University of Utah Spring 2024

MATH 2270-002 Selected Solutions to PSet Problems

Instructor: Alp Uzman

Subject to Change; Last Updated: 2024-04-22 12:12:03-06:00

In this document you may find solutions to selected problems from problem sets. As a general rule of thumb, it is more likely that this document will be updated with the solutions to problems that were graded on accuracy.

Please remember that throughout this class your solutions need not match the level of polish and precision of the below solutions. What's most important is grasping the underlying concepts and steadily improving your problem-solving skills as well as your technical presentation skills.

1 PSet 1, Problem 2

Suppose a, b, c, and d are constants such that a is not zero and the system below is consistent for all possible values of f and g. What can you say about the numbers a, b, c, and d?

$$ax_1 + bx_2 = f$$
$$cx_1 + dx_2 = g$$

Let's first parse the problem: we are given the following:

- For fixed numbers a, b, c, d, the system is consistent (aka has a solution) for any two numbers f, g.
- $a \neq 0$.

The problem is asking for a statement about a, b, c, d. While the formulation is somewhat ambiguous, it is safe to assume that the problem is asking for an equality or inequality involving only the numbers a, b, c, d. For instance certainly it can not be the case that

$$a = b = c = d = 0,$$

for otherwise unless $f = g = 0$, the system would fail to be consistent. This is not the most general statement one can obtain however, and to obtain a more general statement one can switch to the augmented matrix notation and do row reduction.

The augmented matrix corresponding to the system is:

$$\begin{pmatrix} a & b & f \\ c & d & g \end{pmatrix}$$

which one can then row reduce as follows:

$$\begin{pmatrix} a & b & f \\ c & d & g \end{pmatrix} \xrightarrow{R_1 \leftarrow \frac{1}{a} R_1} \begin{pmatrix} 1 & \frac{b}{a} & \frac{f}{a} \\ c & d & g \end{pmatrix}$$
$$\xrightarrow{R_2 \leftarrow R_2 - cR_1} \begin{pmatrix} 1 & \frac{b}{a} & \frac{f}{a} \\ 0 & d - \frac{bc}{a} & g - \frac{fc}{a} \end{pmatrix}$$

Looking at the second row, the only way the system would be inconsistent is if

$$\boxed{d - \frac{bc}{a} = 0} \text{ and simultaneously } \boxed{g - \frac{fc}{a} \neq 0},$$

for fixed a, b, c, d and arbitrary f, g. But one can find f, g such that the second inequality holds, e.g. $f = 0$, $g = 1$, which means that in order for the system to be consistent no matter what f and g are, one must have

$$\boxed{d - \frac{bc}{a} \neq 0}$$

which is equivalent to

$$\boxed{ad - bc \neq 0}.$$

 uzman@math.utah.edu

Follow-up Exercise The assumption that $a \neq 0$ is not needed; without this assumption the final answer is again the same, though the argument requires a case analysis. (Either $a \neq 0$, xor $a = 0$. In the former case the above argument is sufficient; in the latter case argue first that c must be nonzero, and then apply the above argument with c replacing a.)

Follow-up Exercise 2 The problem goes the other way around too: if for some f, g, the system is inconsistent, then $ad - bc = 0$. This is harder to justify, at least using only row reduction.

Follow-up Exercise 3 Same problem, but with 3 equations and 3 unknowns. Using only row reduction the algebra will be quite complicated, but it can be done!

2 PSet 2, Problem 8

Consider the linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ that first reflects points through the x_1-axis and then reflects points through the x_2-axis. Show that T can also be described as a linear transformation that rotates points about the origin by a certain angle. What is the angle of that rotation?

We consider $\mathbb{R}^2$ to be parameterized so that the horizontal axis is labeled by x_1 and the vertical axis is labeled by x_2. Recall the following two key ideas:

- One can think of a linear transformation from the plane to itself as a 2×2 matrix.
- The vector $T(1, 0)$ that the transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ takes $(1, 0)$ to corresponds to the first column of the matrix of T, and similarly $T(0, 1)$ corresponds to the second column of the matrix of T.

Thus one way to attack this problem is to first keep track of where $e_1 = (1, 0)$ and $e_2 = (0, 1)$ are sent under the transformation T to obtain a matrix, and second solve the matrix equation $T = R_\theta$ for θ, where as usual R_θ denotes the matrix of the linear transformation that corresponds to rotating the plane by θ angle in the counterclockwise direction.

Let's first track down where e_1 and e_2 end up at:

- e_1 is on the x_1-axis, so reflecting e_1 along the x_1-axis fixes e_1. Next, reflecting it along the x_2-axis now transforms it to $-e_1 = (-1, 0)$. Thus $T(e_1) = -e_1$.

- Now we need to reflect e_2 first along the x_1-axis and then along the x_2-axis (same order as in the previous item). Thus e_2 first transforms into $-e_2 = (0, -1)$, and then since this new vector is on the x_2-axis one has $T(e_2) = -e_2$.

These considerations thus give that the matrix corresponding to T is

$$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -I$$

Thus we now have the equation $R_\theta = -I$ and we want to solve it for θ. Writing R_θ explicitly we have:

$$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix},$$

that is, we want the angle θ to be so that

$$\cos(\theta) = -1, \sin(\theta) = 0.$$

From trigonometry, one immediate solution to this is $\theta = \pi$. Though note that really the above equations have infinitely many solutions, indeed any angle of the form $\theta = \pi + 2\pi k$, where k is an integer, would work. In this sense one could say that the problem is somewhat misleading by asking for "the" angle.

Follow-up Exercise Describe all pairs of vectors $v, w \in \mathbb{R}^2$ such that the linear transformation of the plane obtained by first reflecting along Span{v} and then reflecting along Span{w} can be represented by a rotation.

Follow-up Exercise 2 For θ a fixed angle, describe the set of pairs of vectors $v, w \in \mathbb{R}^2$ such that the linear transformation of the plane obtained by first reflecting along Span{v} and then reflecting along Span{w} can be represented by R_θ.

Follow-up Exercise 3 So far we were considering reflections in the plane only along lines that pass through the origin. Write a transformation that reflects the plane along an arbitrary line

$$x_2 = mx_1 + b.$$

Is this a linear transformation? Can this transformation be stored in a matrix?

3 PSet 3, Problem 3

Consider the family of linear transformations

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

parameterized by θ.

(a) Verify the identity

$$R_\alpha R_\beta = R_{\alpha+\beta} = R_\beta R_\alpha$$

by using trigonometric identities.

(b) Describe each three sides of the identity in the previous part geometrically.

(c) Verify that the following two identities hold for z and w complex numbers:

$$M(z) + M(w) = M(z + w),$$
$$M(z)M(w) = M(zw),$$

where $M(a + ib) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$.

(d) Verify that $M(\bar{z}) = M(z)^T$ by computing both sides.

Most parts of this problem is a matter of calculating both sides of an equation and displaying that the corresponding entries on each side indeed match.

For the first part we have:

$$
\begin{aligned}
&\mathrm{R}_\alpha \mathrm{R}_\beta \\
&= \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \begin{pmatrix} \cos(\beta) & -\sin(\beta) \\ \sin(\beta) & \cos(\beta) \end{pmatrix} \\
&= \begin{pmatrix} \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta) & -\cos(\alpha)\sin(\beta) - \sin(\alpha)\cos(\beta) \\ \sin(\alpha)\cos(\beta) + \cos(\alpha)\sin(\beta) & -\sin(\alpha)\sin(\beta) + \cos(\alpha)\cos(\beta) \end{pmatrix} \\
&= \begin{pmatrix} \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta) & -[\sin(\alpha)\cos(\beta) + \cos(\alpha)\sin(\beta)] \\ \sin(\alpha)\cos(\beta) + \cos(\alpha)\sin(\beta) & \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta) \end{pmatrix} \\
&= \begin{pmatrix} \cos(\alpha+\beta) & -\sin(\alpha+\beta) \\ \sin(\alpha+\beta) & \cos(\alpha+\beta) \end{pmatrix} \\
&= \mathrm{R}_{\alpha+\beta}.
\end{aligned}
$$

Here in the fourth equality we used the trigonometry formulas, and in the fifth equality we used the definition of R_θ with $\theta = \alpha + \beta$. Since $\alpha+\beta = \beta+\alpha$ we also have that $\mathrm{R}_{\alpha+\beta} = \mathrm{R}_{\beta+\alpha}$, and by changing the roles of α and β in the above calculation we also have that $\mathrm{R}_{\beta+\alpha} = \mathrm{R}_\beta \mathrm{R}_\alpha$, completing the first part.

For the second part, note the following two key points:

- R_θ is an invertible linear transformation of the plane that corresponds to rotating the plane around the origin by θ angle in the counterclockwise direction.
- Matrix products correspond to composite transformations: a syntactic expression AB corresponds to first applying B and then A. (Here we are assuming that we are putting the input vectors to the right of B.)

With these two observations, the answer is clear:

- $R_\alpha R_\beta$ is the composite transformation of first rotating by β degrees and then rotating by α degrees.
- $R_{\alpha+\beta}$ is the transformation that rotates by $\alpha + \beta$ degrees.
- $R_\beta R_\alpha$ is the composite transformation of first rotating by α degrees and then rotating by β degrees.

In the previous part then we really verified that matrix mutliplication matches with our geometric intuition that these three transformations are the same. Note that we could also briefly write $[R_\alpha, R_\beta] = 0$.

The third part extrapolates the correspondence between rotations and certain matrices by noting that $\cos(\theta)$ and $\sin(\theta)$ are really the real and imaginary parts of the complex number $e^{i\theta} = \cos(\theta) + i\sin(\theta)$. Generalizing, the question suggests that one can use the schema

$$z = a + ib \rightsquigarrow M(z) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

to store complex numbers in 2×2 matrices. Note that there is some redundancy here at first glance (after all, we are using four slots to store two numbers), but that redundancy really is what allows matrix multiplication to model properly complex number multiplication.

