equations $q_i = Px_i$ for $i = 1, 2, 3$ via row reduction, as in previous responses. The final expressions obtained are

$$\boxed{q_1 = \frac{1}{5}(p_1 + 2p_2 + p_3) \qquad q_2 = \frac{1}{5}(3p_1 - 4p_2 - 2p_3) \qquad q_3 = \frac{1}{5}(3p_1 + p_2 + 3p_3)}$$

Note that if we combine the coefficients above into a matrix X such that $PX = Q$, then $A = PXP^{-1} = QXQ^{-1}$. $\square$

6. Let $\theta \in [0, 2\pi)$. The rotation through angle θ in the plane is represented by the matrix

$$R(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

Compute its determinant, characteristic polynomial, and eigenvalues. Compute its eigenvectors in $\mathbb{C}^2$. You need to use the Euler formula $e^{i\theta} = \cos\theta + i\sin\theta$. For two angles θ, φ, compute the product $R(\theta)R(\varphi)$ and represent it in terms of $\theta + \varphi$. What is the geometric meaning of this equality?

Proof. The determinant of R is

$$\det R = \cos^2\theta + \sin^2\theta$$

$$\boxed{\det R = 1}$$

The characteristic polynomial of R is

$$\begin{aligned} \chi_R(z) &= \det(R - zI) \\ &= (\cos\theta - z)^2 + \sin^2\theta \\ &= z^2 - 2z\cos\theta + \cos\theta^2 + \sin\theta^2 \end{aligned}$$

$$\boxed{\chi_R(z) = z^2 - 2z\cos\theta + 1}$$

The eigenvalues of R are

$$
\begin{aligned}
0 = \chi_R(\lambda) \\
= (\cos\theta - \lambda)^2 + \sin^2\theta \\
-\sin^2\theta = (\cos\theta - \lambda)^2 \\
\pm i\sin\theta = \pm(\cos\theta - \lambda) \\
\lambda = \cos\theta \pm i\sin\theta \\
\boxed{\lambda = e^{\pm i\theta}}
\end{aligned}
$$

It follows by solving the systems of equations

$$
\begin{aligned}
x^1\cos\theta - x^2\sin\theta = e^{i\theta}x^1 \\
x^1\sin\theta + x^2\cos\theta = e^{i\theta}x^2
\end{aligned}
\qquad\qquad
\begin{aligned}
y^1\cos\theta - y^2\sin\theta = e^{-i\theta}y^1 \\
y^1\sin\theta + y^2\cos\theta = e^{-i\theta}y^2
\end{aligned}
$$

that the eigenvectors are

$$
\boxed{x = \begin{pmatrix} 1 \\ -i \end{pmatrix} \qquad\qquad y = \begin{pmatrix} 1 \\ i \end{pmatrix}}
$$

The product $R(\theta)R(\varphi)$ may be computed as follows.

$$
\begin{aligned}
R(\theta)R(\phi) &= \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} \\
&= \begin{pmatrix} \cos\theta\cos\varphi - \sin\theta\sin\varphi & -\cos\theta\sin\varphi - \sin\theta\cos\varphi \\ \sin\theta\cos\varphi + \cos\theta\sin\varphi & -\sin\theta\sin\varphi + \cos\theta\cos\varphi \end{pmatrix} \\
&= \begin{pmatrix} \cos(\theta+\varphi) & -\sin(\theta+\varphi) \\ \sin(\theta+\varphi) & \cos(\theta+\varphi) \end{pmatrix}
\end{aligned}
$$

$$
\boxed{R(\theta)R(\varphi) = R(\theta + \phi)}
$$

The geometric meaning is that rotating through an angle θ and then through an additional angle φ is the same as rotating through an angle $\theta + \varphi$ all at once. $\qquad\square$

8. Find the algebraic and geometric multiplicities of the eigenvalues of the following matrices.

$$
A = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 3 \end{pmatrix}
\qquad\qquad
B = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 3 \end{pmatrix}
$$

Proof. We tackle A first. A is an upper triangular matrix. Thus, $\chi_A(\lambda) = \det(A - \lambda I)$ can be read directly off of the diagonal:

$$
\chi_A(\lambda) = (1-\lambda)^2(3-\lambda)
$$

Thus, the eigenvalues are $\lambda = 1, 3$ with respective algebraic multiplicities

$$
\boxed{\alpha_1 = 2 \qquad\qquad \alpha_3 = 1}
$$

It follows immediately that

$$
\boxed{\gamma_3 = 1}
$$

and from the observation that $A - 1I$ has 2 linearly independent columns that this 3×3 matrix has a $3 - 2 = 1$ dimensional null space, i.e.,

$$
\boxed{\gamma_1 = 1}
$$

The procedure for B is almost entirely symmetric. Once again, B is upper triangular, so

$$\chi_B(\lambda) = (1 - \lambda)^2(3 - \lambda)$$

implying that

$$\boxed{\alpha_1 = 2 \qquad\qquad\qquad\qquad \alpha_3 = 1}$$

There is a difference with respect to the geometric multiplicities, however. We still have

$$\boxed{\gamma_3 = 1}$$

but since $A - I$ now has only 1 linearly independent column, we have

$$\boxed{\gamma_1 = 2}$$

this time. □

9. Compute the Jordan normal form of the following 2×2 matrices.

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \qquad\qquad\qquad B = \begin{pmatrix} 0 & -1 \\ 1 & -2 \end{pmatrix}$$

Notice that you not only need to find all the Jordan blocks, but also need to find the Jordan basis matrix Q such that $Q^{-1}AQ$ is in Jordan normal form.

Proof. We tackle A first.

Calculate the characteristic polynomial to begin.

$$\begin{aligned} \chi_A(z) &= \det(A - zI) \\ &= z^2 - 4z + 3 \\ &= (1 - z)(3 - z) \end{aligned}$$

It follows that the eigenvalues are

$$\lambda_1 = 1 \qquad\qquad\qquad\qquad \lambda_2 = 3$$

Since these eigenvalues are distinct, we can fully diagonalize this matrix. Indeed, we can determine by inspection that suitable corresponding eigenvectors are

$$v_1 = \begin{pmatrix} -1 \\ 1 \end{pmatrix} \qquad\qquad\qquad\qquad v_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Therefore,

$$\boxed{Q = \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix} \qquad\qquad\qquad Q^{-1}AQ = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}}$$

The procedure for B is very much analogous to the procedure for A.

Characteristic polynomial:

$$\begin{aligned} \chi_B(z) &= \det(B - zI) \\ &= z^2 + 2z + 1 \\ &= (1 + z)^2 \end{aligned}$$

Eigenvalue:

$$\lambda = -1$$

By inspection of $B + I$, we can pick one eigenvector of B:

$$v = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

We now solve $(B + I)u = v$. By inspection, this yields

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

Therefore,

$$\boxed{Q = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \qquad\qquad Q^{-1}BQ = \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix}}$$

$\square$

10. Compute the Jordan normal form of the following 3×3 matrices.

$$A = \begin{pmatrix} 4 & -5 & 2 \\ 5 & -7 & 3 \\ 6 & -9 & 4 \end{pmatrix} \qquad B = \begin{pmatrix} 2 & -1 & -1 \\ 2 & -1 & -2 \\ -1 & 1 & 2 \end{pmatrix} \qquad C = \begin{pmatrix} 2 & 1 & 3 \\ 0 & 2 & -1 \\ 0 & 0 & 2 \end{pmatrix}$$

Notice that you not only need to find all the Jordan blocks, but also need to find the Jordan basis matrix Q such that $Q^{-1}AQ$ is in Jordan normal form. *Hint*: These three matrices represent three different possibilities of nondiagonalizable Jordan normal forms of a 3×3 matrix: A reduces to $(2 \times 2) \oplus (1 \times 1)$ Jordan blocks with different eigenvalues, B reduces to $(2 \times 2) \oplus (1 \times 1)$ Jordan blocks with the same eigenvalue, and C reduces to a 3×3 Jordan block.