We need two complex numbers z and w. For brevity let's say $z = a + ib$ and $w = c + id$. Thus we may write

$$\begin{aligned} z + w &= (a + c) + i(b + d) \\ zw &= (a + ib)(c + id) = (ac - bd) + i(ad + bc). \end{aligned}$$

For the first equation we have

$$
\begin{aligned}
M(z) + M(w) &= M(a + ib) + M(c + id) \\
&= \begin{pmatrix} a & -b \\ b & a \end{pmatrix} + \begin{pmatrix} c & -d \\ d & c \end{pmatrix} \\
&= \begin{pmatrix} a + c & -b - d \\ b + d & a + c \end{pmatrix} \\
&= \begin{pmatrix} a + c & -(b + d) \\ b + d & a + c \end{pmatrix} \\
&= M((a + c) + i(b + d)) = M(z + w).
\end{aligned}
$$

Similarly, for the second equation we have

$$
\begin{aligned}
M(z) + M(w) &= M(a + ib)M(c + id) \\
&= \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \begin{pmatrix} c & -d \\ d & c \end{pmatrix} \\
&= \begin{pmatrix} ac - bd & -ad - bc \\ bc + ad & -bd + ac \end{pmatrix} \\
&= \begin{pmatrix} ac - bd & -(ad + bc) \\ ad + bc & ac - bd \end{pmatrix} \\
&= M((a + ib)(c + id)) = M(zw).
\end{aligned}
$$

Finally for the fourth part we have

$$
M\left(\overline{z}\right) = M\left(\overline{a + ib}\right) = M(a - ib) = \begin{pmatrix} a & b \\ -b & a \end{pmatrix} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}^T = M(z)^T.
$$

Follow-up Exercise One can similarly define a rotation in any dimension. Find all numbers n such that for any two rotations R_α and R_β of $\mathbb{R}^n$, one has that $[R_\alpha, R_\beta] = 0$.

Follow-up Exercise 2 Since one can store a complex number in a 2×2 matrix with real entries, applying this to matrices with complex entries entry by entry one can also store an $m \times n$ matrix A with complex entries in a $(2m) \times (2n)$ matrix $A_\mathbb{R}$ with real entries:

$$A = \begin{pmatrix} z_{11} & z_{12} & \cdots & z_{1n} \\ z_{21} & z_{22} & \cdots & z_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ z_{m1} & z_{m2} & \cdots & z_{mn} \end{pmatrix}$$

$$\rightsquigarrow A_\mathbb{R} = \begin{pmatrix} M(z_{11}) & M(z_{12}) & \cdots & M(z_{1n}) \\ M(z_{21}) & M(z_{22}) & \cdots & M(z_{2n}) \\ \vdots & \vdots & \ddots & \vdots \\ M(z_{m1}) & M(z_{m2}) & \cdots & M(z_{mn}) \end{pmatrix}$$

For instance applying this to the second Pauli matrix one obtains:

$$(\sigma_2)_\mathbb{R} = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}_\mathbb{R} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

Verify that similar to complex number storage one has

$$A_\mathbb{R} + B_\mathbb{R} = (A + B)_\mathbb{R}$$

$$A_\mathbb{R} B_\mathbb{R} = (AB)_\mathbb{R}.$$

4 PSet 4, Problem 8

Consider the $(p, q) \times (p, q)$ block matrix: $M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$.

(a) If A is invertible, verify that M block row reduces to the block echelon form $\begin{pmatrix} I & A^{-1}B \\ 0 & M/A \end{pmatrix}$.

(b) Suppose A is invertible and compute the matrices X and Y such that

$$M = \begin{pmatrix} I & 0 \\ X & I \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & M/A \end{pmatrix} \begin{pmatrix} I & Y \\ 0 & I \end{pmatrix}.$$

(c) Verify that if A is invertible, then the following two formulas hold:

$$\operatorname{rank}(M) = \operatorname{rank}(A) + \operatorname{rank}(M/A),$$
$$\operatorname{nullity}(M) = \operatorname{nullity}(A) + \operatorname{nullity}(M/A).$$

(d) Verify that if A is invertible, then the invertibility of M is equivalent to the invertibility of M/A.

(e) Suppose both M and A are invertible. Find a formula for the block inverse of M.

One can approach this problem as a matrix algebra problem where the entries are not numbers but matrices themselves. We know that the fundamental distinction between algebra with numbers and algebra with matrices is the noncommutativity of multiplication; hence here too one needs to be mindful of the order factors of products are written.

For part (a) we have:

$$\begin{aligned}\begin{pmatrix} A & B \\ C & D \end{pmatrix} &\xrightarrow{R_1 \leftarrow A^{-1}R_1} \begin{pmatrix} I & A^{-1}B \\ C & D \end{pmatrix} \\ &\xrightarrow{R_2 \leftarrow R_2 - CR_1} \begin{pmatrix} I & A^{-1}B \\ 0 & D - CA^{-1}B \end{pmatrix} \\ &= \begin{pmatrix} I & A^{-1}B \\ 0 & M/A \end{pmatrix}.\end{aligned}$$

Note that the dimensions match for each operations, so that everything is syntactic. Note further that these operations are simply shorthand for standard elementary row operations. For instance, the first block row explicitly is:

$$\begin{pmatrix} A & B \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1p} & b_{11} & \cdots & b_{1q} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{p1} & \cdots & a_{pp} & b_{p1} & \cdots & b_{pq} \end{pmatrix}.$$

Recall that A is invertible if and only if its reduced echelon form is the identity matrix. That is, there are finitely many elementary matrice s $L_1, L_2, ..., L_k$ such that

$$L_k \cdots L_2 L_1 A = I,$$

so that $L_k \cdots L_2 L_1 = A^{-1}$. Multiplying the first row from the left by these matrices then has the following effect:

$$
\begin{aligned}
L_k \cdots L_2 L_1 \begin{pmatrix} A & B \end{pmatrix} &= \begin{pmatrix} L_k \cdots L_2 L_1 A & L_k \cdots L_2 L_1 B \end{pmatrix} \\
&= \begin{pmatrix} I & A^{-1}B \end{pmatrix}
\end{aligned}
$$

Finally, each of the L_i's are $p \times p$, but we can turn them into $(p, q) \times (p, q)$ matrices by considering instead

$$
\begin{pmatrix} L_i & 0 \\ 0 & I \end{pmatrix}.
$$

Here the top right 0 is $p \times q$, the bottom left 0 is $q \times p$, and the bottom right I is $q \times q$. Thus the first elementary block row operation above can be written in terms of matrix multiplication as follows:

$$
\begin{aligned}
&\begin{pmatrix} L_k & 0 \\ 0 & I \end{pmatrix} \cdots \begin{pmatrix} L_2 & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} L_1 & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix} \\
&\qquad = \begin{pmatrix} L_k \cdots L_2 L_1 A & L_k \cdots L_2 L_1 B \\ C & D \end{pmatrix} \\
&\qquad = \begin{pmatrix} I & A^{-1}B \\ C & D \end{pmatrix}.
\end{aligned}
$$

While it is important to be able to unpack block matrix operations, one of the benefits of using block matrices is to lump together certain parts of the matrices to simplify and fasten certain calculations.

For part (b), we may continue with block row and block column operations, or alternatively we may block multiply the matrices on the RHS and set the corresponding blocks equal. Let's employ the latter strategy:

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = M = \begin{pmatrix} I & 0 \\ X & I \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & M/A \end{pmatrix} \begin{pmatrix} I & Y \\ 0 & I \end{pmatrix}$$
$$= \begin{pmatrix} A & 0 \\ XA & M/A \end{pmatrix} \begin{pmatrix} I & Y \\ 0 & I \end{pmatrix}$$
$$= \begin{pmatrix} A & AY \\ XA & XAY + M/A \end{pmatrix}$$

Comparing the top right entries, we have that $B = AY$ and comparing the bottom left entries, we have $C = XA$. Solving for X and Y we obtain

$$\boxed{X = CA^{-1}, \quad Y = A^{-1}B}.$$

Thus we have a block factorization of M:

$$M = \underbrace{\begin{pmatrix} I & 0 \\ CA^{-1} & I \end{pmatrix}}_{\Lambda} \underbrace{\begin{pmatrix} A & 0 \\ 0 & M/A \end{pmatrix}}_{\Delta} \underbrace{\begin{pmatrix} I & A^{-1}B \\ 0 & I \end{pmatrix}}_{\Upsilon} = \Lambda \, \Delta \, \Upsilon,$$

with Λ lower block unitriangular, Δ block diagonal, Υ upper block unitriangular. We will use this factorization for the remaining three parts.

For part (c), note that Λ and Υ are invertible:

$$\Lambda^{-1} = \begin{pmatrix} I & 0 \\ -CA^{-1} & I \end{pmatrix}, \quad \Upsilon^{-1} = \begin{pmatrix} I & -A^{-1}B \\ 0 & I \end{pmatrix}.$$

Consequently the rank of M must be the same as the rank of Δ. Next, note that no column of A interacts with a column of M/A (and vice versa), thus the rank of Δ must be the sum of the ranks of A and M/A:

$$\text{rank}(M) = \text{rank}(\Delta) = \text{rank}(A) + \text{rank}(M/A).$$

Of course, since A is invertible we know that it has rank p, but that information is not needed. Using the Fundamental Theorem of Linear Algebra, we obtain the equation for the nullities also:

$$\begin{aligned}
\text{nullity}(M) &= (p+q) - \text{rank}(M) \\
&= (p+q) - (\text{rank}(A) + \text{rank}(M/A)) \\
&= (p - \text{rank}(A)) + (q - \text{rank}(M/A)) \\
&= \text{nullity}(A) + \text{nullity}(M/A).
\end{aligned}$$

For part (d) we assume again that A is invertible. Note that then the only way Δ would be invertible is if M/A is invertible, since the columns of A and M/A may not have any relations. If M/A is invertible, then so is Δ, and since Λ and Υ are always invertible, that would imply that M is invertible, since it is the product of three invertible matrices and invertible matrices constitute a matrix group. Conversely, if M is invertible, then so is $\Delta = \Lambda^{-1} M \Upsilon^{-1}$, which means that M/A must be invertible.