Proof. We tackle A first.

Calculate the characteristic polynomial to begin.

$$\begin{aligned} \chi_A(z) &= \det(A - zI) \\ &= -z^3 + z^2 \\ &= z^2(1 - z) \end{aligned}$$

It follows that the eigenvalues are

$$\lambda_1 = \lambda_2 = 0 \qquad\qquad\qquad \lambda_3 = 1$$

We can solve for an eigenvector v_1 corresponding to $\lambda_1 = \lambda_2 = 0$ using the augmented matrix and row reduction as follows.

$$\left(\begin{array}{ccc|c} 4 & -5 & 2 & 0 \\ 5 & -7 & 3 & 0 \\ 6 & -9 & 4 & 0 \end{array}\right) \cong \left(\begin{array}{ccc|c} 1 & 0 & -\frac{1}{3} & 0 \\ 0 & 1 & -\frac{2}{3} & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Thus, if we choose $v_1^3 = 3$, then the desired eigenvector is

$$v_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

Similarly, we can solve for an eigenvector v_3 corresponding to $\lambda_3 = 1$ using the following. Note that to solve $Ax = 1x$, we row-reduce $(A - I)x = 0$.

$$\left(\begin{array}{ccc|c} 3 & -5 & 2 & 0 \\ 5 & -8 & 3 & 0 \\ 6 & -9 & 3 & 0 \end{array}\right) \cong \left(\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

This yields

$$v_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

We now solve the equation $(A - 0I)u = v_1$ to find a generalized eigenvector u corresponding to $\lambda_1 = \lambda_2 = 0$. This can also be done with an augmented matrix.

$$\left(\begin{array}{ccc|c} 4 & -5 & 2 & 1 \\ 5 & -7 & 3 & 2 \\ 6 & -9 & 4 & 3 \end{array} \right) \cong \left(\begin{array}{ccc|c} 1 & 0 & -\frac{1}{3} & 0 \\ 0 & 1 & -\frac{2}{3} & 0 \\ 0 & 0 & 0 & 0 \end{array} \right)$$

This yields

$$u = \begin{pmatrix} 0 \\ 1 \\ 3 \end{pmatrix}$$

Therefore,

$$Q = \begin{pmatrix} 1 & 0 & 1 \\ 2 & 1 & 1 \\ 3 & 3 & 1 \end{pmatrix} \qquad\qquad Q^{-1}AQ = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

The procedure for B is very much analogous to the procedure for A.

Characteristic polynomial:

$$\begin{aligned} \chi_B(z) &= \det(B - zI) \\ &= -z^3 + 3z^2 - 3z + 1 \\ &= (1 - z)^3 \end{aligned}$$

Eigenvalue:

$$\lambda = 1$$

By inspection of

$$B - I = \begin{pmatrix} 1 & -1 & -1 \\ 2 & -2 & -2 \\ -1 & 1 & 1 \end{pmatrix}$$

we can pick two eigenvectors of B corresponding to λ, i.e., two elements of the null space of the above matrix. In this subcase of the 3×3 case, we always pick the first of these to be an element of the column space of $B - I$, as well. Thus, choose

$$v_1 = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} \qquad\qquad v_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

We now solve $(B - \lambda I)u = v_1$. By inspection, this yields

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

Therefore,

$$Q = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 0 & 1 \\ -1 & 0 & 0 \end{pmatrix} \qquad\qquad Q^{-1}BQ = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

The procedure for C is likewise quite analogous.

The matrix is upper triangular, so the eigenvalues are on the diagonal. It follows that

$$\lambda = 2$$

is the sole eigenvalue. We can solve $(C - 2I)v = 0$ for one eigenvector v by inspection, yielding

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

We can also solve $(C - 2I)u_1 = v$ by inspection to get

$$u_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

One more time, we can solve $(C - 2I)u_2 = u_1$ by inspection to get

$$u_2 = \begin{pmatrix} 0 \\ 3 \\ -1 \end{pmatrix}$$

Therefore,

$$Q = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 3 \\ 0 & 0 & -1 \end{pmatrix} \qquad\qquad Q^{-1}CQ = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$$

$\square$

Bonus Problems

1. You may find the characteristic root method for the second-order equation $y'' + ay' + b = 0$ quite abrupt. This problem helps you see where it comes from. The origin of this method is in fact a comparison with the linear recursive relation

$$y_{n+2} + ay_{n+1} + by_n = 0$$

where a, b are given complex numbers.

(1) The linear recursive relation $y_{n+1} + ay_n = 0$ gives rise to a geometric sequence

$$y_0, y_0(-a), y_0(-a)^2, \ldots$$

We now want to try to reduce the second-order recursive relation $y_{n+2} + ay_{n+1} + by_n = 0$ to a first-order relation. Thus, we look for complex numbers λ, μ such that

$$(y_{n+2} - \lambda y_{n+1}) - \mu(y_{n+1} - \lambda y_n) = 0$$

Then λ, μ should be the roots of the characteristic polynomial

$$X^2 + aX + b$$

Taking λ, μ as known quantities, find the general formula for y_n, regarding y_0, y_1 as known quantities. *Hint*: The sequence of numbers $y_{n+1} - \lambda y_n$ is a geometric sequence with ratio μ. You should also discuss $\mu \neq \lambda$ and $\mu = \lambda$ separately.

Proof. We start with the case $\mu \neq \lambda$ here. Define a sequence of numbers $\{x_n\}$ by

$$x_n := y_{n+1} - \lambda y_n$$

We have from the above that x_n is recursively defined by

$$x_{n+1} - \mu x_n = 0 \iff x_n = x_0 \mu^n$$

Returning the substitution, we have that

$$y_{n+1} - \lambda y_n = (y_1 - \lambda y_0)\mu^n$$

Rearranging, we have that

$$y_{n+1} = \lambda y_n + (y_1 - \lambda y_0)\mu^n$$

In other words, to generate each new term, we multiply through the previous term by λ and add to it a new term. Let's investigate the effect of this procedure on the first few terms past y_0, y_1. These terms are

$$
\begin{aligned}
y_2 &= \lambda y_1 + (y_1 - \lambda y_0)\mu^1 \\
&= (\underbrace{\lambda + \mu}_{p_1(\lambda,\mu)})y_1 - \lambda\mu(\underbrace{1}_{q_1(\lambda,\mu)})y_0 \\
y_3 &= \lambda(\lambda + \mu)y_1 - \lambda\mu \cdot \lambda y_0 + (y_1 - \lambda y_0)\mu^2 \\
&= [\lambda(\lambda + \mu) + \mu^2]y_1 - \lambda\mu \cdot [\lambda + \mu]y_0 \\
&= (\underbrace{\lambda^2 + \lambda\mu + \mu^2}_{p_2(\lambda,\mu)})y_1 - \lambda\mu(\underbrace{\lambda + \mu}_{q_2(\lambda,\mu)})y_0 \\
y_4 &= \lambda(\lambda^2 + \lambda\mu + \mu^2)y_1 - \lambda\mu \cdot \lambda(\lambda + \mu)y_0 + (y_1 - \lambda y_0)\mu^3 \\
&= [\lambda(\lambda^2 + \lambda\mu + \mu^2) + \mu^3]y_1 - \lambda\mu \cdot [\lambda(\lambda + \mu) + \mu^2]y_0 \\
&= (\underbrace{\lambda^3 + \lambda^2\mu + \lambda\mu^2 + \mu^3}_{p_3(\lambda,\mu)})y_1 - \lambda\mu(\underbrace{\lambda^2 + \lambda\mu + \mu^2}_{q_3(\lambda,\mu)})y_0
\end{aligned}
$$