For the final part, we assume that both M and A are invertible; which means that by the previous part that M/A too is invertible. Using the factorization we have

$$
\begin{aligned}
M^{-1} &= (\Lambda\,\Delta\,\Upsilon)^{-1} = \Upsilon^{-1}\Delta^{-1}\Lambda^{-1} \\
&= \begin{pmatrix} I & -A^{-1}B \\ 0 & I \end{pmatrix} \begin{pmatrix} A^{-1} & 0 \\ 0 & (M/A)^{-1} \end{pmatrix} \begin{pmatrix} I & 0 \\ -CA^{-1} & I \end{pmatrix} \\
&= \begin{pmatrix} A^{-1} & -A^{-1}B(M/A)^{-1} \\ 0 & (M/A)^{-1} \end{pmatrix} \begin{pmatrix} I & 0 \\ -CA^{-1} & I \end{pmatrix} \\
&= \begin{pmatrix} A^{-1} + A^{-1}B(M/A)^{-1}CA^{-1} & -A^{-1}B(M/A)^{-1} \\ -(M/A)^{-1}CA^{-1} & (M/A)^{-1} \end{pmatrix}.
\end{aligned}
$$

Thus we have the following block matrix formula for M^{-1}:

$$
\begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1} =
\begin{pmatrix} A^{-1} + A^{-1}B(D - CA^{-1}B)^{-1}CA^{-1} & -A^{-1}B(D - CA^{-1}B)^{-1} \\ -(D - CA^{-1}B)^{-1}CA^{-1} & (D - CA^{-1}B)^{-1} \end{pmatrix}.
$$

Note that in case $p = q = 1$, this formula reduces to the standard formula for inverses of 2×2 matrices, for instance:

$$
\begin{aligned}
A^{-1} + A^{-1}B(D - CA^{-1}B)^{-1}CA^{-1} &= \frac{1}{a} + \frac{1}{a}b\left(d - c\frac{1}{a}b\right)^{-1} c\frac{1}{a} \\
&= \frac{1}{a} + \frac{bc}{a^2}\left(\frac{ad - bc}{a}\right)^{-1} \\
&= \frac{1}{a} + \frac{bc}{a^2}\frac{a}{ad - bc} \\
&= \frac{1}{a} + \frac{bc}{a(ad - bc)} \\
&= \frac{1}{a}\left(1 + \frac{bc}{ad - bc}\right) \\
&= \frac{1}{a}\frac{ad}{ad - bc} \\
&= \frac{d}{ad - bc}.
\end{aligned}
$$

5 PSet 5, Problem 8

Let Q be a 3×3 orthogonal matrix.

(a) Verify that the determinant of Q is either 1 or -1.

(b) Give an example of a 3×3 orthogonal matrix Q such that

- Q is not the identity matrix, and
- $\det(Q) = 1$.

(c) Give an example of a 3×3 orthogonal matrix Q such that

- Q is not a permutation matrix, and
- $\det(Q) = -1$.

(d) Let C be the cube in $\mathbb{R}^3$ determined by the vectors $(1, 0, 0)$, $(0, 1, 0)$ and $(0, 0, 1)$. Verify that the image of C under Q is a cube with each edge of length 1.

For part (a) one can apply two important properties of the determinant the defining equation for a matrix to be orthogonal; namely that the determinant is multiplicative and that it is invariant under taking transposes:

$$1 = \det(I) = \det(QQ^T) = \det(Q)\det(Q^T) = \det(Q)^2 \implies \det(Q) = \pm 1.$$

Note that this does not mean that any 3×3 matrix with determinant ± 1 is orthogonal[1]. For the second and third parts, to construct the desired examples one can use products of matrices that are easily

[1] Indeed this is not true.

verified to be orthogonal. Note that this is a valid strategy since the collection of all 3×3 orthogonal matrices is a matrix group: if Q and R are orthogonal, then so is QR since

$$\begin{aligned}(QR)(QR)^T &= (QR)(R^TQ^T)\\ &= Q(RR^T)Q^T\\ &= QIQ^T = QQ^T = I\end{aligned}$$

and similarly if Q is orthogonal then part (a) its determinant is nonzero, hence Q^{-1} exists and

$$\begin{aligned}Q^{-1}(Q^{-1})^T &= I(Q^{-1}(Q^{-1})^T)\\ &= (Q^TQ)(Q^{-1}(Q^{-1})^T)\\ &= Q^T(QQ^{-1})(Q^{-1})^T\\ &= Q^TI(Q^{-1})^T = Q^T(Q^{-1})^T\\ &= (Q^{-1}Q)^T = I^T = I\end{aligned}$$

We know that permutation matrices are orthogonal, so taking the product of two different permutation matrices corresponding to row swap operations would work for part (b)[2]; for instance

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

Part (c) is slightly more tricky since now the end result can not

[2] Such a product would again be a permutation matrix as permutation matrices too constitute a matrix group; one would say that the group of permutation matrices is a "subgroup" of the group of orthogonal matrices.

be a permutation matrix. Still we can use a permutation matrix as a factor; if this permutatioon matrix has determinant –1, its product with any rotation matrix would have determinant –1. For instance

$$\begin{aligned}
Q &= \begin{pmatrix} R_\theta & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \\
&= \begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \\
&= \begin{pmatrix} \cos(\theta) & 0 & -\sin(\theta) \\ \sin(\theta) & 0 & \cos(\theta) \\ 0 & 1 & 0 \end{pmatrix}
\end{aligned}$$

Choosing the angle for instance $\theta = \frac{\pi}{4}$ thus gives

$$Q = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \end{pmatrix}.$$

Certainly this is not a permutation matrix as it has entries different from 0 and 1, and further by the multiplicativity of determinant it has determinant –1 and since orthogonal matrices is a group the above Q is an orthogonal matrix.

Finally for part (d) one can argue as follows. First let's write the orthogonality equation for an arbitrary orthogonal matrix Q explicitly:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} q_{11} & q_{12} & q_{13} \\ q_{21} & q_{22} & q_{23} \\ q_{31} & q_{32} & q_{33} \end{pmatrix} \begin{pmatrix} q_{11} & q_{12} & q_{13} \\ q_{21} & q_{22} & q_{23} \\ q_{31} & q_{32} & q_{33} \end{pmatrix}^T$$
$$= \begin{pmatrix} q_{11} & q_{12} & q_{13} \\ q_{21} & q_{22} & q_{23} \\ q_{31} & q_{32} & q_{33} \end{pmatrix} \begin{pmatrix} q_{11} & q_{21} & q_{31} \\ q_{12} & q_{22} & q_{32} \\ q_{13} & q_{23} & q_{33} \end{pmatrix}$$

Thus comparing the diagonal entries above gives that for $j = 1, 2, 3$

$$q_{1j}^2 + q_{2j}^2 + q_{3j}^2 = 1$$

Each of these three equations corresponds to a column of Q, and consequently each of them corresponds to one of the three edges of the parallelepided Q(C) connected to the vertex at the origin. Interpreting the equations geometrically, we have that the image parallelepiped Q(C) must have all edges of length 1.

Since the determinant of Q is ± 1 by part (a), we know that the volume of Q(C) is 1. So we have a parallelepiped with each side of length 1 and with volume 1, and we need to argue that it's actually a cube (so that the edges of the parallelepiped are orthogonal to eachother, hence the name "orthogonal matrix")[3]. Taking one face as the base B, the volume is equal to height h times the area of B. Similarly taking one of the edges of B as the base, the area of B is equal to depth d times the length of the base, which based on the previous paragraph is 1. Equating the two expressions for volume we have

$$1 = \text{volume}(\,Q(C)\,) = h\ \text{area}(\,B\,) = hd1 = hd.$$

[3]The rest of the argument is geometric; so it might be useful to draw pictures to follow.

Further, note that both h and d can at most be 1; since they can not be longer than the edges of Q(C). Thus we have

$$\boxed{1 = hd,\ \ 0 < h \leq 1,\ \ 0 < d \leq 1}$$

The only way this can happen is if h = d = 1, which can only happen if Q is actually a cube, so that all edges are perpendicular to each other.

Follow-Up Exercise Describe the collection of all nonorthogonal 3×3 matrices with determinant ± 1.

Follow-up Exercise 2 Let C be the cube in $\mathbb{R}^3$ determined by the vectors $(1, 0, 0)$, $(0, 1, 0)$ and $(0, 0, 1)$. Give an example of a nonorthogonal 3×3 matrix A with determinant 1 and sketch the image A(C).

Follow-Up Exercise 3 Let C be the cube in $\mathbb{R}^3$ determined by the vectors $(1, 0, 0)$, $(0, 1, 0)$ and $(0, 0, 1)$, and let A be a 3×3 matrix. If the image of C under A is a cube with each edge of length 1, then is A orthogonal?

6 PSet 6, Problem 5

Consider the following functions:

I: $f_I(t) = \begin{cases} 1/2 & \text{, if } 0 \leq t \leq 2 \\ -1 & \text{, if } -1 \leq t < 0 \\ 0 & \text{, else} \end{cases}$

II: $f_{II}(t) = e^{2t}$

III: $f_{III}(t) = \cos(3t) + |t|$

IV: $f_{IV}(t) = \begin{cases} e^{2t} + \cos(3t) + 2|t| + 1/2 & \text{, if } 0 \leq t \leq 2 \\ e^{2t} + \cos(3t) + 2|t| - 1 & \text{, if } -1 \leq t < 0 \\ e^{2t} + \cos(3t) + 2|t| & \text{, else} \end{cases}$

Compute the even and odd parts of each of the given functions.