There are a number of interesting things to note here. First of all, see that the first line defining each term comes straight from the recursive relation (i.e., is λ times the previous term plus an auxiliary term). The second line is always a rewrite that combines all of the y_1 terms and all of the y_0 terms. The third line (if it exists; the second line for y_2 serves this purpose as well) is always a final simplification in terms of polynomials of λ and μ. We have taken the liberty of naming each of these final polynomials for purposes that will soon be revealed.

Notice that the polynomials appear to be defined by the recursive relations[1]

$$p_n = \lambda p_{n-1} + \mu^n \qquad\qquad q_n = \lambda q_{n-1} + \mu^{n-1}$$

Also note that we factor $-\lambda\mu$ out of the y_0 term every time so that we have

$$q_n = p_{n-1}$$

These polynomials have a very well-defined form; specifically,

$$q_{n+1} = p_n = \sum_{i=0}^{n} \lambda^{n-i}\mu^i = \frac{\lambda^{n+1} - \mu^{n+1}}{\lambda - \mu}$$

This allows us to explicitly define the following general formula for y_n in the $\mu \neq \lambda$ case.

$$\boxed{y_n = \frac{\lambda^n - \mu^n}{\lambda - \mu}y_1 - \lambda\mu\frac{\lambda^{n-1} - \mu^{n-1}}{\lambda - \mu}y_0}$$

I believe that this is along the lines of what Shao wanted us to do.

---

[1]We can rigorously prove this via induction. Also note the familiar theme of "multiply through by λ and add a term."

Note, however, that there exists an alternate, much more logical way to derive this formula using linear algebra. In particular, we have from the recursive relation and the trivial reflexive equality $y_1 = y_1$ that

$$\begin{pmatrix} y_2 \\ y_1 \end{pmatrix} = \begin{pmatrix} -ay_1 - by_0 \\ y_1 \end{pmatrix}$$
$$= \begin{pmatrix} -a & -b \\ 1 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_0 \end{pmatrix}$$

It follows that

$$\begin{pmatrix} y^{n+1} \\ y^n \end{pmatrix} = \begin{pmatrix} -a & -b \\ 1 & 0 \end{pmatrix}^n \begin{pmatrix} y_1 \\ y_0 \end{pmatrix}$$

Diagonalizing the matrix raised to the nth power above (which we may call A) requires finding the roots of the (of course similarly named) characteristic polynomial

$$0 = \chi_A(z)$$
$$= (-a - z)(-z) + b$$
$$= z^2 + az + b$$

But we have from the problem statement that the roots of this polynomial are the same λ, μ we've been working with this whole time. Thus, we have a new interpretation of these roots: As *eigenvalues*. And to each of these distinct eigenvalues corresponds a distinct eigenvector, which we may find by inspection as follows. Consider λ WLOG. The eigenvector v corresponding to λ must satisfy

$$\begin{pmatrix} -a - \lambda & -b \\ 1 & -\lambda \end{pmatrix} \begin{pmatrix} v^1 \\ v^2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

If we choose $v^1 = \lambda$ and $v^2 = 1$, then we of course have

$$(1)(v^1) + (-\lambda)(v^2) = \lambda - \lambda = 0$$

But we actually also have

$$(-a - \lambda)(v^1) + (-b)(1) = -\lambda^2 - a\lambda - b = -(\lambda^2 + a\lambda + b) = -0 = 0$$

since λ is a root of the characteristic polynomial! Thus, the eigenvectors corresponding to λ, μ are

$$\begin{pmatrix} \lambda \\ 1 \end{pmatrix}, \begin{pmatrix} \mu \\ 1 \end{pmatrix}$$

respectively. Thus, with what we have so far plus an assist from Problem 2, we can diagonalize A to

$$\begin{pmatrix} -a & -b \\ 1 & 0 \end{pmatrix} = \frac{1}{\lambda - \mu} \begin{pmatrix} \lambda & \mu \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix} \begin{pmatrix} 1 & -\mu \\ -1 & \lambda \end{pmatrix}$$