For this problem we need the definitions of the even and odd parts of a function. Recall that for f an arbitrary single variable, real valued function (aka a vector in the vector space $\mathcal{F} = \mathcal{F}(\mathbb{R};\mathbb{R})$), the even and odd parts are defined as follows:

$$f_{even}(t) = \frac{f(t) + f(-t)}{2}, \quad f_{odd}(t) = \frac{f(f) - f(-t)}{2},$$

in other words the even part of a function f is the function that is the image of f under the even projector and likewise the odd part of f is the function that is the image of f under the odd projector. In particular the even and odd parts of a function are unique. Finally the even and odd projectors are linear transformations from $\mathcal{F}$ to itself; this observation will come in handy for the fourth part as the fourth function is a certain linear combination of the previos examples.

Let's first start with part I. This function is piecewise constant and for such functions one can use the formula for the even and odd parts either graphically or algebraically. For most functions given sketching graphs (at least by hand) would not be precise enough, so let's take the algebraic approach:

$$
\begin{aligned}
f_{I,even}(t) &= \frac{1}{2}(f_I(t) + f_I(-t)) \\
&= \frac{1}{2}\left(\begin{cases} 1/2 & \text{, if } 0 \leq t \leq 2 \\ -1 & \text{, if } -1 \leq t < 0 \\ 0 & \text{, else} \end{cases} + \begin{cases} 1/2 & \text{, if } 0 \leq -t \leq 2 \\ -1 & \text{, if } -1 \leq -t < 0 \\ 0 & \text{, else} \end{cases}\right) \\
&= \frac{1}{2}\left(\begin{cases} 1/2 & \text{, if } 0 \leq t \leq 2 \\ -1 & \text{, if } -1 \leq t < 0 \\ 0 & \text{, else} \end{cases} + \begin{cases} 1/2 & \text{, if } -2 \leq t \leq 0 \\ -1 & \text{, if } 0 < t \leq 1 \\ 0 & \text{, else} \end{cases}\right) \\
&= \frac{1}{2}\left(\begin{cases} 1/2 & \text{, if } 1 < t \leq 2 \\ 1/2 & \text{, if } 0 < t \leq 1 \\ 1/2 & \text{, if } t = 0 \\ -1 & \text{, if } -1 \leq t < 0 \\ 0 & \text{, if } -2 \leq t < -1 \\ 0 & \text{, else} \end{cases} + \begin{cases} 0 & \text{, if } 1 < t \leq 2 \\ -1 & \text{, if } 0 < t \leq 1 \\ 1/2 & \text{, if } t = 0 \\ 1/2 & \text{, if } -1 \leq t < 0 \\ 1/2 & \text{, if } -2 \leq t < -1 \\ 0 & \text{, else} \end{cases}\right)
\end{aligned}
$$

$$= \frac{1}{2}\left(\begin{cases} 1/2 & , \text{if } 1 < t \leq 2 \\ -1/2 & , \text{if } 0 < t \leq 1 \\ 1 & , \text{if } t = 0 \\ -1/2 & , \text{if } -1 \leq t < 0 \\ 1/2 & , \text{if } -2 \leq t < -1 \\ 0 & , \text{else} \end{cases}\right).$$

In the above calculation since we are adding two piecewise defined functions we further split the pieces into common pieces to make addition simpler. Thus we have the even part of f_I:

$$\boxed{f_{I,even}(t) = \begin{cases} 1/4 & , \text{if } 1 < t \leq 2 \\ -1/4 & , \text{if } 0 < t \leq 1 \\ 1/2 & , \text{if } t = 0 \\ -1/4 & , \text{if } -1 \leq t < 0 \\ 1/4 & , \text{if } -2 \leq t < -1 \\ 0 & , \text{else} \end{cases} = \begin{cases} 1/4 & , \text{if } 1 < |t| \leq 2 \\ -1/4 & , \text{if } 0 < |t| \leq 1 \\ 1/2 & , \text{if } t = 0 \\ 0 & , \text{else} \end{cases}}$$

Similarly doing the corresponding calculation for the odd part gives this:

$$
\begin{aligned}
f_{I,odd}(t) &= \frac{1}{2}(f_I(t) - f_I(-t)) \\
&= \frac{1}{2}\left(\begin{cases} 1/2 & \text{, if } 0 \leq t \leq 2 \\ -1 & \text{, if } -1 \leq t < 0 \\ 0 & \text{, else} \end{cases} - \begin{cases} 1/2 & \text{, if } 0 \leq -t \leq 2 \\ -1 & \text{, if } -1 \leq -t < 0 \\ 0 & \text{, else} \end{cases}\right) \\
&= \frac{1}{2}\left(\begin{cases} 1/2 & \text{, if } 1 < t \leq 2 \\ 1/2 & \text{, if } 0 < t \leq 1 \\ 1/2 & \text{, if } t = 0 \\ -1 & \text{, if } -1 \leq t < 0 \\ 0 & \text{, if } -2 \leq t < -1 \\ 0 & \text{, else} \end{cases} + \begin{cases} 0 & \text{, if } 1 < t \leq 2 \\ 1 & \text{, if } 0 < t \leq 1 \\ -1/2 & \text{, if } t = 0 \\ -1/2 & \text{, if } -1 \leq t < 0 \\ -1/2 & \text{, if } -2 \leq t < -1 \\ 0 & \text{, else} \end{cases}\right) \\
&= \frac{1}{2}\left(\begin{cases} 1/2 & \text{, if } 1 < t \leq 2 \\ 3/2 & \text{, if } 0 < t \leq 1 \\ 0 & \text{, if } t = 0 \\ -3/2 & \text{, if } -1 \leq t < 0 \\ -1/2 & \text{, if } -2 \leq t < -1 \\ 0 & \text{, else} \end{cases}\right).
\end{aligned}
$$

Thus the odd part of f_I is:

$$\boxed{f_{I,odd}(t) = \begin{cases} 1/4 & \text{, if } 1 < t \leq 2 \\ 3/4 & \text{, if } 0 < t \leq 1 \\ 0 & \text{, if } t = 0 \\ -3/4 & \text{, if } -1 \leq t < 0 \\ -1/4 & \text{, if } -2 \leq t < -1 \\ 0 & \text{, else} \end{cases}}$$

Part II is somewhat shorter, as there are no simplifications at hand:

$$\boxed{f_{II,even}(t) = \frac{e^{2t} + e^{-2t}}{2}, \quad f_{II,odd}(t) = \frac{e^{2t} - e^{-2t}}{2}}.$$

The third part is even faster: the space of even functions is a linear subspace, and f_{III} is the sum of two even functions, hence is even already. Consequently

$$\boxed{f_{III,even} = f_{III}, \quad f_{III,odd} = 0}.$$

Finally for part IV note that

$$f_{IV}(t) = f_I(t) + f_{II}(t) + f_{III}(t) + |t|,$$

and since even and odd projectors are linear operators we have

$$f_{IV,even}(t) = f_{I,even}(t) + f_{II,even}(t) + f_{III}(t) + |t|,$$
$$f_{IV,odd}(t) = f_{I,odd}(t) + f_{II,odd}(t).$$

Writing these out explicitly we obtain

$$f_{IV,even}(t) = \begin{cases} \frac{e^{2t}+e^{-2t}}{2} + \cos(3t) + 2|t| + 1/4 & \text{, if } 1 < t \leq 2 \\ \frac{e^{2t}+e^{-2t}}{2} + \cos(3t) + 2|t| - 1/4 & \text{, if } 0 < t \leq 1 \\ \frac{e^{2t}+e^{-2t}}{2} + \cos(3t) + 2|t| + 1/2 & \text{, if } t = 0 \\ \frac{e^{2t}+e^{-2t}}{2} + \cos(3t) + 2|t| - 1/4 & \text{, if } -1 \leq t < 0 \\ \frac{e^{2t}+e^{-2t}}{2} + \cos(3t) + 2|t| + 1/4 & \text{, if } -2 \leq t < -1 \\ \frac{e^{2t}+e^{-2t}}{2} + \cos(3t) + 2|t| & \text{, else} \end{cases}$$

and

$$f_{IV,odd}(t) = \begin{cases} \frac{e^{2t}-e^{-2t}}{2} + 1/4 & \text{, if } 1 < t \leq 2 \\ \frac{e^{2t}-e^{-2t}}{2} + 3/4 & \text{, if } 0 < t \leq 1 \\ 0 & \text{, if } t = 0 \\ \frac{e^{2t}-e^{-2t}}{2} - 3/4 & \text{, if } -1 \leq t < 0 \\ \frac{e^{2t}-e^{-2t}}{2} - 1/4 & \text{, if } -2 \leq t < -1 \\ \frac{e^{2t}-e^{-2t}}{2} & \text{, else} \end{cases}$$

For graphs of such functions the interactive graphing calculator at the following link may be useful:

https://www.desmos.com/calculator/7dya0enrjg.

7 PSet 7, Problem 2

Let $\mathcal{P}_2$ be the space of polynomial functions of degree at most 2.

(a) Verify that

$$\beta = (\, 1 - t^2 \,,\, t - t^2 \,,\, 2 - 2t + t^2 \,).$$

is a basis of $\mathcal{P}_2$.

(b) Compute the matrix representation of the vector

$$p(t) = 3 + t - 6t^2$$

relative to the basis β.

(c) Construct another basis γ of $\mathcal{P}_2$ such that no vector in β is in γ.

(d) Compute the matrix representation of the vector

$$p(t) = 3 + t - 6t^2$$

relative to the basis γ.

(e) Compute the change-of-basis matrix from β to γ.

(f) Compute the change-of-basis matrix from γ to β.

(g) Compute the matrix representation of differentiation $D : \mathcal{P}_2 \to \mathcal{P}_2$ with β as both the input basis and the output basis.

(h) Compute the matrix representation of twice-differentiation $D^2 : \mathcal{P}_2 \to \mathcal{P}_2$ with β as the input basis and γ as the output basis.

Before working on the separate parts of the problem let's first set up some notation and do some preliminary calculations. For brevity let's use the following abbreviations for the vectors in the proposed basis β:

$$\beta_1(t) = 1 - t^2, \quad \beta_2(t) = t - t^2, \quad \beta_3(t) = 2 - 2t + t^2.$$

Taking an arbitrary linear combination, we obtain:

$$\begin{aligned}
(c_1\beta_1 + c_2\beta_2 + c_3\beta_3)(t) &= c_1\beta_1(t) + c_2\beta_2(t) + c_3\beta_3(t) \\
&= c_1(1 - t^2) + c_2(t - t^2) + c_3(2 - 2t + t^2) \\
&= \underbrace{(c_1 + 2c_3)}_{=d_1} 1 + \underbrace{(c_2 - 2c_3)}_{=d_2} t + \underbrace{(-c_1 - c_2 + c_3)}_{d_3} t^2 \\
&= d_1 + d_2 t + d_3 t^3.
\end{aligned}$$

Writing out the relations between c_1, c_2, c_3 and d_1, d_2, d_3 gives a system of three linear equations with three unknowns:

$$\begin{aligned} c_1 + 2c_3 &= d_1 \\ c_2 - 2c_3 &= d_2 \\ -c_1 - c_2 + c_3 &= d_3, \end{aligned}$$

which reads in matrix form like so:

$$\underbrace{\begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & -2 \\ -1 & -1 & 1 \end{pmatrix}}_{=A} \underbrace{\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}}_{=c} = \underbrace{\begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix}}_{=d};$$

$Ac = d$ in short. Note that the column matrix d stores the coefficients of any polynomial in $\mathcal{P}_2$ relative to the standard basis $\sigma = (\sigma_1(t) = 1, \sigma_2(t) = t, \sigma_3(t) = t^2)$.

Let's first calculate the determinant of A to see if it's invertible. Using the triangular formula we have:

$$\begin{aligned} \det(A) &= \begin{vmatrix} 1 & 0 & 2 \\ 0 & 1 & -2 \\ -1 & -1 & 1 \end{vmatrix} \\ &= [(1.1.1) + (0.(-2).(-1)) + (0.(-1).(2))] \\ &\qquad - [(2.1.(-1)) + (0.0.1) + (1.(-1)(-2))] \\ &= [1] - [-2 + 2] = 1. \end{aligned}$$

Thus A is indeed invertible. We may also compute its inverse by row reducing the augmented matrix $\begin{pmatrix} A & I \end{pmatrix}$:

$$\begin{pmatrix} A & I \end{pmatrix} = \begin{pmatrix} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 1 & -2 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix} \xrightarrow{R_2 \leftarrow R_2 + 2R_3} \begin{pmatrix} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 1 & 0 & 2 & 3 & 2 \\ 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}$$

$$\xrightarrow{R_1 \leftarrow R_1 - 2R_3} \begin{pmatrix} 1 & 0 & 0 & -1 & -2 & -2 \\ 0 & 1 & 0 & 2 & 3 & 2 \\ 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix};$$

so that

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

With these preliminary calculations we can move on to the parts of the problem.

For part (a), β being a basis means that its vectors are both linearly independent and they span $\mathcal{P}_2$. These two properties correspond to A being one-to-one (aka its nullity being zero), and A being onto (aka its rank being three). Since A is a square matrix, it is one-to-one if and only if (iff for short) it is onto iff it is invertible; which is the case based on our preliminary calculations, so that β is indeed a basis for $\mathcal{P}_2$.

Consequently the column matrix c above would store matrix representations of polynomials relative to the basis β, and A would be the change-of-basis matrix from β to the standard basis σ. Accordingly A^{-1} would be the change-of-basis matrix from σ to β. In terms of the universal diagram this means:

$$\begin{tikzcd}
\mathcal{P}_2 \arrow[r, "\mathrm{id}"] \arrow[d, "\mathcal{M}_\beta"'] & \mathcal{P}_2 \arrow[d, "\mathcal{M}_\sigma"] \\
\mathcal{M}(3 \times 1; \mathbb{R}) \arrow[r, "\mathrm{L_A}"'] & \mathcal{M}(3 \times 1; \mathbb{R})
\end{tikzcd}$$

For part (b), based on the previous paragraph, it suffices to compute

$$\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = c = A^{-1}d = \begin{pmatrix} -1 & -2 & -2 \\ 2 & 3 & 2 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ 1 \\ -6 \end{pmatrix} = \begin{pmatrix} 7 \\ 3 \\ -2 \end{pmatrix}.$$

For part (c) we might as well take $\gamma = \sigma$ to be the standard basis. With this choice we are already done parts (d), (e), (f).

For part (g) we'll use the universal diagram:

$$\begin{tikzcd}
\mathcal{P}_2 \arrow[r, "\mathrm{D}"] \arrow[d, "\mathcal{M}_\beta"', "\cong"] & \mathcal{P}_2 \arrow[d, "\mathcal{M}_\beta", "\cong"'] \\
\mathcal{M}(3 \times 1; \mathbb{R}) \arrow[r, "\mathrm{L_B}"'] & \mathcal{M}(3 \times 1; \mathbb{R})
\end{tikzcd}$$

Here the horizontal arrow at the bottom represents left multiplication by the matrix B; B is exactly the matrix representation of differentiation relative to $\beta \leftarrow \beta$. To discover B, all we need to do is to probe it using sufficiently many linearly independent column matrices. For this we'll ultimately need to differentiate some polynomials, and convert them back to column matrices pre- and post-differentiation. We can certainly use the basis vectors in β, but to use our change-of-basis matrix A let's instead use the standard basis vectors and their derivatives:

$$D(\sigma_1)(t) = 1' = 0$$
$$D(\sigma_2)(t) = (t)' = 1 = \sigma_1(t)$$
$$D(\sigma_3)(t) = (t^2)' = 2t = 2\sigma_2(t).$$

This calculation means, in terms of the universal diagram, that for

$$S = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

we have

$$\begin{tikzcd}
\mathcal{P}_2 \arrow[r, "D"] \arrow[d, "\mathcal{M}_\sigma"'] & \mathcal{P}_2 \arrow[d, "\mathcal{M}_\sigma"] \\
\mathcal{M}(3 \times 1; \mathbb{R}) \arrow[r, "L_S"'] & \mathcal{M}(3 \times 1; \mathbb{R})
\end{tikzcd}$$

To compute B, we may use the universal diagrams with A and S at the bottom horizontal arrows as follows:

$$\begin{tikzcd}
\mathcal{P}_2 \arrow[r, "id"] \arrow[d, "\mathcal{M}_\beta"'] \arrow[rrr, dashed, bend left, "D=id\circ D\circ id"] & \mathcal{P}_2 \arrow[r, "D"] \arrow[d, "\mathcal{M}_\sigma"'] & \mathcal{P}_2 \arrow[r, "id"] \arrow[d, "\mathcal{M}_\sigma"] & \mathcal{P}_2 \arrow[d, "\mathcal{M}_\beta"] \\
\mathcal{M}(3 \times 1; \mathbb{R}) \arrow[r, "L_A"'] \arrow[rrr, dashed, bend right, "L_B=L_{A^{-1}SA}=L_{A^{-1}}\circ L_S\circ L_A"'] & \mathcal{M}(3 \times 1; \mathbb{R}) \arrow[r, "L_S"'] & \mathcal{M}(3 \times 1; \mathbb{R}) \arrow[r, "L_{A^{-1}}"'] & \mathcal{M}(3 \times 1; \mathbb{R})
\end{tikzcd}$$

The dashed arrows are obtained by considering the composite transformations obtained by following arrows in series, and the fact that the bottom dashed arrow is left multiplication by B is due to the fact that universal diagrams are universal. Thus to compute the unknown matrix B the following computation is sufficient:

$$\begin{aligned} B &= A^{-1}SA \\ &= \begin{pmatrix} -1 & -2 & -2 \\ 2 & 3 & 2 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & -2 \\ -1 & -1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} -1 & -2 & -2 \\ 2 & 3 & 2 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & -2 \\ -2 & -2 & 2 \\ 0 & 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 4 & 3 & -2 \\ -6 & -4 & 2 \\ -2 & -1 & 0 \end{pmatrix}. \end{aligned}$$

Finally for part (h), again using the universal diagrams with A and S on the bottom horizontal arrows suffices; for C the unknown matrix we have:

$$
\begin{tikzcd}
\mathcal{P}_2 \arrow[r, "\mathrm{id}"] \arrow[d, "\mathcal{M}_\beta"'] \arrow[rrr, bend left, dashed, "\mathrm{D}^2=\mathrm{D}\circ\mathrm{D}\circ\mathrm{id}"] & \mathcal{P}_2 \arrow[r, "\mathrm{D}"] \arrow[d, "\mathcal{M}_\sigma"'] & \mathcal{P}_2 \arrow[r, "\mathrm{D}"] \arrow[d, "\mathcal{M}_\sigma"] & \mathcal{P}_2 \arrow[d, "\mathcal{M}_\sigma"] \\
\mathcal{M}(3\times 1;\mathbb{R}) \arrow[r, "\mathrm{L_A}"'] \arrow[rrr, bend right, dashed, "\mathrm{L_C}=\mathrm{L}_{\mathrm{S}^2\mathrm{A}}=\mathrm{L}_\mathrm{S}^2\circ\mathrm{L_A}"'] & \mathcal{M}(3\times 1;\mathbb{R}) \arrow[r, "\mathrm{L_S}"'] & \mathcal{M}(3\times 1;\mathbb{R}) \arrow[r, "\mathrm{L_S}"'] & \mathcal{M}(3\times 1;\mathbb{R})
\end{tikzcd}
$$

Thus it suffices to perform the following computation:

$$
\begin{aligned}
\mathrm{C} &= \mathrm{S}^2\mathrm{A} \\
&= \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & -2 \\ -1 & -1 & 1 \end{pmatrix} \\
&= \begin{pmatrix} 0 & 0 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & -2 \\ -1 & -1 & 1 \end{pmatrix} \\
&= \begin{pmatrix} -2 & -2 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.
\end{aligned}
$$

Follow-up Exercise Let V be a finite dimensional vector space. Verify the following:

1. Given two bases β, γ of V, there is a unique invertible square matrix that is the change-of-basis matrix from β to γ.
2. Given two bases β, γ of V, there is a unique invertible transformation from V to itself whose matrix is the identity matrix when the input basis is β and the output basis is γ.
3. Given a basis β of V and an invertible linear transformation T from V to itself, then there is a unique basis γ such that the matrix of T is the identity matrix when the input basis is β and the output basis is γ.
4. Given a basis β of V and an invertible square matrix A, there is a unique basis γ such that A is the change-of-basis matrix from β to γ.