It follows that

$$\begin{pmatrix} y^{n+1} \\ y^n \end{pmatrix} = \begin{pmatrix} -a & -b \\ 1 & 0 \end{pmatrix}^n \begin{pmatrix} y_1 \\ y_0 \end{pmatrix}$$
$$= \frac{1}{\lambda - \mu} \begin{pmatrix} \lambda & \mu \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}^n \begin{pmatrix} 1 & -\mu \\ -1 & \lambda \end{pmatrix} \begin{pmatrix} y_1 \\ y_0 \end{pmatrix}$$
$$= \frac{1}{\lambda - \mu} \begin{pmatrix} \lambda & \mu \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \lambda^n & 0 \\ 0 & \mu^n \end{pmatrix} \begin{pmatrix} y_1 - \mu y_0 \\ -(y_1 - \lambda y_0) \end{pmatrix}$$
$$= \frac{1}{\lambda - \mu} \begin{pmatrix} \lambda & \mu \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \lambda^n(y_1 - \mu y_0) \\ -\mu^n(y_1 - \lambda y_0) \end{pmatrix}$$
$$= \frac{1}{\lambda - \mu} \begin{pmatrix} \lambda^{n+1}(y_1 - \mu y_0) - \mu^{n+1}(y_1 - \lambda y_0) \\ \lambda^n(y_1 - \mu y_0) - \mu^n(y_1 - \lambda y_0) \end{pmatrix}$$

From the above, we can read that

$$y_n = \frac{1}{\lambda - \mu}[\lambda^n(y_1 - \mu y_0) - \mu^n(y_1 - \lambda y_0)]$$

$$= \frac{\lambda^n - \mu^n}{\lambda - \mu}y_1 - \frac{\lambda^n\mu - \lambda\mu^n}{\lambda - \mu}y_0$$

$$= \frac{\lambda^n - \mu^n}{\lambda - \mu}y_1 - \lambda\mu\frac{\lambda^{n-1} - \mu^{n-1}}{\lambda - \mu}y_0$$

in agreement with the original answer.

We now treat the case where $\mu = \lambda$. We will do this using linear algebra, alone, since this is a far more rigorous and overall superior method. That being said, a purely sequence-based method would also work here. Let's begin.

The setup proceeds as before. However, this time we only have one distinct eigenvalue. Moreover, this eigenvalue only corresponds to one eigenvector since

$$\begin{pmatrix} -a - \mu & -b \\ 1 & -\mu \end{pmatrix}$$

is a rank 1 matrix, not a rank 0 matrix. Thus, we must delve into JNF and generalized eigenvectors. In particular, we can still choose

$$\begin{pmatrix} \mu \\ 1 \end{pmatrix}$$

as an eigenvector, but we must solve

$$\begin{pmatrix} -a - \mu & -b \\ 1 & -\mu \end{pmatrix}\begin{pmatrix} u^1 \\ u^2 \end{pmatrix} = \begin{pmatrix} \mu \\ 1 \end{pmatrix}$$

to obtain the following additional *generalized* eigenvector.

$$\begin{pmatrix} -1 \\ -2/\mu \end{pmatrix}$$

Therefore, with what we have so far, we can diagonalize A to

$$\begin{pmatrix} -a & -b \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} \mu & -1 \\ 1 & -2/\mu \end{pmatrix}\begin{pmatrix} \mu & 1 \\ 0 & \mu \end{pmatrix}\begin{pmatrix} 2/\mu & -1 \\ 1 & -\mu \end{pmatrix}$$