Follow-up Exercise 2 Let

- V, W be finite dimensional vector spaces,
- n = dim(V), m = dim(W),
- T : V $\to$ W be a linear transformation,
- β be a basis of V and γ be a basis of W,
- A $\in \mathcal{M}$(m $\times$ n; $\mathbb{F}$) be a matrix,
- $L_A : \mathcal{M}(n \times 1; \mathbb{F}) \to \mathcal{M}(m \times 1; \mathbb{F})$ be the linear transformation that is left multiplication by A.

Consider the set X of all tuples $(V, n, \beta; W, m, \gamma; T, A)$ as above. Identify minimal descriptions of the subset Y of X that consist of all those tuples that fit into the universal diagram

$$\begin{tikzcd}
V \arrow[rr, "T"] \arrow[d, "\mathcal{M}_\beta"', "\cong"] & & W \arrow[d, "\cong"', "\mathcal{M}_\gamma"] \\
\mathcal{M}(n \times 1; \mathbb{F}) \arrow[rr, "L_A"'] & & \mathcal{M}(m \times 1; \mathbb{F})
\end{tikzcd} .$$

For instance for $(V, n, \beta; W, m, \gamma; T, A) \in X$, once one knows β, one also knows the number n as well as V, and similarly if one knows V, then one knows n but not necessarily β. The universality of the universal diagram is the statement that for $(V, n, \beta; W, m, \gamma; T, A) \in Y$, if the first seven entries of the tuple are known, then there is a unique last entry.

Follow-up Exercise 3 Let U, V, W be three finite dimensional vector spaces with bases β, γ, δ, respectively. Let $T : U \to V$ and $S : V \to W$ be linear transformations.

1. Denote by C the matrix representation of the composite transformation $S \circ T$ relative to $\delta \leftarrow \beta$, by A the matrix representation of S relative to $\delta \leftarrow \gamma$ and by B the matrix representation of T relative to $\gamma \leftarrow \beta$. Verify that C = AB.

2. Draw the universal diagrams for T and S and interpret the previous part in terms of universal diagrams.

Follow-up Exercise 4 Construct a basis δ of $\mathcal{P}_2$ such that no vector in δ is in β nor in the standard basis σ.

8 PSet 8, Problem 2

Consider the family of matrices $A_\alpha = \begin{pmatrix} \alpha & -1 \\ 1 & 0 \end{pmatrix}$, where α is a parameter. Perform the following tasks.

(a) Compute the eigenvalues of A_α.

(b) Sketch how the spectrum of A_α changes as one varies the parameter α.

(c) For each eigenvalue of A_α, compute a basis for the associated eigenspace.

(d) Sketch each eigenspace of A_α.

(e) Compute the algebraic and geometric multiplicities of each distinct eigenvalue of A_α.

(f) For which values of α is A_α diagonalizable?

(g) For any parameter α such that A_α is diagonalizable, diagonalize A_α.

We have at hand a family of matrices parameterized by α, so our answers will likely need to depend on α also.

For part (a), the characteristic equation is

$$0 = \det(\lambda I - A_\alpha) = \begin{vmatrix} \lambda - \alpha & 1 \\ -1 & \lambda \end{vmatrix} = \lambda(\lambda - \alpha) + 1 = \lambda^2 - \alpha\lambda + 1.$$

Here we think of λ as the variable. Applying the quadratic formula, we have a formula for the eigenvalues of A_α:

$$\lambda_1 = \frac{\alpha + \sqrt{\alpha^2 - 4}}{2}, \qquad \lambda_2 = \frac{\alpha - \sqrt{\alpha^2 - 4}}{2}$$

For part (b), we first note that, depending on the value of the parameter α, the eigenvalues of A_α can be distinct real, or repeated real, or complex conjugate; indeed,

- If $\alpha > 2$, then $\lambda_2 = \dfrac{\alpha - \sqrt{\alpha^2 - 4}}{2} < \lambda_1 = \dfrac{\alpha + \sqrt{\alpha^2 - 4}}{2}$ are distinct real. Further, since

  $$\alpha^2 - 4 < \alpha^2 \implies \sqrt{\alpha^2 - 4} < \alpha \implies 0 < \lambda_2$$

  both eigenvalues are positive and since

  $$\begin{aligned} 2 < \alpha &\implies 8 < 4\alpha \implies (\alpha - 2)^2 < \alpha^2 - 4 \\ &\implies \alpha - 2 < \sqrt{\alpha^2 - 4} \implies \alpha - \sqrt{\alpha^2 - 4} < 2, \end{aligned}$$

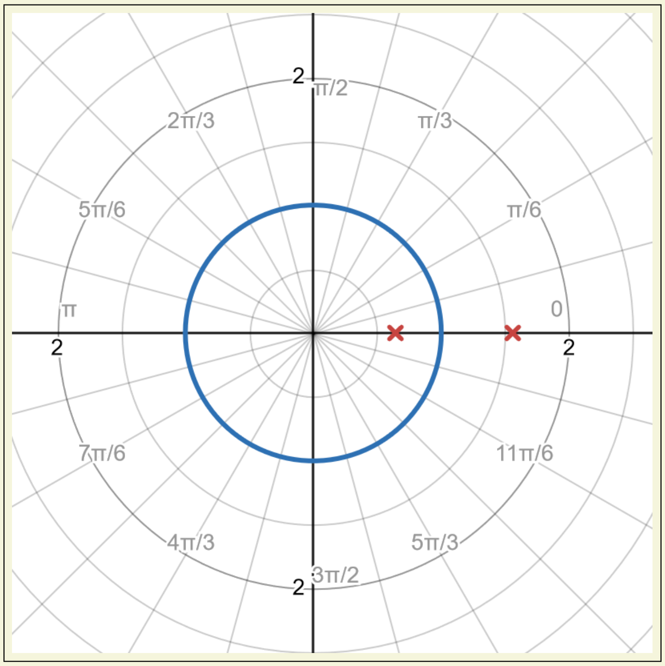
  λ_2 is less than 1 and similarly

  $$2 < \alpha \implies 1 < \frac{\alpha}{2} < \frac{\alpha + \sqrt{\alpha^2 - 4}}{2} = \lambda_1,$$

  so that more accurately in this case we have

  $$\boxed{0 < \lambda_2 < 1 < \lambda_1}.$$

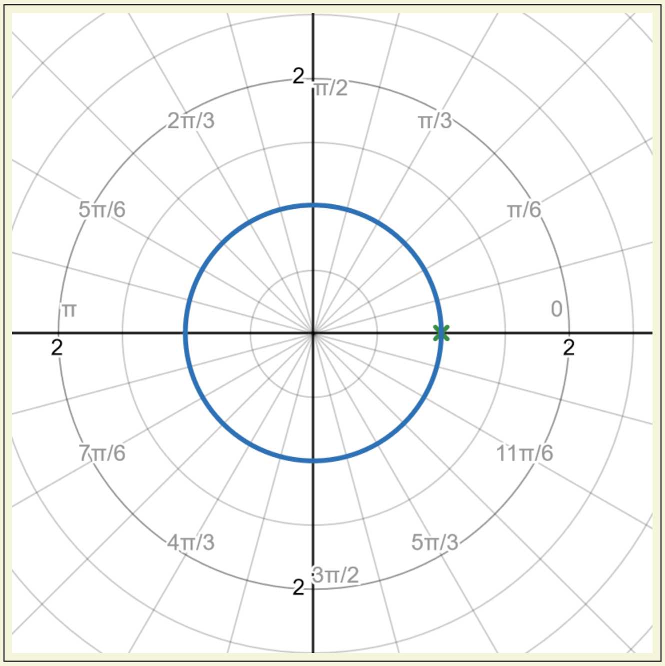
  Using these inequalities we can sketch the spectrum as follows:



- If $\alpha = 2$, then we have repeated real roots and

$$\boxed{\lambda_1 = \lambda_2 = 1};$$

thus in this case we can sketch the spectrum like so:



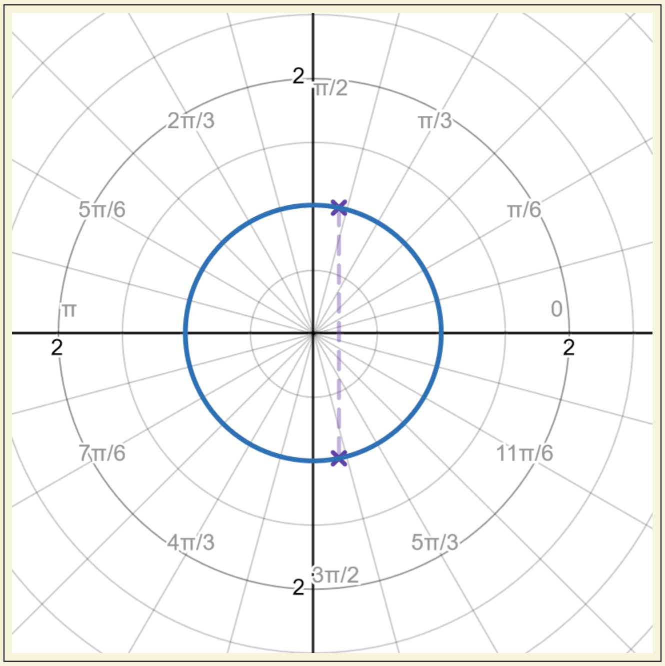
- If $-2 < \alpha < 2$, then $\alpha^2 - 4 < 0$, thus in this case we will have complex conjugate roots:

$$\boxed{\lambda_1 = \frac{\alpha}{2} + i\frac{\sqrt{4-\alpha^2}}{2}, \qquad \lambda_2 = \frac{\alpha}{2} - i\frac{\sqrt{4-\alpha^2}}{2}}.$$

  Note that

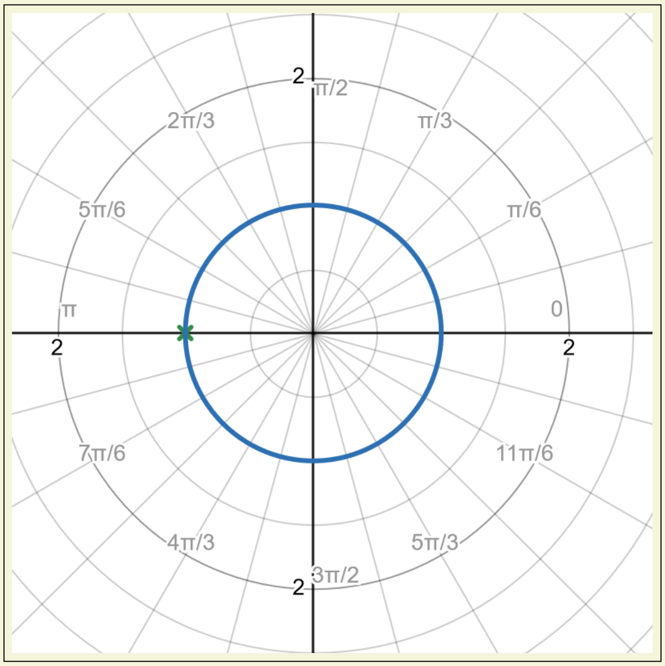
$$|\lambda_i|^2 = \frac{\alpha^2}{4} + \frac{4-\alpha^2}{4} = 1,$$

  so that in this case the eigenvalues will be exactly on the unit circle; here is a representative sketch in this case:



  When $\alpha = 0$, note that the eigenvalues will be exactly $\pm i$.

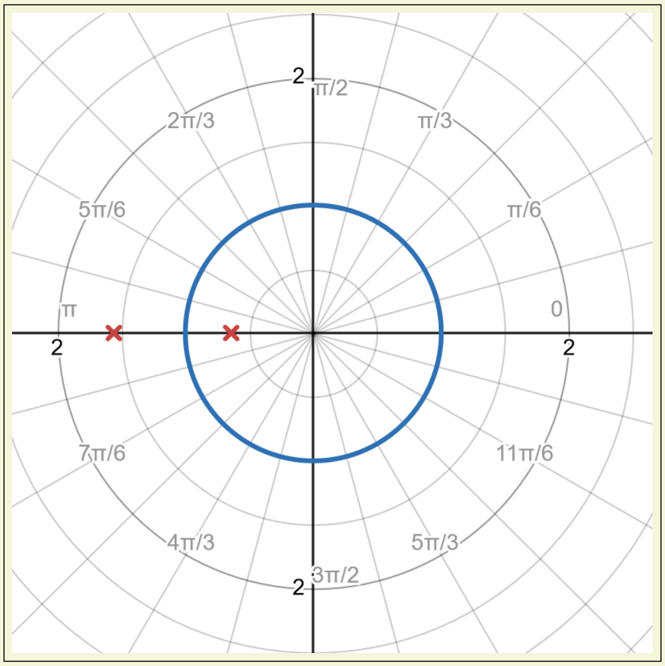
- When $\alpha = -2$, we again have repeated real roots and the sketch for the spectrum looks like so:



- And finally when $\alpha < -2$, arguments similar to the first case guarantee that

$$\lambda_2 < -1 < \lambda_1 < 0,$$

so that a sketch would look as follows:



For an interactive graph that can display all such sketches and further animate how the spectrum varies as the parameter α varies, see

https://www.desmos.com/calculator/g1jzgysgqg.

For part (c), we need eigenvectors for λ_1 and λ_2. For λ_1 we need a vector in the nullspace of

$$
\begin{aligned}
\lambda_1 I - A_\alpha &= \begin{pmatrix} \dfrac{\alpha + \sqrt{\alpha^2 - 4}}{2} - \alpha & 1 \\ -1 & \dfrac{\alpha + \sqrt{\alpha^2 - 4}}{2} \end{pmatrix} \\
&= \begin{pmatrix} -\dfrac{\alpha - \sqrt{\alpha^2 - 4}}{2} & 1 \\ -1 & \dfrac{\alpha + \sqrt{\alpha^2 - 4}}{2} \end{pmatrix} \\
&= \begin{pmatrix} -\lambda_2 & 1 \\ -1 & \lambda_1 \end{pmatrix}
\end{aligned}
$$

so that

$$
\boxed{v_1 = \begin{pmatrix} 1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} 1 \\ \dfrac{\alpha - \sqrt{\alpha^2 - 4}}{2} \end{pmatrix}}
$$

is an eigenvector for λ_1. Similarly for λ_2 we need a vector in the nullspace of

$$\lambda_2 \mathrm{I} - A_\alpha = \begin{pmatrix} \frac{\alpha - \sqrt{\alpha^2 - 4}}{2} - \alpha & 1 \\ -1 & \frac{\alpha - \sqrt{\alpha^2 - 4}}{2} \end{pmatrix}$$
$$= \begin{pmatrix} -\frac{\alpha + \sqrt{\alpha^2 - 4}}{2} & 1 \\ -1 & \frac{\alpha - \sqrt{\alpha^2 - 4}}{2} \end{pmatrix}$$
$$= \begin{pmatrix} -\lambda_1 & 1 \\ -1 & \lambda_2 \end{pmatrix}$$

so that

$$\boxed{\mathrm{v}_2 = \begin{pmatrix} 1 \\ \lambda_1 \end{pmatrix} = \begin{pmatrix} 1 \\ \frac{\alpha + \sqrt{\alpha^2 - 4}}{2} \end{pmatrix}}$$

is an eigenvector for λ_2. At this point one needs to be careful about different regions for α.

- If $|\alpha| > 2$, v_1 and v_2 will be two distinct vectors in $\mathbb{R}^2$; since they are associated to distinct eigenvalues they will be automatically linearly independent.

- If $|\alpha| < 2$, v_1 and v_2 will have complex entries, so they will not be vectors in $\mathbb{R}^2$. Consequently even though there are complex eigenspaces, there are no real eigenspaces in this case.

- If $\alpha = 2$, so that $\lambda_1 = \lambda_2 = 1$, then note that

$$\mathrm{v}_1 = \mathrm{v}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

This is inconclusive as there may be yet another eigenvector linearly independent from v_1. Thus we need to look into the nullspace of $I - A_2$. Since we have the row reduction

$$I - A_2 = \begin{pmatrix} -1 & 1 \\ -1 & 1 \end{pmatrix} \to \begin{pmatrix} -1 & 1 \\ 0 & 0 \end{pmatrix}$$

we can conclude in this case that v_1 is a basis for the eigenspace associated to the eigenvalue 1 and the geometric multiplicity of 1 is 1.

- Finally if $\alpha = -2$, so that $\lambda_1 = \lambda_2 = -1$, we have

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

Again this is inconclusive and we need to look at the nullspace of $-I - A_{-2}$. Since we have the row reduction

$$-I - A_{-2} = \begin{pmatrix} -1 & 1 \\ -1 & 1 \end{pmatrix} \to \begin{pmatrix} -1 & 1 \\ 0 & 0 \end{pmatrix}$$

as before, we again conclude in this case that v_1 is a basis for the eigenspace associated to the eigenvalue -1 and the geometric multiplicity of -1 is 1.

For part (d), we should consider cases:

- For $|\alpha| > 2$, the eigenspace associated to λ_1 is the span of

$$v_1 = \begin{pmatrix} 1 \\ \dfrac{\alpha - \sqrt{\alpha^2 - 4}}{2} \end{pmatrix};$$

consequently the eigenspace associated to λ_1 is the line passing through the origin with slope $\frac{\alpha - \sqrt{\alpha^2 - 4}}{2} = \lambda_2$, and the eigenspace associated to λ_2 is the span of

$$v_2 = \begin{pmatrix} 1 \\ \frac{\alpha + \sqrt{\alpha^2 - 4}}{2} \end{pmatrix}$$

and consequently this eigenspace is the line passing through the origin with slope $\frac{\alpha + \sqrt{\alpha^2 - 4}}{2} = \lambda_1$.

- For $|\alpha| < 2$, as the eigenvectors are complex valued, there is nothing to sketch.
- Finally for $|\alpha| = 2$, we have, based on our answers to part (c), that the eigenspace associated to λ is the span of

$$v_1 = v_2 = \begin{pmatrix} 1 \\ \text{sign}(\alpha) \end{pmatrix},$$

where sign(α) is 1 if $\alpha > 0$ and -1 otherwise.

For part (e), based on our answers to parts (a) and (c), the algebraic multiplicities of the eigenvalues will be 1 if $|\alpha| \neq 2$ and 2 otherwise, whereas the geometric multiplicities of the eigenvalues will be 1 if $|\alpha| \geq 2$ and 0 otherwise.