It follows that

$$\begin{pmatrix} y^{n+1} \\ y^n \end{pmatrix} = \begin{pmatrix} \mu & -1 \\ 1 & -2/\mu \end{pmatrix}\begin{pmatrix} \mu & 1 \\ 0 & \mu \end{pmatrix}^n\begin{pmatrix} 2/\mu & -1 \\ 1 & -\mu \end{pmatrix}\begin{pmatrix} y_1 \\ y_0 \end{pmatrix}$$

$$= \begin{pmatrix} \mu & -1 \\ 1 & -2/\mu \end{pmatrix}\begin{pmatrix} \mu^n & n\mu^{n-1} \\ 0 & \mu^n \end{pmatrix}\begin{pmatrix} 2y_1/\mu - y_0 \\ y_1 - \mu y_0 \end{pmatrix}$$

$$= \begin{pmatrix} \mu & -1 \\ 1 & -2/\mu \end{pmatrix}\begin{pmatrix} (n+2)\mu^{n-1}y_1 - (n+1)\mu^n y_0 \\ \mu^n y_1 - \mu^{n+1}y_0 \end{pmatrix}$$

$$= \begin{pmatrix} (n+1)\mu^n y_1 - n\mu^{n+1}y_0 \\ n\mu^{n-1}y_1 - (n-1)\mu^n y_0 \end{pmatrix}$$

From the above, we can read off the following general formula for y_n in the $\mu = \lambda$ case.

$$\boxed{y_n = n\mu^{n-1}y_1 - (n-1)\mu^n y_0}$$

$\square$

(2) Use the method of part (1) to find the general formula for the linear discursive relation

$$y_{n+2} - 2y_{n+1} + y_n = 0$$

Use the same method to find the general formula for the Fibonacci sequence

$$F_{n+2} = F_{n+1} + F_n$$

Proof. The first part is plug and chug; the Fibonacci sequence I covered in LinAlgNotes, Chapter 6.

Last note: Somebody please explain to me how any of this makes the ODE-solving method make any more sense. I mean they're clearly related, but there's no Duhamel formula for sequences that I'm aware of! □

2. In this exercise, we aim to prove an important theorem in linear algebra:

Complex Hermitian matrices are always diagonalizable.

Here the term "Hermitian" means that the matrix equals its conjugate transpose. In terms of entries, this means that in general, $a_{ij} = \bar{a}_{ji}$. For example,

$$\begin{pmatrix} 2 & 1 & -i \\ 1 & 3 & -2i \\ i & 2i & 1 \end{pmatrix}$$

is Hermitian.

(1) Let $\langle \cdot, \cdot \rangle$ be the standard Hermitian inner product, that is, for $x, y \in \mathbb{C}^n$,

$$\langle x, y \rangle = \sum_{j=1}^{n} x^j \bar{y}^j$$

Show that for any $n \times n$ real matrix,

$$\langle Ax, y \rangle = \langle x, A^* y \rangle$$

for any $x, y \in \mathbb{C}^n$, where A^* denotes the conjugate transpose of A. For example,

$$A = \begin{pmatrix} 1 & 1 & 2i \\ 0 & 3+i & 3 \\ 2 & 0 & 1 \end{pmatrix} \iff A^* = \begin{pmatrix} 1 & 0 & 2 \\ 1 & 3-i & 0 \\ -2i & 3 & 1 \end{pmatrix}$$

Proof. I almost surely have this all written up somewhere in LinAlgNotes, LADRNotes, or MATH20700Notes. □

(2) Suppose now that A is Hermitian. Use part (1) to show that any eigenvalue of A must be a real number. Show further that if x, y are eigenvectors corresponding to different eigenvalues, then $\langle x, y \rangle = 0$, that is, x is orthogonal to y.

(3) Prove that every Hermitian matrix A is diagonalizable. *Hint*: Take any eigenvector v_1 of A. Decompose $\mathbb{C}^n$ into the direct sum of span(v_1) and its orthogonal complement. Show that the orthogonal complement is an invariant subspace for A.