For part (f), based on our answers to part (c), exactly for $|\alpha| > 2$ is A_α diagonalizable[4].

Finally for part (g), for $|\alpha| > 2$, putting

[4] In this class by "diagonalizability" we really mean "diagonalizability using real valued matrices"; if complex valued matrices were allowed the case $|\alpha| < 2$ too would be when A_α is diagonalizable.

$$D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = \begin{pmatrix} \frac{\alpha + \sqrt{\alpha^2 - 4}}{2} & 0 \\ 0 & \frac{\alpha - \sqrt{\alpha^2 - 4}}{2} \end{pmatrix}$$

$$P = \begin{pmatrix} v_1 & v_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ \frac{\alpha - \sqrt{\alpha^2 - 4}}{2} & \frac{\alpha + \sqrt{\alpha^2 - 4}}{2} \end{pmatrix}$$

we have the diagonalization

$$A = PDP^{-1}.$$

9 PSet 10, Problem 3

Consider the matrix $A = \begin{pmatrix} 4 & 4 \\ -3 & 3 \end{pmatrix}$.

1. Compute the singular value decomposition of A.
2. Prof. Gil Strang works through this calculation in the video

   https://www.youtube.com/watch?v=TX_vooSnhm8&t=710s.

   He realizes, however, that his calculation turns out to be incorrect. Find the cause of his calculation error and explain it.

First note that while the question asks for "the" singular value decomposition of A, and indeed this is in common parlance, it would be more accurate to talk about "a" singular value decomposition. Thus we need three matrices U, Σ, V such that

- U is an orthogonal matrix (the columns of U are left singular vectors of A),
- $\widetilde{\Sigma}$ is a diagonal matrix with positive numbers on the diagonal (the diagonal entries of $\widetilde{\Sigma}$ are the singular values of A),
- $\Sigma = \begin{pmatrix} \widetilde{\Sigma} & 0 \\ 0 & 0 \end{pmatrix}$ is a version of $\widetilde{\Sigma}$ padded with zero entries so that Σ has the same dimensions as A,
- V is an orthogonal matrix (the columns of V are the right singular vectors of A),

- $A = U\Sigma V^T$.

For the spectral decomposition of a symmetric matrix, it's good to first compute the eigenvalues and then the orthogonal passive transformation. Similarly, for the singular value decomposition, it's good to first compute the singular values of A, so that one first has the $\widetilde{\Sigma}$ and Σ matrices.

To compute the singular values of A, let's first compute A^TA, as the singular values of A are precisely the square roots of the positive eigenvalues of A^TA:

$$\begin{aligned} A^TA &= \begin{pmatrix} 4 & 4 \\ -3 & 3 \end{pmatrix}^T \begin{pmatrix} 4 & 4 \\ -3 & 3 \end{pmatrix} \\ &= \begin{pmatrix} 4 & -3 \\ 4 & 3 \end{pmatrix} \begin{pmatrix} 4 & 4 \\ -3 & 3 \end{pmatrix} \\ &= \begin{pmatrix} 25 & 7 \\ 7 & 25 \end{pmatrix} \end{aligned}$$

Next, we need the eigenvalues of A^TA. The standard calculation goes like this:

$$\begin{aligned} 0 &= \det(\lambda I - A^TA) \\ &= \begin{vmatrix} \lambda - 25 & -7 \\ -7 & \lambda - 25 \end{vmatrix} \\ &= (\lambda - 25)^2 - 7^2 \\ &= (\lambda - 25 - 7)(\lambda - 25 + 7) \\ &= (\lambda - 32)(\lambda - 18) \end{aligned}$$

Thus the eigenvalues of A^TA in descending order are:

$$\boxed{\lambda_1 = 32 > \lambda_2 = 18}.$$

Note that no matter what A we start with, A^TA will always be a symmetric, positive semidefinite matrix, and indeed both eigenvalues of A^TA are positive numbers, so that in this case not only is A^TA positive semidefinite, it is in fact positive definite.

Taking the square roots of the eigenvalues of A^TA, we obtain the singular values of A:

$$\boxed{\sigma_1 = 4\sqrt{2} > \sigma_2 = 3\sqrt{2}}.$$

Thus we have $\widetilde{\Sigma} = \begin{pmatrix} 4\sqrt{2} & 0 \\ 0 & 3\sqrt{2} \end{pmatrix}$. Since $\widetilde{\Sigma}$ is already has the same dimensions as A, there is no need for padding with zeros, so we also have

$$\boxed{\Sigma = \widetilde{\Sigma} = \begin{pmatrix} 4\sqrt{2} & 0 \\ 0 & 3\sqrt{2} \end{pmatrix}}.$$

Next we'll need eigenvectors for A^TA; these we'll later put into columns of V and they will be right singular vectors of A. For $\lambda_1 = 32$, we need a normalized vector in the nullspace of the following matrix:

$$\begin{aligned} \lambda_1 I - A^TA &= 32I - A^TA \\ &= \begin{pmatrix} 32-25 & -7 \\ -7 & 32-25 \end{pmatrix} \\ &= \begin{pmatrix} 7 & -7 \\ -7 & 7 \end{pmatrix}. \end{aligned}$$

From this we have that the nullspace of this matrix consists precisely of vectors (a, b) with $a = b$, a normalized eigenvector thus is

$$\boxed{v_1 = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}}.$$

Similarly for $\lambda_2 = 18$ we need an eigenvector. Since by the spectral theorem eigenspaces of the symmetric matrix A^TA are orthogonal, and since the eigenspace for λ_1 is the diagonal line $y = x$, we have immediately that a normalized eigenvector for λ_2 is

$$\boxed{v_2 = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix}}.$$

(Alternatively one can of course go through the standard calculations to obtain v_2.)

Considering v_1 and v_2 as the columns of a passive transformation, we obtain the matrix V:

$$\boxed{V = \begin{pmatrix}v_1 & v_2\end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\1 & -1\end{pmatrix}}.$$

(†) Finally we need to compute the matrix U whose columns are the left singular vectors of A. For this, one can compute first the eigenvectors of AA^T, but that does not guarantee the compatibility of V and U implied by the singular value decomposition. Thus instead we reverse engineer the end result to compute a U that would work. From desired factorization $A = U\Sigma V^T$, sending V^T to the other side we obtain

$$AV = U\Sigma.$$

Say the columns of U are u_1 and u_2, so that $U = \begin{pmatrix}u_1 & u_2\end{pmatrix}$. Then we have

$$\begin{pmatrix} Av_1 & Av_2 \end{pmatrix} = AV = U\Sigma = \begin{pmatrix} \sigma_1 u_1 & \sigma_2 u_2 \end{pmatrix}.$$

Thus we need to compute $u_1 = \dfrac{Av_1}{\sigma_1}$ and $u_2 = \dfrac{Av_2}{\sigma_2}$; for U with these vectors as the column vectors, the factorization $A = U\Sigma V^T$ will be ensured, and further by the theory behind SVD U is guaranteed to be an orthogonal matrix. The straightforward calculation goes as follows:

$$u_1 = \frac{Av_1}{\sigma_1} = \frac{1}{4\sqrt{2}}\frac{1}{\sqrt{2}}\begin{pmatrix} 8 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix};$$
$$u_2 = \frac{Av_2}{\sigma_2} = \frac{1}{3\sqrt{2}}\frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ -6 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \end{pmatrix}.$$

Thus we have

$$\boxed{u_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad u_2 = \begin{pmatrix} 0 \\ -1 \end{pmatrix}}$$

and consequently

$$\boxed{U = \begin{pmatrix} u_1 & u_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}}.$$

Thus we have computed the following singular value decomposition for A:

$$\boxed{A = U\Sigma V^{T} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 4\sqrt{2} & 0 \\ 0 & 3\sqrt{2} \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix}}.$$

Let us note that it was coincidental in this case that U and V turned out to be symmetric; in general they will not be symmetric.

Finally for the second part of the problem, the answer is essentially the paragraph above marked with (†). More generally, the orthonormal bases for each eigenspace of AA^T ought to be chosen according to the orthonormal bases for each eigenspace of A^TA already constructed[5]. In this case each eigenspace of A^TA (and likewise AA^T) has one dimension, and a one dimensional vector space has exactly two orthonormal bases (either w or –w for a normalized vector w), corresponding to the fact that the only 1×1 matrices that are orthogonal are $\begin{pmatrix} 1 \end{pmatrix}$ and $\begin{pmatrix} -1 \end{pmatrix}$. If A^TA had an eigenvalue of higher multiplicity, then one could for instance rotate an orthonormal basis of said eigenvalue to obtain a different V, and the corresponding parts of U too would need to be rotated.

[5]In the above algorithm we first considered A^TA; had we considered AA^T first, which we could have done, then the eigenvectors of A^TA would have been constrained instead. The point is that U and V ought to be compatible so that $AV = U\Sigma$ holds.

Follow-up Exercise For $A = \begin{pmatrix} 4 & 4 \\ -3 & 3 \end{pmatrix}$, list each distinct SVD for A. For the purposes of this question, two SVDs $A = U\Sigma V^T$ and $A = U'\Sigma'(V')^T$ are distinct if at least one entry of one factor is different.

Follow-up Exercise 2 In the solution above, we used a certain algorithm. Write this algorithm explicitly.

Follow-up Exercise 3 Write all algorithms that one could use to compute the SVD of a matrix.

Follow-up Exercise 4

1. Construct a 4×5 matrix A of rank 3 such that it has exactly two distinct singular values, and no entry of A is zero.
2. Give explicit formulas for U, Σ, V such that
   (a) The columns of U are orthonormal,
   (b) The columns of V are orthonormal,
   (c) Σ is the invertible diagonal matrix $\widetilde{\Sigma}$ with singular values of A in the diagonal, padded appropriately with zeros,
   (d) $A = U\Sigma V^T$.